

Universidade de Brasília
Instituto de Física

Tese de Doutorado

**Inferência estatística no domínio de Fourier
para o estudo da dinâmica da convergência de
processos difusivos anômalos**

por

Raul Yukihiro Matsushita

Brasília, 3 de agosto de 2012

Inferência estatística no domínio de Fourier para o estudo da dinâmica da convergência de processos difusivos anômalos

por

Raul Yukihiro Matsushita

Tese apresentada ao Instituto de Física da Universidade
de Brasília para obtenção do título de Doutor em Física.
Área de concentração: Física Estatística

Orientador: Prof. Dr. Annibal Dias de Figueiredo Neto

Brasília, 3 de agosto de 2012

Tese submetida ao Instituto de Física da Universidade de Brasília como parte dos requisitos para a obtenção do grau de Doutor em Física.

Aprovada por:

Prof. Annibal Dias de Figueiredo Neto
Orientador, IF/UnB

Prof. Tarcísio Marciano da Rocha Filho
IF/UnB

Prof. Fábio Macêdo Mendes
Gama/UnB

Prof. Pushpa Narayan Rathie
EST/UnB

Prof. Iram Marcelo Gléria
IF/UFAL



“ $P(X_t = 0, Y_t = 0, 2Y_{t-1}) = 0,01,$

$P(X_t = 0,85X_{t-1} + 0,04Y_{t-1}, Y_t = -0,04X_{t-1} + 0,85Y_{t-1} + 1) = 0,85,$

$P(X_t = 0,20X_{t-1} - 0,26Y_{t-1}, Y_t = 0,26X_{t-1} + 0,22Y_{t-1} + 1.8) = 0,07$ e

$P(X_t = -0,15X_{t-1} + 0,28Y_{t-1}, Y_t = 0,26X_{t-1} + 0,24Y_{t-1} + 1) = 0,07,$

em que $X_0 = Y_0 = 1$ e $t \geq 1.$ ”

(Michael Barnsley)

Para Leo, Aline e Miwa,
Yasuko e Masaru

Meus sinceros agradecimentos aos professores do Instituto de Física da UnB, em especial, a Annibal Dias de Figueiredo Neto, Marco Antônio Amato, Tarcísio Marciano da Rocha Filho, Antony Marco Mota Polito, Amilcar Rabelo de Queiroz, Fábio Macêdo Mendes e Ademir Eugênio de Santana; aos prezados professores Iram Marcelo Gléria (Instituto de Física, UFAL), Eraldo Sérgio Barbosa Da Silva (Departamento de Economia, UFSC) e Pushpa Narayan Rathie (Departamento de Estatística, UnB); aos professores do Departamento de Estatística da UnB; e aos colegas André Telles, Regina Fonseca e Márcio de Castro.

Resumo

Sistemas complexos sob regime difusivo anômalo podem ser descritos por distribuições truncadas de Lévy. Problemas de inferência estatística nesse ambiente não gaussiano podem ser abordados via transformadas de Fourier, como as funções características. Este trabalho apresenta uma expansão alternativa da função característica que se mostrou útil para a estimação por máxima verossimilhança dos parâmetros das distribuições sob a hipótese de estabilidade. Para ilustrar, consideramos as séries temporais do índice da Bolsa de Valores de São Paulo, do índice Dow Jones Industrial Average da Bolsa de Valores de Nova Iorque (NYSE) — contemplando o evento denominado *flash crash* ocorrido em 6 de maio de 2010 —, das taxas de câmbio das principais moedas frente ao dólar norte americano, e dos preços de algumas ações negociadas na NYSE que sofreram *mini-flash crashes* em 2011. Em geral, esses dados podem ser modelados por distribuições truncadas, e a lentidão da convergência desses processos para a gaussiana se explica pela dependência serial de curto e de longo alcance. Observamos também que a função característica empírica sofre truncamento devido à finitude da amostra, havendo quebra de *scaling* sempre no mesmo patamar, independentemente da forma da distribuição dos dados. Finalmente, introduzimos um novo método assintótico que permite testar a hipótese de independência entre dois conjuntos de dados. Nosso teste é do tipo Cramér-von Mises, em que o processo empírico é obtido com base na divergência de Kullback-Leibler, e se mostrou estatisticamente poderoso para detectar dependência não linear fora do ambiente gaussiano.

Abstract

Complex systems under anomalous diffusive regime can be approximately described by truncated Lévy flights. Many difficult statistical issues in this non-Gaussian environment can be amenable to solution by the Fourier transform methods, as the characteristic functions. In this work, we put forward an alternative expansion of the characteristic function which proved useful for the maximum likelihood estimation of the parameters under the stability hypothesis. Our approach is exemplified with the Sao Paulo Stock Exchange index time series, the high-frequency data from the Dow Jones Industrial Average index — which encompass the recent episode known as the *flash crash* of May 6, 2010 —, the foreign exchange rate data, and the high-frequency data from stocks listed on the NYSE that recently experienced so-called *mini-flash crashes*. We confirm that the sluggish convergence of the truncated Lévy flights to a Gaussian can be explained by the presence of short range and long range serial dependence in these data. We also investigated the truncation phenomenon of the empirical characteristic function (ECF) due to the sample finitude. Regardless of the distribution shape, the ECF scaling breaks down always at the same level, depending only on the sample size. Finally, we devise a novel asymptotic statistical test to assess independence in bivariate data set. Our approach is based on the Cramér-von Mises test, and proved able to detect nonlinear dependence even if the environment is non-Gaussian.

Sumário

1	Introdução	1
1.1	Considerações iniciais	1
1.2	Uma breve retrospectiva	3
1.3	Objetivos	7
1.4	Dados	9
1.4.1	O IBovespa	9
1.4.2	O Índice DJIA	11
1.4.3	Preços de algumas ações negociadas na NYSE	11
1.4.4	Taxas de câmbio	14
1.5	Esboço do trabalho	19
2	A função característica e as distâncias entre distribuições	23
2.1	Introdução	23
2.2	A função característica	24
2.2.1	Propriedades	25
2.2.2	Uma expansão alternativa	26
2.2.3	Distribuições simétricas em torno de zero	28
2.2.4	Relações com respeito ao vetor de parâmetros	29
2.3	Distâncias entre duas distribuições	30
2.3.1	A distância L_2	30
2.3.2	A divergência de Kullback-Leibler	31

2.3.3	Expansão da função $\omega(q; \boldsymbol{\theta})$	32
2.3.4	Relação com a medida de informação de Fisher	35
2.4	Considerações	35
3	As distribuições infinitamente divisíveis e as estáveis	37
3.1	Introdução	37
3.2	Distribuições infinitamente divisíveis	38
3.3	O processo de Lévy	39
3.4	A distribuição estável	40
3.4.1	O processo de Lévy estável	45
3.5	O polinômio característico	45
3.6	Considerações	58
4	A função característica empírica	59
4.1	Introdução	59
4.2	Definição e algumas propriedades da FCE	60
4.3	O polinômio característico empírico	65
4.4	A FCE truncada	67
4.5	A FCE de uma soma de variáveis aleatórias	69
4.6	Estudo por simulações de Monte Carlo	70
4.7	Ilustração: dados do IBovespa	74
4.8	Considerações	77
5	Estimação por funções características	85
5.1	Introdução	85
5.2	O Método da Máxima Verossimilhança	87
5.3	A Equação de MV no Domínio de Fourier	88
5.4	Estudo com dados financeiros	92
5.4.1	IBovespa	92
5.4.2	Taxas de câmbio	94

5.4.3	Índice DJIA	100
5.5	Considerações	101
6	Teste de independência	111
6.1	Introdução	111
6.2	A FC multivariada e independência	113
6.3	O teste de independência	114
6.4	Valores críticos assintóticos	119
6.5	Validação e o poder do teste	120
6.6	Ilustrações	123
6.6.1	Ações da bolsa de Nova Iorque	123
6.6.2	Taxas de câmbio	124
6.7	Discussão	126
7	Considerações finais	129
7.1	Perspectivas para trabalhos futuros	133
7.1.1	Representação em séries	133
7.1.2	Estudo da origem do agrupamento de volatilidades e das correlações de longo alcance	135
7.1.3	A FCE multivariada e outras medidas de independência	137
7.1.4	Extensões do teste de independência	138
A	<i>Addendum</i> matemático	141
A.1	Integrais	141
A.2	Função gama	142
A.3	Derivadas da função delta	143
A.4	Coefficientes binomiais	143
B	Uma relação binomial da função escore	145
C	A fórmula de Lévy-Khinchine	149

D	A informação de Fisher	153
E	Simulação de uma VA estável	155
F	Detalhamento do Cap. 6	157
	F.1 Forma geral da função característica de B	157
	F.2 Solução do problema de autovalores	159

Abreviações e Siglas

AAS	amostra aleatória simples
Cov	covariância
CV	validação cruzada
DJIA	(índice) <i>Dow Jones Industrial Average</i>
FC	função característica
FCE	função característica empírica
FD	função de densidade
FDA	função de distribuição acumulada
H_0	hipótese nula
H_1	hipótese alternativa
HBKR	(teste de independência de) Hoeffding, Blum, Kiefer e Rosenblatt
IBovespa	índice da Bolsa de Valores de São Paulo
IID	independentes e identicamente distribuídas
MMV	método da máxima verossimilhança
MSE	<i>mean squared errors</i>
MV	máxima verossimilhança
NYSE	<i>New York Stock Exchange</i>
<i>P-value</i>	nível descritivo de um teste de hipóteses
TLF	<i>truncated Lévy flights</i>
VA	variável aleatória
Var	variância
ns	nível de significância

Lista de Símbolos e Notações

b_k	coeficiente, Eq. (2.30)
c_k	coeficiente, Eq. (2.12)
d_k	coeficiente, Eq. (5.17)
\hat{c}_k	estimativa de c_k
$f(x; \boldsymbol{\theta})$	função de densidade de probabilidade no ponto x parametrizada por $\boldsymbol{\theta}$
f	$= f(0; \boldsymbol{\theta})$
$f^{(k)}(x; \boldsymbol{\theta})$	$= \frac{d^k}{dx^k} f(x; \boldsymbol{\theta})$
$f^{(0)}(x; \boldsymbol{\theta})$	$= f(x; \boldsymbol{\theta})$
$f^{(k)}$	$= f^{(k)}(0; \boldsymbol{\theta}) = \frac{d^k}{dx^k} f(x; \boldsymbol{\theta}) \Big _{x=0}$
$\hat{f}^{(k)}$	estimativa de $f^{(k)}$
$g_j(x; \boldsymbol{\theta})$	$= \frac{d}{d\theta_j} f(x; \boldsymbol{\theta})$, em que $\theta_j \in \boldsymbol{\theta}$
$g_j^{(k)}$	$= g_j^{(k)}(0; \boldsymbol{\theta})$
$h_j(q; \boldsymbol{\theta})$	$= \frac{d}{d\theta_j} \phi_X(q; \boldsymbol{\theta})$
$h_{j,2}(q; \boldsymbol{\theta})$	parte real de $h_j(q; \boldsymbol{\theta})$
$h_{j,1}(q; \boldsymbol{\theta})$	parte imaginária de $h_j(q; \boldsymbol{\theta})$
i	unidade imaginária, $i^2 = -1$
i, j, k, l, s, t	índices
n	tamanho de uma amostra
p	dimensão de $\boldsymbol{\theta}$
q	ponto do suporte de uma FC
$s_j(x; \boldsymbol{\theta})$	função escore, (2.41)

(continua)

(continuação da lista de símbolos e notações)

u, v, x, y, z	possíveis realizações de variáveis aleatórias
$w_{\Delta t}(q)$	ver Eq. (1.7)
\bar{z}	complexo conjugado de um número $z \in \mathbb{C}$
z_{ns}	quantil relativo ao nível de significância ns, i.e., $\text{ns} = \text{P}(Z > z_{\text{ns}})$
$A_{1,k}$	ver Eq. (3.49)
$A_{2,k}$	ver Eq. (3.50)
$B_{2,0}$	fator de inflação, Eq. (3.60)
B	estatística do teste de independência
C^+ e C^-	coeficientes de assimetria
D_q^k	operador diferencial com respeito a q , $D_q^k = \frac{d^k}{dq^k}$
D_{KL}	distância de Kullback-Leibler, Eq. (2.37)
$F(x; \boldsymbol{\theta})$	função de distribuição acumulada
$\hat{F}(x)$	estimativa da função de distribuição acumulada
$F_n(x)$	função de distribuição acumulada empírica
H	expoente de Hurst
$H(\cdot)$	entropia
$I(\cdot)$	função indicadora
$\text{IF}(\boldsymbol{\theta})$	medida de informação de Fisher
$K(u)$	função Kernel
L_2	distância L_2 , Eq. (2.35)
$M(u)$	Eq. (3.3)
$\text{P}(\cdot)$	probabilidade
Q	intervalo ou faixa de operação da FCE
$R_j(q; \boldsymbol{\theta})$	ver Eq.(2.54)
R_t	$= \ln(W_t) - \ln(W_{t-1})$, retorno logarítmico no instante t

(continua)

(continuação da lista de símbolos e notações)

$S_{\Delta t}$	$= X_1 + \dots + X_{\Delta t}$, passeio aleatório
$S_{\Delta t}^R$	passeio aleatorizado Eq. (4.48)
$S_{\Delta t}^0$	passeio não aleatorizado Eq. (4.49)
W_t	valor de um índice (ou indicador ou preço de um ativo) no instante t
X_t	$= R_t - \mu$, retorno centrado
X, Y, Z	variáveis aleatórias (letras maiúsculas)
α	índice de estabilidade, $0 < \alpha \leq 2$
β	parâmetro de assimetria, $ \beta \leq 1$
β_α	$= \beta \tan \frac{\pi\alpha}{2}$ (assimetria efetiva)
$\delta(\cdot)$	função delta de Dirac
$\phi(q; \boldsymbol{\theta})$	função característica (FC), Eq. (2.1)
$\hat{\phi}(q)$	função característica empírica (FCE)
$\phi_2(q; \boldsymbol{\theta})$	parte real da FC
$\phi_1(q; \boldsymbol{\theta})$	parte imaginária da FC
$\phi_{0,n}$	ponto crítico da FCE
$\hat{\phi}_2(q)$	parte real da FCE
$\hat{\phi}_1(q)$	parte imaginária da FCE
$\hat{\phi}_*(q)$	FCE truncada
$\phi^{(k)}(q; \boldsymbol{\theta})$	k -ésima derivada da FC com respeito a q
$\varphi(q; \boldsymbol{\theta})$	$= \ln \phi(q; \boldsymbol{\theta})$
γ	parâmetro de escala, $\gamma > 0$
γ_*	$= \gamma \cdot \sqrt{1 + \beta_\alpha^2}$
$\gamma_{*,0}$	parâmetro de escala inflacionado, Eq. (3.60)
η	constante de normalização
λ_j	j -ésimo autovalor

(continua)

(continuação da lista de símbolos e notações)

μ	parâmetro de locação, $\mu \in \mathbb{R}$
$\boldsymbol{\theta}$	vetor de parâmetros, $\boldsymbol{\theta} \in \mathbb{R}^p$
$\hat{\boldsymbol{\theta}}$	estimativa de $\boldsymbol{\theta}$
θ_j	elemento do vetor $\boldsymbol{\theta}$
$\vartheta(u)$	medida de Lévy, Eq. (3.3)
ρ	correlação linear de Pearson
σ	$= \gamma^{\frac{1}{\alpha}}$
$\omega_j(q; \boldsymbol{\theta})$	transformada inversa do escore, (2.44)
ζ	ponto de truncamento da TLF
Δt	tamanho do passeio aleatório
$\Delta(q; \boldsymbol{\theta})$	Eqs. (4.17) e (4.18)
$\Phi(D_q; \boldsymbol{\theta})$	polinômio característico, Eq. (2.16)
$\Phi_2(D_q; \boldsymbol{\theta})$	parte real do polinômio característico
$\Phi_1(D_q; \boldsymbol{\theta})$	parte imaginária do polinômio característico
$\hat{\Phi}(D_q)$	polinômio característico empírico
$\Gamma(\cdot)$	função gamma
Σ	matriz de covariâncias
$\Psi(D_q; \boldsymbol{\theta})$	Eq. (5.16)
$\langle X \rangle$	valor esperado de X
\square	sinaliza o final de um exemplo

Capítulo 1

Introdução

1.1 Considerações iniciais

Um sistema econômico pode ser considerado como um sistema complexo aberto, em que há inúmeras formas de interação entre seus componentes [17, 114]. A dinâmica que rege esse sistema ainda não é completamente conhecida, o que atrai muitos pesquisadores para o desafio de desvendá-la aos poucos mediante estudos empíricos. Em finanças, as regularidades estatísticas observadas empiricamente em séries temporais de retornos financeiros denominam-se *fatos estilizados* [38, 105, 119]. A não gaussianidade e a presença de agrupamentos de volatilidades, por exemplo, se encontram entre os fatos mais conhecidos. A partir dessas observações empíricas, modelos teóricos podem ser sugeridos para se descrever o comportamento desse sistema [17, 75, 114, 118].

Na década de 1960, B. Mandelbrot observou que as distribuições das variações de preços (como a do algodão) não se ajustavam a uma distribuição gaussiana, pois elas apresentavam excesso de curtose e caudas mais pesadas [69, 70, 71, 72]. Ele também notou que a distribuição X das variações diárias se relacionava com a das variações mensais mediante transformações de escala. Ou seja, X seguia aproximadamente uma lei de potência (*power law*) na forma $f(\gamma^{-1/\alpha}x) = \gamma^{1/\alpha}f(x)$, em que $f(x)$ representa a função de densidade, x é uma possível realização de X , $\gamma > 0$ é o parâmetro de escala e $0 < \alpha \leq 2$

é o índice de estabilidade. Assim, Mandelbrot observou que as distribuições infinitamente divisíveis e estáveis eram possíveis modelos candidatos para descrever suas descobertas.

Uma distribuição X é infinitamente divisível se, para qualquer $n \geq 1$, existir uma distribuição X_n tal que X é a convolução de n cópias independentes de X_n [47, 48, 57, 103, 104]. E, em particular, uma distribuição infinitamente divisível X é estável se — a menos de um parâmetro de locação $\mu \in \mathbb{R}$ e de escala $\gamma > 0$ — as propriedades distribucionais são preservadas após convoluções de cópias independentes de X [3, 64, 103, 104]. Por exemplo, se X_1 e X_2 são cópias independentes de uma variável aleatória estável X , então X se relaciona com suas cópias X_1 e X_2 mediante uma convolução na forma $\gamma X + \mu = \gamma_1 X_1 + \gamma_2 X_2$, em que $\gamma_1, \gamma_2 > 0$ também são parâmetros de escala.

As caudas de uma distribuição estável seguem uma lei de potência na forma $f(|x|) \propto |x|^{-(\alpha+1)}$ ($0 < \alpha \leq 2$), e, além disso, $\langle |X|^q \rangle = \infty$, se $q \geq \alpha$, enquanto $\langle |X|^q \rangle < \infty$, se $q < \alpha$ [2, 48, 61, 103]. Assim, um fenômeno descrito por uma distribuição estável com $\alpha < 2$ não possui escala característica nem segundo momento; e, se $\alpha < 1$, tampouco a média existe. Desse modo, a teoria das distribuições estáveis, introduzida entre 1924 e 1936 por P. Lévy e A. Khinchine [64, 103, 104], remete naturalmente a um teorema limite central generalizado, já que uma distribuição estável se relaciona com uma soma de variáveis aleatórias independentes com variâncias não necessariamente finitas.

Apesar das descobertas de Mandelbrot, as distribuições estáveis de Lévy foram mantidas à margem da área principal em finanças [17, 105]. Entre as possíveis razões, a inexistência do desvio padrão como medida de volatilidade da distribuição é um inconveniente, pois ele representa uma medida de risco financeiro. Por exemplo, uma grande variação média de uma série de retornos em certo período de tempo indica maior exposição do investidor a perdas ou ganhos consideráveis.

Em meados da última década do séc. XX, porém, R. Mantegna e H. Stanley [73, 75] propuseram uma nova perspectiva para o estudo dos fenômenos financeiros. Eles observaram que as leis de escala no comportamento dos retornos do índice Standard & Poor's 500 da Bolsa de Valores de Nova Iorque (NYSE) eram compatíveis com as propriedades

de uma distribuição de Lévy simétrica. A novidade, no entanto, foi a observação de quebras nas leis de escala sugeridas por Mandelbrot, de modo que esses dados não poderiam ser de fato estáveis nem possuir momentos infinitos. Assim, esses autores propuseram a distribuição de Lévy truncada (*truncated Levy flights*, TLF) para contemplar esse novo fato estilizado. Trabalhos subsequentes mostraram resultados similares em diversas outras séries financeiras, como as do índice da Bolsa de Valores de São Paulo [46], as dos índices de outras bolsa de valores [45, 88, 109] e as das taxas de câmbio [30, 90].

1.2 Uma breve retrospectiva

Considere o passeio aleatório

$$S_{\Delta t} = X_1 + \cdots + X_{\Delta t}, \quad (1.1)$$

em que $\{X_k\}_{k=1, \dots, \Delta t}$ constitui uma amostra aleatória retirada de uma distribuição X , estável e simétrica em torno de zero, cujos parâmetros são representados pelo vetor $\boldsymbol{\theta}_s = (\alpha, \gamma)'$. Nessa situação, a função de densidade de $S_{\Delta t}$ no ponto $u \in \mathbb{R}$ é [2, 30, 75, 103]

$$f_{S_{\Delta t}}(u; \boldsymbol{\theta}_s) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma \Delta t q^\alpha} \cos(qu) dq, \quad (1.2)$$

e sua função característica no ponto $q \in \mathbb{R}$ é

$$\phi_{S_{\Delta t}}(q; \boldsymbol{\theta}_s) = \langle e^{iqS_{\Delta t}} \rangle = e^{-\gamma \Delta t q^\alpha}. \quad (1.3)$$

Pela estabilidade, se $\Delta t = 1$, as expressões acima representam a distribuição de X . Agora, considere uma soma de variáveis aleatórias não necessariamente independentes

$$S'_{\Delta t} = X'_1 + \cdots + X'_{\Delta t}, \quad (1.4)$$

em que cada X'_k segue uma distribuição abruptamente truncada (TLF) na forma [73, 75]

$$f_{X'_k}(u; \boldsymbol{\theta}_{TLF}) = \eta \mathbf{I}(|u| \leq \zeta) f_X(u; \boldsymbol{\theta}_s), \quad (1.5)$$

com $\boldsymbol{\theta}_{TLF} = (\alpha, \gamma, \eta, \zeta)'$, $\eta > 0$ é o parâmetro de normalização, $\zeta > 0$ é o ponto de truncamento, e $\mathbf{I}(|u| \leq \zeta) = 1$, se $|u| \leq \zeta$, e $\mathbf{I}(|u| \leq \zeta) = 0$, se $|u| > \zeta$. Para $\Delta t \rightarrow 1$,

embora seja truncado, o processo $S'_{\Delta t}$ pode ser aproximadamente descrito pelas formas estáveis correspondentes (1.2) ou (1.3). Porém, espera-se que $S'_{\Delta t}$ se aproxime de uma distribuição gaussiana à medida que Δt aumenta, já que a TLF não é estável e possui momentos finitos [30]. Considere então a variável reduzida (ou padronizada)

$$\bar{S}'_{\Delta t} = \frac{S'_{\Delta t} - \langle S'_{\Delta t} \rangle}{\sigma_{\Delta t}}, \quad (1.6)$$

em que $\sigma_{\Delta t}^2 = \langle (S'_{\Delta t})^2 \rangle - \langle S'_{\Delta t} \rangle^2$ representa a variância do processo truncado $S'_{\Delta t}$. Nesse caso, a função característica da variável reduzida pode ser representada na forma [30, 31, 64]

$$\phi_{\bar{S}'_{\Delta t}}(q) = e^{-q^2(1+w_{\Delta t}(q))/2}, \quad (1.7)$$

em que $w_{\Delta t}(q)$ é uma função tal que $w(0) = 0$.

Se $\{X'_k\}$ for uma sequência de cópias independentes de uma distribuição com média μ e variância σ^2 , então $\langle S'_{\Delta t} \rangle = \Delta t \mu$ e $\sigma_{\Delta t}^2 = \Delta t \sigma^2$. Nesse caso,

$$\begin{aligned} \bar{S}'_{\Delta t} &= \frac{S'_{\Delta t} - \Delta t \mu}{\sqrt{\Delta t} \sigma} \\ &= \frac{1}{\sqrt{\Delta t}} \sum_{k=1}^{\Delta t} \bar{X}'_k. \end{aligned}$$

Assim,

$$\begin{aligned} \phi_{\bar{S}'_{\Delta t}}(q) &= \langle e^{iq\bar{S}'_{\Delta t}} \rangle = \langle e^{i\frac{q}{\sqrt{\Delta t}} \sum_{k=1}^{\Delta t} \bar{X}'_k} \rangle = \langle e^{i\frac{q}{\sqrt{\Delta t}} \bar{X}'} \rangle^{\Delta t} \\ &= e^{-q^2(1+w_1(q/\Delta t))/2}, \end{aligned}$$

de modo que $\phi_{\bar{S}'_{\Delta t}}(q) \rightarrow e^{-q^2/2}$ à medida que $\Delta t \rightarrow \infty$, em que $e^{-q^2/2}$ representa a função característica da gaussiana padronizada.

Mas, se $\{X'_k\}$ não for uma sequência de variáveis aleatórias independentes, há redução na velocidade de convergência de $w_{\Delta t}(q)$ para zero à medida que Δt aumenta [30, 31, 32, 33, 34, 45]. Assim, enquanto houver memória serial significativa, o termo $w_{\Delta t}(q)$ pode variar lentamente, de modo que

$$w_{\Delta t}(q) \approx w(q) \quad (1.8)$$

para algum intervalo $\Delta t_0 \leq \Delta t \leq \Delta t_1$. Essa estabilidade momentânea foi denominada *quase-estabilidade* por A. Figueiredo e seus colaboradores [30, 31, 45], tendo sido observada empiricamente em séries dos retornos de taxas de câmbio e de bolsas de valores. Se os retornos X_t de determinado ativo financeiro fossem independentes e identicamente distribuídos (IID) segundo uma distribuição de Lévy simétrica truncada, pelo teorema limite central, as somas parciais desses retornos, $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$, deveriam convergir rapidamente para a gaussiana. Porém, havendo correlações, observou-se que há um intervalo $\Delta t_0 \leq \Delta t \leq \Delta t_1$ em que o processo $\{S_{\Delta t}\}$ é aproximadamente estável pela lentidão da convergência para a gaussiana. Mesmo que as autocorrelações lineares em uma série temporal financeira se encontrem no nível de ruído, formas não lineares de autocorrelação bem como tipos particulares de não estacionariedade também podem contribuir para a permanência de $S_{\Delta t}$ no regime de Lévy [32, 35].

Se houver quase-estabilidade, a região modal da densidade empírica pode ser aproximadamente descrita por uma distribuição estável. Fora da região modal, porém, o comportamento empírico das caudas pode se desviar do que se espera de uma distribuição estável. Assim, sob a hipótese de que os processos reais são limitados pela finitude dos recursos [73, 75], outras formas de truncamento da distribuição de Lévy podem ser sugeridas, como o truncamento suave [92], o gradual [50, 51] e o exponencialmente amortecido [84, 85, 45]. Essas modificações resultam em distribuições não estáveis com momentos finitos, e permitem explicar, por exemplo, a presença de *multiscaling* nos momentos absolutos das somas parciais $S_{\Delta t}$.

Como a lei de potência descoberta por Mandelbrot implica ausência de escala típica, naturalmente é possível associar o fenômeno em estudo à geometria fractal (posteriormente, Mandelbrot considerou os modelos multifractais para contemplar a dependência serial [71, 72]). Com respeito ao caos determinístico em séries temporais financeiras, se houver, não é fácil identificá-lo, possivelmente pela dificuldade de se distinguir os padrões caóticos dos estocásticos, ou simplesmente porque esses sistemas são de elevada complexidade [15, 44]. A taxa de câmbio da moeda chinesa frente ao dólar americano é um exemplo

à parte. Os retornos dessa taxa de câmbio apresentam uma estrutura fractal típica de um jogo caótico conhecido como triângulo de Sierpinski, em que as regras determinísticas coexistem com as estocásticas [81, 83, 113, 118]. Do ponto de vista estocástico, a dimensão fractal D de um processo se relaciona com a dependência de longo alcance medida com base no expoente H de Hurst [5, 16, 48, 72, 82], definido como

$$\sigma_{\Delta t} \sim \Delta t^H = \Delta t^{2-D}.$$

Para um passeio aleatório com incrementos independentes, tem-se $H = 0,5$. O expoente de Hurst pode ser estimado com base na estatística R/S (*rescaled range analysis*, [5, 14, 72]), no método DFA (*detrended fluctuation analysis*, [124]) ou DMA (*detrended moving average*, [12, 82, 124]). Em finanças, o expoente de Hurst — e conseqüentemente a dimensão fractal — permite avaliar a hipótese do mercado eficiente. Segundo essa hipótese, com base em um conjunto de informações publicamente disponíveis à comunidade financeira, um investidor não é capaz de obter, sistematicamente, rendimentos superiores à média do mercado [17], e, assim, $H = 0,5$. Estudos empíricos, no entanto, mostram resultados que enfraquecem essa hipótese, em que $H < 0,5$ [16, 17, 82].

Por exemplo, às vésperas de uma quebra na bolsas de valores (*crash*) ou de uma crise econômica em grande escala, os agentes que compram e vendem ativos podem seguir um comportamento coletivo em massa (o que, em parte, ajuda a explicar a presença de dependência serial no período que antecede uma crise). Em fenômenos de ruptura, D. Sornette observou um padrão log-periódico na forma $x_t \sim \cos \ln t$ [111], em que t é o tempo e x_t é uma variável do sistema. E assim, evidências de log-periodicidade foram encontradas em índices de bolsas de valores [110, 112], em taxa de câmbio [79] e no índice Dow Jones [78].

A eficiência de mercado também pode ser estudada sob a perspectiva da complexidade de Kolmogorov [10] — uma abordagem que permite quantificar uma informação contida em uma sequência de dígitos binários (*string*). Define-se a complexidade algorítmica de um *string* como o tamanho do menor algoritmo computacional necessário para gerar esse mesmo *string*. Assim, por exemplo, a complexidade de um *string* é máxima se o menor

algoritmo computacional disponível para gerá-lo é tão grande quanto ao próprio *string*. A diferença entre o tamanho de um *string* e o menor algoritmo possível representa o seu grau de compressibilidade. Desse modo, um *string* de baixa complexidade é altamente compressível, enquanto um *string* de dígitos binários aleatórios é incompressível. Essa abordagem permite, por exemplo, descrever e classificar os mercados com base no algoritmo de compressão de dados de Lempel-Ziv [16, 41, 42, 43].

Dada a abrangência do tema, este trabalho se restringe aos aspectos da inferência estatística via funções características, deixando à margem diversos assuntos como complexidade, criticalidade, dependência de longo alcance, caos e logperiodicidade.

1.3 Objetivos

Nos estudos anteriores, a convergência do processo $S'_{\Delta t}$ para a gaussiana foi avaliada com base no comportamento da função $w_{\Delta t}(q)$ (Eq. (1.8)). Considerando que a quase estabilidade remete aproximadamente ao regime de Lévy, que a distribuição para $\Delta t \rightarrow 1$ se assemelha a uma distribuição estável (pelo menos na região modal da distribuição), e que a distância entre o processo empírico e o hipotético pode ser medida com base nas funções características empírica e hipotética [26, 27, 95, 97, 117, 125, 120], este trabalho propõe um estudo da dinâmica da convergência dos processos sob a perspectiva das distribuições quase estáveis. Ou seja, no caso simétrico, por exemplo, em lugar da função característica (1.7), propõe-se que o processo $S'_{\Delta t}$ seja descrito pela função característica na forma

$$\phi_{S'_{\Delta t}}(q; \boldsymbol{\theta}_s) \approx \phi_{S_{\Delta t}}(q; \boldsymbol{\theta}_s) = \langle e^{iqS_{\Delta t}} \rangle = \exp(-\gamma_{\Delta t} \Delta t q^{\alpha_{\Delta t}}).$$

Assim, há estabilidade se $\alpha_{\Delta t} = \alpha$ for constante para todo Δt e $\gamma_{\Delta t} = \Delta t \gamma$. Porém, nas condições do teorema limite central, se não houver estabilidade, espera-se que $\alpha_{\Delta t} \rightarrow 2$ à medida que Δt aumenta. E, se $\alpha_{\Delta t} \approx \alpha$ em algum intervalo $\Delta t_0 \leq \Delta t \leq \Delta t_1$, então há quase estabilidade. E, ainda, o efeito da dependência temporal no parâmetro de escala pode ser avaliado com base no comportamento de $\gamma_{\Delta t}$ *versus* Δt . Essa abordagem requer estimação de $\alpha_{\Delta t}$ e de $\gamma_{\Delta t}$ para cada Δt desejado. O método da máxima verossimi-

lhança (MMV) fornece estimativas com boas propriedades estatísticas como consistência, eficiência e normalidade assintótica das distribuições amostrais [89, 98, 101]. Porém, o fato de a função de densidade (FD) da distribuição estável não possuir forma fechada para $\alpha \neq 1$ e 2 [97, 99] motivou a busca por diferentes outros métodos de estimação [21, 75, 103, 82, 87, 97], embora eles sejam menos eficientes do que o MMV. Ao contrário da FD, a função característica (FC) da distribuição estável possui forma fechada. Por causa da correspondência entre a FD e a FC, espera-se que seja possível obter estimativas de máxima verossimilhança (MV) com base em funções características [125]. As equações de verossimilhança — que formam o sistema de equações para a determinação das estimativas de MV — se relacionam com a divergência de Kullback-Leibler entre a distribuição empírica e a hipotética [26, 27].

Assim, o primeiro objetivo deste trabalho é desenvolver uma equação de verossimilhança com base em funções características, considerando-se as distribuições (aproximadamente) estáveis simétricas e as assimétricas.

Como essa inferência estatística depende da função característica empírica (FCE), o segundo objetivo trata do estudo do truncamento natural dessa função. Por causa da finitude do tamanho da amostra, estatisticamente, por exemplo, a hipótese $\phi_{S_{\Delta t}}(q; \boldsymbol{\theta}_s) = 0$ não poderia ser rejeitada caso sua estimativa ($\hat{\phi}(q)$) se encontre no nível de ruído.

Considerando que a dependência serial produz quase estabilidade no processo $S'_{\Delta t}$, o terceiro objetivo é propor um novo teste de hipóteses para a detecção de dependência não linear [80]. O coeficiente de correlação e a função de autocorrelação não são consistentes para os casos em que há dependência não linear nos dados. Para distribuições que não possuem momento finito ou que apresentam dependência não linear, espera-se que um teste elaborado com base na divergência de Kullback-Leibler forneça resultados consistentes [80, 89, 98, 101].

1.4 Dados

Para as ilustrações apresentadas neste trabalho, consideramos a série temporal do índice diário da Bolsa de Valores de São Paulo (IBovespa), a série *intraday* (minuto a minuto) do índice Dow Jones Industrial Average (DJIA) da Bolsa de Valores de Nova Iorque (NYSE), as séries *intraday* (minuto a minuto) dos preços das ações de algumas empresas negociadas na NYSE, e as das taxas diárias de câmbio de algumas moedas (Tab. 1.2) frente ao dólar americano. A seguir, descrevemos brevemente essas séries financeiras.

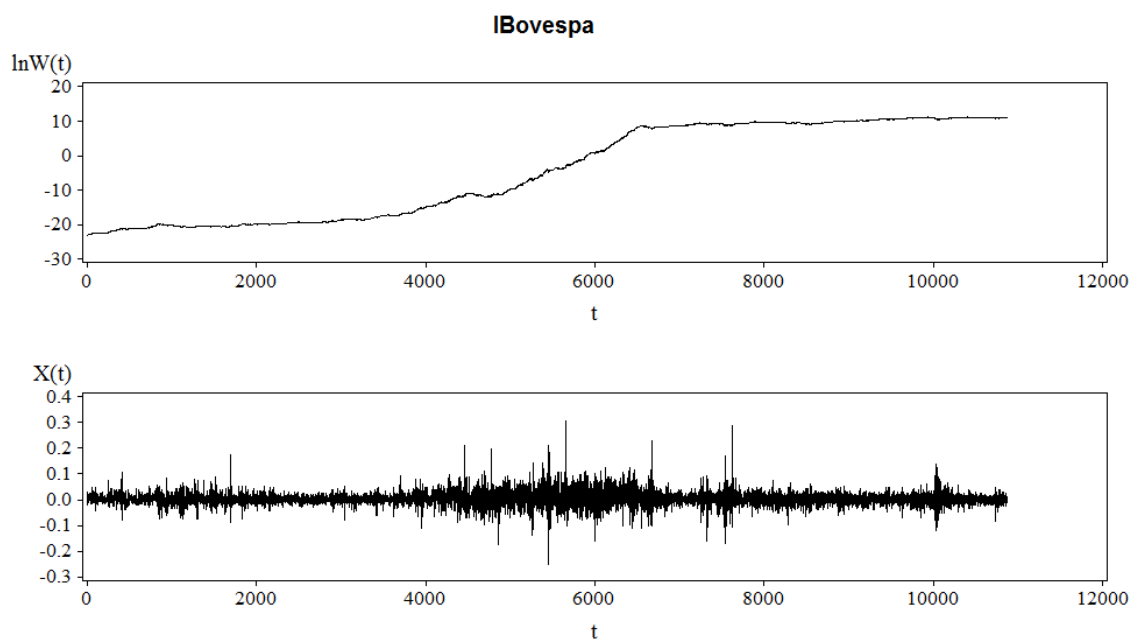


Figura 1.1: Observações diárias do logaritmo natural do IBovespa, $\ln W_t$ (painel superior), e seus retornos X_t (painel inferior), de 2 de janeiro de 1968 a 29 de fevereiro de 2012. O instante $t = 6.500$ corresponde a 04/07/1994, três dias após o dia em que o Plano Real entrou em vigor.

1.4.1 O IBovespa

O Índice da Bolsa de Valores de São Paulo — IBovespa — é um importante indicador do desempenho médio das cotações do mercado brasileiro de ações. Ele retrata o comportamento dos principais papéis negociados na BM&FBOVESPA, e sua metodologia de

cálculo se manteve a mesma desde sua implementação em 1968. O painel superior da Fig. 1.1 mostra a série histórica do logaritmo da pontuação de fechamento do Ibovespa de 2 de janeiro de 1968 a 29 de fevereiro de 2012, perfazendo o total de 10.870 observações. Considerando que W_t representa a pontuação do Ibovespa ao final do dia t (ignorando-se feriados e finais de semana), define-se o retorno logarítmico como

$$R_t = \ln(W_t) - \ln(W_{t-1}), \quad (1.9)$$

e o retorno centrado na média histórica dos retornos é dado por

$$X_t = R_t - \mu, \quad (1.10)$$

em que $\mu = \langle R_t \rangle$. O painel inferior da Fig. 1.1 mostra a evolução temporal da série dos retornos X_t .

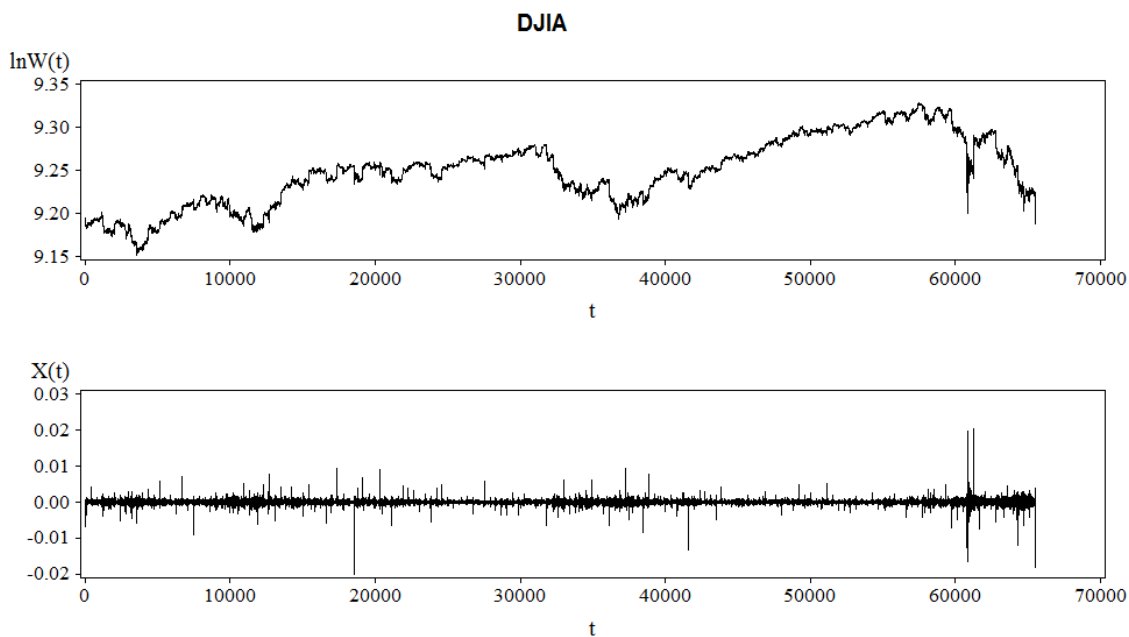


Figura 1.2: Observações minuto a minuto do logaritmo natural do DJIA, $\ln W_t$ (painel superior), e seus retornos X_t (painel inferior), de 15h09 do dia 18 de setembro de 2009 a 10h09 do dia 25 de maio de 2010. O *flash crash* ocorreu em 6 de maio de 2010 ($60.491 \leq t \leq 60.881$).

1.4.2 O Índice DJIA

O painel superior da Fig. 1.2 mostra a evolução minuto a minuto do logaritmo natural do índice DJIA (*Dow Jones Industrial Average*) da bolsa de valores de Nova Iorque, a partir de 15h09 do dia 18 de setembro de 2009 até 10h09 do dia 25 de maio de 2010, perfazendo o total de 65.535 observações. Nessa série temporal, um episódio conhecido como *flash crash* [78] marcou o dia 6 de maio de 2010 (na Fig. 1.2 esse dia corresponde ao intervalo $60.491 \leq t \leq 60.881$). Nessa quinta-feira negra, repentinamente, o índice sofreu uma queda abrupta de 998.5 pontos. A queda ocorreu principalmente entre 14h40 e 15h00, e nesse período o preço da ação da empresa de consultoria Accenture, por exemplo, despencou de US\$ 60,00 para US\$ 0,01. Essa quebra foi provocada por uma ordem de venda de contratos futuros feita por um operador que utilizou uma plataforma automatizada para suas negociações. De acordo com o órgão regulador Securities & Exchange Commission, essa ordem automatizada vendeu, em apenas 20 minutos, 75 mil contratos futuros E-mini do S&P 500, com valor estimado em US\$ 4,1 milhões. A rapidez da execução dessa ordem provou um choque no mercado, e o declínio que se seguiu nos índices de futuros alarmou os demais operadores. A fuga massiva desses operadores produziu a queda em poucos minutos (já que a ordem de proteção contra perdas na negociação de futuros também é automatizada).

1.4.3 Preços de algumas ações negociadas na NYSE

A Tab. 1.1 descreve as séries temporais minuto a minuto dos preços das ações de algumas empresas negociadas na Bolsa de Valores de Nova Iorque (NYSE) que experimentaram dias de extrema volatilidade (ou *mini flash crashes*) entre 2010 e 2011. Essas empresas sofreram quedas dramáticas e repentinas em um curto período de tempo. No dia 27/04/2011, o preço das ações da empresa Jazz Pharmaceuticals caiu de US\$ 33,59 para US\$ 23,50, mas fechou o dia em US\$ 32,93. Em 11 de maio de 2011, a RLJ Lodging Trust entrou na NYSE com uma oferta pública inicial (*initial public offering*) no valor de US\$ 17,25, mas um grande volume de negociações a US\$ 0,0001 em poucos segundos. Em 13

de maio de 2011, a seguradora Enstar viu suas ações despencarem de US\$ 100,00 para zero, e segundos depois, de zero para US\$ 100,00. Já os laboratórios Pfizer and Abbott experimentaram *mini flash crashes* na direção oposta. No dia 2 de maio de 2011, as ações da Abbott saltaram de US\$ 50,00 para US\$ 250,00, e as da Pfizer de US\$ 27,60 para US\$ 88,71, em menos de um segundo.

A Fig. 1.3 mostra a evolução temporal do logaritmo dos preços (painel superior) e dos retornos (painel inferior) das ações RLJ e ESGR, do dia 13 de junho de 2011 a 13 de outubro de 2011. Esse período não contempla os *mini flash crashes*, pois esses episódios geralmente resultam de erros nas transações, e, se forem detectados, tais operações devem ser canceladas. A rapidez dos mercados automatizados, no entanto, permite uma instantânea propagação de erros para o resto do sistema, produzindo falsos sinais de alerta.

Tabela 1.1: Descrição das séries de preços de ações

ação	sigla	data inicial:	data final:	data do	tamanho da
		13 Jun 2011	13 Out 2011	<i>flash crash</i>	amostra
Jazz Pharmaceuticals	JAZZ	10h30	11h03	27 Abr 2011	31.393
RLJ Lodging	RLJ	10h30	11h02	11 Mai 2011	19.932
Enstar	ESGR	10h30	10h53	13 Mai 2011	4.902
Pfizer	PFE	10h30	11h04	2 Mai 2011	33.748
Abbott Labs	ABT	10h30	11h04	2 Mai 2011	33.777
Progress Energy	PGN	10h30	11h04	27 Set 2010	33.179
Citigroup	C	10h30	11h03	29 Jun 2010	33.754
Washington Post Company	WPO	10h36	11h01	16 Jun 2010	11.946
Micron	MU	10h30	11h04	5 Ago 2010	33.660
Cisco Systems	CSCO	10h30	11h04	29 Jul 2010	33.661
Core Molding	CORE	10h30	10h47	26 Ago 2010	12.037

Fonte: Bloomberg

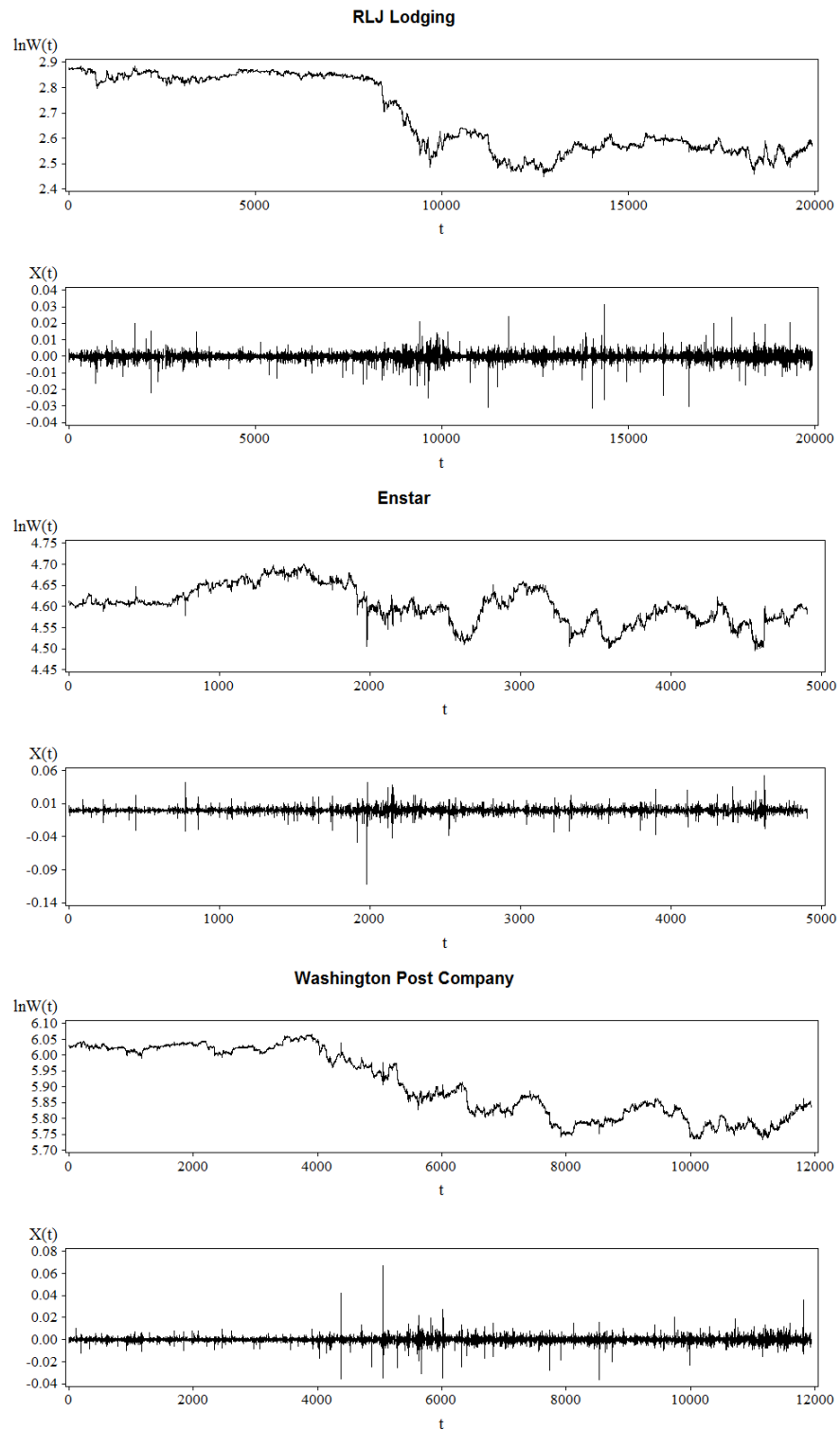


Figura 1.3: Evolução temporal *intraday* dos logaritmos dos preços $\ln W_t$ (painel superior) e dos retornos X_t (painel inferior) das ações das empresas RLJ Lodging, Enstar e Washington Post Company.

1.4.4 Taxas de câmbio

A Tab. 1.2 descreve as taxas de câmbio de algumas moedas, para compra ao meio-dia cotadas pelo Federal Reserve Bank of New York, expressas em unidades monetárias por dólar americano. Por exemplo, a taxa de câmbio R\$/US\$ para compra ao meio-dia em 16 de março de 2012 foi de R\$ 1,8025 por US\$ 1,00. As Figs. 1.4, 1.5, 1.6 e 1.7 mostram as evoluções temporais das taxas diárias de câmbio (painel superior) e de seus retornos (painel inferior) das moedas da África do Sul, Austrália, Brasil, Canadá, Índia, Japão, Suíça e Reino Unido com respeito ao dólar americano.

Tabela 1.2: Descrição das séries de taxas de câmbio

país	moeda	data inicial:	data final:	tamanho da amostra
África do Sul	Rand	4 jan 1971	16 mar 2012	10.316
Austrália	Dólar australiano	4 jan 1971	16 mar 2012	10.336
Brasil	Real	2 jan 1995	16 mar 2012	4.325
Canadá	Dólar canadense	4 jan 1971	16 mar 2012	10.349
Coréia do Sul	Won	13 abr 1981	16 mar 2012	7.727
Dinamarca	Coroa dinamarquesa	4 jan 1971	16 mar 2012	10.342
Índia	Rúpia	2 jan 1973	16 mar 2012	9.835
Japão	Yen	4 jan 1971	16 mar 2012	10.337
México	Peso	8 nov 1993	16 mar 2012	4.611
Nova Zelândia	Dólar neozelandês	4 jan 1971	16 mar 2012	10.327
Noruega	Coroa norueguesa	4 jan 1971	16 mar 2012	10.342
Reino Unido	Libra Esterlina	4 jan 1971	16 mar 2012	10.343
Singapura	Dólar de Singapura	2 jan 1981	16 mar 2012	7.842
Suécia	Coroca Sueca	4 jan 1971	16 mar 2012	10.342
Suíça	Franco Suíço	4 jan 1971	16 mar 2012	10.343
Tailândia	Baht	2 jan 1981	16 mar 2012	7.739
Taiwan	Dólar de Taiwan	30 out 1983	16 mar 2012	6.856
Zona do Euro	Euro	4 jan 1999	16 mar 2012	3.322

Fonte: Federal Reserve

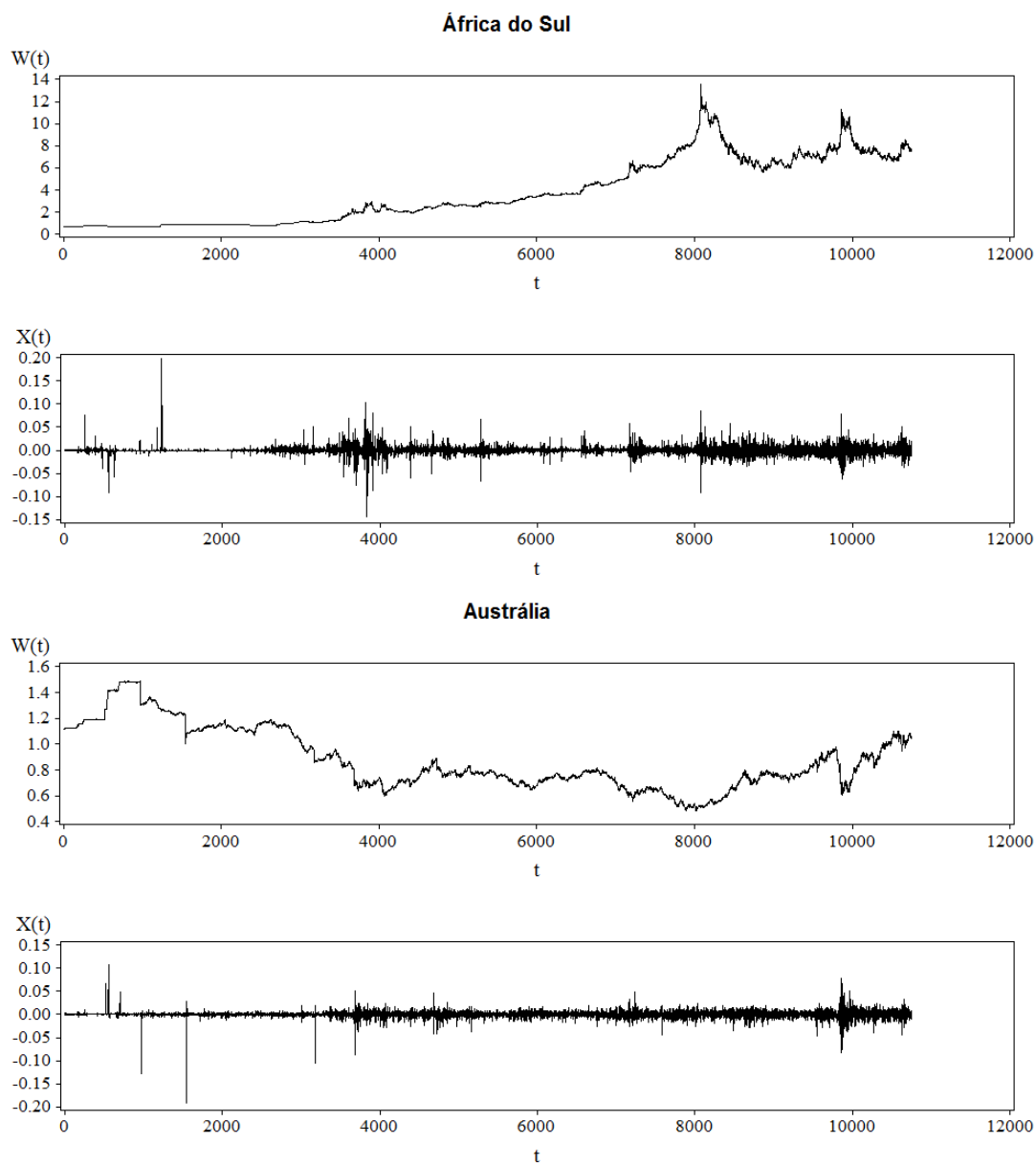


Figura 1.4: Evoluções diárias das taxas de câmbio W_t do rand e do dólar australiano frente ao dólar americano (painel superior), e dos respectivos retornos X_t (painel inferior).

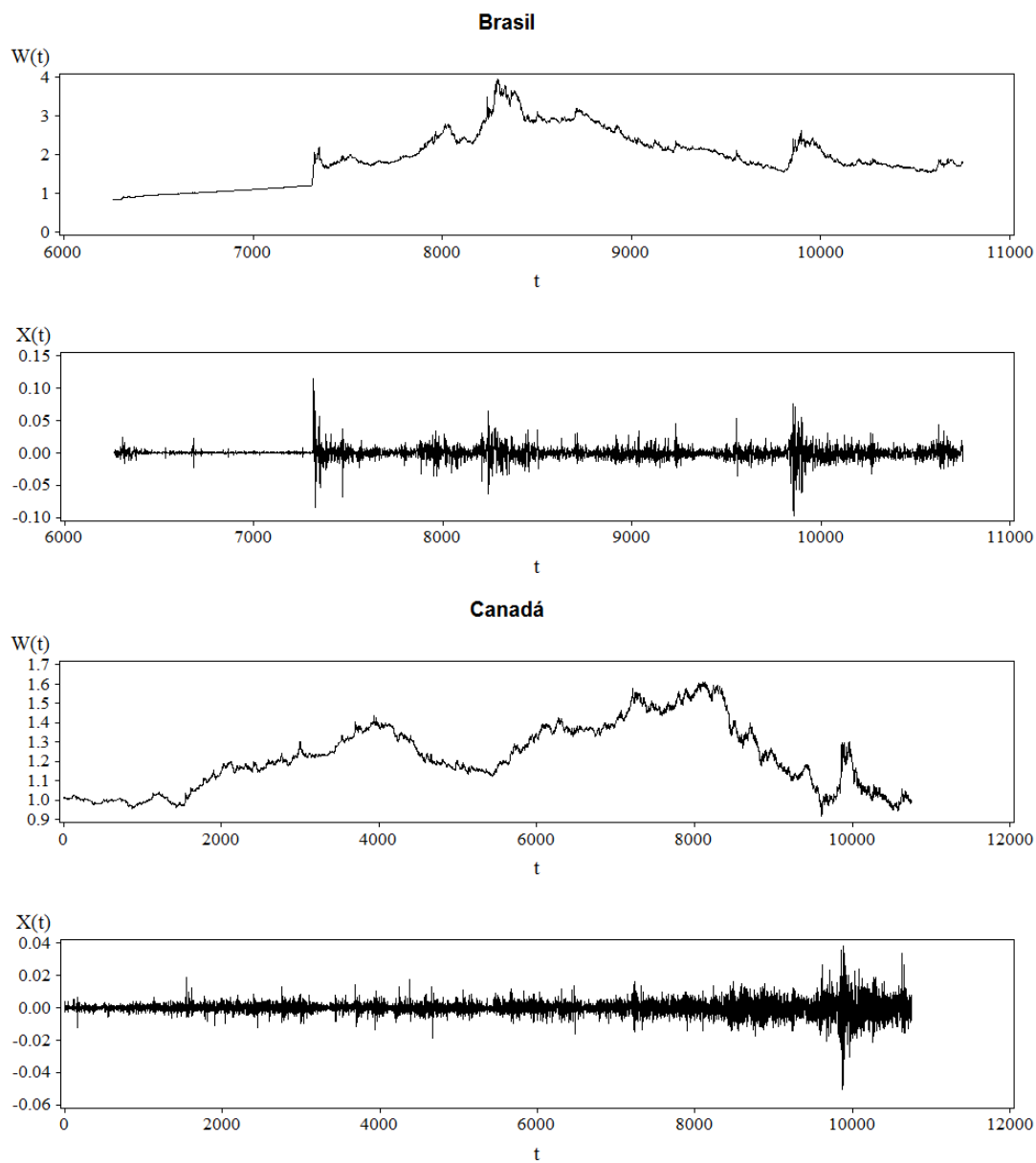


Figura 1.5: Evoluções diárias das taxas de câmbio W_t do real e do dólar canadense frente ao dólar americano (painel superior), e dos respectivos retornos X_t (painel inferior).

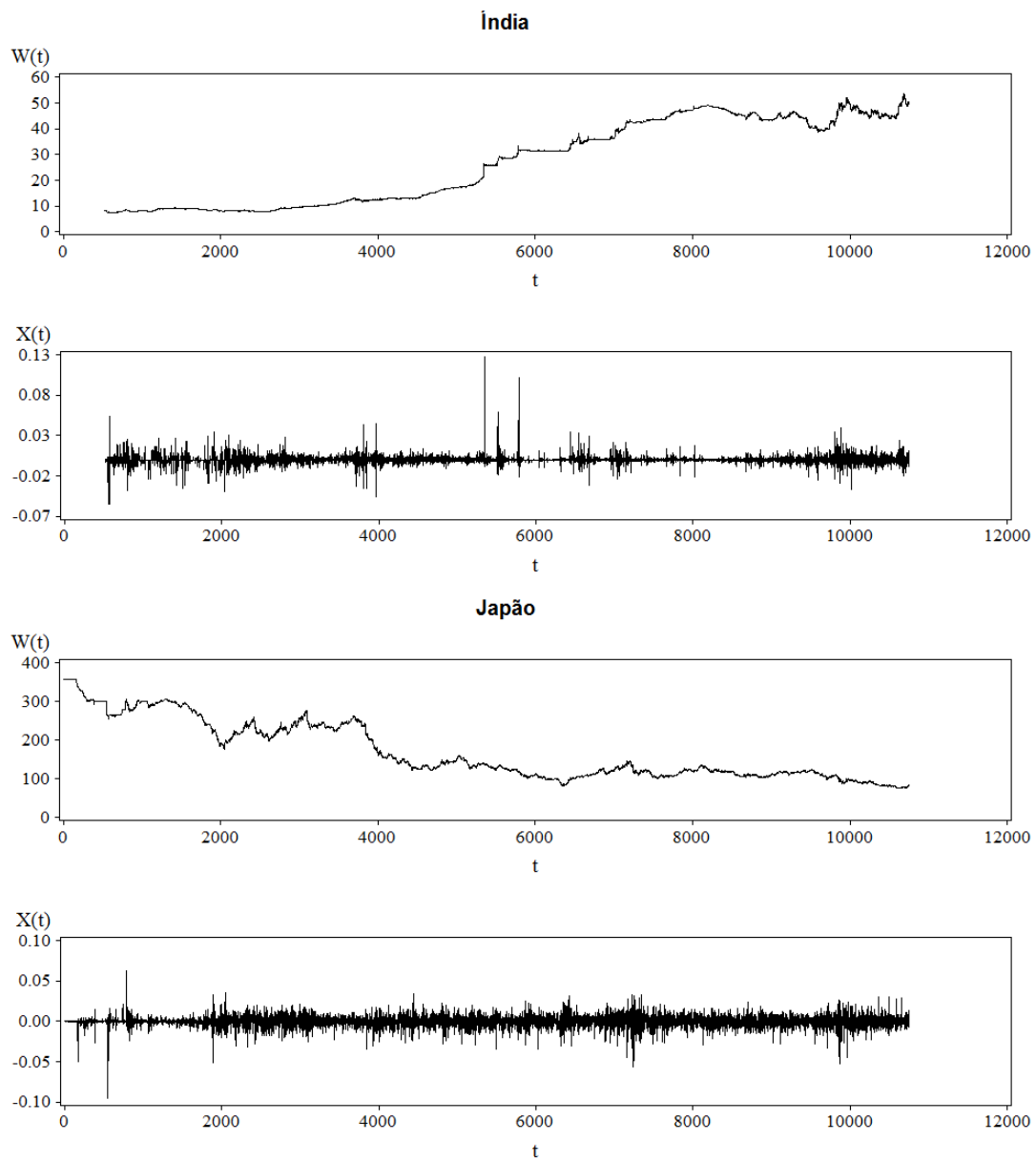


Figura 1.6: Evoluções diárias das taxas de câmbio W_t da rúpia e do yen frente ao dólar americano (painel superior), e dos respectivos retornos X_t (painel inferior).

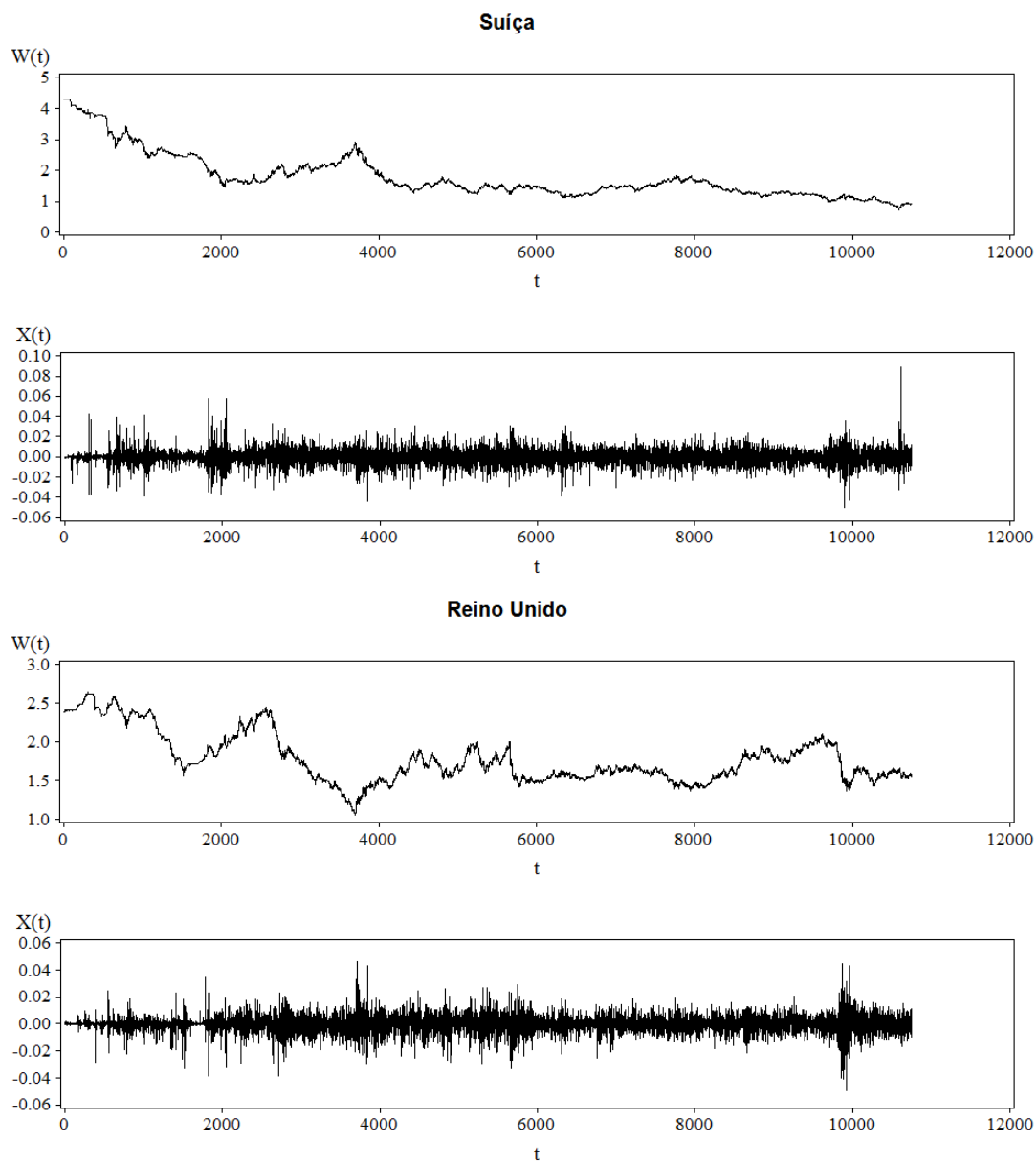


Figura 1.7: Evoluções diárias das taxas de câmbio W_t do franco suíço e da libra esterlina frente ao dólar americano (painel superior), e dos respectivos retornos X_t (painel inferior).

1.5 Esboço do trabalho

O Cap. 2 apresentará algumas propriedades da função característica de uma distribuição X parametrizada por um vetor $\boldsymbol{\theta} \in \mathbb{R}^p$. Para contemplar os casos em que $\langle |X|^k \rangle = \infty$, $k > 0$, será proposta uma expansão alternativa da função característica na forma $\phi_X(q; \boldsymbol{\theta}) = \Phi(D_q; \boldsymbol{\theta})\delta(q)$, em que $\delta(q)$ é a função delta de Dirac, e $\Phi(D_q; \boldsymbol{\theta})$ é uma forma polinomial do operador diferencial $D_q = \frac{d}{dq}$, cujos coeficientes dependem de $\boldsymbol{\theta}$. Essa expansão não se encontra nas obras clássicas que versam acerca da matéria [47, 66, 120]. Mostraremos que a mínima divergência de Kullback-Leibler entre duas distribuições X e Y com respeito ao parâmetro $\theta_j \in \boldsymbol{\theta}$, $j = 1, \dots, p$, satisfaz à equação

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \left\{ \phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0, \quad (1.11)$$

em que

$$\omega_j(q; \boldsymbol{\theta}) = \overline{\{\Phi(D_q; \boldsymbol{\theta})\}^{-1} h_j(q; \boldsymbol{\theta})} \quad (1.12)$$

e

$$h_j(q; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \phi_X(q; \boldsymbol{\theta}). \quad (1.13)$$

Além disso,

$$\text{IF}_j(\boldsymbol{\theta}) = \int_{-\infty}^{+\infty} \overline{\{\Phi(D_q; \boldsymbol{\theta})\}^{-1} |h_j(q, \boldsymbol{\theta})|^2} dq, \quad (1.14)$$

em que $\text{IF}_j(\boldsymbol{\theta})$ é a medida de informação de Fisher (Apêndice D).

O Cap. 3 apresentará o conceito de divisibilidade infinita introduzido por B. de Finetti [18, 19]. Com base nesse conceito, define-se uma classe de distribuições geradas por somas de variáveis aleatórias IID da qual as distribuições estáveis de Lévy fazem parte [47, 103, 104]. No caso não gaussiano das distribuições estáveis, nem todos os momentos existem [103]. A teoria das distribuições estáveis remete a um teorema limite central generalizado, já que ela permite descrever a forma assintótica da distribuição de uma soma de variáveis aleatórias independentes com variâncias não necessariamente finitas [103, 104]. Assim, com base nessa teoria, é possível modelar fenômenos sujeitos a grandes flutuações. Nesse capítulo será estudada a forma particular do polinômio característico das distribuições estáveis.

O Cap. 4 tratará sobre a função característica empírica (FCE). Ela é uma importante ferramenta estatística para o estudo de observações que seguem distribuições estáveis [26, 95, 97, 125] ou quase estáveis [30, 31, 33, 35], pois a representação da distribuição dos dados por funções características é mais simples do que a por funções de densidade ou funções de distribuição de probabilidade. Entre outras aplicações, as estimativas dos parâmetros da distribuição hipotética podem ser obtidas via FCE ($\hat{\phi}(q)$) [26, 97]. Além das propriedades estatísticas básicas de $\hat{\phi}(q)$ [27, 120], discutiremos sobre o fenômeno de truncamento da FCE devido à finitude do tamanho da amostra, conforme o teorema de Glivenko-Cantelli [120]. Esse truncamento permite explicar as quebras abruptas de *scaling* de outras características distribucionais, como o momentos absolutos das distribuições quase estáveis [84]. Com base na distribuição amostral da FCE e na teoria de testes de hipóteses [101], espera-se que haja truncamento natural da FCE quando sua magnitude for inferior à do nível de ruído. Com respeito às propriedades de *scaling* da FCE, considerando um passeio aleatório $S_{\Delta t} = X_1 + X_2 + \dots + X_{\Delta t}$, avaliaremos em que ponto q ocorre a quebra do comportamento esperado de $\ln(-\ln|\hat{\phi}_{S_{\Delta t}}(q)|)$ versus $\ln(q)$. A estimativa da FD no ponto zero (ou a "probabilidade de retorno à origem") permite o estudo do *scaling* de uma distribuição de dados [75]. No entanto, as estimativas produzidas pelo método do Kernel [107, 108] — o que inclui o Kernel triangular utilizado em trabalhos anteriores [21, 46, 75] — são tendenciosas e inconsistentes [115]. Como alternativa, a FD no ponto zero pode ser estimada com base na transformada inversa de Fourier da FCE.

O Cap. 5 tratará do método de estimação por máxima verossimilhança (EMV). Esse método proporciona estimadores com boas propriedades estatísticas como consistência — convergência em probabilidade para os respectivos alvos —, eficiência — variância mínima em comparação com outros estimadores não tendenciosos — e normalidade assintótica das distribuições amostrais. Discutiremos que o princípio da máxima verossimilhança (MV) é equivalente à minimização da divergência de Kullback-Leibler entre uma distribuição hipotética X e a empírica Y . Desse modo, as Eqs. (1.11), (1.12) e (1.13) formam o

sistema de equações de MV, e sua solução $\hat{\theta}$ são as estimativas de MV. A avaliação da estabilidade ou quase estabilidade de um passeio aleatório $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$ pode ser feita estudando-se o comportamento do índice $\alpha_{S_{\Delta t}}$. Se, por exemplo, $X_1, \dots, X_{\Delta t}$ for uma sequência de VA estáveis IID com parâmetros α e γ , então $\alpha_{S_{\Delta t}} = \alpha$ e $\gamma_{S_{\Delta t}} = \Delta t \gamma$, para qualquer Δt . Mas se essa sequência IID não for de VA estáveis, então $\alpha_{S_{\Delta t}} \rightarrow 2$ à medida que Δt cresce e $\gamma_{S_{\Delta t}} = \Delta t \gamma$. E, se as VA não forem IID nem estáveis, então é possível que $\alpha_{S_{\Delta t}} \approx \alpha$ enquanto $\Delta t_0 \leq \Delta t \leq \Delta t_1$; e agora, $\gamma_{S_{\Delta t}} \neq \Delta t \gamma$ por causa da autocorrelação serial.

O Cap. 6 introduzirá um novo teste estatístico para a detecção de dependência não linear [80]. Essa ferramenta é importante, considerando o papel da dependência não linear na dinâmica da convergência de uma soma de VA [33, 35]. Embora haja muitos métodos estatísticos para a avaliação da independência entre duas VA, poucos deles são capazes de detectar as diversas formas de dependência não linear [6]. Entre esses poucos métodos estatísticos, encontra-se o teste de HBKR (Hoeffding, Blum, Kiefer e Rosenblatt) [9, 54]. Esse teste é do tipo Cramér-von Mises e considera um processo aleatório gerado pela diferença entre a distribuição conjunta empírica e o produto correspondente entre as marginais empíricas. Embora haja testes mais recentes que envolvem outras caracterizações de processos aleatórios, eles são, essencialmente, equivalentes ao teste de HBKR para o caso bivariado [40]. É possível, no entanto, propor um teste assintótico com maior poder estatístico para se rejeitar a hipótese de independência. Para isso, nosso teste será construído com base na estatística χ^2 da razão de verossimilhança, o que equivale a minimizar a divergência de Kullback-Leibler entre duas distribuições empíricas. Nosso teste não deverá requerer suposições acerca da distribuição dos dados (*distribution free*), e deverá ser aplicável para dados não gaussianos com observações extremas, por exemplo.

Para concluir, as considerações finais e as perspectivas para trabalhos futuros serão apresentados no Cap. 7.

Finalmente, alguns resultados matemáticos úteis e esboços de demonstrações são apresentados nos apêndices. O Apêndice A lista algumas integrais e propriedades das funções

gama e delta. O Apêndice B expõe uma relação binomial da função escore (aplicada no Cap. 2), o Apêndice C esboça a demonstração da forma geral da função característica de uma distribuição infinitamente divisível (fórmula de Lévy-Khinchine), o Apêndice D trata brevemente sobre a informação de Fisher, o Apêndice E apresenta o método de Janaki-Weron para a geração computacional de realizações de distribuições estáveis, e alguns cálculos referentes ao desenvolvimento do Cap. 6 são detalhados no Apêndice F.

Capítulo 2

A função característica e as distâncias entre distribuições

2.1 Introdução

Em teoria de probabilidade, uma variável aleatória contínua X é aquela que possui uma função de densidade $f(x)$ tal que $P(X \leq x) = \int_{-\infty}^x f(u)du$ para qualquer número real x [47, 89, 98]. Define-se como função característica (FC) de X a transformada de Fourier de $f(x)$, e assim, a forma da distribuição X também pode ser equivalentemente representada segundo essa transformada [66, 98, 120]. Enquanto $f(x)$ define a forma da distribuição dos possíveis valores x e da probabilidade acumulada $P(X \leq x)$, a FC permite descrever X no espaço dos momentos estatísticos. Por isso, as propriedades distribucionais de X podem ser estudadas tanto sob a perspectiva da função de densidade como a da FC, dependendo da conveniência. Por exemplo, por causa da simplicidade da forma geral da FC de uma soma de variáveis aleatórias independentes, A. Lyapunov e P. Lévy a utilizaram largamente no estudo de teoremas limites ao longo da primeira metade do séc. XX [36, 37, 48, 64, 120]. Logo em seguida, a teoria das funções características se consolidou [24, 47, 57, 66, 98, 120], sendo de grande utilidade para o desenvolvimento de métodos estatísticos para o estudo da dinâmica dos fenômenos físicos [*e.g.*, 29, 68, 96, 117, 121].

Além das propriedades básicas das FC, o propósito deste capítulo é apresentar uma relação entre as funções características de duas distribuições X e Y e a divergência de Kullback-Leibler entre essas distribuições [63]. Por exemplo, Y pode ser uma distribuição empírica, enquanto X representa uma distribuição hipotética. Nesse caso, como veremos no Cap. 5, a minimização da divergência de Kullback-Leibler com respeito aos parâmetros da distribuição hipotética se relaciona com o princípio da máxima verossimilhança (MV) para a estimação de parâmetros [26, 101]. Desse modo, o resultado deste capítulo será útil para se fazer inferências estatísticas sobre os parâmetros de distribuições de Levy estáveis (ou quase estáveis) com base em funções características (Cap. 5). Isso porque a forma da função característica de uma variável aleatória hipoteticamente estável (Cap. 3) é mais simples do que a da função de densidade, pois esta não possui forma fechada. No Cap. 6, o princípio da MV será considerado para se contruir um novo teste de independência entre duas variáveis aleatórias.

Na Seção 2.2 são apresentadas as propriedades das funções características que são relevantes para o desenvolvimento dos capítulos subsequentes. Propõe-se ainda uma expansão alternativa da FC para o caso em que $f(x)$ possui representação em série de Maclaurin. Essa expansão é aplicável para o caso em que $\langle X^k \rangle = \infty$, $k \in \mathbb{N}^+$. A distância L_2 e a divergência de Kullback-Leibler são apresentadas na Seção 2.3, e nela, mostramos que minimização da divergência de Kullback-Leibler depende de uma função peso $\omega(\cdot)$, cuja forma apresentada na Eq. (2.58) constitui o ápice deste capítulo. As considerações concernentes a este capítulo são apresentadas no final (Seção 2.4).

2.2 A função característica

Seja X uma variável aleatória (VA) real com função de distribuição acumulada (FDA) absolutamente contínua $F(x; \boldsymbol{\theta}) = \int_{-\infty}^x f(u; \boldsymbol{\theta}) du$, em que $f(x; \boldsymbol{\theta}) \geq 0$ é a função de densidade (FD) e $\boldsymbol{\theta} \in \mathbb{R}^p$ representa seu vetor de parâmetros. Define-se a função característica

(FC) de X como a transformada de Fourier [47, 66]

$$\phi(q; \boldsymbol{\theta}) = \langle e^{iqX} \rangle \quad (2.1)$$

$$= \int_{-\infty}^{+\infty} e^{iqx} dF(x; \boldsymbol{\theta}) \quad (2.2)$$

$$= \langle \cos(qX) \rangle + i \langle \sin(qX) \rangle \quad (2.3)$$

$$= \phi_2(q; \boldsymbol{\theta}) + i\phi_1(q; \boldsymbol{\theta}), \quad (2.4)$$

em que $q \in \mathbb{R}$. Por outro lado, se $\phi(q; \boldsymbol{\theta})$ for uma função absolutamente integrável, a FDA correspondente é absolutamente contínua, e a FD pode ser obtida mediante a transformada inversa

$$f(x; \boldsymbol{\theta}) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(q; \boldsymbol{\theta}) e^{-iqx} dq. \quad (2.5)$$

2.2.1 Propriedades

Com base na definição (2.1), (2.2) ou (2.3), conclui-se que a FC possui as seguintes propriedades básicas: *a*) $\phi(q; \boldsymbol{\theta})$ é uniformemente contínua; *b*) $\phi(0; \boldsymbol{\theta}) = 1$; *c*) $|\phi(q; \boldsymbol{\theta})| \leq 1$; *d*) $\phi(-q; \boldsymbol{\theta}) = \overline{\phi(q; \boldsymbol{\theta})}$; e *e*) se a distribuição de X for simétrica em torno de zero, então $\phi(q; \boldsymbol{\theta}) = \phi(-q; \boldsymbol{\theta}) \in \mathbb{R}$. Em particular, para as distribuições absolutamente contínuas, tem-se que [66, 120]

$$\lim_{|q| \rightarrow \infty} \phi(q; \boldsymbol{\theta}) = 0, \quad (2.6)$$

enquanto para as discretas,

$$\lim_{|q| \rightarrow \infty} \phi(q; \boldsymbol{\theta}) = 1. \quad (2.7)$$

Além dessas propriedades básicas, tem-se que

- o complementar $\overline{\phi(q; \boldsymbol{\theta})} = \phi(q; \boldsymbol{\theta})$ é FC de $-X$;
- $\phi^k(q; \boldsymbol{\theta})$, em que $k \in \mathbb{N}^+$, é a FC da convolução de n cópias independentes de X ;
- a parte real da FC de uma VA absolutamente contínua X , $\phi_2(q; \boldsymbol{\theta})$, é FC de uma VA cuja função de distribuição é dada por $\frac{1}{2}(1 + F(x) - F(-x))$;

- $|\phi(q; \boldsymbol{\theta})|^2$ é a FC da diferença $X_1 - X_2$ (simetrização), em que X_1 e X_2 são cópias independentes de X .

Considerando-se que a k -ésima derivada da FC com respeito a q pode ser escrita como

$$\phi^{(k)}(q; \boldsymbol{\theta}) = i^k \int_{-\infty}^{+\infty} x^k e^{iqx} dF(x; \boldsymbol{\theta}), \quad (2.8)$$

se $|\phi^{(k)}(0; \boldsymbol{\theta})| < +\infty$, então o k -ésimo momento da distribuição X existe [66, 67], e pode ser obtido mediante a operação $\langle X^k \rangle = (-i)^k \phi^{(k)}(0; \boldsymbol{\theta})$. Assim, a existência de todos os momentos de ordem k permite uma expansão na forma

$$\begin{aligned} \phi(q; \boldsymbol{\theta}) &= \int_{-\infty}^{+\infty} e^{iqx} dF(x; \boldsymbol{\theta}) \\ &= \int_{-\infty}^{+\infty} \sum_{k=0}^{+\infty} \frac{(iqx)^k}{k!} dF(x; \boldsymbol{\theta}) \\ &= \sum_{k=0}^{+\infty} \frac{(iq)^k \langle X^k \rangle}{k!}. \end{aligned} \quad (2.9)$$

O resultado (2.9) é bastante conhecido e importante, pois estabelece uma relação entre a FC e os momentos da distribuição. Entretanto, ele não é aplicável se $\langle X^k \rangle = \infty$ para algum k .

2.2.2 Uma expansão alternativa

Se a FD possui uma representação em série de Taylor em $x = 0$, considerando que $f^{(k)} = f^{(k)}(0; \boldsymbol{\theta})$, tem-se a seguinte expansão alternativa:

$$\begin{aligned} \phi(q; \boldsymbol{\theta}) &= \int_{-\infty}^{+\infty} e^{iqx} f(x; \boldsymbol{\theta}) dx \\ &= \int_{-\infty}^{+\infty} \sum_{k=0}^{+\infty} \frac{f^{(k)} x^k}{k!} e^{iqx} dx \\ &= \sum_{k=0}^{+\infty} \frac{2\pi f^{(k)} (-i)^k \delta^{(k)}(q)}{k!} \end{aligned} \quad (2.10)$$

$$\begin{aligned} &= \sum_{k=0}^{+\infty} c_k \frac{d^k}{dq^k} \delta(q) \\ &= \sum_{k=0}^{+\infty} c_k D_q^k \delta(q), \end{aligned} \quad (2.11)$$

em que $D_q^k \doteq \frac{d^k}{dq^k}$ é o operador diferencial na notação de Euler, $\delta(q)$ é a função delta de Dirac,

$$c_k = c_k(\boldsymbol{\theta}) = \frac{2\pi f^{(k)}(-i)^k}{k!}, \quad (2.12)$$

e, pela Eq. (2.5),

$$f^{(k)}(x; \boldsymbol{\theta}) = \frac{(-i)^k}{2\pi} \int_{-\infty}^{+\infty} \phi(q; \boldsymbol{\theta}) q^k e^{-iqx} dq, \quad (2.13)$$

de modo que

$$f^{(k)} = f^{(k)}(0; \boldsymbol{\theta}) = \frac{(-i)^k}{2\pi} \int_{-\infty}^{+\infty} \phi(q; \boldsymbol{\theta}) q^k dq. \quad (2.14)$$

Assim, com base na expansão (2.11), a FC pode ser escrita como

$$\phi(q; \boldsymbol{\theta}) = \Phi(D_q; \boldsymbol{\theta}) \delta(q), \quad (2.15)$$

em que

$$\Phi(D_q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} c_k D_q^k \quad (2.16)$$

denomina-se *polinômio característico* de $\phi(q; \boldsymbol{\theta})$ — já que esse polinômio caracteriza a forma da FC e, conseqüentemente, da distribuição da variável aleatória X .

Embora seja naturalmente uma forma alternativa à Eq. (2.9), a expansão (2.15) não consta em obras clássicas [e.g., 47, 66, 120] que abordam sobre as funções características. Em nosso trabalho, essa expansão será fundamental para se obter a equação de máxima verossimilhança para a estimação dos parâmetros das distribuições estáveis (Cap. 5).

Quanto à k -ésima derivada $\phi^{(k)}(q; \boldsymbol{\theta})$, uma forma alternativa à Eq. (2.8), obtida com base em (2.15), pode ser escrita como

$$\begin{aligned} \phi^{(k)}(q; \boldsymbol{\theta}) &= D_q^k \sum_{j=0}^{+\infty} c_j \delta^{(j)}(q) \\ &= \sum_{j=0}^{+\infty} c_j \delta^{(j+k)}(q) \\ &= \Phi(D_q^k) \delta(q). \end{aligned} \quad (2.17)$$

2.2.3 Distribuições simétricas em torno de zero

Se a variável aleatória X for simétrica em torno de zero, tem-se que $\phi(q; \boldsymbol{\theta}) = \phi(-q; \boldsymbol{\theta}) \in \mathbb{R}$ e $f(x; \boldsymbol{\theta}) = f(-x; \boldsymbol{\theta})$. Nessa situação, com respeito à k -ésima derivada da FD em $x = 0$, os termos de ordem par e ímpar da Eq. (2.14) são, respectivamente,

$$\begin{aligned} f^{(2k)} &= \frac{(-1)^k}{2\pi} \int_{-\infty}^{+\infty} \phi(q; \boldsymbol{\theta}) q^{2k} dq \\ &= \frac{(-1)^k}{\pi} \int_0^{+\infty} \phi(q; \boldsymbol{\theta}) q^{2k} dq, \end{aligned} \quad (2.18)$$

e

$$\begin{aligned} f^{(2k+1)} &= -\frac{i^{2k+1}}{2\pi} \int_{-\infty}^{+\infty} \phi(q; \boldsymbol{\theta}) q^{2k+1} dq \\ &= -\frac{i^{2k+1}}{2\pi} \int_0^{+\infty} \phi(q; \boldsymbol{\theta}) q^{2k+1} dq + \frac{i^{2k+1}}{2\pi} \int_0^{+\infty} \phi(q; \boldsymbol{\theta}) q^{2k+1} dq \\ &= 0. \end{aligned} \quad (2.19)$$

O polinômio característico (2.16) pode ser escrito como

$$\Phi(D_q; \boldsymbol{\theta}) \delta(q) = \Phi_2(D_q; \boldsymbol{\theta}) + \Phi_1(D_q; \boldsymbol{\theta}), \quad (2.20)$$

em que $\Phi_2(D_q; \boldsymbol{\theta})$ e $\Phi_1(D_q; \boldsymbol{\theta})$ são respectivamente as partes real (par) e imaginária (ímpar) do polinômio característico, ou seja,

$$\Phi_2(D_q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} c_{2k} D_q^{2k}, \quad (2.21)$$

$$\Phi_1(D_q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} c_{2k+1} D_q^{2k+1}, \quad (2.22)$$

já que $c_{2k} \in \mathbb{R}$ e $c_{2k+1} \in \mathbb{C}$. Assim, a FC pode ser escrita como

$$\phi(q; \boldsymbol{\theta}) = \phi_2(q; \boldsymbol{\theta}) + i\phi_1(q; \boldsymbol{\theta}), \quad (2.23)$$

em que $\phi_2(q; \boldsymbol{\theta}) = \Phi_2(D_q; \boldsymbol{\theta}) \delta(q)$ e $i\phi_1(q; \boldsymbol{\theta}) = \Phi_1(D_q; \boldsymbol{\theta}) \delta(q)$ (exemplos serão apresentados no próximo capítulo). Portanto, em caso de simetria de X em torno de zero, como

$f^{(2k+1)} = 0$ e $\phi(q; \boldsymbol{\theta}) \in \mathbb{R}$, tem-se que

$$\begin{aligned}\phi(q; \boldsymbol{\theta}) &= \phi_2(q; \boldsymbol{\theta}) \\ &= \Phi_2(D_q; \boldsymbol{\theta})\delta(q),\end{aligned}\tag{2.24}$$

e a k -ésima derivada de $\phi(q; \boldsymbol{\theta})$ tem a forma

$$\begin{aligned}\phi^{(k)}(q; \boldsymbol{\theta}) &= \sum_{j=0}^{+\infty} c_{2j} \delta^{(2j+2k)}(q) \\ &= \Phi_2(D_q^k) \delta(q).\end{aligned}\tag{2.25}$$

2.2.4 Relações com respeito ao vetor de parâmetros

Considerando-se agora que o vetor de parâmetros $\boldsymbol{\theta}$ possa variar, as derivadas da FC e FD com respeito ao j -ésimo elemento, $\theta_j \in \boldsymbol{\theta}$, são respectivamente denotadas como:

$$h_j(q; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \phi(q; \boldsymbol{\theta}),\tag{2.26}$$

$$g_j(x; \boldsymbol{\theta}) = \frac{d}{d\theta_j} f(x; \boldsymbol{\theta}).\tag{2.27}$$

Assim, por (2.2) tem-se

$$h_j(q; \boldsymbol{\theta}) = \int_{-\infty}^{+\infty} e^{iqx} g_j(x; \boldsymbol{\theta}) dx,\tag{2.28}$$

de modo que se obtém por analogia ao resultado (2.10),

$$h_j(q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} \frac{2\pi g_j^{(k)}(-i)^k \delta^{(k)}(q)}{k!}\tag{2.29}$$

$$= \sum_{k=0}^{+\infty} b_k \delta^{(k)}(q)\tag{2.30}$$

$$= h_{j,2}(q; \boldsymbol{\theta}) + h_{j,1}(q; \boldsymbol{\theta}),\tag{2.31}$$

em que $g_j^{(k)} = g_j^{(k)}(0; \boldsymbol{\theta})$, $b_k = \frac{2\pi g_j^{(k)}(-i)^k}{k!}$ e

$$h_{j,2}(q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} b_{2k} \delta^{(2k)}(q),\tag{2.32}$$

$$h_{j,1}(q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} b_{2k+1} \delta^{(2k+1)}(q).\tag{2.33}$$

Em particular, para distribuições simétricas em torno de zero,

$$h_j(q; \boldsymbol{\theta}) = h_{j,2}(q; \boldsymbol{\theta}). \quad (2.34)$$

Os resultados apresentados nesta seção serão aplicados no estudo das distâncias entre duas distribuições. O assunto da seção a seguir se relaciona com a questão do ajuste ou da estimação de parâmetros com relação a uma distribuição de referência (Cap. 5), e também será útil para se testar a hipótese de independência entre duas distribuições (Cap. 6).

2.3 Distâncias entre duas distribuições

Considere duas distribuições X e Y , cujas FD e FC correspondentes são representadas por $f_X(x; \boldsymbol{\theta})$, $f_Y(y; \boldsymbol{\theta}')$, $\phi_X(q; \boldsymbol{\theta})$ e $\phi_Y(q'; \boldsymbol{\theta}')$.

2.3.1 A distância L_2

Define-se a distância L_2 entre essas densidades como [120]

$$L_2(X, Y; \boldsymbol{\theta}, \boldsymbol{\theta}') = \int (f_Y(x; \boldsymbol{\theta}') - f_X(x; \boldsymbol{\theta}))^2 dx. \quad (2.35)$$

Assim,

$$\begin{aligned} L_2(X, Y; \boldsymbol{\theta}, \boldsymbol{\theta}') &= \\ &= \frac{1}{(2\pi)^2} \int \left| \int (\phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta})) e^{-iqx} dq \right|^2 dx \\ &= \frac{1}{(2\pi)^2} \int_{x \in \mathbb{R}} \int_{q \in \mathbb{R}} \int_{q' \in \mathbb{R}} (\phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta})) (\overline{\phi_Y(q'; \boldsymbol{\theta}')} - \overline{\phi_X(q'; \boldsymbol{\theta})}) e^{-i(q-q')x} dq dq' dx \\ &= \frac{1}{2\pi} \int_{q \in \mathbb{R}} \int_{q' \in \mathbb{R}} (\phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta})) (\overline{\phi_Y(q'; \boldsymbol{\theta}')} - \overline{\phi_X(q'; \boldsymbol{\theta})}) \delta(q - q') dq dq' \\ &= \frac{1}{2\pi} \int |\phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta})|^2 dq, \end{aligned}$$

de modo que se tem a identidade [120]

$$\int (f_Y(x; \boldsymbol{\theta}') - f_X(x; \boldsymbol{\theta}))^2 dx = \frac{1}{2\pi} \int |\phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta})|^2 dq. \quad (2.36)$$

Portanto, a distância L_2 entre duas FD é equivalente à distância L_2 entre duas FC. Essa relação é útil para se medir distâncias em situações nas quais a forma funcional da FC é mais simples do que a da FD, como é o caso das distribuições estáveis (Cap. 3).

2.3.2 A divergência de Kullback-Leibler

Embora não seja uma distância propriamente dita, a divergência de Kullback-Leibler pode ser considerada como uma medida de dissimilaridade entre duas distribuições, relacionando-se com o princípio da máxima verossimilhança [98, 101]. Ela é definida como [63]

$$D_{KL}(X, Y; \boldsymbol{\theta}, \boldsymbol{\theta}') = \int f_Y(x; \boldsymbol{\theta}') \ln \frac{f_Y(x; \boldsymbol{\theta}')}{f_X(x; \boldsymbol{\theta})} dx \quad (2.37)$$

$$= H(Y; \boldsymbol{\theta}') - H(Y, X; \boldsymbol{\theta}', \boldsymbol{\theta}), \quad (2.38)$$

em que $H(Y; \boldsymbol{\theta}')$ é a entropia de Y e $H(Y, X; \boldsymbol{\theta}', \boldsymbol{\theta})$ é a entropia cruzada entre Y e X.

Agora, considere o problema da determinação da menor distância entre X e Y, supondo-se que a entropia $H(Y; \boldsymbol{\theta}')$ seja constante. Isto é, $\boldsymbol{\theta}'$ é fixo enquanto $\boldsymbol{\theta}$ pode variar, o que permite ajustar a FD $f_X(x; \boldsymbol{\theta})$ em relação à densidade de referência $f_Y(x; \boldsymbol{\theta}')$. Nessa situação, a distância D_{KL} pode ser minimizada com respeito a um elemento $\theta_j \in \boldsymbol{\theta}$ fazendo-se

$$\frac{d}{d\theta_j} D_{KL}(\boldsymbol{\theta}, \boldsymbol{\theta}') = -\frac{d}{d\theta_j} H(Y, X; \boldsymbol{\theta}', \boldsymbol{\theta}) \quad (2.39)$$

$$\begin{aligned} &= -\int_{-\infty}^{+\infty} \frac{d}{d\theta_j} \ln f_X(x; \boldsymbol{\theta}) f_Y(x; \boldsymbol{\theta}') dx = \\ &= -\int_{-\infty}^{+\infty} s_j(x; \boldsymbol{\theta}) f_Y(x; \boldsymbol{\theta}') dx = 0. \end{aligned} \quad (2.40)$$

A função

$$s_j(x; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \ln f(x; \boldsymbol{\theta}), \quad (2.41)$$

conhecida como score eficiente [98], indica a sensibilidade relativa de $f(x; \boldsymbol{\theta})$ a variações de θ_j . Como

$$\int s_j(x; \boldsymbol{\theta}) dF(x; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \int f(x; \boldsymbol{\theta}) dx = 0,$$

a equação (2.40) pode ser equivalentemente representada por

$$\int_{-\infty}^{+\infty} s_j(x; \boldsymbol{\theta}) (f_Y(x; \boldsymbol{\theta}') - f_X(x; \boldsymbol{\theta})) dx = 0. \quad (2.42)$$

Com base em (2.2) e (2.5), tem-se que

$$\begin{aligned} \int s_j(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) dx &= \frac{1}{2\pi} \int s_j(x; \boldsymbol{\theta}) \int \phi_X(q; \boldsymbol{\theta}) e^{-iqx} dq dx \\ &= \int \left\{ \int \frac{1}{2\pi} s_j(x; \boldsymbol{\theta}) e^{-iqx} dx \right\} \phi_X(q; \boldsymbol{\theta}) dq \\ &= \int \omega_j(q; \boldsymbol{\theta}) \phi_X(q; \boldsymbol{\theta}) dq, \end{aligned} \quad (2.43)$$

em que

$$\omega_j(q; \boldsymbol{\theta}) = \frac{1}{2\pi} \int s_j(x; \boldsymbol{\theta}) e^{-iqx} dx \quad (2.44)$$

é uma transformada inversa do escore $s_j(x; \boldsymbol{\theta})$. Analogamente, tem-se

$$\int s_j(x; \boldsymbol{\theta}) f_Y(x; \boldsymbol{\theta}') dx = \int \omega_j(q; \boldsymbol{\theta}) \phi_Y(q; \boldsymbol{\theta}') dq. \quad (2.45)$$

Substituindo-se (2.43) e (2.45) em (2.42), conclui-se que

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \left\{ \phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0. \quad (2.46)$$

Portanto, o valor de θ_j que minimiza a distância D_{KL} entre f_X e a referência f_Y é solução da equação (2.42) ou da equação no domínio de Fourier (2.46).

2.3.3 Expansão da função $\omega(q; \boldsymbol{\theta})$

Para situações em que não se dispõe de uma fórmula exata simples para a função escore — como é o caso de algumas distribuições estáveis (Cap. 3)—, pode-se usar sua expansão de Taylor

$$s_j(x; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} \frac{x^k}{k!} s_j^{(k)}(0; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} \frac{x^k}{k!} s_j^{(k)}. \quad (2.47)$$

Com base nessa expansão, a função peso pode ser escrita como

$$\begin{aligned} \omega_j(q; \boldsymbol{\theta}) &= \frac{1}{2\pi} \int s_j(x; \boldsymbol{\theta}) e^{-iqx} dx \\ &= \frac{1}{2\pi} \int \sum_{k=0}^{+\infty} \frac{x^k}{k!} s_j^{(k)} e^{-iqx} dx \\ &= \sum_{k=0}^{+\infty} \frac{s_j^{(k)}}{k!} i^k \delta^{(k)}(q). \end{aligned} \quad (2.48)$$

Para obtermos uma expressão para as derivadas da função escore no ponto zero, $s_j^{(k)}$, primeiramente reescrevemos a equação (2.41) como

$$s_j(x; \boldsymbol{\theta}) = \frac{g_j(x; \boldsymbol{\theta})}{f(x; \boldsymbol{\theta})}, \quad (2.49)$$

em que $g_j(x; \boldsymbol{\theta}) = \frac{d}{d\theta_j} f(x; \boldsymbol{\theta})$; em seguida, usando a relação (ver Ap. B)

$$g_j^{(k)}(x; \boldsymbol{\theta}) = \sum_{l=0}^k \binom{k}{l} s_j^{(k-l)}(x; \boldsymbol{\theta}) f^{(l)}(x; \boldsymbol{\theta}), \quad (2.50)$$

temos a relação (para $k \geq 1$)

$$s_j^{(k)} = \frac{g_j^{(k)}}{f} - \sum_{l=1}^k \binom{k}{l} \frac{s_j^{(k-l)} f^{(l)}}{f}, \quad (2.51)$$

em que $f^{(l)} = f^{(l)}(0; \boldsymbol{\theta})$ e $g_j^{(l)} = g_j^{(l)}(0; \boldsymbol{\theta})$.

Em particular, se X for simétrica em torno de zero ($k \geq 1$),

$$s_j^{(2k)} = \frac{g_j^{(2k)}}{f} - \sum_{l=1}^k \binom{2k}{2l} \frac{s_j^{(2k-2l)} f^{(2l)}}{f}, \quad (2.52)$$

pois $f^{(2k+1)} = 0$, $g_j^{(2k+1)} = 0$ e $s_j^{(2k+1)} = 0$.

Substituindo-se o resultado (2.51) na expansão (2.48), e considerando-se (2.29),

$$\begin{aligned} \omega_j(q; \boldsymbol{\theta}) &= \sum_{k=0}^{+\infty} \frac{s_j^{(k)}}{k!} i^k \delta^{(k)}(q) \\ &= s_j \delta(q) + \sum_{k=1}^{+\infty} \frac{s_j^{(k)}}{k!} i^k \delta^{(k)}(q) \\ &= \frac{g_j}{f} \delta(q) + \sum_{k=1}^{+\infty} \frac{i^k \delta^{(k)}(q)}{k!} \left\{ \frac{g_j^{(k)}}{f} - \sum_{l=1}^k \binom{k}{l} \frac{s_j^{(k-l)} f^{(l)}}{f} \right\} \\ &= \frac{g_j}{f} \delta(q) + \frac{1}{f} \sum_{k=1}^{+\infty} \frac{i^k g_j^{(k)} \delta^{(k)}(q)}{k!} - \frac{1}{f} \sum_{k=1}^{+\infty} \sum_{l=1}^k \binom{k}{l} \frac{s_j^{(k-l)} f^{(l)} i^k \delta^{(k)}(q)}{k!} \\ &= \frac{1}{2\pi f} \sum_{k=0}^{+\infty} \overline{b_k} \delta^{(k)}(q) - \frac{R_j(q; \boldsymbol{\theta})}{f} \\ &= \frac{\overline{h_j(q; \boldsymbol{\theta})}}{2\pi f} - \frac{R_j(q; \boldsymbol{\theta})}{f}, \end{aligned} \quad (2.53)$$

em que $\overline{b_k} = \frac{2\pi g_j^{(k)} i^k}{k!}$ e

$$R_j(q; \boldsymbol{\theta}) \doteq \sum_{k=1}^{+\infty} \sum_{l=1}^k \binom{k}{l} \frac{s_j^{(k-l)} f^{(l)} i^k \delta^{(k)}(q)}{k!}. \quad (2.54)$$

Se X for simétrica em torno de zero,

$$\omega_j(q; \boldsymbol{\theta}) = \frac{1}{2\pi f} h_{j,2}(q; \boldsymbol{\theta}) - \frac{R_j(q; \boldsymbol{\theta})}{f}. \quad (2.55)$$

Desenvolvendo agora (2.54), obtemos

$$\begin{aligned} R_j(q; \boldsymbol{\theta}) &\doteq \sum_{k=1}^{+\infty} \sum_{l=1}^k \binom{k}{l} \frac{s_j^{(k-l)} f^{(l)} i^k \delta^{(k)}(q)}{k!} \\ &= \sum_{k=0}^{+\infty} \sum_{l=1}^{+\infty} \binom{k+l}{l} \frac{s_j^{(k)} f^{(l)} i^{k+l} \delta^{(k+l)}(q)}{(k+l)!} \\ &= \sum_{k=0}^{+\infty} \sum_{l=1}^{+\infty} \frac{(k+l)! s_j^{(k)} f^{(l)} i^{k+l} \delta^{(k+l)}(q)}{l! k! (k+l)!} \\ &= \sum_{l=1}^{+\infty} \frac{f^{(l)} i^l}{l!} \sum_{k=0}^{+\infty} \frac{s_j^{(k)} i^k \delta^{(k+l)}(q)}{k!} \\ &= \sum_{l=1}^{+\infty} \frac{f^{(l)} i^l}{l!} \frac{d^l}{dq^l} \sum_{k=0}^{+\infty} \frac{s_j^{(k)} i^k \delta^{(k)}(q)}{k!} \\ &= \sum_{l=1}^{+\infty} \frac{f^{(l)} i^l \omega_j^{(l)}(q; \boldsymbol{\theta})}{l!}. \end{aligned} \quad (2.56)$$

Portanto, substituindo (2.56) em (2.53), temos

$$\begin{aligned} \omega_j(q; \boldsymbol{\theta}) &= \frac{\overline{h_j(q; \boldsymbol{\theta})}}{2\pi f} - \frac{1}{f} \sum_{l=1}^{+\infty} \frac{f^{(l)} i^l \omega_j^{(l)}(q; \boldsymbol{\theta})}{l!} \\ \frac{\overline{h_j(q; \boldsymbol{\theta})}}{2\pi f} &= \omega_j(q; \boldsymbol{\theta}) + \frac{1}{f} \sum_{l=1}^{+\infty} \frac{f^{(l)} i^l \omega_j^{(l)}(q; \boldsymbol{\theta})}{l!} \\ \overline{h_j(q; \boldsymbol{\theta})} &= \overline{\Phi(D_q; \boldsymbol{\theta})} \omega_j(q; \boldsymbol{\theta}). \end{aligned} \quad (2.57)$$

Portanto, finalmente, a função peso pode ser representada como

$$\omega_j(q; \boldsymbol{\theta}) = \{\overline{\Phi(D_q; \boldsymbol{\theta})}\}^{-1} \overline{h_j(q; \boldsymbol{\theta})}. \quad (2.58)$$

Para distribuições simétricas em torno de zero, a identidade acima se reduz a

$$h_{j,2}(q; \boldsymbol{\theta}) = \Phi_2(D_q; \boldsymbol{\theta}) \omega_j(q; \boldsymbol{\theta}), \quad (2.59)$$

de modo que

$$\omega_j(q; \boldsymbol{\theta}) = \Phi_2^{-1}(D_q; \boldsymbol{\theta})h_{j,2}(q; \boldsymbol{\theta}). \quad (2.60)$$

2.3.4 Relação com a medida de informação de Fisher

Enquanto $\omega_j(q; \boldsymbol{\theta})$ é uma transformada inversa da função escore de Fisher $s_j(x, \boldsymbol{\theta})$, a função $h_j(q; \boldsymbol{\theta})$ pode ser expressa como

$$\begin{aligned} h_j(q; \boldsymbol{\theta}) &= \int \left\{ \frac{d}{d\theta_j} \ln(f(x; \boldsymbol{\theta})) \right\} f(x; \boldsymbol{\theta}) e^{iqx} dx \\ &= \int s_j(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) e^{iqx} dx. \end{aligned} \quad (2.61)$$

Assim,

$$\begin{aligned} \int_{-\infty}^{+\infty} \omega_j(q, \boldsymbol{\theta}) h_j(q, \boldsymbol{\theta}) dq &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s_j(x'; \boldsymbol{\theta}) s_j(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) e^{iq(x-x')} dq dx dx' \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s_j(x'; \boldsymbol{\theta}) s_j(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) \delta(x-x') dx dx' \\ &= \int_{-\infty}^{+\infty} s_j^2(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) dx = \text{IF}_j(\boldsymbol{\theta}), \end{aligned} \quad (2.62)$$

em que $\text{IF}_j(\boldsymbol{\theta}) = \langle s_j^2(x; \boldsymbol{\theta}) \rangle$ é a conhecida medida de informação de Fisher (Ap. D).

2.4 Considerações

Este capítulo apresentou algumas propriedades das funções características relevantes para o desenvolvimento dos capítulos subsequentes. Um olhar mais aprofundado nesse assunto requer uma visita às clássicas obras, com as de Lukacs [66], Ushakov [120], Ibragimov e Linnik [57], Feller [24] e Gnedenko e Kolmogorov [48]. Curiosamente, não encontramos nessas obras a expansão alternativa da FC na forma

$$\phi(q; \boldsymbol{\theta}) = \Phi(D_q; \boldsymbol{\theta})\delta(q),$$

em que $\Phi(D_q; \boldsymbol{\theta})$ é o polinômio característico definido em (2.16). A partir dessa expansão, mostramos que a distribuição X possui mínima divergência de Kullback-Leibler relativamente à distribuição Y se os parâmetros $\theta_j \in \boldsymbol{\theta}$ de X , $j = 1, \dots, p$, satisfazem ao sistema de equações

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \left\{ \phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0,$$

$j = 1, \dots, p$, em que

$$\omega_j(q; \boldsymbol{\theta}) = \{ \overline{\Phi(D_q; \boldsymbol{\theta})} \}^{-1} \overline{h_j(q; \boldsymbol{\theta})},$$

$h_j(q; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \phi(q; \boldsymbol{\theta})$. Além disso, mostramos na Subseção 2.3.4 que a medida de informação de Fisher (Ap. D) se relaciona com $\Phi(D_q; \boldsymbol{\theta})$ e $h_j(q; \boldsymbol{\theta})$. Se considerarmos Y como uma distribuição empírica (dados) e X como uma distribuição hipotética (modelo), veremos no Cap. 5 que o resultado acima é imediatamente aplicável para o problema de estimação de parâmetros por máxima verossimilhança. E assim será possível medir eficientemente a distância entre uma distribuição de dados e a hipótese de (quase) estabilidade dessa distribuição (Cap. 5); e também a distância entre os dados e a hipótese de independência (Cap. 6). O capítulo que se segue se destina à apresentação das distribuições estáveis e quase estáveis.

Capítulo 3

As distribuições infinitamente divisíveis e as estáveis

3.1 Introdução

O conceito de divisibilidade infinita introduzido por B. de Finetti em 1924 [18, 19] permite definir uma classe de distribuições geradas por somas de variáveis aleatórias IID. Fazem parte dessa classe várias distribuições como por exemplo a Binomial, a de Poisson, a gama, a χ^2 , a gaussiana e as distribuições de Lévy estáveis. Em particular, uma distribuição infinitamente divisível X é estável se — a menos de uma transposição e da escala — as propriedades distribucionais são preservadas após convoluções de cópias independentes de X [3, 103, 104].

Uma propriedade marcante das distribuições estáveis é que suas caudas seguem uma lei de potência na forma $f(|x|) \propto |x|^{-(\alpha+1)}$ [2, 61, 103], em que $0 < \alpha \leq 2$ e $|x|$ é um valor extremo da distribuição. Como consequência, se $\alpha < 2$, $\langle |X|^q \rangle = \infty$, se $q \geq \alpha$, enquanto $\langle |X|^q \rangle < \infty$, se $q < \alpha$. Portanto, nesse caso, fenômenos descritos por distribuições estáveis não possuem escala característica nem segundo momento; e se $\alpha < 1$ tampouco a média existe. Por outro lado, se $\alpha = 2$, a distribuição estável é gaussiana e, assim, $\langle |X|^q \rangle < \infty$ para qualquer $q \in \mathbb{R}^+$.

Por isso, a teoria das distribuições estáveis remete a um teorema limite central generalizado, já que a distribuição de Lévy estável é a forma assintótica de uma soma de variáveis aleatórias independentes com variâncias não necessariamente finitas [21]. Enquanto o teorema limite central clássico permite descrever a distribuição limite de uma soma (ponderada ou não) dos elementos de uma amostra aleatória retirada de qualquer distribuição que possua momentos finitos, a teoria das distribuições estáveis permite contemplar as distribuições cujos momentos absolutos de ordem q não são necessariamente finitos. Assim, essa teoria permite descrever e modelar fenômenos sujeitos a grandes flutuações.

Inicialmente, na Seção 3.2, apresentamos o conceito de divisibilidade infinita e, em seguida, com base nesse conceito, introduzimos o processo de Lévy (Seção 3.3) e as distribuições estáveis (Seção 3.4). A forma particular do polinômio característico de uma distribuição estável é discutida na Seção 3.5. As considerações relativas aos assuntos deste capítulo são apresentadas na Seção 3.6.

3.2 Distribuições infinitamente divisíveis

Uma distribuição F é infinitamente divisível se, para qualquer $n \in \mathbb{N}$, existir uma distribuição F_n tal que [57, 104]

$$F = \underbrace{F_n * F_n * \cdots * F_n}_{n \text{ vezes}}. \quad (3.1)$$

Assim, para cada n inteiro, uma variável infinitamente divisível pode ser representada como

$$X = X_{1,n} + X_{2,n} + \cdots + X_{n,n}, \quad (3.2)$$

em que $\{X_{j,n}\}_{j=1,\dots,n}$ é uma sequência de VA independentes e identicamente distribuídas (IID).

Para que $\varphi(q; \boldsymbol{\theta}) = \ln \phi(q; \boldsymbol{\theta})$ seja o expoente característico de uma variável aleatória infinitamente divisível S , é necessário e suficiente que [47, 57, 104]

$$\varphi(q; \boldsymbol{\theta}) = i\mu q - \gamma q^2 + \int_{-\infty}^{+\infty} (e^{iqu} - 1 - iquM(u))\vartheta(u)du, \quad (3.3)$$

em que $M(u)$ é uma função limitada que satisfaz:

$$M(u) = O(1/|u|), \text{ para } |u| \rightarrow \infty; \quad (3.4)$$

$$M(u) = 1 + o(|u|), \text{ para } |u| \rightarrow 0, \quad (3.5)$$

e $\vartheta(u)$, chamada medida de Lévy, é tal que $\vartheta(0) = 0$ e

$$\int (|u|^2 \wedge 1) \vartheta(u) du < \infty. \quad (3.6)$$

A Eq. (3.3) é denominada fórmula de Lévy-Khintchine e um esboço da sua demonstração se encontra no Ap. C [47]. A tripla (μ, γ, ϑ) denomina-se *característica* da distribuição de probabilidade da variável aleatória X . A escolha da função $M(u)$ depende da conveniência; entre as formas encontradas na literatura temos, por exemplo, $M(u) = 1/(1 + x^2)$ [47, 57, 103], $M(u) = I(|u| \leq 1)$ [3, 104] e $M(u) = (\text{sen } x)/x$ [24, 104]. Pela simplicidade, para o desenvolvimento subsequente deste capítulo escolhemos a função indicadora $M(u) = I(|u| \leq 1)$.

Exemplo 3.2.1.

- Se a característica de X for $(\mu, \sigma^2/2, 0)$, então X é gaussiana com média μ e desvio padrão σ . Outra possibilidade de caracterização será apresentada na seção 3.4.
- Considerando a tripla $(\lambda, 0, \lambda\delta(u - 1))$, tem-se que X segue uma distribuição de Poisson cuja média é λ .

□

Como a distribuição infinitamente divisível pode ser definida como uma soma de VA IID (3.2), é natural associá-la com o processo estocástico de incrementos estacionários e independentes apresentado a seguir.

3.3 O processo de Lévy

Considere que $X(t)$ representa um processo estocástico em tempo contínuo, $t \geq 0$, cujo espaço de estados é real; e que a diferença $X(t) - X(s)$, em que $0 \leq s \leq t < \infty$,

representa um incremento do processo. Um processo de Lévy é aquele que satisfaz as seguintes propriedades [3, 103, 104]:

1. Estacionariedade dos incrementos: as distribuições dos incrementos são invariantes a transposições no intervalo de tempo $(s, t) \rightarrow (s + h, t + h)$, $h > 0$, ou seja, $P(X(t + h) - X(s + h) \leq x) = P(X(t) - X(s) \leq x) = P(X(t - s) - X(0) \leq x)$, $x \in \mathbb{R}$.
2. Independência dos incrementos: dada uma malha temporal arbitrária, $t_0 < t_1 < \dots < t_n$, os incrementos $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ são independentes.
3. $X(0) = 0$ com probabilidade 1.
4. Continuidade estocástica: dado $\epsilon > 0$, tem-se $\lim_{t \rightarrow s} P(|X(t) - X(s)| > \epsilon) = 0$ para todo $s \geq 0$.

Assim, de um modo geral, um processo de Lévy é aquele cujos incrementos são estacionários e independentes. Os processos gaussiano e de Poisson são exemplos desse tipo de processo.

Por construção, uma VA $X(t)$ definida segundo um processo de Lévy é infinitamente divisível. Por exemplo, considere que $t_k = kt/n$, ($0 \leq k \leq n$), de modo que o intervalo de tempo $[0, t]$ seja dividido em n subintervalos comprimentos iguais a t/n . Desse modo, os incrementos $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ são IID. Logo, $X(t) = X(t_1) - X(t_0) + X(t_2) - X(t_1) + \dots + X(t_n) - X(t_{n-1})$, em que $X(t_0) = 0$ e $X(t_n) = X(t)$, é uma soma de VA IID, o que permite concluir que $X(t)$ é infinitamente divisível.

3.4 A distribuição estável

Uma distribuição estável X é uma distribuição infinitamente divisível cujas propriedades distribucionais são preservadas após convoluções de cópias independentes de X , a menos de um parâmetro de locação $\mu \in \mathbb{R}$ e de escala $\gamma > 0$. Se X_1 e X_2 são cópias independentes

de uma variável aleatória estável X , então $\gamma X + \mu = \gamma_1 X_1 + \gamma_2 X_2$, em que $\gamma_1 > 0$ e $\gamma_2 > 0$ são parâmetros de escala.

A distribuição estável X é caracterizada pela tripla $(\mu, 0, \vartheta_L(u))$, em que $\vartheta_L(u) = \frac{C^+}{|u|^{\alpha+1}}I(u > 0) + \frac{C^-}{|u|^{\alpha+1}}I(u < 0)$, com $0 < \alpha \leq 2$, $C^+ > 0$ e $C^- > 0$ [3, 103, 104]. Assim, com a ajuda dos resultados matemáticos apresentados no Ap. A, deduziremos a seguir a forma do seu expoente característico.

Se $0 < \alpha < 1$,

$$\begin{aligned} \varphi(q; \boldsymbol{\theta}) &= i\mu q + \int_{-\infty}^{+\infty} (e^{iqu} - 1 - iquI(|u| \leq 1))\vartheta_L(u)du \\ &= i\mu q + \int_{-\infty}^{+\infty} (e^{iqu} - 1)\vartheta_L(u)du - \int_{-1}^{+1} iqu\vartheta_L(u)du \\ &= i\mu q + |q|^\alpha \Gamma(-\alpha) \{ (C^+ + C^-) \cos(\pi\alpha/2) - i(C^+ - C^-) \operatorname{sgn}(q) \operatorname{sen}(\pi\alpha/2) \} + 0 \\ &= i\mu q + |q|^\alpha \Gamma(-\alpha) (C^+ + C^-) \cos(\pi\alpha/2) \left\{ 1 - i \frac{C^+ - C^-}{C^+ + C^-} \operatorname{sgn}(q) \tan(\pi\alpha/2) \right\} \\ &= i\mu q - \gamma |q|^\alpha \{ 1 - i\beta \operatorname{sgn}(q) \tan(\pi\alpha/2) \}, \end{aligned}$$

em que $\gamma > 0$, pois $\Gamma(-\alpha) < 0$ e $(C^+ + C^-) \cos(\pi\alpha/2) > 0$, $\beta = \frac{C^+ - C^-}{C^+ + C^-}$ e

$$\operatorname{sgn}(q) = \begin{cases} +1 & \text{se } q > 0, \\ 0 & \text{se } q = 0, \\ -1 & \text{se } q < 0. \end{cases} \quad (3.7)$$

Se $1 < \alpha < 2$,

$$\begin{aligned} \varphi(q; \boldsymbol{\theta}) &= i\mu q + \int_{-\infty}^{+\infty} (e^{iqu} - 1 - iquI(|u| \leq 1))\vartheta_L(u)du \\ &= i\mu q + \int_{-\infty}^{+\infty} (e^{iqu} - 1 - iqu)\vartheta_L(u)du + iC^+ \int_1^{+\infty} \frac{qu}{|u|^{\alpha+1}} du + iC^- \int_{-\infty}^{-1} \frac{qu}{|u|^{\alpha+1}} du \\ &= i\mu q + |q|^\alpha \Gamma(-\alpha) \{ (C^+ + C^-) \cos \frac{\pi\alpha}{2} - i(C^+ - C^-) \operatorname{sgn}(q) \operatorname{sen} \frac{\pi\alpha}{2} \} + iq \frac{C^+ + C^-}{\alpha - 1} \\ &= i\mu^* q - \gamma |q|^\alpha \left\{ 1 - i\beta \operatorname{sgn}(q) \tan \frac{\pi\alpha}{2} \right\}, \end{aligned}$$

em que $\gamma > 0$ e $\mu^* = \mu + (C^+ + C^-)/(\alpha - 1)$ é o parâmetro de locação com *drift*.

Finalmente, se $\alpha = 1$,

$$\begin{aligned}
 \varphi(q; \boldsymbol{\theta}) &= i\mu q + \int_{-\infty}^{+\infty} (e^{iqu} - 1 - iquI(|u| \leq 1))\vartheta_L(u)du \\
 &= i\mu q + (i\mu_0 q - \frac{\pi|q|}{2} - i|q|\operatorname{sgn}(q) \ln |q|)C^+ + (i\mu_0 q - \frac{\pi|q|}{2} + i|q|\operatorname{sgn}(q) \ln |q|)C^- \\
 &= i\{\mu + \mu_0(C^+ + C^-)\}q - \frac{\pi(C^+ + C^-)}{2}|q| - i|q|\operatorname{sgn}(q) \ln |q|(C^+ - C^-) \\
 &= i\mu^{**} - \gamma|q|(1 + i\beta\frac{2}{\pi}\operatorname{sgn}(q) \ln |q|),
 \end{aligned}$$

em que $\gamma > 0$, $\mu_0 = \int_1^{+\infty} u^{-2}\operatorname{sen}u du + \int_0^1 u^{-2}(\operatorname{senu} - u)du$ e $\mu^{**} = \mu + \mu_0(C^+ + C^-)$.

Portanto, uma VA estável X é caracterizada pelo vetor de parâmetros $\boldsymbol{\theta} = (\alpha, \gamma, \beta, \mu)'$ e possui uma função característica na forma [3, 75, 103, 104]

$$\phi(q; \boldsymbol{\theta}) = \begin{cases} \exp \left\{ -\gamma|q|^\alpha \left(1 - i\beta\operatorname{sgn}(q) \tan \frac{\pi\alpha}{2}\right) + i\mu_c q \right\} & \text{se } \alpha \neq 1, \\ \exp \left\{ -\gamma|q| \left(1 + i\beta\frac{2}{\pi}\operatorname{sgn}(q) \ln |q|\right) + i\mu_c q \right\} & \text{se } \alpha = 1. \end{cases} \quad (3.8)$$

O parâmetro α ($0 < \alpha \leq 2$) é denominado índice de estabilidade (ou parâmetro de forma), γ é o parâmetro de escala ($\gamma > 0$), β representa o parâmetro de assimetria ($|\beta| \leq 1$) e μ_c é o parâmetro de locação (ou deslocamento ou centralidade). Caso $\alpha = 2$, a distribuição é gaussiana, o parâmetro γ corresponde à metade da variância da distribuição, μ_c representa a média e $\beta \equiv 0$ (pois este torna-se irrelevante). Se $\alpha = 1$ e $\beta = 0$, a distribuição resultante é a de Cauchy.

Padronização

Considere a transformação de escala e centralização em zero

$$Z = \frac{X - \mu_c}{\gamma^{\frac{1}{\alpha}}}. \quad (3.9)$$

Para $\alpha \neq 1$, a função característica resultante tem a forma da distribuição estável com $\gamma = 1$ e $\mu_c = 0$, pois

$$\phi_Z(q; \boldsymbol{\theta}_Z) = \phi_X(\gamma^{-\frac{1}{\alpha}}q; \boldsymbol{\theta})e^{-iq\mu_c} \quad (3.10)$$

$$= \exp \left\{ -|q|^\alpha \left(1 - i\beta\operatorname{sgn}(q) \tan \frac{\pi\alpha}{2}\right) \right\}, \quad (3.11)$$

em que $\boldsymbol{\theta}_Z = (\alpha, \beta)'$. Nesse caso, a função de densidade da variável padronizada Z se relaciona com a distribuição X mediante a transformação de escala

$$f_X(x; \boldsymbol{\theta}) = \gamma^{-\frac{1}{\alpha}} f_Z(\gamma^{-\frac{1}{\alpha}}(x - \mu_c); \boldsymbol{\theta}_Z). \quad (3.12)$$

Agora, para $\alpha = 1$,

$$\phi_Z(q; \boldsymbol{\theta}_Z) = \phi_X(\gamma^{-1}q; \boldsymbol{\theta}) e^{-iq\mu_c} \quad (3.13)$$

$$= \exp \left\{ -|q| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(q) \ln |\gamma^{-1}q| \right) \right\} \quad (3.14)$$

$$= \exp \left\{ -|q| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(q) \ln |\gamma^{-1}q| \right) + i\beta \frac{2}{\pi} \ln |\gamma|q \right\}, \quad (3.15)$$

em que $\boldsymbol{\theta}_Z = (1, 1, \beta, \beta \frac{2}{\pi} \ln |\gamma|)'$. Assim, se $\beta \neq 0$ (Cauchy assimétrica), a padronização curiosamente produz um *drift* que depende do coeficiente de assimetria β e do parâmetro de escala γ .

Simetrização

Caso seja de interesse, é possível destacar apenas os parâmetros α e γ mediante simetrização. A função característica de uma variável simetrizada é dada por $|\phi(q; \boldsymbol{\theta})|^2$. Para uma distribuição estável, temos

$$|\phi(q; \boldsymbol{\theta})| = \exp\{-\gamma|q|^\alpha\}. \quad (3.16)$$

Logo, $|\phi(q; \boldsymbol{\theta})|$ é a função característica de uma distribuição estável simétrica em torno da origem parametrizada por $\boldsymbol{\theta}_s = (\alpha, \gamma)'$.

Assimetria efetiva

Se $\alpha \neq 1$, define-se a assimetria efetiva como

$$\beta_\alpha = \beta \tan \frac{\pi\alpha}{2}, \quad (3.17)$$

pois o efeito do coeficiente de assimetria β na forma da distribuição depende de α . Por exemplo, se $\alpha = 2$, tem-se que $\beta_{\alpha=2} = 0$, e à medida que $\alpha \rightarrow 1$, a forma da distribuição

torna-se mais sensível ao parâmetro β . O painel superior da Fig. 3.1, que ilustra o comportamento da assimetria efetiva para o caso $\beta = -1$, mostra que o efeito de β_α é bastante significativo à medida que $\alpha \rightarrow 1$. O painel inferior da Fig. 3.1 mostra o comportamento de β_α para $\alpha = 1, 1 a 2$ e $-1 \leq \beta \leq 0$ (com incrementos iguais a 0,1).

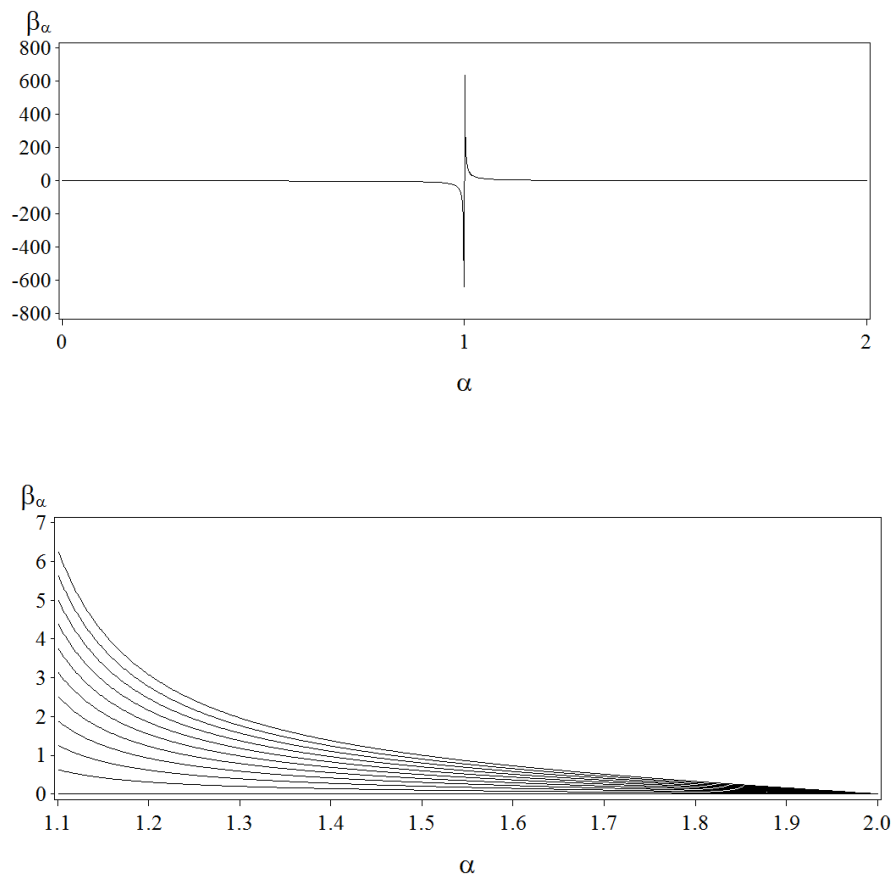


Figura 3.1: Painel superior: comportamento da assimetria efetiva $\beta_\alpha = \beta \tan \frac{\pi\alpha}{2}$ para $\beta = -1$ e $\alpha = 0,001$ a 2 (com incrementos iguais a 0,001). Painel inferior: comportamento da assimetria efetiva β_α para $\beta = -1$ (curva superior) a 0 (com incrementos iguais a 0,1) e $\alpha \geq 1,1$.

3.4.1 O processo de Lévy estável

Considere um processo de Lévy $\{X(t)\}$, em que $X(t)$ segue uma distribuição estável com $\theta_0 = (\alpha, \gamma t, \beta)'$ definida por uma integral estocástica na forma

$$X(t) = \int_0^t Y(u)du, \quad (3.18)$$

cuja FC é dada por

$$\phi_t(q; \theta_0) = \begin{cases} \exp \left\{ -\gamma t |q|^\alpha (1 - i\beta_\alpha \operatorname{sgn}(q)) \right\} & \text{se } \alpha \neq 1, \\ \exp \left\{ -\gamma t |q| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(q) \ln |q| \right) \right\} & \text{se } \alpha = 1. \end{cases} \quad (3.19)$$

Como esse processo é infinitamente divisível, tomando-se uma malha temporal de n intervalos igualmente espaçados, pode-se representá-lo como uma soma na forma

$$\begin{aligned} X(t) &= \int_0^{t/n} Y(u)du + \int_{t/n}^{2t/n} Y(u)du + \cdots + \int_{(n-1)t/n}^t Y(u)du \\ &= X_1(t/n) + X_2(t/n) + \cdots + X_n(t/n), \end{aligned} \quad (3.20)$$

em que $X_1(t/n), X_2(t/n), \dots, X_n(t/n)$ são cópias IID da VA $X(t/n)$ cuja FC é dada por

$$\phi_{t/n}(q; \theta_0) = \begin{cases} \exp \left\{ -\gamma \frac{t}{n} |q|^\alpha (1 - i\beta_\alpha \operatorname{sgn}(q)) \right\} & \text{se } \alpha \neq 1, \\ \exp \left\{ -\gamma \frac{t}{n} |q| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(q) \ln |q| \right) \right\} & \text{se } \alpha = 1. \end{cases} \quad (3.21)$$

Uma característica importante deste processo é sua autosimilaridade, ou seja, $\{X(t/n)\}$ e $\{(1/n)^{\frac{1}{\alpha}} X(t)\}$ são estocasticamente idênticos [2]. Por exemplo, com $t = \Delta t$ tem-se um passeio aleatório

$$S_{\Delta t} = X_1 + X_2 + \cdots + X_{\Delta t}, \quad (3.22)$$

cujas componentes $X_1, X_2, \dots, X_{\Delta t}$ são cópias IID de $X(1)$. Nesse caso, $n^{\frac{1}{\alpha}} X(\Delta t) \stackrel{D}{=} X(1)$.

3.5 O polinômio característico

Nesta seção, discute-se sobre o polinômio característico das distribuições estáveis. Conforme o Cap. 2, se a FD f de uma variável aleatória X for indefinidamente derivável no

ponto zero, a FC pode ser representada como

$$\phi(q; \boldsymbol{\theta}) = \Phi(D_q; \boldsymbol{\theta})\delta(q), \quad (3.23)$$

em que $D_q^k = \frac{d^k}{dq^k}$ representa o operador diferencial,

$$\Phi(D_q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} c_k D_q^k \quad (3.24)$$

é denominado polinômio característico de $\phi(q; \boldsymbol{\theta})$ com

$$c_k = c_k(\boldsymbol{\theta}) = \frac{2\pi(-i)^k f^{(k)}}{k!}. \quad (3.25)$$

Estudaremos agora o comportamento dos coeficientes c_k para os casos que X segue uma distribuição estável.

Caso simétrico

Na situação em que $\mu_c = 0$ e $\beta = 0$, ou seja, $\boldsymbol{\theta}_s = (\alpha, \gamma)'$, a função característica assume a forma

$$\phi(q; \boldsymbol{\theta}_s) = e^{-\gamma|q|^\alpha}, \quad (3.26)$$

e sua função de densidade, obtida mediante a transformação inversa (2.5), é dada por

$$f(x; \boldsymbol{\theta}_0) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma|q|^\alpha} \cos(qx) dq, \quad (3.27)$$

cujas derivadas de ordem ímpar e par com respeito a x (Eqs. (2.18) e (2.19)) são, respectivamente,

$$f^{(2k+1)} = 0, \quad (3.28)$$

$$f^{(2k)} = \frac{(-1)^k \Gamma\left(\frac{2k+1}{\alpha}\right)}{\pi \alpha \gamma^{\frac{2k+1}{\alpha}}}, \quad (3.29)$$

para $k = 0, 1, 2, \dots$. Como forma alternativa, a FC pode ser escrita na forma (2.24) como

$$\phi(q; \boldsymbol{\theta}_s) = \Phi_2(D_q; \boldsymbol{\theta}_s)\delta(q), \quad (3.30)$$

em que

$$\Phi_2(D_q; \boldsymbol{\theta}_s) = \sum_{k=0}^{+\infty} c_{2k} D_q^{2k}, \quad (3.31)$$

com

$$c_{2k}(\boldsymbol{\theta}_s) = \frac{2}{\alpha \gamma^{\frac{2k+1}{\alpha}}} \frac{\Gamma(\frac{2k+1}{\alpha})}{\Gamma(2k+1)}. \quad (3.32)$$

Logo, a razão entre as duas funções Gamma na Eq. (3.32), essencialmente, caracteriza a distribuição estável simétrica em torno de zero. Desse modo, para um passeio aleatório $S_{\Delta t}$, o coeficiente c_{2k} segue uma lei de potência na forma $\Delta t^{-\frac{2k+1}{\alpha}}$. O caso $k = 0$ foi discutido por Mantegna e Stanley [75].

Exemplo 3.5.1. Para o caso lorentziano ($\alpha = 1$) tem-se

$$c_{2k}(1, \gamma) = \frac{2}{\gamma^{2k+1}}. \quad (3.33)$$

Daí, assumindo-se que $\left| \frac{D_q^2}{\gamma^2} \right| < 1$, o polinômio característico $\Phi_2(D_q; \boldsymbol{\theta}_s)$ é uma progressão geométrica convergente, de modo que

$$\Phi_2(D_q; \boldsymbol{\theta}_s) = \frac{2\gamma}{\gamma^2 - D_q^2}. \quad (3.34)$$

Já para o caso Gaussiano ($\alpha = 2$),

$$\begin{aligned} c_{2k}(2, \gamma) &= \frac{1}{\gamma^{k+\frac{1}{2}}} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(2k+1)} \\ &= \frac{\sqrt{\pi}}{\gamma^{k+\frac{1}{2}} 4^k k!}. \end{aligned} \quad (3.35)$$

Nesse caso, o coeficiente c_{2k} representa um termo da expansão da função exponencial, de modo que o polinômio característico pode ser representado como

$$\Phi_2(D_q; \boldsymbol{\theta}_s) = \sqrt{\frac{\pi}{\gamma}} \exp \frac{D_q^2}{4\gamma}. \quad (3.36)$$

Para as distribuições estáveis simétricas, tem-se $c_{2k} > 0$. Para a distribuição de Cauchy, o logaritmo de c_{2k} apresenta um padrão linear, ou seja, $\ln(c_{2k}) = \ln 2 - (2k+1) \ln \gamma$, enquanto para o caso gaussiano, $\ln(c_{2k}) = \frac{1}{2} \ln \pi - (k + \frac{1}{2}) \ln \gamma - \ln k! - k \ln 4$.

□

Para os demais casos, considerando o resultado (A.18), pode-se aproximar assintoticamente o logaritmo da razão $\Gamma(\frac{2k+1}{\alpha})/\Gamma(2k+1)$ como

$$\begin{aligned} \ln \frac{\Gamma(\frac{2k+1}{\alpha})}{\Gamma(2k+1)} &\approx \left(\frac{2k+1}{\alpha} - \frac{1}{2}\right) \ln \frac{2k+1}{\alpha} - \frac{2k+1}{\alpha} - \left(2k + \frac{1}{2}\right) \ln(2k+1) + 2k+1 \\ &\approx -(2k+1)\left(1 - \frac{1}{\alpha}\right) \ln(2k+1) - \left(\frac{2k+1}{\alpha} - \frac{1}{2}\right) \ln \alpha + (2k+1)\left(1 - \frac{1}{\alpha}\right), \end{aligned}$$

o que resulta em

$$c_{2k}(\alpha, \gamma) \approx \frac{2}{\sqrt{\alpha}} \left(\frac{e}{2k+1}\right)^{2k+1} \left(\frac{2k+1}{e\alpha\gamma}\right)^{\frac{2k+1}{\alpha}}. \quad (3.37)$$

A Figura 3.2 mostra uma comparação entre os logaritmos de c_{2k} e da forma (3.37), indicando que a aproximação sugerida é satisfatória. Assim, para o caso geral, $\ln c_{2k} \propto \frac{(2k+1)}{\alpha}((1-\alpha) \ln(2k+1) - \ln \gamma)$. A Fig. 3.3 mostra o comportamento de c_{2k} para $\gamma = 0, 5$ e 50 e $\alpha \geq 1$. Embora a sequência $\{c_{2k}\}$ possa divergir, é possível obter coeficientes convergentes mediante transformação de escala. Com base na Eq. (3.32) ou na sua forma aproximada (3.37), tem-se

$$c_{2k}(\alpha, \gamma) = \gamma_0^{\frac{2k+1}{\alpha}} c_{2k}(\alpha, \gamma\gamma_0), \quad (3.38)$$

em que γ_0 é uma constante tal que se tenha $c_{2k}(\alpha, \gamma\gamma_s) \rightarrow 0$.

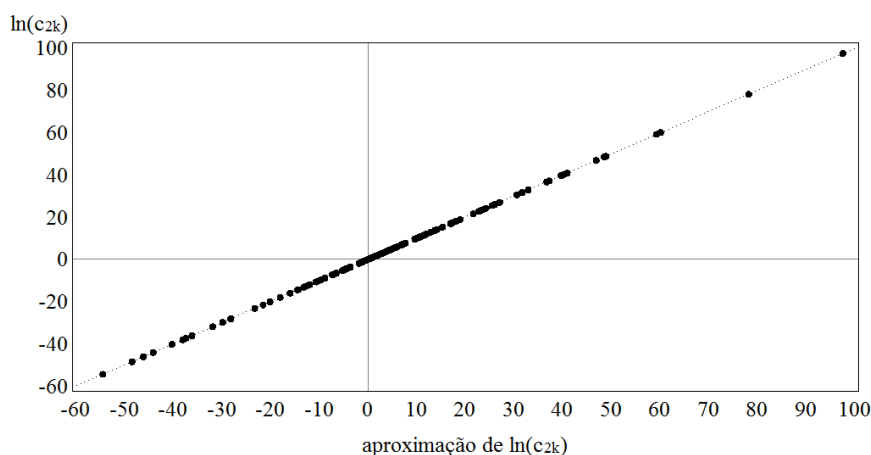


Figura 3.2: Comparação entre $\ln c_{2k}$ e suas respectivas aproximações ($k = 0, \dots, 5$) para $\gamma = 0.05, 0.5, 1, 1.5, 5, 50$ e $0 < \alpha \leq 2.0$.

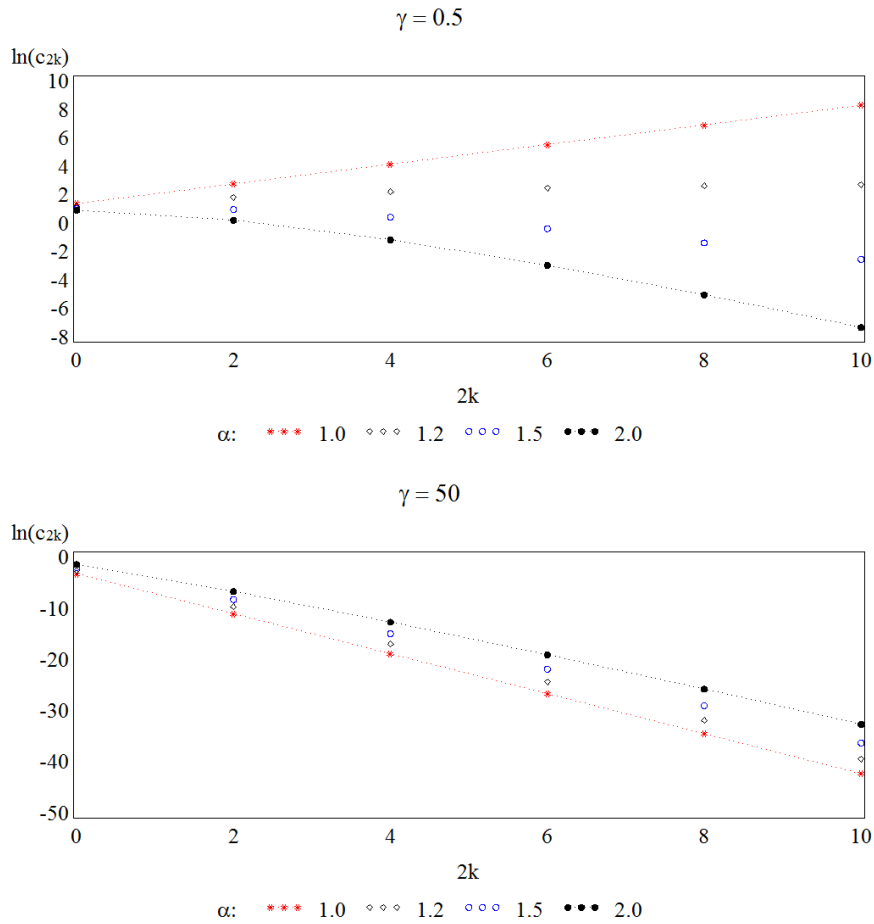


Figura 3.3: logaritmo natural dos coeficientes c_{2k} , $k = 0, \dots, 5$, para $\gamma = 0, 5$ e 50 , $\alpha \geq 1$. As linhas pontilhadas são referenciais obtidos com base nos casos gaussiano ($\alpha = 2$) e lorentziano ($\alpha = 1$).

Caso assimétrico com $\mu_c = 0$ e $\alpha = 1$

Para o caso lorentziano assimétrico, tem-se que

$$\phi(q; \boldsymbol{\theta}) = \exp \left\{ -\gamma |q| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(q) \ln |q| \right) \right\}, \quad (3.39)$$

e agora a FD é dada por

$$\begin{aligned}
 f(x; \boldsymbol{\theta}) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|(1+i\beta\frac{2}{\pi}\text{sgn}(q)\ln|q|)} e^{-iqx} dq \\
 &= \frac{1}{2\pi} \left\{ \int_0^{+\infty} e^{-\gamma|q|(1+i\beta\frac{2}{\pi}\ln|q|)} e^{-iqx} dq + \int_0^{+\infty} e^{-\gamma|q|(1-i\beta\frac{2}{\pi}\ln|q|)} e^{+iqx} dq \right\} \\
 &= \frac{1}{2\pi} \int_0^{+\infty} e^{-\gamma|q|} \left\{ e^{-i\gamma|q|\beta\frac{2}{\pi}\ln|q|} e^{-iqx} + e^{+i\gamma|q|\beta\frac{2}{\pi}\ln|q|} e^{+iqx} \right\} dq \\
 &= \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma|q|} \cos\left(\gamma|q|\beta\frac{2}{\pi}\ln|q| + qx\right) dq.
 \end{aligned}$$

As derivadas de ordem ímpar e par com respeito a x em zero são, respectivamente,

$$f^{(2k+1)} = \frac{(-1)^{k+1}}{\pi} \int_0^{+\infty} q^{2k+1} e^{-\gamma|q|} \text{sen}\left(\gamma|q|\beta\frac{2}{\pi}\ln|q|\right) dq \quad (3.40)$$

e

$$f^{(2k)} = \frac{(-1)^k}{\pi} \int_0^{+\infty} q^{2k} e^{-\gamma|q|} \cos\left(\gamma|q|\beta\frac{2}{\pi}\ln|q|\right) dq, \quad (3.41)$$

de modo que

$$c_{2k+1} = i \frac{2}{\Gamma(2(k+1))\gamma^{2(k+1)}} \int_0^{+\infty} q^{2k+1} e^{-q} \text{sen}\left(q\beta\frac{2}{\pi}\ln\frac{q}{\gamma}\right) dq \quad (3.42)$$

e

$$c_{2k} = \frac{2}{\Gamma(2k+1)\gamma^{2k+1}} \int_0^{+\infty} q^{2k} e^{-q} \cos\left(q\beta\frac{2}{\pi}\ln\frac{q}{\gamma}\right) dq. \quad (3.43)$$

As integrais acima devem ser calculadas numericamente. Neste trabalho, porém, não consideraremos o caso $\alpha = 1$, já que situações em que $\alpha \approx 1$ são contempladas no caso que se segue.

Caso assimétrico com $\mu_c = 0$ e $\alpha \neq 1$

Nessa situação,

$$\phi(q; \boldsymbol{\theta}_0) = \exp\left\{-\gamma|q|^\alpha(1-i\beta_\alpha\text{sgn}(q))\right\}, \quad (3.44)$$

em que $\boldsymbol{\theta}_0 = (\alpha, \gamma, \beta)'$, e a FD é dada por

$$\begin{aligned}
 f(x; \boldsymbol{\theta}_0) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|^\alpha(1-i\beta_\alpha \operatorname{sgn}(q))} e^{-iqx} dq, \\
 &= \frac{1}{2\pi} \int_0^{+\infty} \left\{ e^{-\gamma|q|^\alpha} e^{i\gamma|q|^\alpha \beta_\alpha} e^{-iqx} + e^{-\gamma|q|^\alpha} e^{-i\gamma|q|^\alpha \beta_\alpha} e^{+iqx} \right\} dq \\
 &= \frac{1}{2\pi\alpha} \int_0^{+\infty} q^{\frac{1}{\alpha}-1} e^{-\gamma q} \left\{ e^{i\gamma\beta_\alpha q} e^{-iq\frac{1}{\alpha}x} + e^{-i\gamma\beta_\alpha q} e^{+iq\frac{1}{\alpha}x} \right\} dq \\
 &= \frac{1}{\pi\alpha} \int_0^{+\infty} q^{\frac{1}{\alpha}-1} e^{-\gamma q} \cos\left(q\gamma\beta_\alpha - q\frac{1}{\alpha}x\right) dq. \tag{3.45}
 \end{aligned}$$

As derivadas de ordem ímpar e par com respeito a x em zero são, respectivamente,

$$\begin{aligned}
 f^{(2k+1)} &= \frac{(-1)^k}{\pi\alpha} \int_0^{+\infty} q^{\frac{2(k+1)}{\alpha}-1} e^{-\gamma q} \operatorname{sen}(q\gamma\beta_\alpha) dq \\
 &= \frac{(-1)^k \Gamma\left(\frac{2(k+1)}{\alpha}\right)}{\pi\alpha\gamma^{\frac{2(k+1)}{\alpha}}} \cdot \frac{\operatorname{sen}\left(\frac{2(k+1)}{\alpha} \arctan(\beta_\alpha)\right)}{(1 + \beta_\alpha^2)^{\frac{k+1}{\alpha}}} \\
 &= \frac{(-1)^k \Gamma\left(\frac{2k+2}{\alpha}\right)}{\pi\alpha\gamma_*^{\frac{2k+2}{\alpha}}} \cdot A_{1,k} \tag{3.46}
 \end{aligned}$$

e

$$\begin{aligned}
 f^{(2k)} &= \frac{(-1)^k}{\pi\alpha} \int_0^{+\infty} q^{\frac{2k+1}{\alpha}-1} e^{-\gamma q} \cos(q\gamma\beta_\alpha) dq \\
 &= \frac{(-1)^k \Gamma\left(\frac{2k+1}{\alpha}\right)}{\pi\alpha\gamma^{\frac{2k+1}{\alpha}}} \cdot \frac{\cos\left(\frac{2k+1}{\alpha} \arctan(\beta_\alpha)\right)}{(1 + \beta_\alpha^2)^{\frac{2k+1}{2\alpha}}} \\
 &= \frac{(-1)^k \Gamma\left(\frac{2k+1}{\alpha}\right)}{\pi\alpha\gamma_*^{\frac{2k+1}{\alpha}}} \cdot A_{2,k}, \tag{3.47}
 \end{aligned}$$

em que $k = 0, 1, 2, \dots$,

$$\gamma_* = \gamma \cdot \sqrt{1 + \beta_\alpha^2}, \tag{3.48}$$

$$A_{1,k} \equiv A_1(k; \alpha, \beta) = \operatorname{sen}\left\{\frac{2k+2}{\alpha} \arctan(\beta_\alpha)\right\}, \tag{3.49}$$

e

$$A_{2,k} \equiv A_2(k; \alpha, \beta) = \cos\left\{\frac{2k+1}{\alpha} \arctan(\beta_\alpha)\right\}. \tag{3.50}$$

Com base em (2.12), os coeficientes do polinômio característico são

$$c_{2k+1}(\boldsymbol{\theta}_0) = i \frac{2\Gamma\left(\frac{2(k+1)}{\alpha}\right)}{\Gamma(2(k+1))\alpha\gamma_*^{\frac{2(k+1)}{\alpha}}} \cdot A_{1,k} \quad (3.51)$$

e

$$c_{2k}(\boldsymbol{\theta}_0) = \frac{2\Gamma\left(\frac{2k+1}{\alpha}\right)}{\Gamma(2k+1)\alpha\gamma_*^{\frac{2k+1}{\alpha}}} \cdot A_{2,k}. \quad (3.52)$$

As componentes $A_{1,k}$ e $A_{2,k}$ definem o padrão assimétrico da distribuição e não dependem do parâmetro de escala γ . Assim, a partir dessas equações tem-se que os coeficientes ímpares e pares da distribuição padronizada ($\gamma = 1$) se relacionam com os da distribuição X , respectivamente, mediante as transformações de escala

$$c_{2k+1}(\alpha, \gamma, \beta) = \gamma_0^{\frac{2k+2}{\alpha}} c_{2k+1}(\alpha, \gamma\gamma_0, \beta). \quad (3.53)$$

$$c_{2k}(\alpha, \gamma, \beta) = \gamma_0^{\frac{2k+1}{\alpha}} c_{2k}(\alpha, \gamma\gamma_0, \beta), \quad (3.54)$$

em que $\gamma_0 > 0$.

Além disso, é interessante observar que $c_{2k}(\boldsymbol{\theta}_0)$ se relaciona com o caso simétrico $c_{2k}(\boldsymbol{\theta}_s)$, em que $\boldsymbol{\theta}_s = (\alpha, \gamma)'$; isto é,

$$c_{2k}(\alpha, \gamma, \beta) = c_{2k}(\alpha, \gamma) \frac{A_{2,k}}{(1 + \beta^2_{\alpha})^{\frac{2k+1}{2\alpha}}}. \quad (3.55)$$

Para o caso particular em que X segue uma distribuição simétrica e estável em torno de zero ($\beta = 0$), o parâmetro γ_* coincide com o de escala γ , $A_{1,k} = 0$ e $A_{2,k} = 1$ para todo $k \geq 0$. Nesse caso,

$$f^{(2k+1)}(0; \boldsymbol{\theta}_s) = 0 \quad (3.56)$$

e

$$f^{(2k)}(0; \boldsymbol{\theta}_s) = \frac{(-1)^k \Gamma\left(\frac{2k+1}{\alpha}\right)}{\pi \alpha \gamma^{\frac{2k+1}{\alpha}}}. \quad (3.57)$$

Mas para o caso assimétrico, temos $\gamma_* > \gamma$ e $A_{1,k}$ e $A_{2,k}$ não são funções constantes. Isso sugere que se o coeficiente de assimetria for ignorado em uma análise de dados assimétricos, a estimativa do fator de escala poderá ser superestimada. O estudo sobre o comportamento

dessas funções pode ajudar a descrever o efeito da assimetria nas leis de potência dos passeios aleatórios. As Figuras 3.4, 3.5 e 3.6 exemplificam, respectivamente, gráficos de $A_{2,0}$, $A_{2,1}$ e $A_{1,0}$ em função de β para alguns valores α . Eles mostram que $A_{2,0}$ e $A_{2,1}$ são funções pares com respeito a β , enquanto $A_{1,0}$ é uma função ímpar. A função $A_{2,0}$ é não negativa para todo $0 < \alpha \leq 2$, enquanto $A_{2,1} > 0$ para qualquer $|\beta| \leq 1$ apenas se $1,5 \leq \alpha \leq 2$.

Como $0 \leq A_{2,0} \leq 1$, é possível representar $f^{(0)}$ como

$$f^{(0)}(0; \boldsymbol{\theta}_0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\pi\alpha(\gamma_{*,0})^{\frac{1}{\alpha}}}, \quad (3.58)$$

em que

$$\gamma_{*,0} = \frac{\gamma_*}{A_{2,0}^\alpha} \quad (3.59)$$

$$= \gamma \cdot \frac{\sqrt{1 + \beta_\alpha^2}}{A_{2,0}^\alpha} \equiv \gamma \cdot B_{2,0}(\alpha, \beta). \quad (3.60)$$

Embora não sejam equivalentes, o aspecto da Eq. (3.58) é semelhante ao da equação correspondente no caso simétrico, Eq. (3.57) com $k = 0$. Por isso, se a assimetria da distribuição for ignorada, o resultado de uma análise de dados poderia levar a uma conclusão inflacionada acerca do parâmetro de escala, já que $\gamma_{*,0}$ é um parâmetro de escala inflacionado por $B_{2,0}$.

O fator de inflação $B_{2,0}$ é inversamente proporcional a $A_{2,0}$. De acordo com a Fig. 3.4, a amplitude de $A_{2,0}$ tende a aumentar à medida que α diminui, e função $A_{2,0}$ pode decrescer para zero se $\alpha < 1$. Portanto, o fator de inflação $B_{2,0}$ pode atingir grandes magnitudes se $\alpha < 1$. Mas para $1,5 \leq \alpha \leq 2$, o efeito pode ser pequeno, dependendo de β . A Fig. 3.7 ilustra o comportamento do fator de inflação $B_{2,0}$ em função de β para alguns valores de α . Para $1,5 \leq \alpha \leq 2$, observa-se que $B_{2,0} \leq 1,8$. Há uma tendência de aumento considerável à medida que $\alpha \rightarrow 1$ e $|\beta| \rightarrow 1$. Assim, por exemplo, se $\beta = 0,5$ e $\alpha = 1,7$, $\gamma_{*,0}$ é aproximadamente 5% superior ao parâmetro de escala γ , enquanto se $\beta = 0,5$ e $\alpha = 1,2$, $\gamma_{*,0}$ é quase 3 vezes γ . Pela Figura 3.7, nota-se também que o aumento de $B_{2,0}$ tende a ser explosivo à medida que $\alpha \rightarrow 1$ para o caso em que $\alpha < 1$. Por exemplo, para $\beta = 0,5$ e $\alpha = 1,05$, $\gamma_{*,0} \approx 31\gamma$ enquanto para $\beta = 0,5$ e $\alpha = 0,95$, $\gamma_{*,0} \approx 70\gamma$.

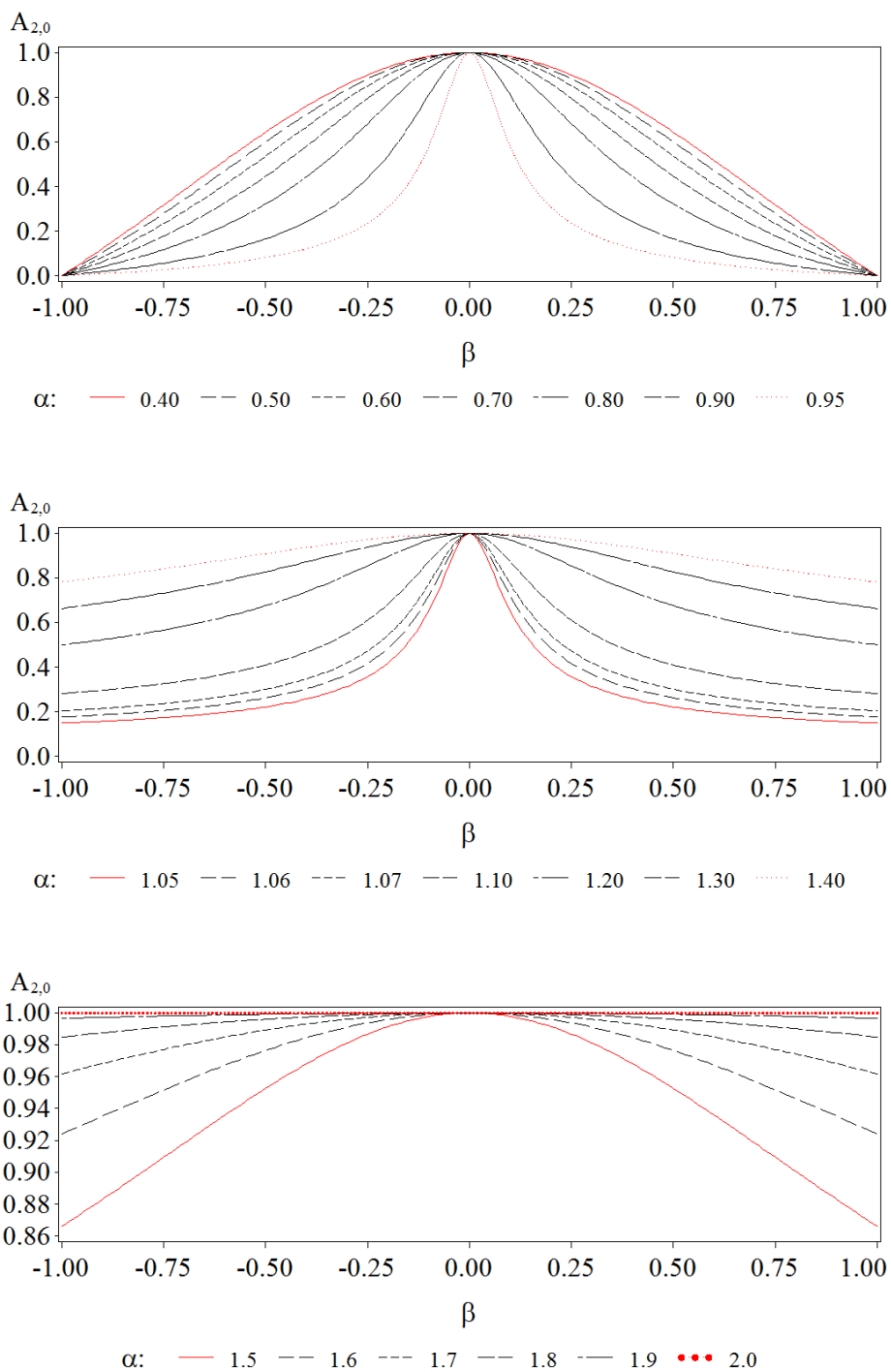


Figura 3.4: Comportamento de $A_{2,0}$ em função de β para alguns valores de α .

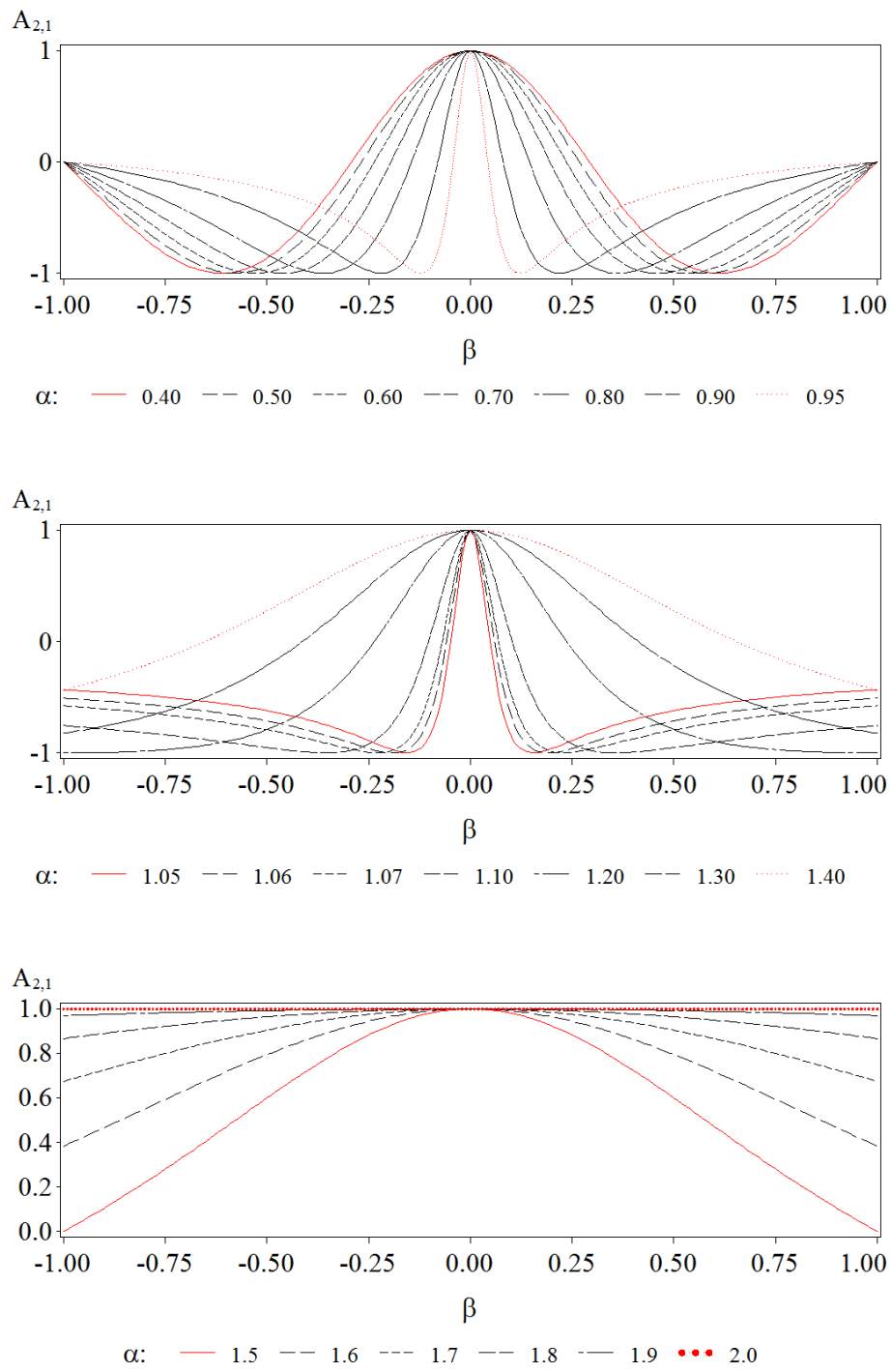


Figura 3.5: Comportamento de $A_{2,1}$ em função de β para alguns valores de α .

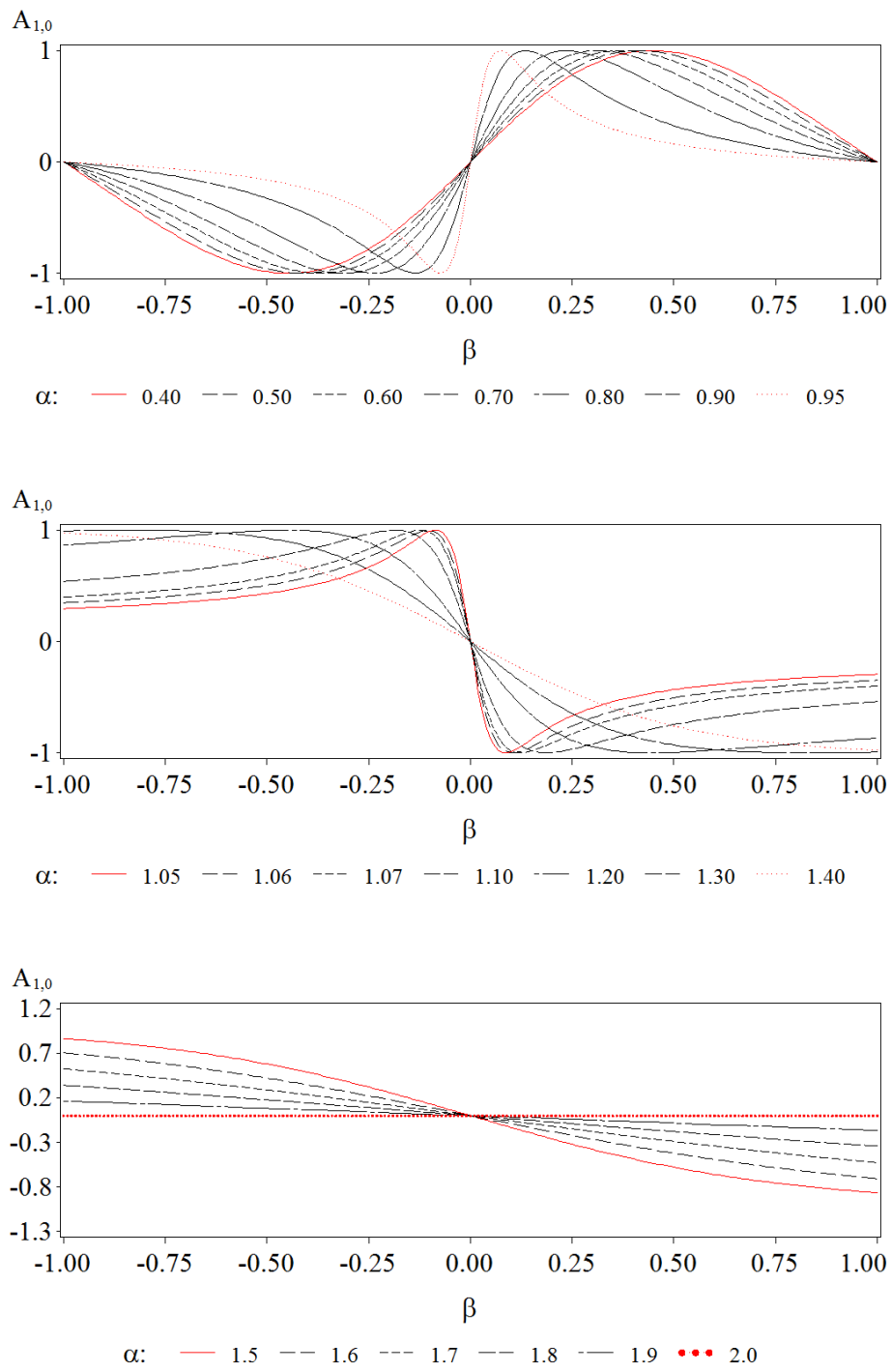


Figura 3.6: Comportamento de $A_{1,0}$ em função de β para alguns valores de α .

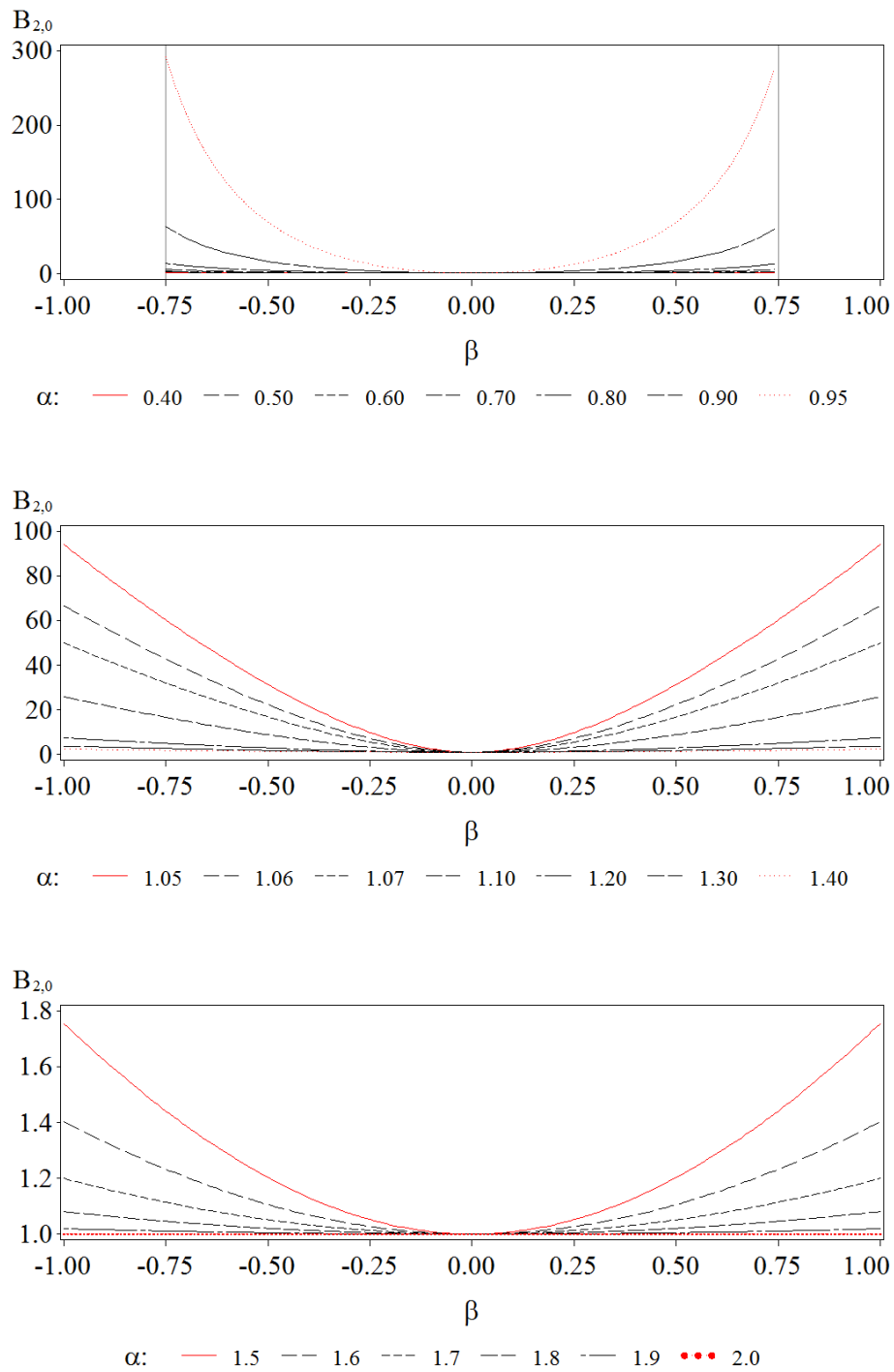


Figura 3.7: Comportamento de $B_{2,0}$ em função de β para alguns valores de α . As linhas verticais em $\beta = \pm 0.75$ representam truncamentos.

Como $B_{2,0}$ é função par com respeito a β , o sinal desse coeficiente pode ser avaliado com base em $f^{(1)}$ ou $A_{1,0}$, conforme a Fig. 3.6 ou o quadro que se segue.

Tabela 3.1: Relações entre α , β , $f^{(1)}$ e $A_{1,0}$

$0 < \alpha < 1$	$\beta > 0$	$f^{(1)} > 0$	$A_{1,0} > 0$
$0 < \alpha < 1$	$\beta < 0$	$f^{(1)} < 0$	$A_{1,0} < 0$
$1 < \alpha < 2$	$\beta < 0$	$f^{(1)} > 0$	$A_{1,0} > 0$
$1 < \alpha < 2$	$\beta > 0$	$f^{(1)} < 0$	$A_{1,0} < 0$

3.6 Considerações

Neste capítulo apresentamos as distribuições estáveis e descrevemos o comportamento do seu polinômio característico. Como veremos no Cap. 5, os coeficientes desse polinômio serão úteis para a estimação dos parâmetros de distribuições estáveis por máxima verossimilhança. A avaliação da estabilidade ou quase estabilidade de um passeio aleatório $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$ será feita com base no comportamento do índice de estabilidade $\alpha_{S_{\Delta t}}$. Se, por exemplo, $X_1, \dots, X_{\Delta t}$ for uma sequência de VA estáveis IID com parâmetros α e γ , então $\alpha_{S_{\Delta t}} = \alpha$ e $\gamma_{S_{\Delta t}} = \Delta t \gamma$, para qualquer Δt . Porém, se essa sequência IID não for de VA estáveis, então $\alpha_{S_{\Delta t}} \rightarrow 2$ à medida que Δt cresce e $\gamma_{S_{\Delta t}} = \Delta t \gamma$. Caso as VA não sejam IID nem estáveis, define-se quase estabilidade [30, 31, 32, 33] se houver um intervalo $\Delta t_0 \leq \Delta t \leq \Delta t_1$ em que $\alpha_{S_{\Delta t}} \approx \alpha$.

Como o procedimento de estimação a ser proposto no Cap. 5 depende da função característica empírica (FCE), uma discussão acerca de suas propriedades e limitações será apresentada no próximo capítulo.

Capítulo 4

A função característica empírica

4.1 Introdução

A função característica empírica ($\hat{\phi}(q)$) é uma importante ferramenta estatística para o estudo de observações que seguem distribuições de Lévy estáveis [26, 95, 97, 125, 100] ou quase estáveis [30, 31, 33, 35]. Nesses casos, a representação da distribuição dos dados por funções características é mais simples do que a por funções de densidade ou funções de distribuição de probabilidade. Além disso, no caso não gaussiano da distribuição de Lévy estável, nem todos os momentos existem [21, 103], o que impede o uso de medidas como o desvio padrão e o coeficiente de correlação. Por exemplo, a função característica empírica (FCE) pode ser aplicada para se testar a hipótese de gaussianidade ($\alpha = 2$) contra a hipótese ($\alpha < 2$) — ou outros testes de aderência (*goodness-of-fit*) — [52, 56, 60]; para se testar a dependência contemporânea e a serial em processos não gaussianos [25, 55, 56]; para se avaliar a assimetria dos dados [28, 56], e para se estimar parâmetros [4, 26, 77, 95, 97, 125]. O Cap. 5 abordará sobre o método de estimação por máxima verossimilhança via FCE.

Além das propriedades estatísticas básicas do estimador $\hat{\phi}(q)$ [13, 27, 47, 120], este capítulo trata acerca do seu truncamento natural devido à finitude do tamanho da amostra (n). Com base na distribuição amostral da FCE e na teoria de testes de hipóteses [89, 101],

espera-se que haja truncamento da parte real de $\hat{\phi}(q)$ quando seu valor for inferior a $z/\sqrt{2n}$, em que z é um quantil apropriado da distribuição normal padrão. Analogamente, trunca-se $|\hat{\phi}(q)|$ caso seu valor se encontre abaixo de um ponto crítico $\phi_{0,n} = \sqrt{z_2^2/2n}$, em que z_2^2 é um quantil apropriado da distribuição χ^2 com dois graus de liberdade.

Com respeito às propriedades de *scaling*, considerando um passeio aleatório na forma $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$, verifica-se que há uma quebra no padrão linear esperado de $\ln(-\ln|\hat{\phi}_{S_{\Delta t}}(q)|)$ versus $\ln(q)$. Essa quebra de *scaling* ocorre sempre no mesmo patamar, independentemente do valor Δt e dos parâmetros da distribuição. Assim, o truncamento da FCE relaciona-se também com as quebras de *scaling* de outras características distribucionais, como o momentos absolutos das distribuições quase estáveis [84].

A estimativa da FD no ponto zero (ou a "probabilidade de retorno à origem") permite estudar as propriedades de *scaling* de uma distribuição de dados [75, 46]. No entanto, as estimativas produzidas pelo método do Kernel [107, 108] — o que inclui o Kernel triangular utilizado em trabalhos anteriores [21, 75, 46] — são tendenciosas e inconsistentes do ponto de vista estatístico [115]. Por outro lado, boas estimativas da FD no ponto zero podem ser obtidas com base na FCE truncada mediante transformada inversa de Fourier.

A próxima seção define e apresenta as principais propriedades da FCE. A Seção 4.3 trata do polinômio característico empírico e descreve procedimentos para se estimar a FD e suas derivadas. A Seção 4.4 trata do fenômeno de truncamento da FCE. A Seção 4.5 aborda acerca do truncamento da FCE de um passeio aleatório simples, e as quebras de *scaling* são ilustradas na Seção 4.6 por meio de simulações de Monte Carlo, e também na Seção 4.7 com dados do IBovespa. A Seção 4.8 apresenta algumas considerações sobre este capítulo.

4.2 Definição e algumas propriedades da FCE

Seja $\{X_j\}_{j=1,\dots,n}$ uma amostra aleatória retirada de uma distribuição contínua X com FD, FDA e FC respectivamente representadas por $f(x; \boldsymbol{\theta})$, $F(x; \boldsymbol{\theta}) = \int_{-\infty}^x f(u; \boldsymbol{\theta}) du$ e $\phi(q; \boldsymbol{\theta}) = \phi_2(q; \boldsymbol{\theta}) + i\phi_1(q; \boldsymbol{\theta})$, em que $\boldsymbol{\theta} \in \mathbb{R}^p$ é o vetor de parâmetros da distribuição e

$q \in \mathbb{R}$. Define-se a FCE como [97, 120]

$$\hat{\phi}(q) = \int e^{iqx} d\hat{F}(x) \quad (4.1)$$

$$= \frac{1}{n} \sum_{j=1}^n e^{iqX_j}, \quad (4.2)$$

$$= \sum_{j=1}^n \frac{\cos(qX_j)}{n} + i \sum_{j=1}^n \frac{\text{sen}(qX_j)}{n} \quad (4.3)$$

$$= \hat{\phi}_2(q) + i\hat{\phi}_1(q), \quad (4.4)$$

em que

$$\hat{F}(x) = \frac{1}{n} \sum_{j=1}^n I(x - X_j) \quad (4.5)$$

é a função de distribuição acumulada empírica e $I(y) = 1$ se $y \geq 0$ e $I(y) = 0$ se $y < 0$.

Propriedades

A FCE é um estimador não viciado da FC, pois

$$\begin{aligned} \langle \hat{\phi}(q) \rangle &= \left\langle \sum_{j=1}^n \frac{\cos(qX_j)}{n} \right\rangle + i \left\langle \sum_{j=1}^n \frac{\text{sen}(qX_j)}{n} \right\rangle \\ &= \phi_2(q; \boldsymbol{\theta}) + i\phi_1(q; \boldsymbol{\theta}) = \phi(q; \boldsymbol{\theta}). \end{aligned}$$

Com respeito ao segundo momento da parte real de $\hat{\phi}(q)$, tem-se que

$$\begin{aligned} \langle \hat{\phi}_2^2(q) \rangle &= \frac{1}{n^2} \left\langle \sum_{j=1}^n \sum_{k=1}^n \cos(qX_j) \cos(qX_k) \right\rangle \\ &= \frac{1}{n^2} \left\langle \sum_{j=1}^n \cos^2(qX_j) \right\rangle + \frac{1}{n^2} \left\langle \sum_{j \neq k} \cos(qX_j) \cos(qX_k) \right\rangle \\ &= \frac{1}{n} \langle \cos^2(qX) \rangle + \frac{n(n-1)}{n} \langle \cos(qX) \rangle^2 \\ &= \frac{1}{n} \langle \cos^2(qX) \rangle + \frac{n-1}{n} \phi_2^2(q; \boldsymbol{\theta}), \end{aligned} \quad (4.6)$$

de modo que sua variância é

$$\text{Var}(\hat{\phi}_2(q)) = \frac{1}{n} \left\{ \langle \cos^2(qX) \rangle - \phi_2^2(q; \boldsymbol{\theta}) \right\}. \quad (4.7)$$

Por analogia, o segundo momento e a variância da parte imaginária de $\hat{\phi}(q)$ são, respectivamente,

$$\langle \hat{\phi}_1^2(q) \rangle = \frac{1}{n} \langle \text{sen}^2(qX) \rangle + \frac{n-1}{n} \phi_1^2(q; \boldsymbol{\theta}), \quad (4.8)$$

e

$$\text{Var}(\hat{\phi}_1(q)) = \frac{1}{n} \left\{ \langle \text{sen}^2(qX) \rangle - \phi_1^2(q; \boldsymbol{\theta}) \right\}. \quad (4.9)$$

Considerando-se as Eqs. (4.7) e (4.9), a média dos erros quadráticos (*mean squared errors*) de $\hat{\phi}(q)$ é

$$\text{MSE}(q; n) = \langle |\hat{\phi}(q) - \phi(q; \boldsymbol{\theta})|^2 \rangle \quad (4.10)$$

$$= \langle (\hat{\phi}_2(q) - \phi(q; \boldsymbol{\theta}))^2 \rangle + \langle (\hat{\phi}_1(q) - \phi(q; \boldsymbol{\theta}))^2 \rangle \quad (4.11)$$

$$= \text{Var}(\hat{\phi}_2^2(q)) + \text{Var}(\hat{\phi}_1^2(q)) \quad (4.12)$$

$$= \frac{1 - |\phi(q; \boldsymbol{\theta})|^2}{n}. \quad (4.13)$$

Com base no resultado acima, para um valor fixo q , conclui-se que $\hat{\phi}(q)$ é um estimador consistente, pois $\text{MSE}(q; n) \rightarrow 0$ à medida que n aumenta.

O problema é que o valor n necessário para haver consistência depende de q [13, 120]. Considerando que n seja fixo e $|q| \rightarrow \infty$, como $\hat{\phi}(q)$ é FC de uma distribuição discreta, tem-se $\limsup_{|q| \rightarrow \infty} |\hat{\phi}(q)| = 1$ (Eq. (2.7)). Por outro lado, $\phi(q; \boldsymbol{\theta})$ é FC de uma distribuição absolutamente contínua e, assim, $\limsup_{|q| \rightarrow \infty} |\phi(q; \boldsymbol{\theta})| = 0$ (Eq. (2.6)). Então, pela desigualdade triangular,

$$\limsup_{|q| \rightarrow \infty} |\hat{\phi}(q) - \phi(q; \boldsymbol{\theta})| \leq \limsup_{|q| \rightarrow \infty} |\hat{\phi}(q)| + \limsup_{|q| \rightarrow \infty} |\phi(q; \boldsymbol{\theta})| = 1, \quad (4.14)$$

de modo que o evento $\limsup_{|q| \rightarrow \infty} |\hat{\phi}(q) - \phi(q; \boldsymbol{\theta})| > 0$, com n fixo, ocorre com probabilidade 1. A discussão acerca das implicações desse fato se encontra na Seção 4.4.

Distribuição amostral

Para se obter a distribuição amostral do par $(\hat{\phi}_2(q), \hat{\phi}_1(q))'$, é preciso determinar sua função de covariância $\text{Cov}(\hat{\phi}_2(q), \hat{\phi}_1(q))$. Para isso, tem-se que

$$\begin{aligned}
 \phi(2q; \boldsymbol{\theta}) &= \langle e^{i2qX} \rangle \\
 &= \langle (\cos(qX) + i\text{sen}(qX))^2 \rangle \\
 &= \langle \cos^2(qX) \rangle + 2i \langle \cos(qX) \rangle \langle \text{sen}(qX) \rangle - \langle \text{sen}^2(qX) \rangle \\
 &= \langle \cos^2(qX) \rangle - \langle \text{sen}^2(qX) \rangle + 2i\phi_2(q; \boldsymbol{\theta})\phi_1(q; \boldsymbol{\theta}) \\
 &= \langle 2 \cos^2(qX) \rangle - 1 + 2i\phi_2(q; \boldsymbol{\theta})\phi_1(q; \boldsymbol{\theta}).
 \end{aligned} \tag{4.15}$$

Logo,

$$|\phi(2q; \boldsymbol{\theta})|^2 = (2 \langle \cos^2(qX) \rangle - 1)^2 + 4\phi_2^2(q; \boldsymbol{\theta})\phi_1^2(q; \boldsymbol{\theta}), \tag{4.16}$$

de modo que

$$\langle \cos^2(qX) \rangle = \frac{1 + \sqrt{\Delta(q; \boldsymbol{\theta})}}{2}, \tag{4.17}$$

e, analogamente,

$$\langle \text{sen}^2(qX) \rangle = \frac{1 - \sqrt{\Delta(q; \boldsymbol{\theta})}}{2}, \tag{4.18}$$

em que

$$\Delta(q; \boldsymbol{\theta}) \doteq |\phi(2q; \boldsymbol{\theta})|^2 - 4\phi_2^2(q; \boldsymbol{\theta})\phi_1^2(q; \boldsymbol{\theta}). \tag{4.19}$$

Quanto ao segundo momento cruzado,

$$\begin{aligned}
 \langle \hat{\phi}_2(q)\hat{\phi}_1(q) \rangle &= \frac{1}{n^2} \left\langle \sum_{j=1}^n \sum_{k=1}^n \cos(qX_j)\text{sen}(qX_k) \right\rangle \\
 &= \frac{1}{n^2} \left\langle \sum_{j=1}^n \cos(qX_j)\text{sen}(qX_j) \right\rangle + \frac{1}{n^2} \left\langle \sum_{j \neq k} \cos(qX_j)\text{sen}(qX_k) \right\rangle \\
 &= \frac{1}{n} \langle \cos(qX)\text{sen}(qX) \rangle + \frac{n-1}{n} \phi_2(q; \boldsymbol{\theta})\phi_1(q; \boldsymbol{\theta}) \\
 &= \frac{1}{2n} \langle \text{sen}(2qX) \rangle + \frac{n-1}{n} \phi_2(q; \boldsymbol{\theta})\phi_1(q; \boldsymbol{\theta}) \\
 &= \frac{1}{2n} \phi_1(2q; \boldsymbol{\theta}) + \frac{n-1}{n} \phi_2(q; \boldsymbol{\theta})\phi_1(q; \boldsymbol{\theta}).
 \end{aligned} \tag{4.20}$$

Assim, a função de covariância entre as partes real e imaginária da FCE é

$$\text{Cov}(\hat{\phi}_2(q), \hat{\phi}_1(q)) = \frac{1}{n} \left\{ \frac{1}{2} \phi_1(2q; \boldsymbol{\theta}) - \phi_2(q; \boldsymbol{\theta}) \phi_1(q; \boldsymbol{\theta}) \right\}. \quad (4.21)$$

Se a distribuição X for simétrica em torno de zero, então $\phi_1(q; \boldsymbol{\theta}) = 0$ e $\phi(q; \boldsymbol{\theta}) = \phi_2(q; \boldsymbol{\theta})$, de maneira que (4.17) e (4.18) se reduzem a

$$\langle \cos^2(qX) \rangle = \frac{1 + \phi(2q; \boldsymbol{\theta})}{2} \quad (4.22)$$

e

$$\langle \text{sen}^2(qX) \rangle = \frac{1 - \phi(2q; \boldsymbol{\theta})}{2}. \quad (4.23)$$

Conseqüentemente, havendo simetria em torno de zero, $\text{Cov}(\phi_{2,n}(q), \phi_{1,n}(q)) = 0$ e

$$\text{Var}(\hat{\phi}_2^2(q)) = \text{Var}(\hat{\phi}_1^2(q)) = \frac{1 - \phi_2^2(q; \boldsymbol{\theta})}{2n}. \quad (4.24)$$

Dado um valor q e n suficientemente grande, mediante aplicação do Teorema Limite Central é possível concluir que as estatísticas $\hat{\phi}_2(q)$ e $\hat{\phi}_1(q)$ — que são somas de variáveis aleatórias IID — são assintoticamente gaussianas, ou seja,

$$\frac{\hat{\phi}_k(q) - \phi_k(q; \boldsymbol{\theta})}{\sqrt{\text{Var}(\phi_k(q; \boldsymbol{\theta}))}} \sim N(0, 1), \quad (4.25)$$

em que $k = 1$ ou 2 . Logo, a distribuição conjunta $(\hat{\phi}_2(q), \hat{\phi}_1(q))'$ segue assintoticamente uma distribuição normal bivariada [76] cujo vetor de médias é $(\phi_2(q), \phi_1(q))'$ e cujos elementos da matriz de covariância são (4.7), (4.9) e (4.21).

Covariâncias cruzadas

Os momentos cruzados $\langle \hat{\phi}_2(q) \hat{\phi}_2(r) \rangle$, $\langle \hat{\phi}_1(q) \hat{\phi}_1(r) \rangle$ e $\langle \hat{\phi}_2(q) \hat{\phi}_1(r) \rangle$ podem ser desenvolvidos com base na relação

$$\langle \cos(qX) \cos(rX) \rangle = \frac{\langle \cos((q+r)X) \rangle + \langle \cos((q-r)X) \rangle}{2}. \quad (4.26)$$

Assim,

$$\begin{aligned} \text{Cov}(\hat{\phi}_2(q), \hat{\phi}_2(r)) &= \left\langle (\hat{\phi}_2(q) - \phi_2(q; \boldsymbol{\theta}))(\hat{\phi}_2(r) - \phi_2(r; \boldsymbol{\theta})) \right\rangle \\ &= \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \langle \cos(qX_j) \cos(rX_k) \rangle - \phi_2(q; \boldsymbol{\theta})\phi_2(r; \boldsymbol{\theta}) \\ &= \frac{\phi_2(q+r; \boldsymbol{\theta}) + \phi_2(q-r; \boldsymbol{\theta}) - 2\phi_2(q; \boldsymbol{\theta})\phi_2(r; \boldsymbol{\theta})}{2n}. \end{aligned} \quad (4.27)$$

Semelhantemente, conclui-se que

$$\text{Cov}(\hat{\phi}_1(q), \hat{\phi}_1(r)) = \frac{\phi_2(q-r; \boldsymbol{\theta}) - \phi_2(q+r; \boldsymbol{\theta}) - 2\phi_1(q; \boldsymbol{\theta})\phi_1(r; \boldsymbol{\theta})}{2n} \quad (4.28)$$

e

$$\text{Cov}(\hat{\phi}_2(q), \hat{\phi}_1(r)) = \frac{\phi_1(q+r; \boldsymbol{\theta}) - \phi_1(q-r; \boldsymbol{\theta}) - 2\phi_2(q; \boldsymbol{\theta})\phi_1(r; \boldsymbol{\theta})}{2n}. \quad (4.29)$$

4.3 O polinômio característico empírico

Com base na expansão (2.15) da FC, o polinômio característico $\Phi(D_q; \boldsymbol{\theta})$ foi introduzido (Eq. (2.16)) como uma forma alternativa que permite caracterizar uma distribuição, e o Cap. 3 mostrou o papel desse operador para o estudo de distâncias entre distribuições. Com relação a análise descritiva de uma amostra $\{X_j\}_{j=1, \dots, n}$, define-se polinômio característico empírico como

$$\hat{\Phi}(D_q) = \sum_{k=0}^{+\infty} \hat{c}_k D_q^k, \quad (4.30)$$

em que $\hat{c}_k = 2\pi(-i)^k \hat{f}^{(k)}/k!$ e $\hat{f}^{(k)}$ representa uma estimativa da FD no ponto zero ($k = 0$) ou da k -ésima derivada da FD no ponto zero ($k \geq 1$).

A estimativa da FD em qualquer ponto x , $\hat{f}(x) \equiv \hat{f}^{(0)}(x)$, e as estimativas das derivadas da FD, $\hat{f}^{(k)}(x)$, $k \geq 1$, podem ser obtidas pelo método do Kernel [107] ou pelo método da transformada inversa da FCE.

O estimador de $f^{(k)}(x; \boldsymbol{\theta})$ pelo método do Kernel [107, 108] é

$$\hat{f}^{(k)}(x) = \frac{1}{nh^{k+1}} \sum_{j=1}^n K^{(k)}\left(\frac{x - X_j}{h}\right), \quad (4.31)$$

em que o kernel $K(u)$ é uma função não negativa tal que

$$\int_{-\infty}^{+\infty} K(u)du = 1, \quad (4.32)$$

e h é a largura da janela de suavização (*bandwidth* ou parâmetro de suavização). Há várias possibilidades de escolha para a função $K(u)$ e, em geral, essas diferentes funções produzem resultados práticos semelhantes [107]. Uma possível escolha, por exemplo, é o kernel gaussiano

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp(-u^2/2). \quad (4.33)$$

Já a determinação do parâmetro de suavização h é um aspecto crítico, pois o resultado é sensível à escolha do valor h . Entre os vários métodos disponíveis para a determinação de um valor ótimo, o da validação cruzada (CV) sugere que h seja tal que minimize a função [53]

$$CV_k(h) = \frac{(-1)^k}{nh^{2k+1}} \left\{ \sum_{i,j} (K*K)^{(2k)}\{(X_i-X_j)/h\} - \frac{2}{n-1} \sum_{i \neq j} K^{(2k)}\{(X_i-X_j)/h\} \right\}, \quad (4.34)$$

em que $K * K$ representa a convolução da função Kernel.

Como alternativa, propõe-se estimar $f^{(k)}(x; \theta)$ pela transformação inversa da FCE,

$$\hat{f}^{(k)}(x) = \frac{(-i)^k}{2\pi} \int_{q \in Q} \hat{\phi}(q) q^k e^{-iqx} dq, \quad (4.35)$$

em que Q é uma faixa de operação da FCE a ser discutida na próxima seção. Embora o método do Kernel forneça, em geral, bons resultados, as estimativas no ponto zero, $\hat{f}^{(k)}, k \geq 1$, são particularmente tendenciosas [115]. Por outro lado, o estimador alternativo (4.35) produziu resultados satisfatórios em nossos estudos (Seção 4.6). Assim, sugere-se o seguinte procedimento para a estimação da FD e suas derivadas:

1. calcular $\hat{f}^{(k)} = \hat{f}^{(k)}(0) = \frac{(-i)^k}{2\pi} \int_{q \in Q} \hat{\phi}(q) q^k dq$;
2. obter h tal que as estimativas por Kernel e por FCE sejam coincidentes;
3. calcular $\hat{f}^{(k)}(x)$ com base no parâmetro de suavização determinado no passo anterior.

Como ilustração, um estudo de Monte Carlo será apresentado na Seção 4.6. Antes, porém, introduziremos a FCE truncada na seção que se segue.

4.4 A FCE truncada

Ao mesmo tempo que $\hat{\phi}(q)$ é um estimador não tendencioso e consistente da FC teórica à medida que n aumenta, observamos na seção anterior que, pela finitude do tamanho da amostra, $P(\limsup_{|q| \rightarrow \infty} |\hat{\phi}(q) - \phi(q; \boldsymbol{\theta})| > 0) = 1$. Segundo o teorema de Glivenko-Cantelli [120], se

$$\lim_{n \rightarrow \infty} \ln \frac{q_n}{n} = 0, \quad (4.36)$$

então, quase certamente para qualquer $\phi(q; \boldsymbol{\theta})$, temos

$$\limsup_{|q| \leq q_n} |\hat{\phi}(q) - \phi(q; \boldsymbol{\theta})| = 0. \quad (4.37)$$

De acordo com esse teorema, $\hat{\phi}(q)$ é um estimador consistente enquanto $|q| \leq q_n$, em que $q_n \rightarrow \infty$ mais lentamente do que $n \rightarrow \infty$, o que sugere um truncamento da FCE na forma

$$\hat{\phi}_*(q) = \begin{cases} \hat{\phi}(q) & \text{se } |q| \leq q_n, \\ 0 & \text{se } |q| > q_n. \end{cases} \quad (4.38)$$

Assim, sugere-se que, caso $q \notin Q = [-q_n, q_n]$, em que o intervalo Q denomina-se *intervalo de operação*, as estimativas da FCE se encontram dentro do nível de ruído e, portanto, devem ser desprezadas. O problema é determinar o valor do ponto crítico q_n (ou *primeiro zero positivo* [68, 120]).

Alternativamente, propõe-se que o truncamento seja estabelecido com base na significância estatística de $\hat{\phi}(q)$. Já que $\lim_{|q| \rightarrow \infty} \phi(q; \boldsymbol{\theta}) = 0$ para o caso de uma distribuição absolutamente contínua, para um dado valor q pode-se efetuar o teste de hipóteses

$$H_0 : \phi_k(q; \boldsymbol{\theta}) = 0 \text{ versus } H_1 : \phi_k(q; \boldsymbol{\theta}) \neq 0, \quad (4.39)$$

em que $k = 1$ ou 2 . Caso a estimativa $\hat{\phi}_k(q)$ seja inferior a determinada margem de erro, não haveria evidências estatísticas contra a hipótese nula H_0 e, portanto, essa hipótese não poderia ser rejeitada. De acordo com (4.25), a distribuição amostral da FCE sob a hipótese H_0 é

$$\frac{\hat{\phi}_k(q)}{\frac{1}{\sqrt{2n}}} \sim N(0, 1). \quad (4.40)$$

Logo, considera-se que $\hat{\phi}_2(q)$ e $\hat{\phi}_1(q)$ não são estimativas nulas para um nível de significância ns se

$$|\sqrt{2n}\hat{\phi}_k(q)| > z_{ns}, \quad (4.41)$$

em que $ns = P(|Z| > z_{ns})$ e $Z \sim N(0,1)$. Por exemplo, para $ns = 0,3\%$, tem-se aproximadamente

$$|\sqrt{2n}\hat{\phi}_k(q)| > 3. \quad (4.42)$$

Dessa forma, as limitações do instrumento estatístico $\hat{\phi}_k(q)$ permitem inferências acerca de $\phi_k(q; \theta)$ desde que $|\hat{\phi}_k(q)| > \frac{3}{\sqrt{2n}}$.

Na Subseção 2.2.1, observamos que $|\phi(q; \theta)|$ é a FC de uma distribuição simetrizada. Quanto à inferência acerca dessa distribuição, sob a hipótese nula $\phi(q; \theta) = 0$, tem-se que os estimadores $\hat{\phi}_1(q)$ e $\hat{\phi}_2(q)$ são independentes. Logo, a estatística $2n\{\hat{\phi}_1^2(q) + \hat{\phi}_2^2(q)\}$ segue a distribuição χ^2 com dois graus de liberdade, de modo que módulo da FCE se encontra no nível de ruído se $|\hat{\phi}(q)| < \sqrt{z_2^2/2n}$, em que z_2^2 é um quantil apropriado da distribuição χ^2 .

Com base nessas considerações, propomos um truncamento na forma

$$\hat{\phi}_*(q) = \begin{cases} \hat{\phi}(q) & \text{se } |\hat{\phi}(q)| \geq \phi_{0,n}, \\ 0 & \text{se } |\hat{\phi}(q)| < \phi_{0,n}, \end{cases} \quad (4.43)$$

em que $\phi_{0,n} = \sqrt{z_2^2/2n}$ e z_2^2 é o quantil desejado da distribuição χ^2 com dois graus de liberdade. A Fig. 4.1 ilustra o comportamento do ponto crítico $\phi_{0,n}$ conforme o tamanho da amostra ($500 \leq n \leq 10.000$) e os níveis de significância 0,1%, 0,3% e 1% (que, respectivamente, correspondem aos quantís 13,8155, 11,6183 e 9,2103 da distribuição χ^2 com dois graus de liberdade. A função $\phi_{0,n}$ decresce lentamente para zero à medida que n aumenta.

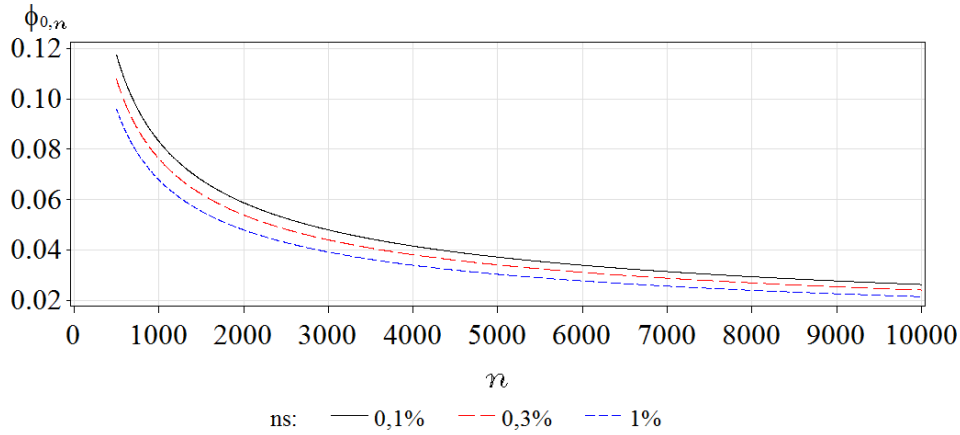


Figura 4.1: Comportamento do ponto crítico $\phi_{0,n} = \sqrt{z_2^2/2n}$ segundo o tamanho da amostra ($500 \leq n \leq 10.000$) e os níveis de significância $ns = 0,1\%$ ($z_2^2 = 13,8155$), $0,3\%$ ($z_2^2 = 11,6183$) e 1% ($z_2^2 = 9,2103$).

4.5 A FCE de uma soma de variáveis aleatórias

Considerando que $\{X_j\}_{j=1,\dots,n}$ seja uma sequência de VA IID, a FCE de um passeio aleatório $S_{\Delta t} = X_1 + X_2 + \dots + X_{\Delta t}$ pode ser escrita como

$$\hat{\phi}_{S_{\Delta t}}(q) = \prod_{j=1}^{\Delta t} \hat{\phi}_X(q) = (\hat{\phi}_X(q))^{\Delta t}. \quad (4.44)$$

Logo, se houver independência, a estimativa $(\hat{\phi}_{S_{\Delta t}}(q))^{\frac{1}{\Delta t}}$ deve ser equivalente à $\hat{\phi}_X(q)$ ($\Delta t = 1$).

Há outras propriedades de *scaling* se X_j for estável, pois, nesse caso, $X_j = \Delta t^{\frac{1}{\alpha}} S_{\Delta t}$. Assim, para uma soma de VA estáveis IID, tem-se

$$\hat{\phi}_{S_{\Delta t}}(q) = \hat{\phi}_X(q^{-\frac{1}{\alpha}}), \quad (4.45)$$

ou

$$\hat{\phi}_X(q) = \hat{\phi}_{S_{\Delta t}}(q^{\frac{1}{\alpha}}), \quad (4.46)$$

isto é, a equivalência também ocorre mediante transformação de escala no suporte da FCE. Além disso, se $S_{\Delta t}$ for estável, espera-se um *scaling* na forma

$$\ln(-\ln |\hat{\phi}_{S_{\Delta t}}(q)|) = \alpha \ln |q| + \ln \Delta t + \ln \gamma. \quad (4.47)$$

Assim, teoricamente, para $\Delta t = 1, 2, 3, \dots, N$, o gráfico de $\ln(-\ln |\hat{\phi}_{S_{\Delta t}}(q)|)$ versus $\ln |q|$ deve apresentar N retas paralelas cujas inclinações são iguais a α , com espaçamento vertical igual a $\ln \Delta t$ em relação à reta inicial com $\Delta t = 1$. Entretanto, pela finitude da amostra, espera-se também uma quebra desse padrão linear a partir de algum valor próximo do limite estatístico $\phi_{0,n}$. Na próxima seção, observaremos que esse valor independe de Δt e de quaisquer outros parâmetros da distribuição em estudo.

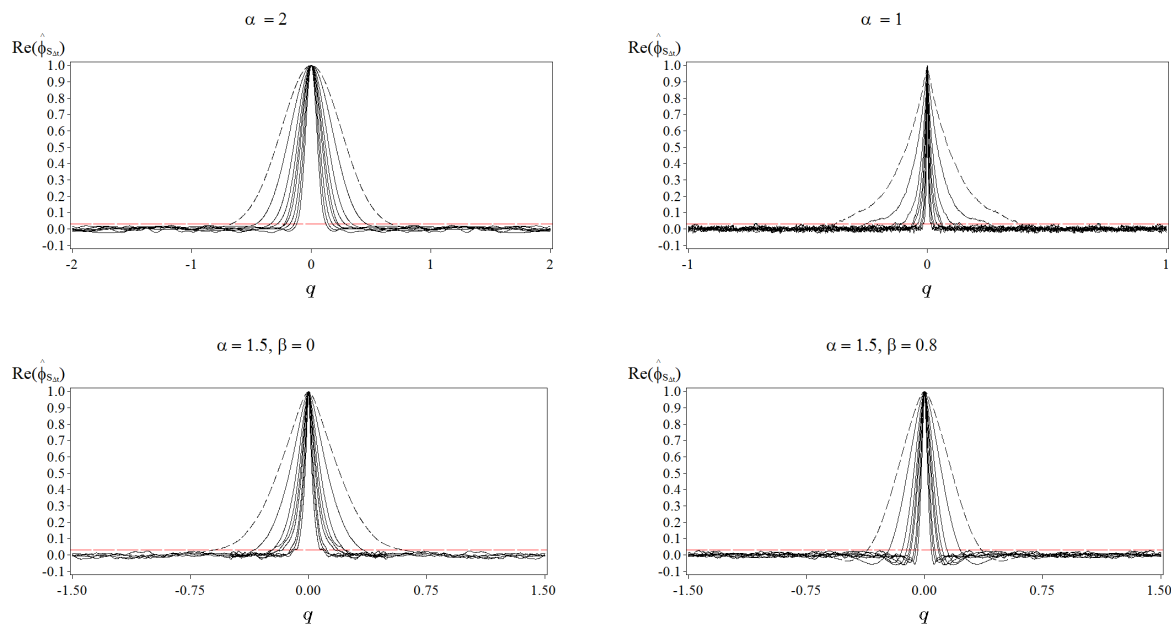


Figura 4.2: Parte real das funções características empíricas obtidas com base em $n = 5.000$ replicações de passeios aleatórios com $\Delta t = 1$ (curva tracejada), 2, 4, 6, 8, 10, 15, 20, 30 gerados por distribuições estáveis com $\alpha = 2$ (gaussiana), $\alpha = 1$ (lorentziana), $\alpha = 1.5$ (simétrica) e $\alpha = 1.5$ (assimétrica). A linha horizontal tracejada representa o ponto crítico $3/\sqrt{2n}$ ($ns = 0, 3\%$).

4.6 Estudo por simulações de Monte Carlo

Esta seção apresenta um experimento de Monte Carlo [102] para ilustrar o fenômeno do truncamento natural da FCE, a quebra do seu padrão esperado de *scaling* e o procedimento para estimação de densidades e suas derivadas. Esse estudo considera amostragens de distribuições estáveis com $\alpha = 1$ (lorentziana), $\alpha = 1.5$ e $\alpha = 2$ (gaussiana). Para

$\alpha = 1.5$, consideramos o caso simétrico ($\beta = 0$) e um caso assimétrico ($\beta = 0.8$). O parâmetro de escala considerado foi $\gamma = 8$. Com base nesses parâmetros, foram geradas $n = 5.000$ replicações de passeios aleatórios com $\Delta t = 1, \dots, 20$.

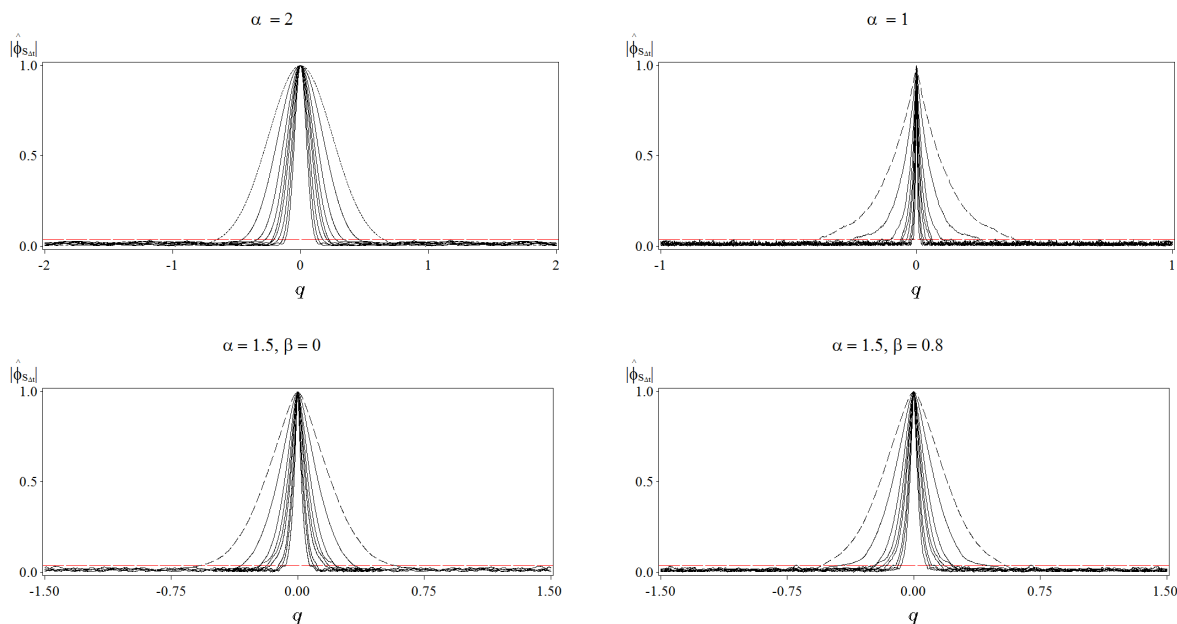


Figura 4.3: Módulos das FCE. A linha horizontal tracejada representa o ponto crítico $\sqrt{13.8155/2n}$ ($ns = 0,1\%$).

A Fig. 4.2 mostra as estimativas da FC para cada série simulada (a curva tracejada representa o caso $\Delta t = 1$). A linha horizontal tracejada representa o ponto crítico $3/\sqrt{2n}$ associado com o nível de significância de $0,3\%$. Assim, as estimativas abaixo dessa linha são consideradas ruído e, portanto, são estatisticamente nulas. Quanto ao módulo da FCE (que desconsidera a parte assimétrica da distribuição, se for o caso), os resultados são apresentados na Fig. 4.3.

A Fig. 4.4 mostra que a quebra do padrão esperado de *scaling* do gráfico $\ln(-\ln|\hat{\phi}(q)|)$ versus $\ln(q)$ ocorre no patamar $\ln(-\ln\sqrt{13.8155/2n})$, independentemente da forma da distribuição. Logo, a ocorrência natural do truncamento da FCE não permite que se estabeleça uma malha de valores q arbitrariamente extensa. Antes da quebra, as inclinações das retas são aproximadamente iguais a α , e os espaçamentos verticais com relação à reta

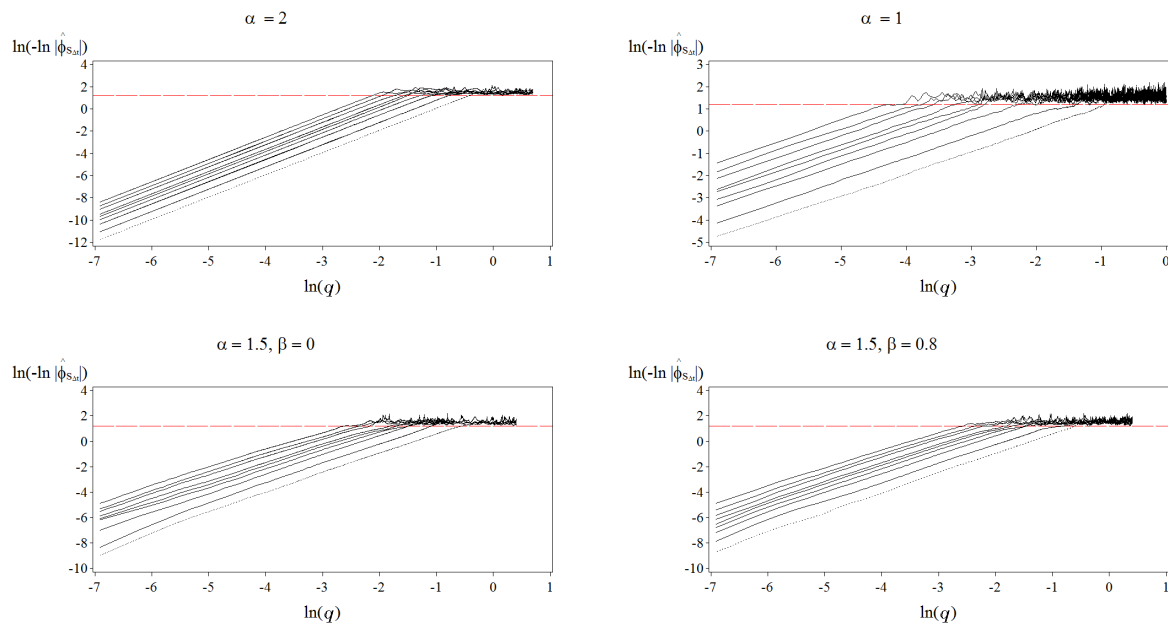


Figura 4.4: Dupla transformação logarítmica dos módulos das FCE da Fig. 4.3. A linha horizontal tracejada representa a transformação do ponto crítico, $\ln(-\ln \sqrt{13.8155/2n})$, relativo a $ns = 0,1\%$.

tracejada ($\Delta t = 1$) são aproximadamente iguais a $\ln \Delta t$. A Fig. 4.5 mostra que $\hat{\phi}_2(q)$ colapsam para o caso $\Delta t = 1$ mediante transformação de escala do suporte $\Delta t^{1/\alpha}q$.

As Figs. 4.6 e 4.7 mostram relações lineares entre $\ln f(0)$ e $\ln f(2)$ versus $\ln \Delta t$ com inclinações respectivamente iguais a $-1/\alpha$ e $-3/\alpha$. Com base no procedimento descrito na Seção 4.3, a Fig. 4.8 mostra as densidades empíricas para os casos simétricos e assimétricos com $\alpha = 1.5$. Por causa da auto-similaridade da distribuição de Levy estável, a transformação de escala da FD $\hat{f}_{S_1}(z) = \hat{f}_{S_{\Delta t}}(x)\Delta t^{1/\alpha}$, em que $z = x\Delta t^{-1/\alpha}$, gera distribuições colapsadas em $\Delta t = 1$.

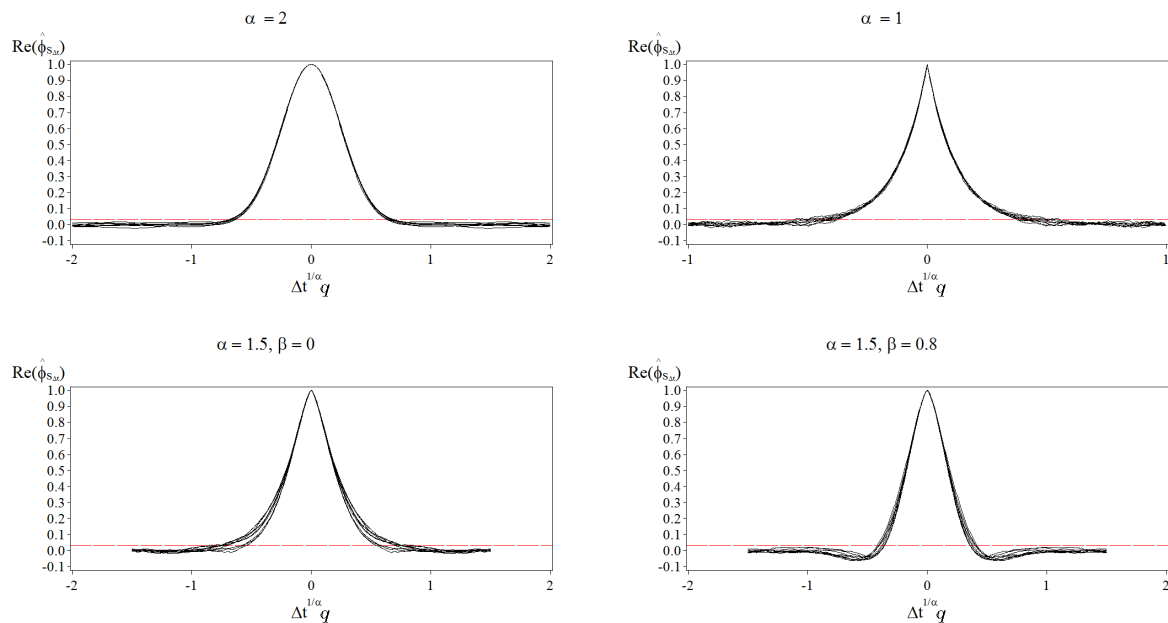


Figura 4.5: *Rescaling*: parte real das FCE apresentadas na Fig. 4.2 *versus* $\Delta t^{1/\alpha} q$. As estimativas colapsam para o caso $\Delta t = 1$ mediante transformação de escala de q .

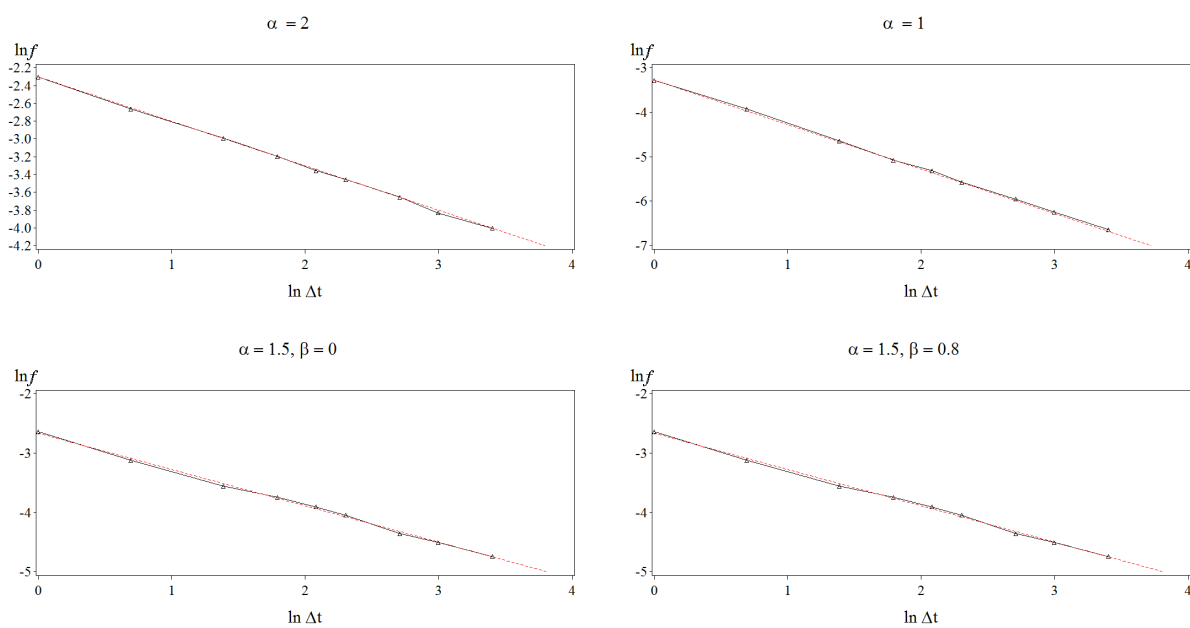


Figura 4.6: Estimativas do logaritmo da FD no ponto zero *versus* $\ln \Delta t$. As retas tracejadas possuem inclinações iguais a $1/\alpha$.

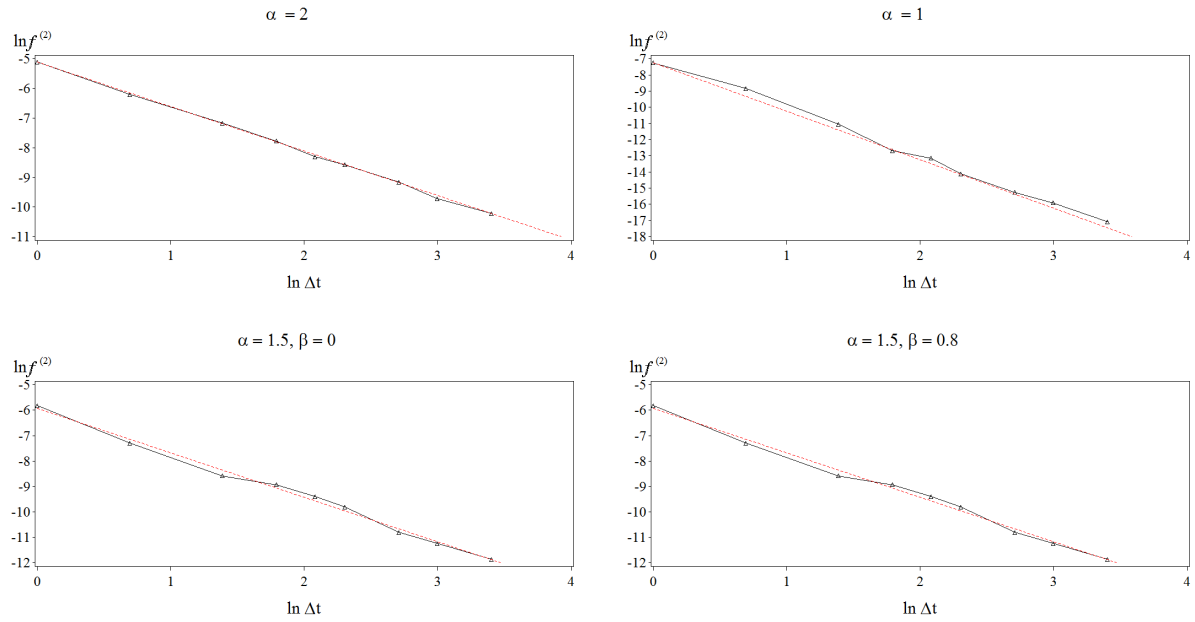


Figura 4.7: Estimativas do logaritmo da segunda derivada da FD no ponto zero *versus* $\ln \Delta t$. As retas tracejadas possuem inclinações iguais a $3/\alpha$.

4.7 Ilustração: dados do IBovespa

Considere os dados do IBovespa descritos no Cap. 1, e que

$$S_{\Delta t}^R = X_1^* + X_2^* + \cdots + X_{\Delta t}^* \quad (4.48)$$

representa a soma de Δt retornos amostrados aleatoriamente, com reposição, do conjunto de dados $\{X_t\}_{t=1, \dots, 10.870}$ — isto é, uma reamostragem *bootstrap* [23, 106] de retornos centrados na média. Isso permite que os tamanhos das amostras para diferentes valores $\Delta t \geq 2$ sejam iguais ao tamanho da amostra original ($\Delta t = 1$).

Considere também o passeio aleatório

$$S_{\Delta t}^0 = X_1 + X_2 + \cdots + X_{\Delta t} \quad (4.49)$$

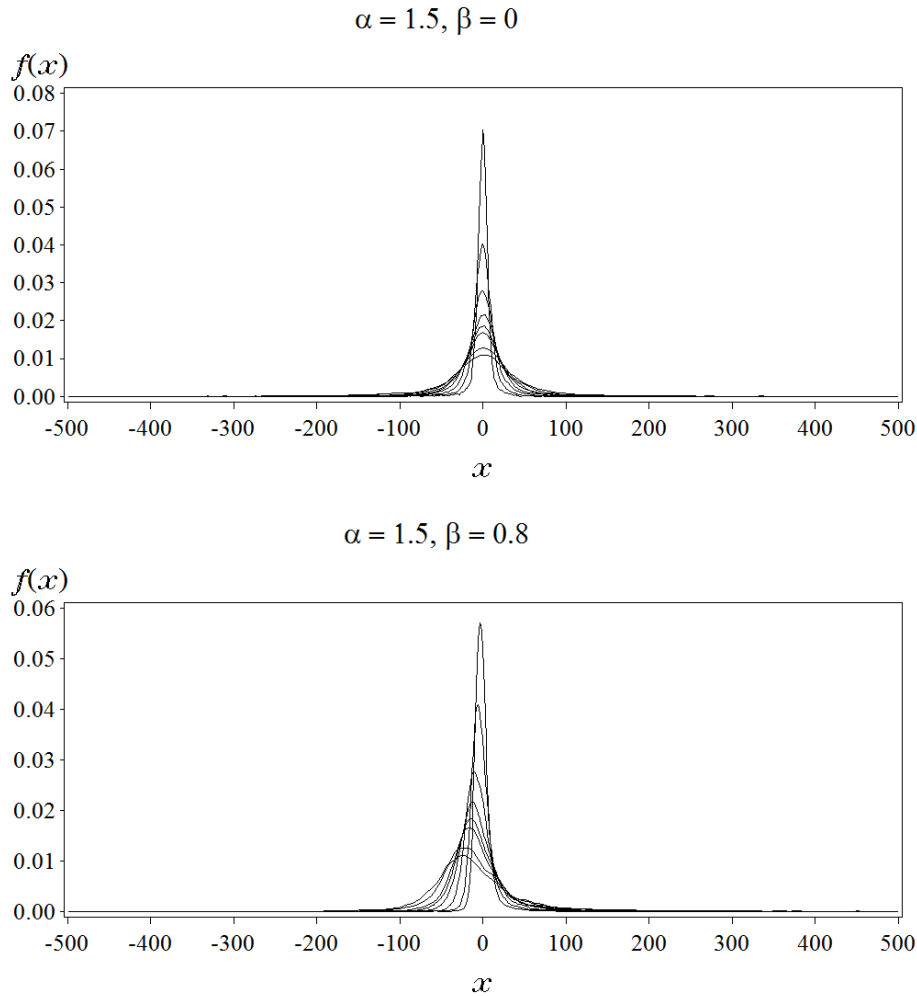


Figura 4.8: Densidades empíricas obtidas pelo método do Kernel para o caso $\alpha = 1.5$.

construído de modo a preservar a estrutura serial original. Ou seja, se $\Delta t = 2$, forma-se um conjunto de dados cujos elementos são as somas parciais adjacentes $D_2 = \{X_1 + X_2, X_3 + X_4, \dots, X_{10.869} + X_{10.870}\}$, a partir do qual será extraída uma amostra *bootstrap* de 10.870 observações.

Para cada $\Delta t = 1, 2, \dots, 10$, foram obtidas 10.870 replicações das variáveis aleatórias $S_{\Delta t}^R$ e $S_{\Delta t}^0$, e a Fig. 4.10 mostra as estimativas da parte real das funções características correspondentes. Nota-se que as estimativas da FC são significativas sobre determinado intervalo $Q \in [-q, q]$, mas fora dele os valores se confundem com o ruído (i.e., não são estatisticamente significativos).

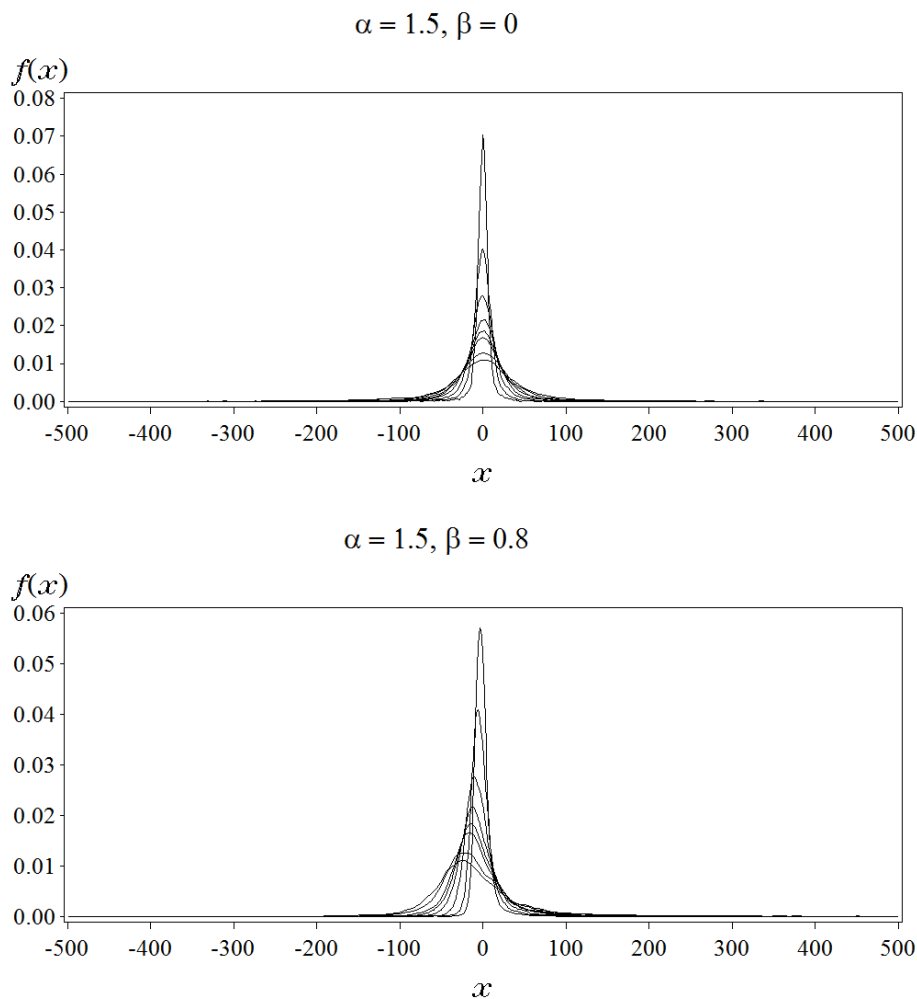


Figura 4.9: Densidades empíricas da Fig. 4.8 colapsadas para o caso $\Delta t = 1$ mediante transformação de escala $z = x(\Delta t)^{-1/1.5}$ e $f(z) = f(x)(\Delta t)^{1/1.5}$.

Para as distribuições estáveis, as estimativas da parte real da FCE dependem da forma (α), da escala (γ) e da assimetria (β). Para que o foco seja concentrado apenas na forma e na escala, considera-se a FC simetrizada $|\hat{\phi}_{S_{\Delta t}}(q)|$ (Fig. 4.11). A Fig. 4.12 apresenta um comportamento de $\ln(-\ln |\hat{\phi}_{S_{\Delta t}}(q)|)$ versus $\ln(q)$ compatível com a hipótese de esses passeios aleatórios serem estáveis ou quase estáveis. Nela, percebe-se um padrão *multiscaling* da FCE, e é interessante notar que a perda de significância estatística produz quebras da tendência linear no mesmo patamar (linha tracejada horizontal). Esse patamar relaciona-se com o limite de confiança estatístico, ou seja, $\ln(-\ln \sqrt{13.8/(2 \times 10870)}) \approx$

1.3, em que 13.8 é o valor do percentil 99,9% da distribuição χ^2 com dois graus de liberdade.

Para $\Delta t = 1$, a reta mostrada na Fig. 4.12 possui inclinação aproximadamente igual a 1.60, e as retas subsequentes para $\Delta t = 2, \dots, 10$ são paralelas a ela. Assim, com base na estimativa $\alpha \approx 1.60$, a Fig. 4.13 mostra que as FCE colapsam para a estimativa $\hat{\phi}_{S_1}(q)$ mediante a transformação de escala $\Delta t^{1/\alpha}q$.

As Figs. 4.14 e ?? mostram as estimativas de $\hat{f}_{S_{\Delta t}}^0$ e $\hat{f}_{S_{\Delta t}}^2$ obtidas mediante a transformação inversa da FCE, ou seja,

$$\hat{f}_{S_{\Delta t}}^k = \frac{(-i)^k}{2\pi} \int_{q \in Q} q^k \hat{\phi}_{S_{\Delta t}}(q) dq. \quad (4.50)$$

A Fig. 4.14 mostra que o comportamento de $\ln \hat{f}_{S_{\Delta t}}^0$ é compatível com a hipótese de estabilidade, enquanto $\ln \hat{f}_{S_{\Delta t}}^R$ se desvia dessa hipótese, principalmente após $\Delta t = 20$ (em que a inclinação da curva tende para -0.5 , ou seja, $\alpha \approx 2$). A Fig. 4.15 mostra que, à medida que Δt aumenta, $\hat{f}_{S_{\Delta t}}^{(1)}$ tende para zero mais rapidamente do que se espera sob a hipótese de estabilidade. Com base no procedimento descrito na Seção 4.3, a Fig. 4.16 mostra as estimativas da FD para os valores de Δt considerados no estudo. Na Fig. 4.17 observamos que a transformação de escala remete cada distribuição para $\Delta t = 1$, pelo menos na região modal da distribuição, apesar das evidências de não estabilidade de $S_{\Delta t}^R$.

4.8 Considerações

Neste capítulo, observamos que o truncamento abrupto da FCE é uma consequência natural da finitude do tamanho da amostra. Esse truncamento é persistente para grandes amostras, pois $\phi_{0,n} \propto n^{-\frac{1}{2}}$ e, assim, $\phi_{0,n} \rightarrow 0$ mais lentamente do que $n \rightarrow \infty$. Esse fenômeno está associado com as quebras abruptas de *scaling* dos momentos absolutos observadas em trabalhos anteriores [84]. Assim, como uma variação do teorema de Glivenko-Cantelli [120], definimos uma FCE truncada como

$$\hat{\phi}_*(q) = \begin{cases} \hat{\phi}(q) & \text{se } |\hat{\phi}(q)| \geq \phi_{0,n}, \\ 0 & \text{se } |\hat{\phi}(q)| < \phi_{0,n}. \end{cases}$$

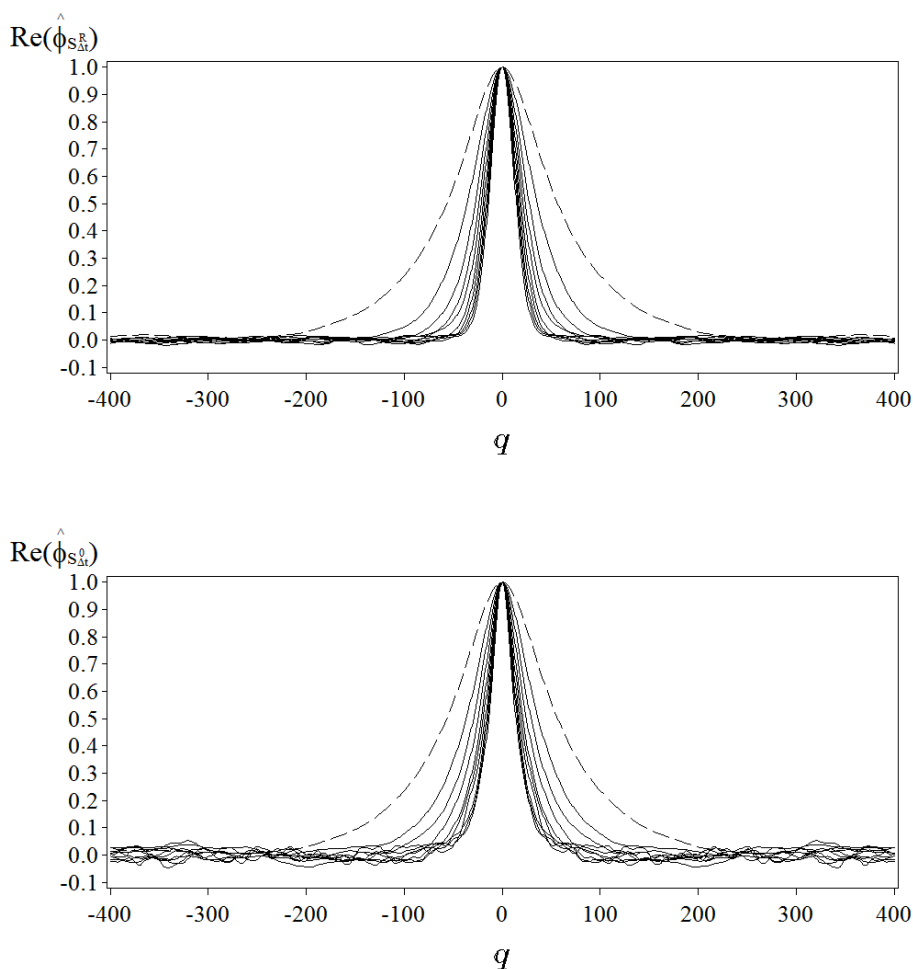


Figura 4.10: Ibovespa: estimativas da parte real da FC de $S_{\Delta t}^R$ (painel superior) e $S_{\Delta t}^0$ (painel inferior) para $\Delta t = 1$ (linha tracejada), 2, \dots , 10.

Embora uma FCE truncada não seja uma função característica propriamente dita [120], é possível associar o truncamento abrupto da FCE com o truncamento não necessariamente abrupto [50, 51, 61, 85] da função de densidade de uma distribuição de dados.

Os resultados mostrados na ilustração com os dados do Ibovespa são consistentes com aqueles observados em nossos trabalhos anteriores [46, 30]. De fato, os dados reais se desviam da hipótese de estabilidade, embora a presença de autocorrelação serial possa retardar a atração para o domínio gaussiano [30, 33, 45]. Por isso, como os dados reais

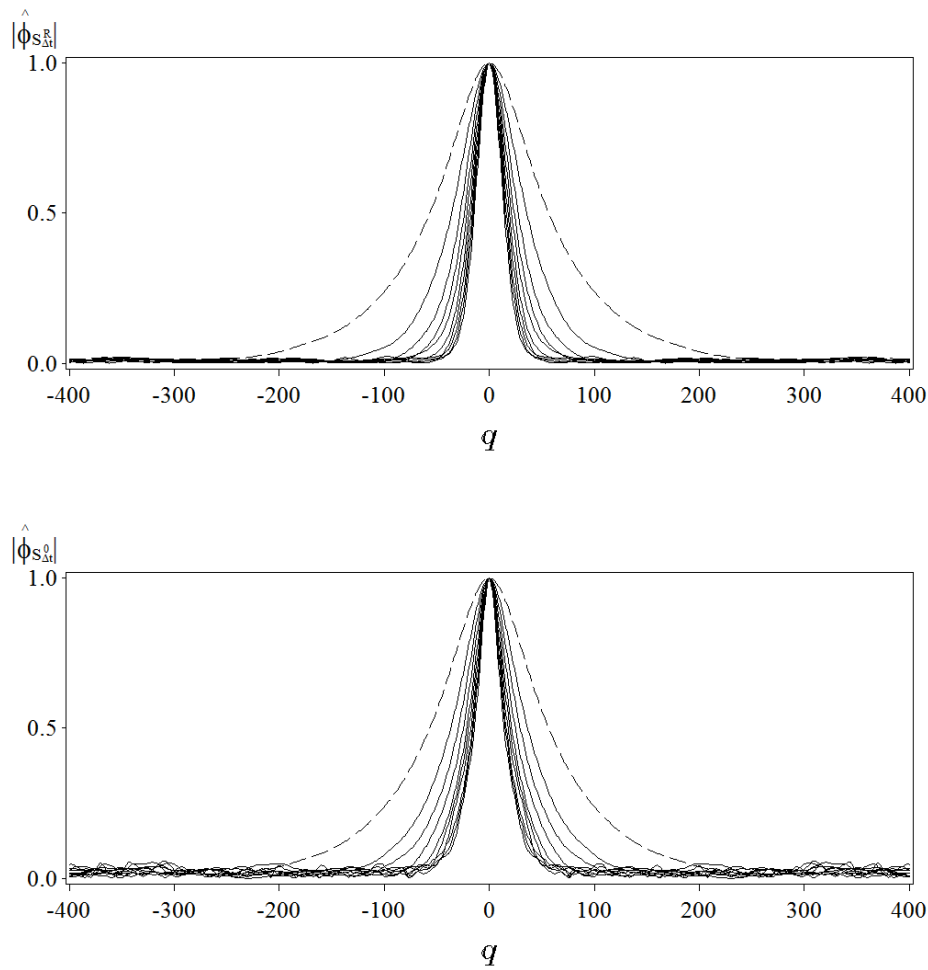


Figura 4.11: IBovespa: estimativas do módulo da FC de $S_{\Delta t}^R$ (painel superior) e $S_{\Delta t}^0$ (painel inferior) para $\Delta t = 1$ (linha tracejada), 2, \dots , 10.

não são IID, o processo pode ser governado por dois regimes distintos: o de Lévy (que rege uma faixa $\Delta t_0 \leq \Delta t \leq \Delta t_1$) e o gaussiano (para $\Delta t > \Delta t_1$).

O próximo capítulo trata do método de máxima verossimilhança (MV) no domínio de Fourier para a estimação dos parâmetros sob a hipótese de estabilidade. Isso porque a estimação de MV, por ser estatisticamente consistente, permite medir e avaliar adequadamente a atração (ou não atração) de um passeio aleatório $S_{\Delta t}$ para o domínio gaussiano à medida de Δt aumenta.

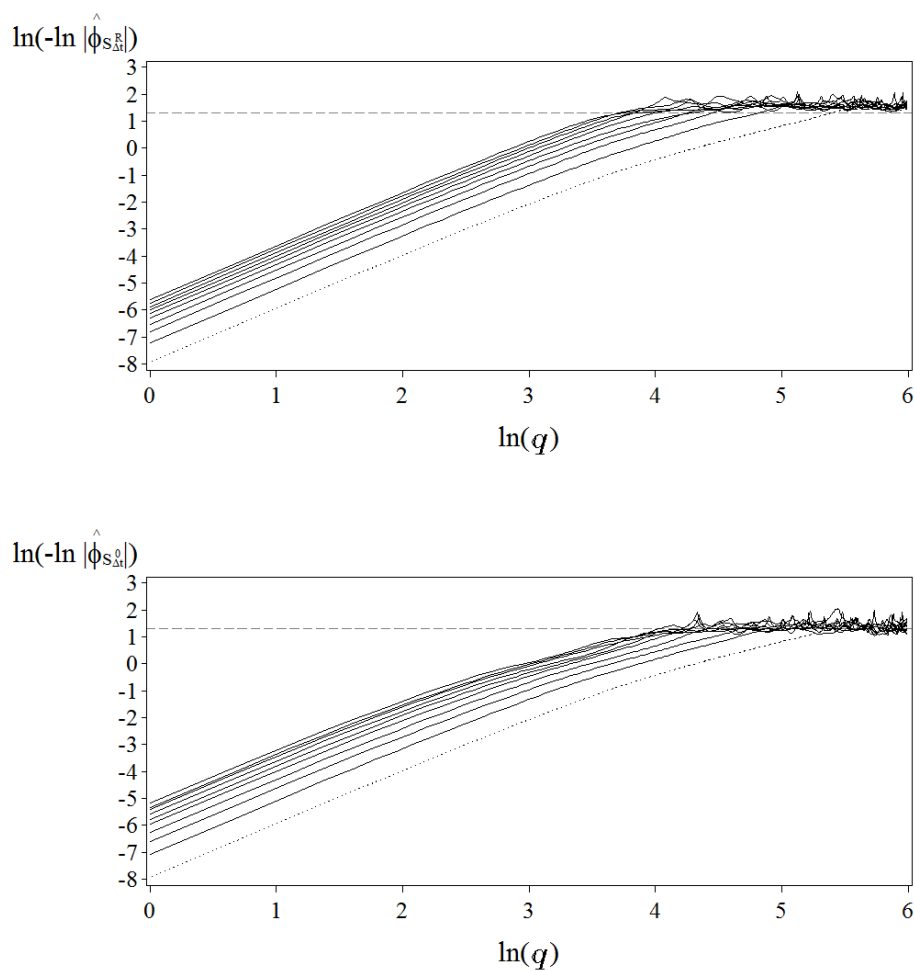


Figura 4.12: IBovespa: estimativas do logaritmo duplo do módulo da FC de $S_{\Delta t}^R$ (painel superior) e $S_{\Delta t}^0$ (painel inferior) versus $\ln(q)$, para $\Delta t = 1$ (linha pontilhada), 2, \dots , 10.

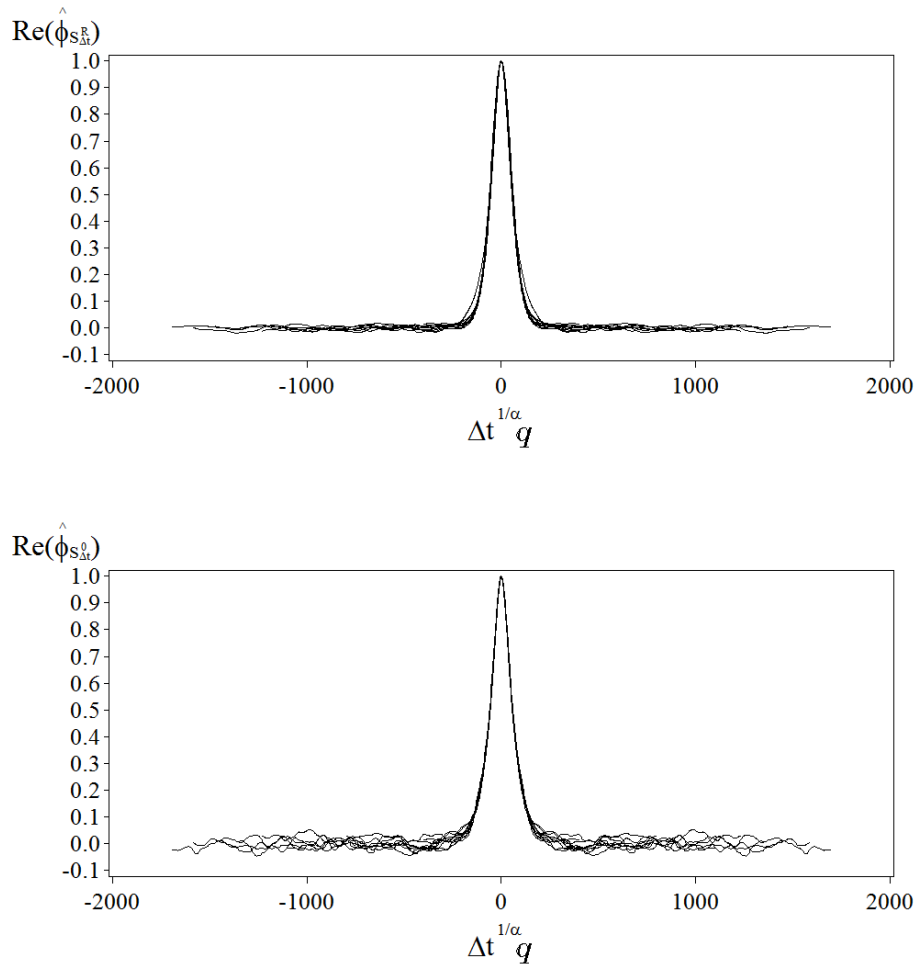


Figura 4.13: IBovespa: *reescaling* — estimativas da parte real da FC de $S_{\Delta t}^R$ (painel superior) e $S_{\Delta t}^0$ (painel inferior) versus $\Delta t^{1/\alpha} q$ ($\Delta t = 1, \dots, 10$).

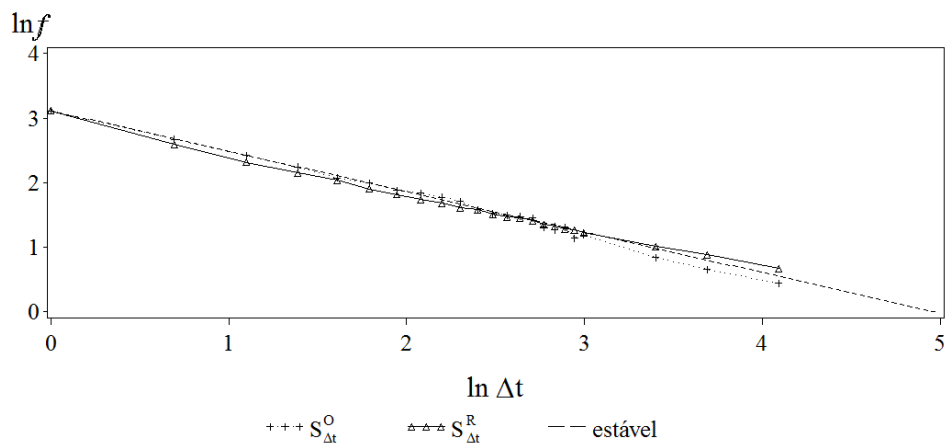


Figura 4.14: IBovespa: estimativas do logaritmo da FD de $S_{\Delta t}^R$ e $S_{\Delta t}^0$ no ponto zero *versus* $\ln \Delta t$ ($\Delta t = 1, \dots, 10$). A linha tracejada é a reta esperada sob a hipótese de a distribuição do passeio aleatório ser estável. O comportamento referente ao processo $S_{\Delta t}^0$ é compatível com a hipótese de estabilidade, enquanto o de $S_{\Delta t}^R$ tende a se desviar dessa hipótese, principalmente após $\Delta t = 20$.

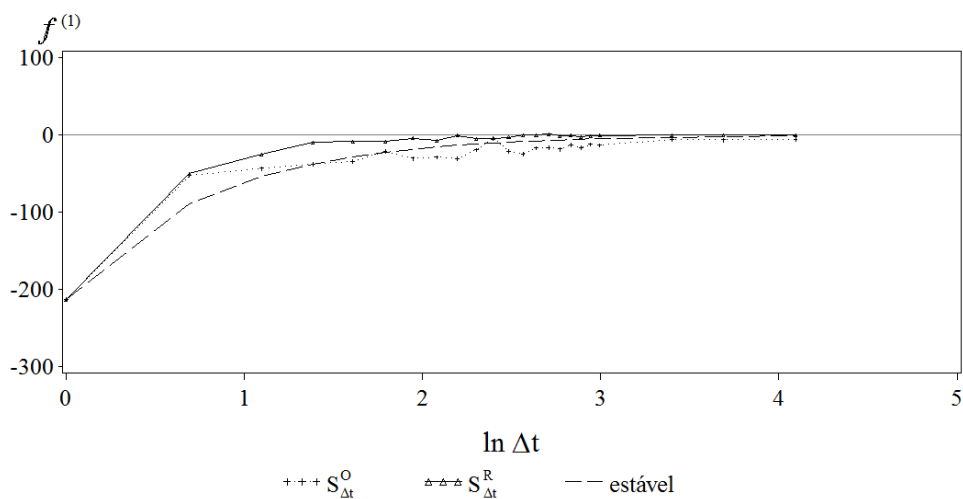


Figura 4.15: IBovespa: estimativas da primeira derivada da FD de $S_{\Delta t}^R$ e $S_{\Delta t}^0$ no ponto zero *versus* $\ln \Delta t$ ($\Delta t = 1, \dots, 10$). A linha tracejada é a curva esperada sob a hipótese de a distribuição do passeio aleatório ser estável.

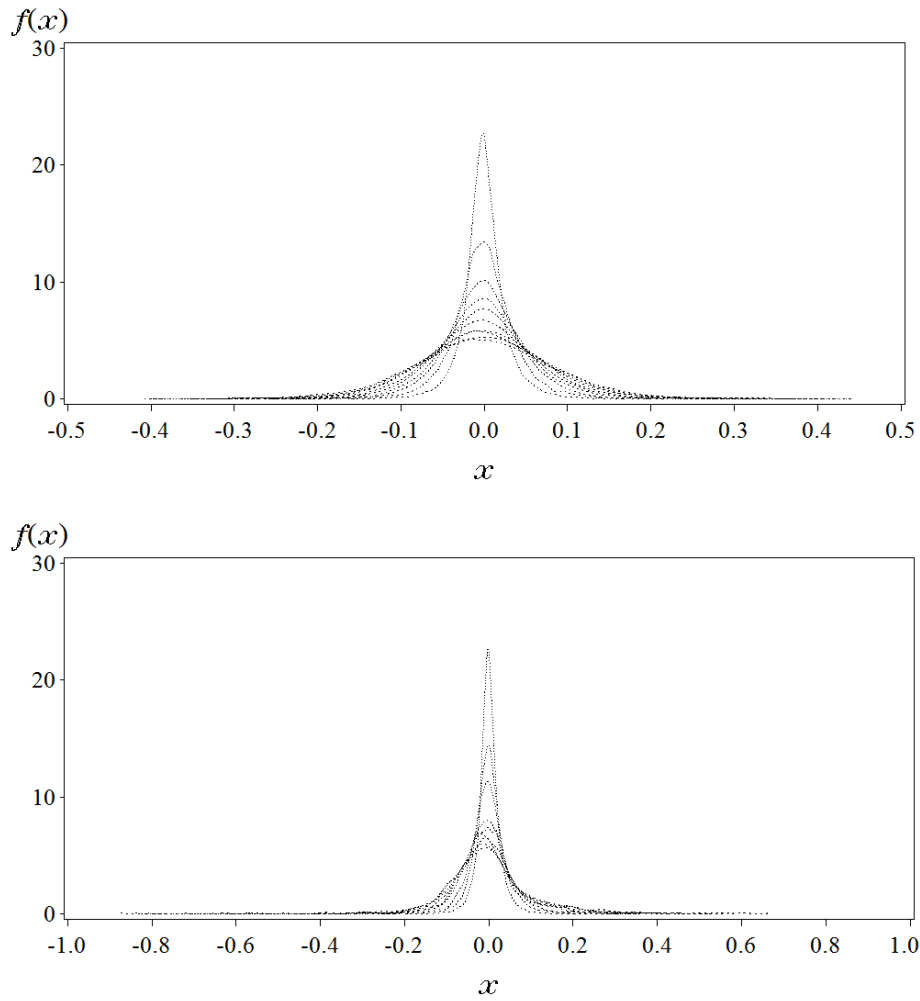


Figura 4.16: IBovespa: estimativas Kernel da FD de $S_{\Delta t}^R$ e $S_{\Delta t}^0$ para $\Delta t = 1, \dots, 10$.

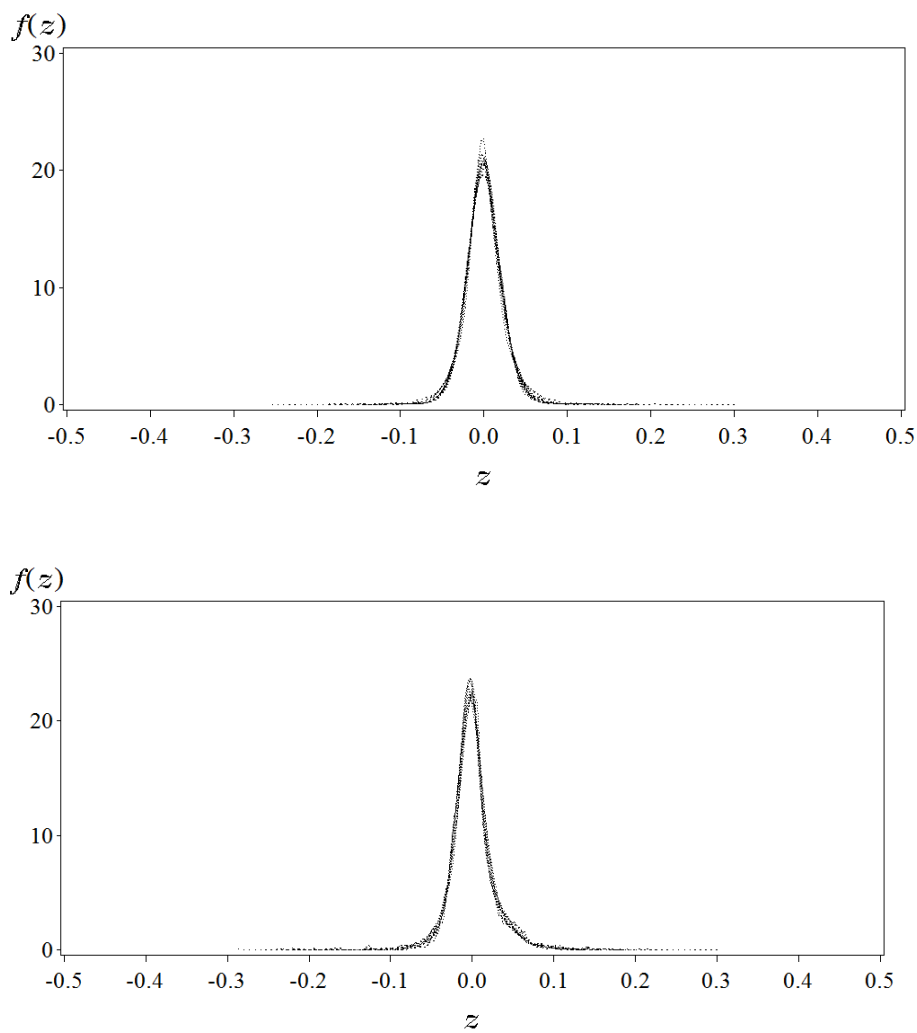


Figura 4.17: IBovespa: estimativas Kernel da FD de $S_{\Delta t}^R$ e $S_{\Delta t}^0$ para $\Delta t = 2, \dots, 10$ colapsadas para $\Delta t = 1$ mediante transformação de escala $z = x(\Delta t)^{-1/1.6}$ e $f(z) = f(x)(\Delta t)^{1/1.6}$.

Capítulo 5

Estimação por funções características

5.1 Introdução

O estudo da estabilidade ou quase estabilidade de um passeio aleatório $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$ pode ser feito com base no comportamento de $\alpha_{S_{\Delta t}}$. Se, por exemplo, $S_{\Delta t}$ for uma soma de VA estáveis IID, então $\alpha_{S_{\Delta t}} = \alpha$ e $\gamma_{S_{\Delta t}} = \Delta t \gamma$, para qualquer Δt . Porém, se $S_{\Delta t}$ não for estável, então $\alpha_{S_{\Delta t}} \rightarrow 2$ à medida que Δt cresce e $\gamma_{S_{\Delta t}} = \Delta t \gamma$. E, se $S_{\Delta t}$ for quase estável [30], então $\alpha_{S_{\Delta t}} \approx \alpha$ enquanto $\Delta t_0 \leq \Delta t \leq \Delta t_1$, e depois $\alpha_{S_{\Delta t}} \rightarrow 2$; e agora, $\gamma_{S_{\Delta t}} \neq \Delta t \gamma$ por causa da presença de autocorrelação serial.

Para esse estudo, há diversos métodos para se estimar os parâmetros de distribuições sob hipótese de estabilidade. Pode-se estimar α com base no gráfico $\ln \hat{P}(X \leq x)$ versus $\ln x$, já que as caudas das leis estáveis seguem o comportamento de Pareto [103]

$$\lim_{x \rightarrow \infty} x^\alpha P(X > x) = \eta \alpha (1 + \beta) \gamma, \quad (5.1)$$

em que η é um fator de normalização, β é o parâmetro de assimetria e γ é o de escala. Assim, se os dados forem estáveis, um padrão linear de inclinação $-\alpha$ deve ser encontrado no gráfico log-log. Também é possível estimar α com base na ordenação dos dados. Considerando que $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ são as estatísticas de ordem da amostra

aleatória $\{X_j\}_{j=1,\dots,n}$ [21, 87],

$$H_{n,k} = \frac{1}{k} \sum_{j=n-k+1}^n \ln \frac{X^{(j)}}{X^{(n-k)}} \quad (5.2)$$

converge em probabilidade para $1/\alpha$ à medida que n aumenta. A fórmula $H_{n,k}$ é chamada estimador de Hill e o inverso multiplicativo de α é chamado *índice caudal* ou *índice de valor extremo*. Outra possibilidade é obter estimativas da FD no ponto zero de $S_{\Delta t}$ e obter α com base na reta que emerge no gráfico log-log de $\hat{f}_{S_{\Delta t}}(0)$ contra Δt [75]. Resultados consistentes podem ser obtidos combinando-se o estimador de Hill com transformações de escala de passeios aleatórios $S_{\Delta t}$ [21]. Os parâmetros também podem ser estimados mediante ajustamento da FCE por regressão não linear [4, 45, 84, 97], i.e., determinar $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$ de modo que se minimize a distância L_2

$$L_2(\boldsymbol{\theta}) = \int (\hat{\phi}(q) - \phi(q; \boldsymbol{\theta}))^2 dx. \quad (5.3)$$

Muitos desses métodos foram propostos frente a limitações computacionais impostas pela forma não fechada da FD das distribuições de Lévy. Com os avanços dos métodos computacionais essas limitações foram quebradas, o que tornou viável o uso do método da máxima verossimilhança (MMV) para a estimação dos parâmetros [94], cuja apresentação é feita na Seção 5.2. O MMV proporciona estimadores com boas propriedades estatísticas como consistência — convergência em probabilidade para os respectivos alvos —, eficiência — variância mínima em comparação com outros estimadores não tendenciosos — e normalidade assintótica das distribuições amostrais [89, 98, 101]. Além disso, o MMV é equivalente à minimização da divergência de Kullback-Leibler entre uma distribuição hipotética e a empírica [26]. Na Seção 5.3, mostramos que a estimação por MV pode ser feita via FCE, o que permite reduzir o esforço computacional para a solução das equações de verossimilhança. A Seção 5.4 apresenta um estudo com dados do IBovespa, das taxas de câmbio frente ao dólar americano e o índice Dow Jones Industrial (DJIA), e a Seção 5.5 expõe algumas observações importantes deste capítulo.

5.2 O Método da Máxima Verossimilhança

Seja $F(x; \boldsymbol{\theta}) = \int_{-\infty}^x f(u; \boldsymbol{\theta}) du$ a função de distribuição acumulada de uma variável aleatória X , em que $f(x; \boldsymbol{\theta})$ é a função de densidade e $\boldsymbol{\theta} \in \mathbb{R}^p$ representa o vetor de parâmetros desconhecidos. Para a estimação de $\boldsymbol{\theta}$ com base em um conjunto $\{X_j\}_{j=1, \dots, n}$ cujos elementos são cópias independentes de X — isto é, uma amostra aleatória simples (AAS) —, o MMV consiste na maximização da função de verossimilhança

$$l(\boldsymbol{\theta}; \mathbf{x}) = \sum_{j=1}^n \ln f(x_j; \boldsymbol{\theta}) \quad (5.4)$$

com respeito ao vetor de parâmetros $\boldsymbol{\theta}$, em que $\mathbf{x} = \{x_1, \dots, x_n\}$ representa uma realização da amostra aleatória. Se $\ln f(x_j; \boldsymbol{\theta})$ for diferenciável com respeito a $\theta_j \in \boldsymbol{\theta}$, o estimador de MV de $\boldsymbol{\theta}$ será a solução o sistema de equações de verossimilhança

$$\begin{aligned} \frac{d}{d\theta_j} l(\boldsymbol{\theta}; \mathbf{x}) &= \frac{1}{n} \sum_{k=1}^n \frac{d}{d\theta_j} \ln f(x_k; \boldsymbol{\theta}) = \\ &= \int_{-\infty}^{+\infty} \frac{d}{d\theta_j} \ln f(x; \boldsymbol{\theta}) dF_n(x) = \\ &= \int_{-\infty}^{+\infty} s_j(x; \boldsymbol{\theta}) dF_n(x) = 0, \end{aligned} \quad (5.5)$$

em que $j = 1, \dots, p$ e $F_n(x)$ é a função de distribuição acumulada empírica e a função

$$s_j(x; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \ln f(x; \boldsymbol{\theta}), \quad (5.6)$$

conhecida como score eficiente [98], indica a sensibilidade relativa de $f(x; \boldsymbol{\theta})$ a variações de $\boldsymbol{\theta}$. Como $\int s(x; \boldsymbol{\theta}) dF(x; \boldsymbol{\theta}) = \frac{d}{d\boldsymbol{\theta}} \int f(x; \boldsymbol{\theta}) dx = 0$, a equação (5.5) é equivalente a

$$\int_{-\infty}^{+\infty} s_j(x; \boldsymbol{\theta}) (dF_n(x) - dF(x; \boldsymbol{\theta})) = 0. \quad (5.7)$$

Naturalmente, a solução da equação de verossimilhança (5.5) requer certo esforço computacional se $f(x; \boldsymbol{\theta})$, e conseqüentemente $s(x; \boldsymbol{\theta})$, não assumir uma expressão em forma fechada — como é o caso de algumas distribuições estáveis [94, 22, 86, 116]. As equações (5.5) e (5.7) são bastante conhecidas e estão presentes em livros textos de inferência estatística [89, 98, 101]. Porém, há uma representação alternativa menos conhecida — a equação de MV no domínio de Fourier [26] — que apresentaremos a seguir.

5.3 A Equação de MV no Domínio de Fourier

Considerando a definição (2.2) da função característica (FC) e a função de densidade (FD) como uma transformada inversa de Fourier, Eq. (2.5), obtém-se a relação

$$\begin{aligned} \int s_j(x; \boldsymbol{\theta}) f(x; \boldsymbol{\theta}) dx &= \frac{1}{2\pi} \int s_j(x; \boldsymbol{\theta}) \int \phi_X(q; \boldsymbol{\theta}) e^{-iqx} dq dx \\ &= \int \left\{ \int \frac{1}{2\pi} s_j(x; \boldsymbol{\theta}) e^{-iqx} dx \right\} \phi_X(q; \boldsymbol{\theta}) dq \\ &= \int \omega_j(q; \boldsymbol{\theta}) \phi_X(q; \boldsymbol{\theta}) dq, \end{aligned} \quad (5.8)$$

em que

$$\omega_j(q; \boldsymbol{\theta}) = \frac{1}{2\pi} \int s_j(x; \boldsymbol{\theta}) e^{-iqx} dx \quad (5.9)$$

é uma transformada inversa do escore $s_j(x; \boldsymbol{\theta})$. Analogamente, considerando a definição (4.4) da FC empírica, temos

$$\int s_j(x; \boldsymbol{\theta}) dF_n(x) dx = \int \omega_j(q; \boldsymbol{\theta}) \hat{\phi}(q) dq. \quad (5.10)$$

Logo, substituindo-se (5.8) e (5.10) em (5.7), a equação original de máxima verossimilhança pode ser escrita como [26]:

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \left\{ \hat{\phi}(q) - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0. \quad (5.11)$$

A forma (5.11) é a equação de verossimilhança no domínio de Fourier [26], pela qual é possível obter estimadores de MV via funções características. Esta é a mesma equação obtida no Cap. 2 com respeito à mínima divergência de Kullback-Leibler. Considerando, porém, o truncamento da FCE discutido no Cap.4, propõe-se a equação

$$\int_{q \in Q} \omega_j(q; \boldsymbol{\theta}) \left\{ \hat{\phi}(q) - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0, \quad (5.12)$$

em que Q é o intervalo de operação da FCE.

A função peso $\omega_j(q; \boldsymbol{\theta})$

Os exemplos a seguir ilustram a aplicação da função peso para a determinação de estimadores de MV para os parâmetros de escala das distribuições gaussiana e de Cauchy, respectivamente.

Exemplo 5.3.1. Considere uma sequência IID $\{X_j\}_{j=1,\dots,n}$ retirada de uma distribuição $N(0, 2\gamma)$, em que $\gamma > 0$ é o parâmetro desconhecido. Nesse caso, a função densidade e a função característica de X_j são, respectivamente, $f(x; \gamma) = \frac{1}{\sqrt{2\pi\gamma}} e^{-\frac{x^2}{4\gamma}}$ e $\phi(q; \gamma) = e^{-\gamma q^2}$. A equação de verossimilhança é $\frac{d}{d\gamma} l(\gamma, \{X_j\}) = -\frac{n}{2\gamma} + \sum \frac{X_i^2}{4\gamma^2} = 0$. Portanto, o estimador de máxima verossimilhança de γ é $\hat{\gamma} = \frac{1}{2n} \sum X_i^2$. A função score é $s(x; \gamma) = \frac{d}{d\gamma} \ln f(x; \gamma) = -\frac{1}{2\gamma} + \frac{X^2}{4\gamma^2}$, de modo que a função peso para esse exemplo é

$$\begin{aligned} \omega(q; \gamma) &= \frac{1}{2\pi} \int s(x; \gamma) e^{-iqx} dx \\ &= \frac{1}{2\pi} \int \left(-\frac{1}{2\gamma} + \frac{x^2}{4\gamma^2} \right) e^{-iqx} dx \\ &= \frac{1}{4\pi\gamma} \left\{ \frac{1}{2\gamma} \int x^2 e^{-iqx} dx - \int e^{-iqx} dx \right\} \\ &= -\frac{1}{4\pi\gamma} \left\{ \frac{1}{2\gamma} \frac{d^2}{dq^2} \int e^{-iqx} dx + \int e^{-iqx} dx \right\} \\ &= -\frac{1}{2\gamma} \left\{ \frac{1}{2\gamma} \delta''(q) + \delta(q) \right\}. \end{aligned} \tag{5.13}$$

Substituindo (5.13) na equação (5.11), e considerando que $\int_{-\infty}^{+\infty} \delta^{(n)}(q) f(q) dq = (-1)^n f^{(n)}(0)$, $\phi'(q; \gamma) = -2q\gamma e^{-\gamma q^2}$, $\phi''(q; \gamma) = (-2\gamma + 4q^2\gamma^2) e^{-\gamma q^2}$, $\phi'_n(q) = -\frac{1}{n} \sum X_i \text{sen}(qX_i)$, $\phi''_n(q) = -\frac{1}{n} \sum X_i^2 \cos(qX_i)$, obtém-se a equação de verossimilhança

$$\begin{aligned} \int \left\{ \frac{\delta''(q)}{2\gamma} + \delta(q) \right\} (\phi_n(q) - \phi(q; \gamma)) dq &= \\ \frac{1}{2\gamma} (\phi''_n(0) - \phi''(0; \gamma)) + (\phi_n(0) - \phi(0; \gamma)) &= \\ -\frac{\sum X_i^2}{n} + 2\gamma &= 0, \end{aligned}$$

ou seja, $\hat{\gamma} = \frac{1}{2n} \sum X_i^2$.

□

Exemplo 5.3.2. Considere agora que $\{X_j\}_{j=1,\dots,n}$ seguem uma distribuição de Cauchy com função densidade e FC, respectivamente, iguais a $f(x; \gamma) = \frac{1}{\pi\gamma\{1+(\frac{x}{\gamma})^2\}}$ e $\phi(q; \gamma) = e^{-\gamma|q|}$. A função score pode ser escrita como

$$s(x; \gamma) = -f(x; \gamma) \pi \left\{ 1 - \left(\frac{x}{\gamma} \right)^2 \right\}, \tag{5.14}$$

e a função peso correspondente é

$$\begin{aligned}
 \omega(q; \gamma) &= \frac{1}{2\pi} \int s(x; \gamma) e^{-iqx} dx \\
 &= -\frac{\pi}{2\pi} \left\{ \int f(x; \gamma) e^{-iqx} dx - \frac{1}{\gamma^2} \int x^2 e^{-iqx} f(x; \gamma) dx \right\} \\
 &= -\frac{1}{2} \left\{ e^{-\gamma|q|} + \frac{\gamma^2}{\gamma^2} e^{-\gamma|q|} \right\} \\
 &= -\frac{1}{2} \left\{ e^{-\gamma|q|} + e^{-\gamma|q|} \right\} \\
 &= -e^{-\gamma|q|}.
 \end{aligned} \tag{5.15}$$

□

Embora a função peso $\omega_j(q; \boldsymbol{\theta})$, de um modo geral, não possua forma fechada para a distribuição de Lévy estável, no Cap. 2, mostramos que a função peso pode ser representada como

$$\omega_j(q; \boldsymbol{\theta}) = \{ \overline{\Phi(D_q; \boldsymbol{\theta})} \}^{-1} \overline{h_j(q; \boldsymbol{\theta})},$$

em que $h_j(q; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \phi(q; \boldsymbol{\theta})$ e $\overline{\Phi(D_q; \boldsymbol{\theta})} = \sum_{k=0}^{+\infty} \overline{c_k} D_q^k$ é o complexo conjugado do polinômio característico cujos coeficientes são $\overline{c_k} = \frac{2\pi i^k f^{(k)}}{k!}$.

O inverso multiplicativo de $\overline{\Phi(D_q; \boldsymbol{\theta})}$

O polinômio $\overline{\Phi(D_q; \boldsymbol{\theta})}$ pode ser representado como

$$\begin{aligned}
 \overline{\Phi(D_q; \boldsymbol{\theta})} &= \sum_{k=0}^{+\infty} \overline{c_k} D_q^k \\
 &= c_0 \sum_{k=0}^{+\infty} \overline{d_k} D_q^k \\
 &= c_0 \left\{ 1 + \sum_{k=1}^{+\infty} \overline{d_k} D_q^k \right\} \\
 &= c_0 \{ 1 - \Psi(D_q; \boldsymbol{\theta}) \},
 \end{aligned} \tag{5.16}$$

em que $d_k \doteq \frac{c_k}{c_0}$ — lembrando-se que $c_{2k} \in \mathbb{R}$ e $c_{2k+1} \in \mathbb{C}$ —, e $\Psi(D_q; \boldsymbol{\theta}) = -\sum_{k=1}^{+\infty} \bar{d}_k D_q^k$.

Admitindo-se a existência do inverso multiplicativo de $\overline{\Phi(D_q; \boldsymbol{\theta})}$, pode-se escrever

$$\begin{aligned} \{\overline{\Phi(D_q; \boldsymbol{\theta})}\}^{-1} &= \frac{1}{c_0\{1 - \Psi(D_q; \boldsymbol{\theta})\}} \\ &= \frac{1}{c_0} \sum_{k=0}^{+\infty} \Psi^k(D_q; \boldsymbol{\theta}) \\ &= \frac{1}{c_0} \sum_{k=0}^{+\infty} \bar{d}'_k D_q^k. \end{aligned} \tag{5.17}$$

Por exemplo, para obtermos os quatro primeiros coeficientes \bar{d}'_k , tem-se

- $k = 0$, $\Psi^0(D_q; \boldsymbol{\theta}) = 1$,
- $k = 1$, $\Psi^1(D_q; \boldsymbol{\theta}) = -\sum_{k=1}^{+\infty} \bar{d}_k D_q^k$,
- $k = 2$, $\Psi^2(D_q; \boldsymbol{\theta}) = \sum_{j,k=1}^{+\infty} \bar{d}_j \bar{d}_k D_q^{j+k}$,
- $k = 3$, $\Psi^3(D_q; \boldsymbol{\theta}) = -\sum_{j,k,l=1}^{+\infty} \bar{d}_j \bar{d}_k \bar{d}_l D_q^{j+k+l}$,
- $k = 4$, $\Psi^4(D_q; \boldsymbol{\theta}) = \sum_{j,k,l,m=1}^{+\infty} \bar{d}_j \bar{d}_k \bar{d}_l \bar{d}_m D_q^{j+k+l+m}$.

Em seguida, após agrupar e somar os coeficientes correspondentes a cada D_q^k , obtemos:

$$\bar{d}'_0 = 1, \tag{5.18}$$

$$\bar{d}'_1 = -\bar{d}_1, \tag{5.19}$$

$$\bar{d}'_2 = \bar{d}_1^2 - \bar{d}_2, \tag{5.20}$$

$$\bar{d}'_3 = -(\bar{d}_1^3 - 2\bar{d}_1\bar{d}_2 + \bar{d}_3), \tag{5.21}$$

$$\bar{d}'_4 = \bar{d}_1^4 - 3(\bar{d}_1)^2(\bar{d}_2) + 2\bar{d}_1\bar{d}_3 - \bar{d}_4 + \bar{d}_2^2. \tag{5.22}$$

Se X for simétrica em torno de zero, $c_{2k+1} = 0$ para todo $k \geq 1$. Logo, nessa situação particular, $d_{2k+1} = 0$ e

$$d'_0 = 1, \tag{5.23}$$

$$d'_2 = -d_2, \tag{5.24}$$

$$d'_4 = d_2^2 - d_4, \tag{5.25}$$

Nessa situação, para a estimação de parâmetros de distribuições simétricas em torno de zero pode-se considerar a aproximação

$$\{\overline{\Phi(D_q; \boldsymbol{\theta})}\}^{-1} \approx \sum_{k=0}^4 \overline{d_k} D_q^k. \quad (5.26)$$

Em particular, caso X seja estável com parâmetros $\boldsymbol{\theta}_0 = (\alpha, \beta, \gamma)'$, tem-se que a assimetria efetiva (Cap. 3) é dada por $\beta_\alpha = \beta \tan \frac{\pi\alpha}{2}$, e as derivadas da função de densidade no ponto zero são

$$f^{(2j+1)}(0; \boldsymbol{\theta}_0) = \frac{(-1)^j \Gamma\left(\frac{2j+2}{\alpha}\right)}{\pi \alpha \gamma_*^{\frac{2j+2}{\alpha}}} \cdot A_{1,j}$$

e

$$f^{(2j)}(0; \boldsymbol{\theta}_0) = \frac{(-1)^j \Gamma\left(\frac{2j+1}{\alpha}\right)}{\pi \alpha \gamma_*^{\frac{2j+1}{\alpha}}} \cdot A_{2,j},$$

em que $j = 0, 1, 2, \dots$, $\gamma_* = \gamma \cdot \sqrt{1 + \beta_\alpha^2}$, $A_{1,j} = \sin\{(2j + 2)(\arctan \beta_\alpha)/\alpha\}$, e $A_{2,j} = \cos\{(2j + 1)(\arctan \beta_\alpha)/\alpha\}$.

Dada uma estimativa preliminar $(\alpha_0, \beta_0, \gamma_0)'$, os coeficientes $\overline{d_k}$ e as derivadas de $h_j(q; \boldsymbol{\theta})$ com respeito a q podem ser obtidos numericamente. Com base nessa aproximação inicial de $\omega_j(q; \boldsymbol{\theta})$, a solução das equações de verossimilhança pode ser obtida numericamente, utilizando-se pacotes estatísticos como o SAS (www.sas.com) e o R (www.r-project.org). Uma programa executável específico para estimação por MV foi desenvolvido por Nolan [94] e se encontra disponível em <http://academic2.american.edu/~jpnolan/>.

5.4 Estudo com dados financeiros

5.4.1 IBovespa

Com respeito aos retornos do IBovespa, $\{X_t\}$, a Tab. 5.1 apresenta as estimativas de máxima verossimilhança (EMV) dos parâmetros da sua distribuição, sob hipótese de estabilidade. Considerando-se a soma $S_{\Delta t}^R$ (aleatorizada) e a $S_{\Delta t}^0$ (não aleatorizada), conforme

as Eqs. (4.48) e (4.49), a Fig. 5.1 mostra as estimativas correspondentes de α para cada Δt considerado ($1 \leq \Delta t \leq 100$).

Considere que $\alpha_{\Delta t}^R$ e $\alpha_{\Delta t}^0$ sejam as EMV para os processos $S_{\Delta t}^R$ e $S_{\Delta t}^0$, respectivamente. Enquanto $\alpha_{\Delta t}^R \rightarrow 2$, o que sugere que os retornos não são estáveis, observa-se que $1,40 \leq \alpha_{\Delta t}^0 \leq 1,61$. Assim, levando-se em conta as autocorrelações, o processo $S_{\Delta t}^0$ permanece aproximadamente dentro do regime de Lévy [30, 31], apesar de os retornos não serem estáveis.

Tabela 5.1: Distribuição dos retornos do IBovespa sob hipótese de estabilidade ($\Delta t = 1$): estimativas de MV \pm intervalos de confiança de 95%

α	β	γ	$\sigma = \gamma^{\frac{1}{\alpha}}$
1.566 ± 0.030	0.132 ± 0.062	0.001199 ± 0.000003	0.013650 ± 0.000266

A Fig. 5.2 mostra as estimativas de $\gamma_{\Delta t}$ (painel superior) e de $\sigma_{\Delta t} = \gamma_{\Delta t}^{1/\alpha}$ (painel inferior), em que $\sigma_{\Delta t}$ corresponde ao desvio padrão no caso gaussiano. Para o processo $S_{\Delta t}^R$, tem-se $\gamma_{\Delta t}^R \approx 0.0004\Delta t$, enquanto $\gamma_{\Delta t}^0 \approx 0.0012\Delta t$ para o processo $S_{\Delta t}^0$. Observa-se também que $\sigma_{\Delta t}^R$ e $\sigma_{\Delta t}^0$ se distanciam substancialmente a partir do intervalo $\Delta t = 20$ (ou $\ln \Delta t \approx 3$). Esse intervalo coincide com aquele em que $S_{\Delta t}^R$ entra no regime gaussiano (Fig. 5.1).

Esse desvio, que se relaciona com a dependência serial de longo alcance [14], pode ser avaliado com base no expoente de Hurst (H) [5, 16, 48, 72, 82], definido como

$$\sigma_{\Delta t} \sim \Delta t^H. \quad (5.27)$$

Por exemplo, para um passeio aleatório gaussiano com incrementos independentes tem-se $H = 0,5$. Mas se houver dependência de longo alcance com autocorrelações positivas, então $H > 0,5$. A Fig. 5.2 apresenta os ajustes por regressão não linear da forma (5.27) (curvas tracejadas). Como se espera, a curva ajustada para o processo $S_{\Delta t}^R$ foi $\sigma_{\Delta t}^R \approx (0,0004\Delta t)^{0,5}$, ou seja, $H = 0,5$. Já para $S_{\Delta t}^0$, a curva estimada foi $\sigma_{\Delta t}^0 \approx (0,0012\Delta t)^{0,67}$, o que sugere a presença de dependência de longo alcance na série de retornos do IBovespa com $H \approx 0,67$. Conforme o gráfico da função de autocorrelação amostral mostrada na

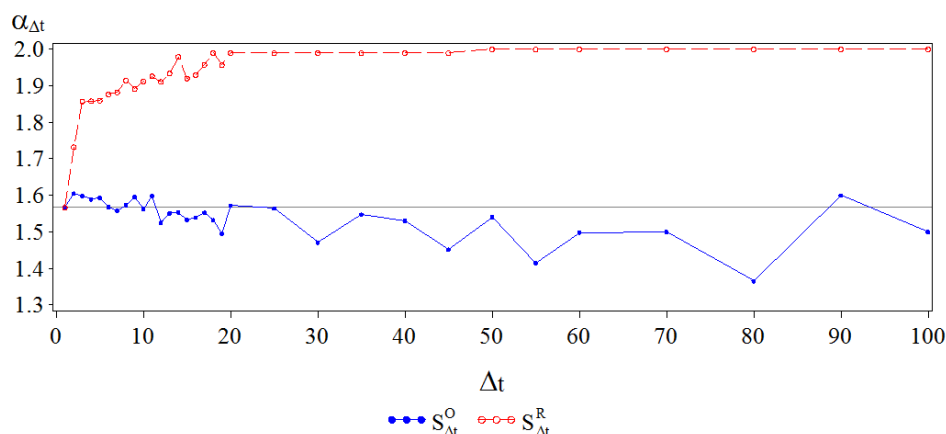


Figura 5.1: IBovespa: estimativas de máxima verossimilhança de $\alpha_{\Delta t}$ para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$. O processo aleatorizado $S_{\Delta t}^R$ entra no regime gaussiano após $\Delta t = 20$.

Fig. 5.3, o expoente $H > 0,5$ é consistente com a presença de autocorrelações persistentes e predominantemente positivas.

A partir do instante Δt em que $S_{\Delta t}^R$ entra no domínio gaussiano (i.e., para Δt suficientemente grande), a diferença

$$C_{\Delta t} = \ln \sigma_{\Delta t}^0 - \ln \sigma_{\Delta t}^R \tag{5.28}$$

permite medir o expoente H , pois $C_{\Delta t} \sim (H - \frac{1}{2}) \ln \Delta t$. Na Fig. 5.4, a reta ajustada a partir de $\Delta t = 20$ foi $C_{\Delta t} \approx (0,67 - 0,5) \ln \Delta t$, e, assim, obtemos novamente $H = 0,67$.

A Fig. 5.5 mostra o comportamento da assimetria efetiva $\beta_{\alpha, \Delta t} = \beta_{\Delta t} \tan \frac{\pi \alpha_{\Delta t}}{2}$. Para $S_{\Delta t}^R$, observa-se $\beta_{\alpha, \Delta t}^R \rightarrow 0$, mas para $S_{\Delta t}^O$, a assimetria se distancia de zero à medida que Δt aumenta. Logo, de um modo geral, além de produzir resistência para que o processo entre no regime gaussiano, a autocorrelação serial também inflaciona os parâmetros de escala e de assimetria.

5.4.2 Taxas de câmbio

Considerando-se as taxas de câmbio descritas no Cap. 1, foram geradas replicações dos passeios aleatorizados $S_{\Delta t}^R$ e das somas parciais $S_{\Delta t}^O$, conforme o procedimento descrito na

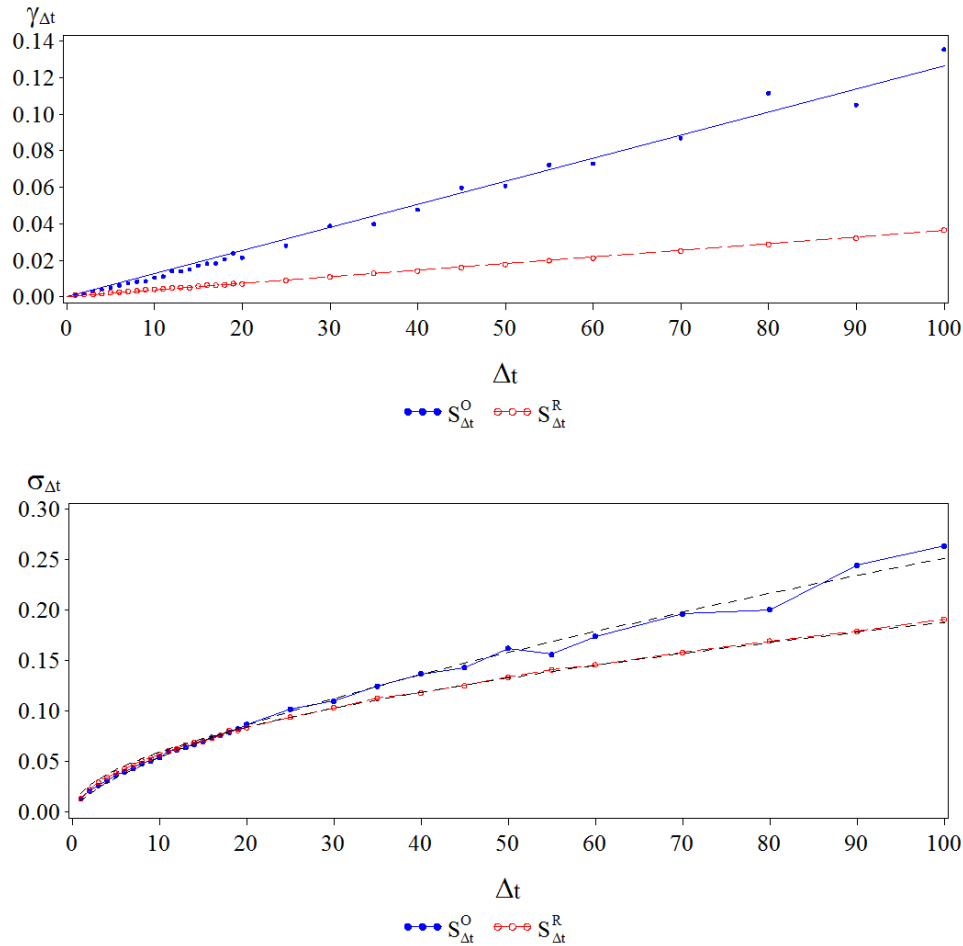


Figura 5.2: Ibovespa: estimativas de máxima verossimilhança de $\gamma_{\Delta t}$ (painel superior) e $\sigma_{\Delta t} = \gamma_{\Delta t}^{\frac{1}{2}}$ (painel inferior) para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt . As curvas tracejadas representam os ajustes $\sigma_{\Delta t} \approx (0,0004\Delta t)^{0,5}$ (para $S_{\Delta t}^R$) e $\sigma_{\Delta t} \approx (0,0012\Delta t)^{0,67}$ (para $S_{\Delta t}^O$), o que evidencia a dependência de longo alcance no Ibovespa.

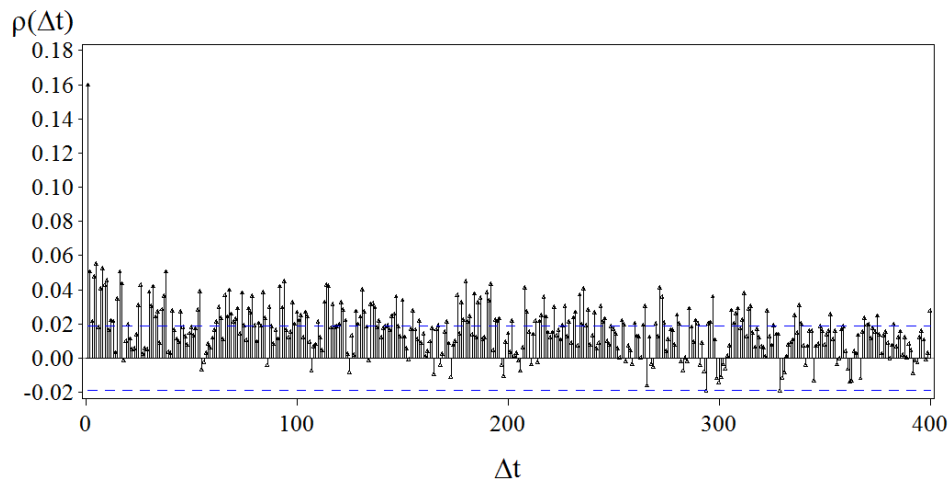


Figura 5.3: Ibovespa: gráfico da função de autocorrelação amostral dos retornos do Ibovespa, em que $\rho(\Delta t) = \text{Corr}(X_t, X_{t+\Delta t})$, $\Delta t \geq 1$. As autocorrelações são persistentes e predominantemente positivas, o que indica que há dependência de longo alcance entre os retornos do Ibovespa.

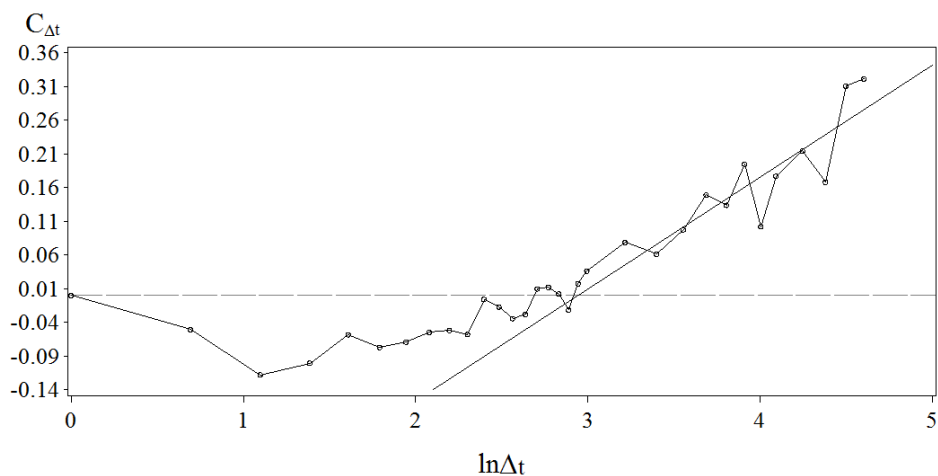


Figura 5.4: Estimativas de $C_{\Delta t} = \ln \sigma_{\Delta t}^0 - \ln \sigma_{\Delta t}^R$ versus $\ln \Delta t$. Para $\ln \Delta t > 3$, a reta apresenta inclinação igual a $H - 0,5 \approx 0,17$, i.e., $H \approx 0,67$.

Seção 4.7 do Cap. 4. Para $\Delta t = 1$, sob a hipótese de os retornos serem estáveis, a Tab. 5.2 mostra as estimativas de máxima verossimilhança (EMV) de α , β , γ e σ . Como se espera, de um modo geral, os índices de estabilidade dos processos $S_{\Delta t}^R$ tendem para 2

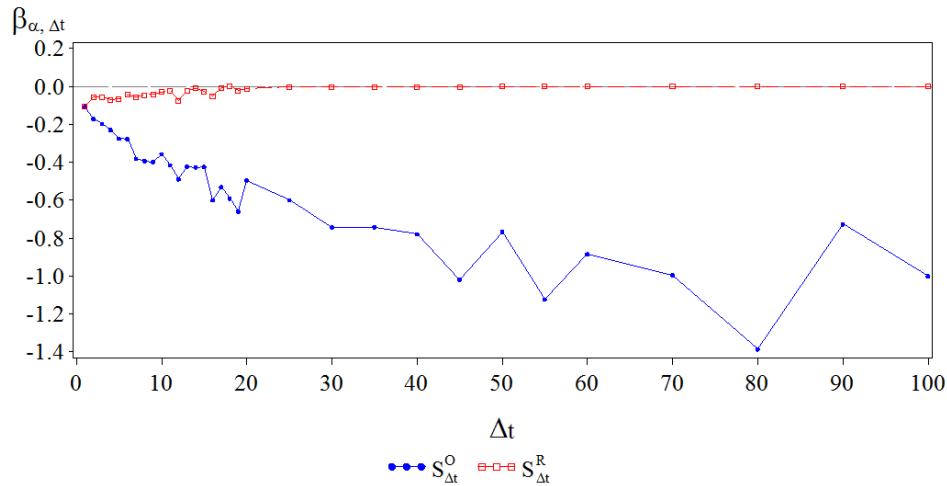


Figura 5.5: IBovespa: estimativas de máxima verossimilhança para a assimetria efetiva $\beta_{\alpha, \Delta t} = \beta_{\Delta t} \tan \frac{\pi \alpha \Delta t}{2}$ para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt .

mais rapidamente do que os dos processos $S_{\Delta t}^O$. A Fig. 5.6 mostra os comportamentos das estimativas de $\alpha_{\Delta t}$ para o real, a libra esterlina e o franco suíço.

Tabela 5.2: Taxas de câmbio: estimativas de MV e intervalos de confiança de 95% ($\Delta t = 1$)

país	moeda	α	β	γ	σ
África do Sul	Rand	0.880 ± 0.023	$+0.028 \pm 0.031$	0.004100 ± 0.000191	0.001940 ± 0.000060
Austrália	Dólar australiano	1.333 ± 0.030	-0.021 ± 0.048	0.000357 ± 0.000002	0.002593 ± 0.000059
Brasil	Real	1.285 ± 0.045	$+0.003 \pm 0.070$	0.000739 ± 0.000010	0.003651 ± 0.000132
Canadá	Dólar canadense	1.448 ± 0.031	$+0.001 \pm 0.055$	0.000102 ± 0.000001	0.001756 ± 0.000038
Coréia do Sul	Won	0.840 ± 0.026	-0.022 ± 0.034	0.002621 ± 0.000164	0.000847 ± 0.000031
Dinamarca	Coroa dinamarquesa	1.664 ± 0.030	-0.022 ± 0.078	0.000088 ± 0.000001	0.003652 ± 0.000069
Índia	Rúpia	1.090 ± 0.030	$+0.093 \pm 0.046$	0.000622 ± 0.000003	0.001145 ± 0.000057
Japão	Yen	1.580 ± 0.030	-0.099 ± 0.065	0.000125 ± 0.000001	0.003377 ± 0.000067
México	Peso	1.503 ± 0.046	$+0.215 \pm 0.084$	0.000193 ± 0.000001	0.003377 ± 0.000095
Nova Zelândia	Dólar neozelandês	1.325 ± 0.030	$+0.003 \pm 0.047$	0.000407 ± 0.000003	0.002761 ± 0.000063
Noruega	Coroa norueguesa	1.604 ± 0.031	-0.001 ± 0.069	0.000115 ± 0.000001	0.003499 ± 0.000069
Reino Unido	Libra Esterlina	1.641 ± 0.030	-0.089 ± 0.074	0.000086 ± 0.000001	0.003326 ± 0.000064
Singapura	Dólar de Singapura	1.564 ± 0.035	$+0.011 \pm 0.075$	0.000045 ± 0.000001	0.001655 ± 0.000038
Suécia	Coroa sueca	1.603 ± 0.031	$+0.010 \pm 0.069$	0.000113 ± 0.000001	0.003457 ± 0.000068
Suíça	Franco suíço	1.704 ± 0.029	-0.090 ± 0.083	0.000091 ± 0.000001	0.004251 ± 0.000079
Tailândia	Baht	0.934 ± 0.012	-0.012 ± 0.037	0.001354 ± 0.000058	0.000846 ± 0.000029
Taiwan	Dólar de Taiwan	1.102 ± 0.033	-0.016 ± 0.046	0.000372 ± 0.000008	0.000774 ± 0.000025
Zona do Euro	Euro	1.863 ± 0.042	-0.133 ± 0.261	0.000038 ± 0.000001	0.004253 ± 0.000124

Porém, os retornos das taxas de câmbio apresentam concentração (ou excesso) de zeros, pois $P(X_t = 0) > 0$. A Tab. 5.3 mostra as estimativas das incidências de retornos nulos ($\hat{P}(X = 0)$) nesses dados. Na série dos retornos da rúpia, por exemplo, aproximadamente

28,5% dos valores são nulos. Essa concentração de zeros indica que sua distribuição é mista, o que afeta o comportamento de $\alpha_{\Delta t}$. A Fig. 5.7 mostra as estimativas de $\alpha_{\Delta t}$ para o dólar australiano, a rúpia (Índia) e o rand (África do Sul). Para o rand, os índices de estabilidade do processo $S_{\Delta t}^R$ tendem para 2 menos rapidamente do que os dos mesmos processos para as moedas com menor incidência de zeros. Para a rúpia e o dólar australiano, os índices não tendem para 2, o que sugere que as distribuições das somas dos retornos correspondentes podem ser estáveis.

Uma distribuição X que possui concentração de zeros pode ser definida como

$$X = Y \cdot I, \quad (5.29)$$

em que Y é uma VA contínua e I segue uma distribuição de Bernoulli definida por $P(I = 0) = P(X = 0)$ e $P(I = 1) = P(X \neq 0) = 1 - P(X = 0)$. Nesse caso, a FC de X pode ser escrita como

$$\phi_X(q; \theta_X) = P(X = 0) + P(X \neq 0) \cdot \phi_Y(q; \theta_Y). \quad (5.30)$$

Assim, $\lim_{|q| \rightarrow \infty} \phi_X(q; \theta_X) = P(X = 0)$ e, se o objeto de estimação for θ_Y , a FCE de Y pode ser obtida com base na FCE de X ,

$$\hat{\phi}_Y(q) = \frac{\hat{\phi}_X(q) - \hat{P}(X = 0)}{1 - \hat{P}(X = 0)}. \quad (5.31)$$

Se Y for uma distribuição estável, o excesso de zeros produz leptocurtose. Como ilustração, a Fig. 5.8 mostra a FCE da série temporal original dos retornos da rúpia ($\hat{\phi}_X(q)$) e a da mesma moeda sem os retornos nulos $\hat{\phi}_Y(q)$. Nessa figura, a curva referente a $\hat{\phi}_Y(q)$ representa tanto a FCE obtida via Eq. (5.31) como aquela obtida com base nos dados sem os retornos nulos (mediante aplicação da Eq. (4.4)). Como aproximadamente 28,5% dos retornos da rúpia são nulos, tal concentração de zeros eleva substancialmente as caudas da FCE. Empiricamente, nota-se que $\hat{\phi}_X(q) \rightarrow 0,285$ à medida que $|q|$ aumenta. Logo, considerando que todas as moedas possuem excesso de zeros, as estimativas de α mostradas na Tab. 5.2 devem ser recalculadas.

Excluindo-se os retornos nulos, a Tab. 5.4 mostra as novas EMV, sob hipótese de estabilidade, para $\Delta t = 1$. Com essa exclusão, as estimativas de α aumentam. Por

Tabela 5.3: Taxas de câmbio: quantidade de retornos nulos ($X_t = 0$).

moeda	n	quantidade de zeros	(%)
Rand	10.315	1.935	18,76
Dólar australiano	10.335	859	8,31
Real	4.324	177	4,09
Dólar Canadense	10.348	203	1,96
Won	7.726	1.053	13,63
Coroa dinamarquesa	10.341	242	2,34
Rúpia	9.834	2.800	28,47
Yen	10.336	359	3,47
Peso mexicano	4.610	284	6,16
Dólar neozelandês	10.326	860	8,33
Coroa Norueguesa	10.341	283	2,74
Libra esterlina	10.342	148	1,43
Dólar de Singapura	7.841	358	4,57
Coroa sueca	10.341	260	2,51
Franco Suíço	10.342	160	1,55
Baht	7.738	1.723	22,27
Dólar de Taiwan	6.855	1.187	17,32
Euro	3.321	28	0,84

exemplo, para a série de retornos da rúpia, a estimativa do índice de estabilidade aumenta de 1,09 para 1,36. A Fig.5.9 mostra a tendência do aumento relativo de α em função da incidência de zeros na amostra.

Com respeito ao parâmetro de escala, como a retirada de zeros modifica a estrutura de dependência serial, apenas descreveremos os resultados para alguns casos em que há baixa incidência de zeros, sem excluí-los da série. A Fig. 5.10 mostra as estimativas de $\sigma_{\Delta t}^R$ e de $\sigma_{\Delta t}^0$ para as taxas de câmbio do real, do franco suíço e da libra esterlina. Para o processo aleatorizado, respectivamente, as curvas estimadas foram $\sigma_{\Delta t} \approx (0,000102\Delta t)^{0,57}$, $\sigma_{\Delta t} \approx (0,000031\Delta t)^{0,51}$ e $\sigma_{\Delta t} \approx (0,000024\Delta t)^{0,52}$. Para o processo $\sigma_{\Delta t}^0$, as curvas correspondentes foram $\sigma_{\Delta t} \approx (0,000099\Delta t)^{0,62}$, $\sigma_{\Delta t} \approx (0,000071\Delta t)^{0,58}$ e $\sigma_{\Delta t} \approx (0,000033\Delta t)^{0,55}$. Quanto à assimetria efetiva, $\beta_{\alpha,\Delta t}$ da variável $S_{\Delta t}^R$ tende para zero mais rapidamente do

Tabela 5.4: Taxas de câmbio: estimativas de MV e intervalos de confiança de 95% ($\Delta t = 1$), excluindo-se os casos em que os retornos são nulos

país	moeda	α	β	γ	σ
África do Sul	Rand	1.354 ± 0.033	$+0.018 \pm 0.054$	0.000000 ± 0.000000	0.003829 ± 0.000096
Austrália	Dólar australiano	1.471 ± 0.032	-0.077 ± 0.058	0.000000 ± 0.000000	0.003062 ± 0.000067
Brasil	Real	1.378 ± 0.048	$+0.020 \pm 0.079$	0.000000 ± 0.000000	0.004044 ± 0.000142
Canadá	Dólar canadense	1.468 ± 0.031	$+0.000 \pm 0.057$	0.000000 ± 0.000000	0.001809 ± 0.000039
Coréia do Sul	Won	1.049 ± 0.032	-0.024 ± 0.044	0.000000 ± 0.000000	0.001342 ± 0.000045
Dinamarca	Coroa dinamarquesa	1.686 ± 0.030	-0.021 ± 0.082	0.000000 ± 0.000000	0.003757 ± 0.000071
Índia	Rúpia	1.364 ± 0.036	$+0.061 \pm 0.059$	0.000000 ± 0.000000	0.002032 ± 0.000055
Japão	Yen	1.631 ± 0.030	-0.123 ± 0.073	0.000000 ± 0.000000	0.003558 ± 0.000070
México	Peso	1.542 ± 0.047	$+0.253 \pm 0.092$	0.000000 ± 0.000000	0.003314 ± 0.000103
Nova Zelândia	Dólar neozelandês	1.454 ± 0.032	-0.022 ± 0.058	0.000000 ± 0.000000	0.003246 ± 0.000072
Noruega	Coroa norueguesa	1.644 ± 0.030	$+0.011 \pm 0.076$	0.000000 ± 0.000000	0.003641 ± 0.000071
Reino Unido	Libra Esterlina	1.659 ± 0.030	-0.097 ± 0.077	0.000000 ± 0.000000	0.003390 ± 0.000065
Singapura	Dólar de Singapura	1.610 ± 0.036	$+0.015 \pm 0.082$	0.000000 ± 0.000000	0.001756 ± 0.000040
Suécia	Coroa sueca	1.639 ± 0.031	$+0.014 \pm 0.075$	0.000000 ± 0.000000	0.003583 ± 0.000070
Suíça	Franco suíço	1.721 ± 0.029	-0.095 ± 0.088	0.000000 ± 0.000000	0.004332 ± 0.000080
Tailândia	Baht	1.268 ± 0.038	0.017 ± 0.058	0.000000 ± 0.000000	0.001546 ± 0.000048
Taiwan	Dólar de Taiwan	1.365 ± 0.040	0.029 ± 0.067	0.000000 ± 0.000000	0.001160 ± 0.000035
Zona do Euro	Euro	1.864 ± 0.042	-0.133 ± 0.261	0.000000 ± 0.000000	0.004280 ± 0.000125

que a da soma $S_{\Delta t}^0$ (Fig. 5.11).

5.4.3 Índice DJIA

Sob hipótese de estabilidade, a Tab. 5.5 mostra as estimativas de máxima verossimilhança (EMV) dos parâmetros da distribuição dos retornos do Índice Dow Jones Industrial. Os comportamentos das estimativas dos índices de estabilidade do processo aleatorizado $S_{\Delta t}^R$ e do processo não aleatorizado $S_{\Delta t}^0$ são apresentados na Fig. 5.12, para $1 \leq \Delta t \leq 100$. Ao contrário do IBovespa, apesar da aleatorização, o processo $S_{\Delta t}^R$ não converge para a gaussiana (aproximadamente, observa-se que $\alpha_{\Delta t} \rightarrow 1.87$). E para $S_{\Delta t}^0$, tem-se $1,4 \leq \alpha_{\Delta t} \leq 1,5$.

Tabela 5.5: Distribuição dos retornos do DJIA sob hipótese de estabilidade ($\Delta t = 1$): estimativas de MV \pm intervalos de confiança de 95%

α	β	γ	σ
1.482 ± 0.012	-0.009 ± 0.023	$3.22E - 06 \pm 3.58E - 09$	0.000197 ± 0.000002

A Fig. 5.13 apresenta as estimativas de $\gamma_{\Delta t}$ (painel superior) e de $\sigma_{\Delta t} = \gamma_{\Delta t}^{1/\alpha}$ (painel

inferior). Para o processo $S_{\Delta t}^R$, tem-se $\gamma_{\Delta t}^R \approx 2,4 \cdot 10^{-7} \Delta t$, enquanto $\gamma_{\Delta t}^0 \approx 1,5 \cdot 10^{-6} \Delta t$ para o processo $S_{\Delta t}^0$. Com respeito a $\sigma_{\Delta t}$, a curva relativa ao processo Δt^R se encontra acima da curva do processo não aleatorizado. Os ajustes da forma (5.27) por regressão não linear (curvas tracejadas) mostram que $\sigma_{\Delta t}^R \approx (2,3 \cdot 10^{-7} \Delta t)^{0,54}$ e $\sigma_{\Delta t}^0 \approx (8,0 \cdot 10^{-8} \Delta t)^{0,52}$.

Aqui, porém, é necessário reinterpretar o expoente H , pois o índice de estabilidade do processo aleatorizado se encontra no patamar $\alpha \approx 1,858$ após $\Delta t = 20$ (ou $1/\alpha \approx 0,54$). Ou seja, não há gaussianização e, logo, não há sentido em afirmar que $H = 0,5$ indica ausência de memória de longo alcance. Se o processo $S_{\Delta t}^R$ for estável com $\alpha < 2$, seu parâmetro de escala será $\gamma_{\Delta t} = \Delta t \cdot \gamma$, ou seja, $\sigma_{\Delta t} = \gamma_{\Delta t}^{\frac{1}{\alpha}} = \sigma(\Delta t)^{\frac{1}{\alpha}}$. Portanto, agora, há ausência de dependência de longo alcance quando $H = 1/\alpha$. Assim, o ajuste $\sigma_{\Delta t}^R \approx (2,3E - 7\Delta t)^{0,54}$ proposto na Fig. 5.13 é consistente com a hipótese de não gaussianidade de $S_{\Delta t}^R$.

Desse modo, o expoente $H = 0,52 < \alpha^{-1}$ para o processo $S_{\Delta t}^0$ significa que há um tipo de dependência de longo alcance em que as autocorrelações positivas e negativas se alternam, o que se confirma com base no gráfico da função de autocorrelação amostral mostrada na Fig. 5.14.

5.5 Considerações

Neste capítulo, mostramos que a estimação por máxima verossimilhança dos parâmetros da distribuição de Lévy estável pode ser efetuada por meio de funções características. Essas estimativas permitem avaliar de forma consistente o comportamento da estabilidade ou quase estabilidade de passeios aleatórios $S_{\Delta t}$. O parâmetro de estabilidade $\alpha_{S_{\Delta t}}$ permite avaliar a convergência para a gaussiana, de modo que, se os dados forem estáveis, $\alpha_{S_{\Delta t}} = \alpha$ e $\gamma_{S_{\Delta t}} = \Delta t \gamma$ (ou $\sigma_{S_{\Delta t}} = (\Delta t)^{1/\alpha} \sigma$), para qualquer Δt . Porém, se os dados não forem estáveis, então $\alpha_{S_{\Delta t}} \rightarrow 2$ à medida que Δt cresce, embora tenhamos ainda $\gamma_{S_{\Delta t}} = \Delta t \gamma$. E, se houver autocorrelação serial, então é possível que $\alpha_{S_{\Delta t}} \approx \alpha$ enquanto $\Delta t_0 \leq \Delta t \leq \Delta t_1$. Mas o comportamento do parâmetro de escala será diferente do caso IID, pois $\gamma_{S_{\Delta t}} \neq \Delta t \gamma$ por causa da estrutura de autocorrelação serial.

As ilustrações com dados do IBovespa e de algumas taxas de câmbio mostraram que $S_{\Delta t}^R$ converge rapidamente para a gaussiana, enquanto $S_{\Delta t}^0$ tende a permanecer no domínio da distribuição de Levy por conta das autocorrelações seriais [30]. A dependência de longo alcance pode ser examinada com base na comparação entre $\sigma_{S_{\Delta t}^0}$ e $\sigma_{S_{\Delta t}^R}$.

Para outros dados, como o índice DJIA, porém, não se observou gaussianização do passeio aleatorizado $S_{\Delta t}^R$. Isso sugere que tais retornos podem ser estáveis, e o expoente de Hurst deve ser reinterpretado, tendo como referência α^{-1} em lugar do valor usual 0,5.

Também mostramos que a incidência de retornos nulos produz uma tendência de superestimação do índice de estabilidade e, por isso, o excesso de zeros deve ser removido para a correta avaliação de α .

Como tanto a dependência serial linear como não linear produzem efeitos na convergência para a gaussiana, o próximo capítulo apresenta um teste que possui poder estatístico para a detecção de vários tipos de dependência não linear.

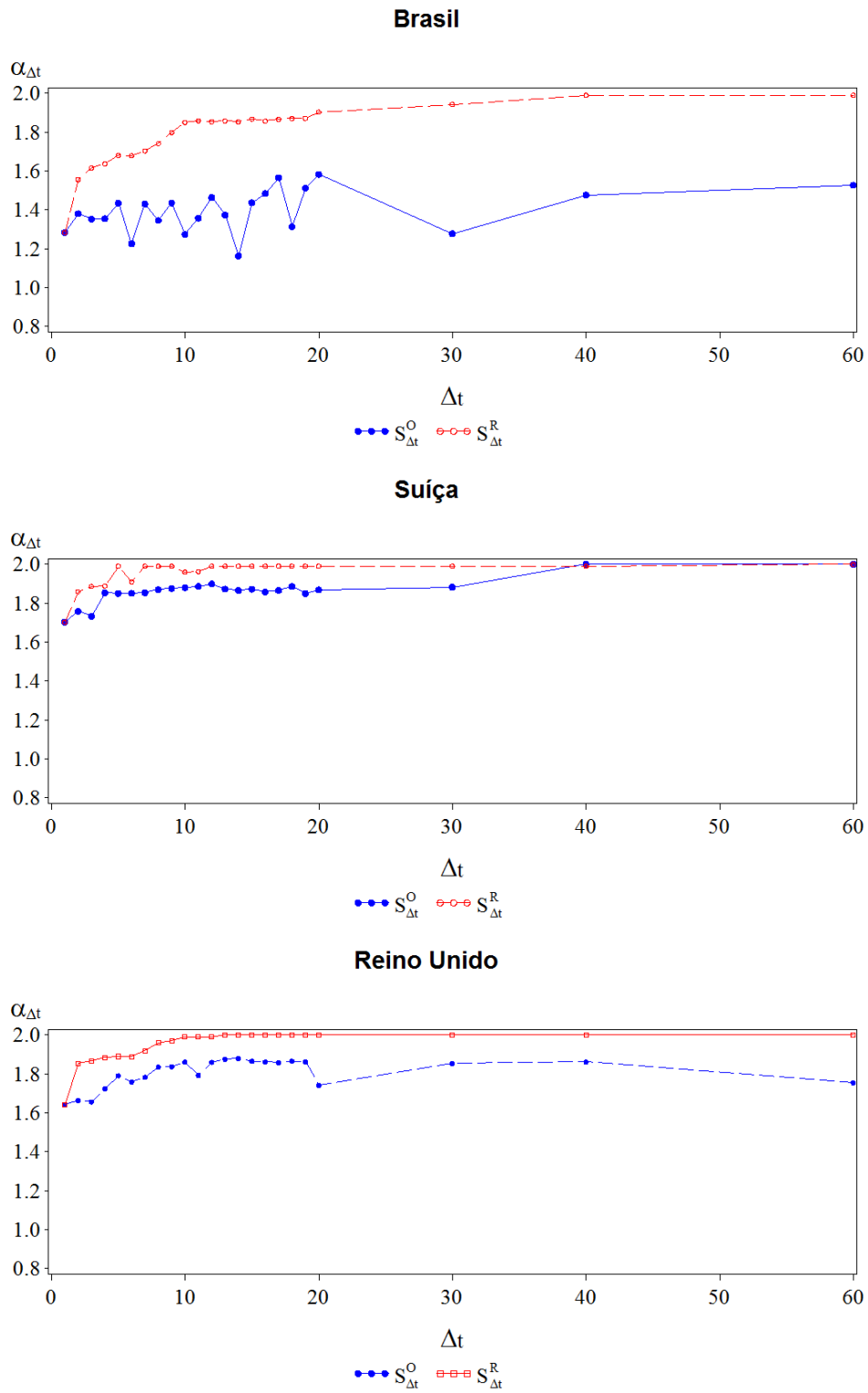


Figura 5.6: Real, franco suíço e a libra esterlina: estimativas de máxima verossimilhança de $\alpha_{\Delta t}$ para os processos $S_{\Delta t}^R$ e $S_{\Delta t}^O$.

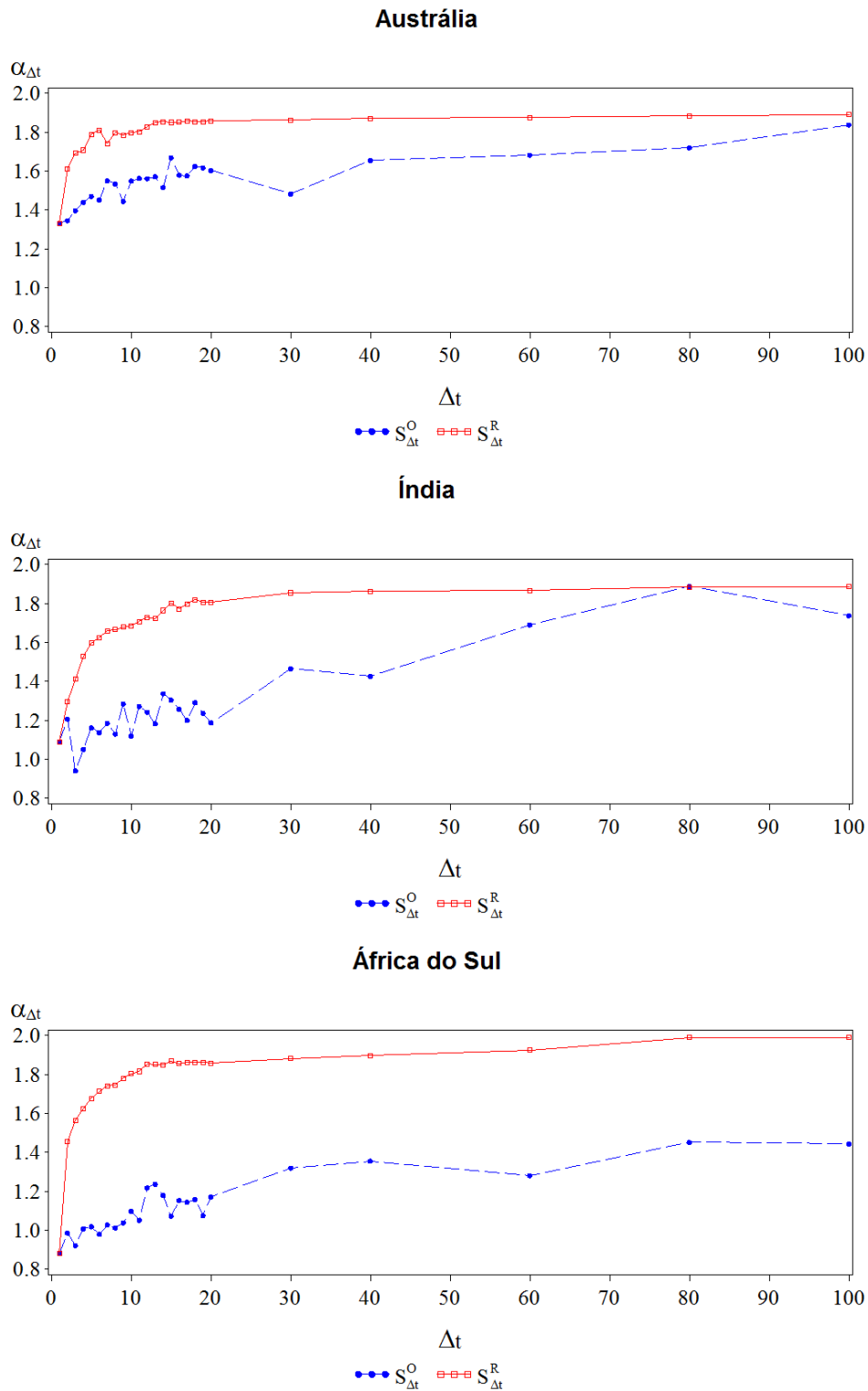


Figura 5.7: Dólar australiano, rúpia (Índia) e rand (África do Sul): estimativas de máxima verossimilhança de $\alpha_{\Delta t}$ para os processos $S_{\Delta t}^R$ e $S_{\Delta t}^O$.

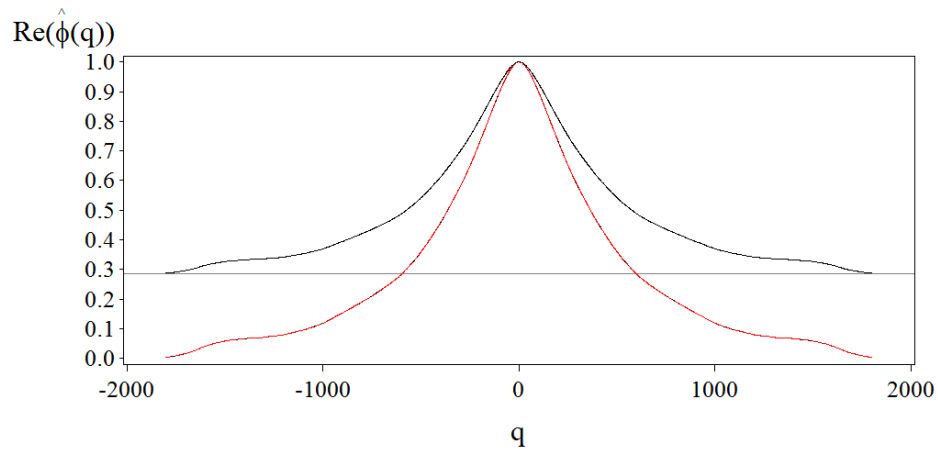


Figura 5.8: Rúpia: parte real da FCE da série temporal original ($\hat{\phi}_2(q) > 0.3$) e a da série sem os retornos nulos.

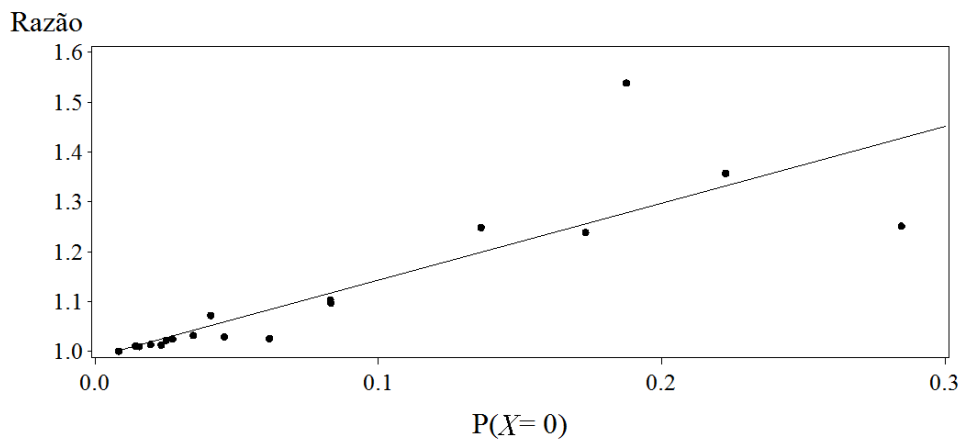


Figura 5.9: Comportamento da razão $\frac{\alpha^*}{\alpha}$ em função da incidência de zeros, em que α^* representa a estimativa obtida desconsiderando-se os retornos nulos (Tab. 5.4) e α representa a estimativa obtida com base nos dados originais (Tab. 5.2).

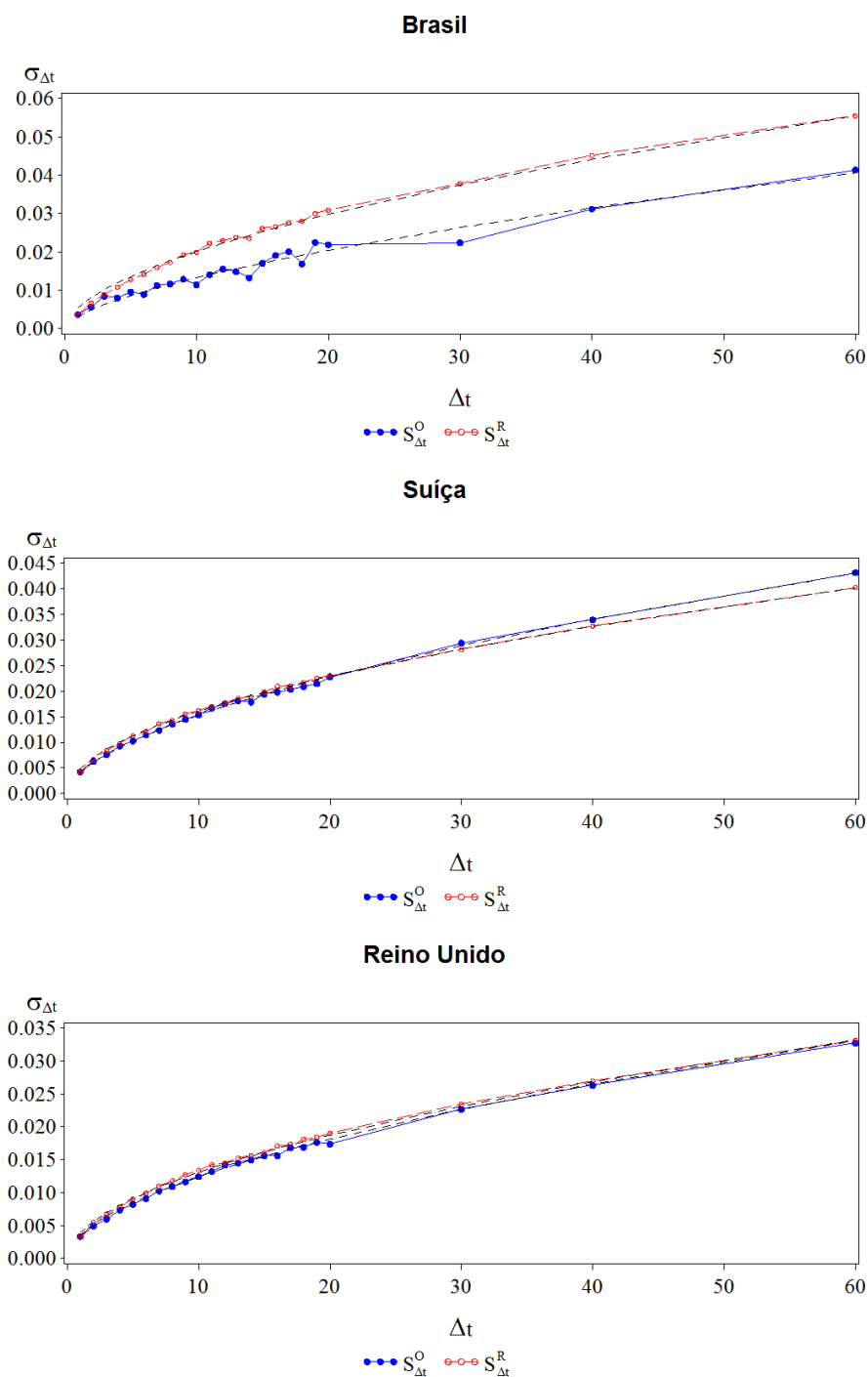


Figura 5.10: Estimativas de máxima verossimilhança de $\sigma_{\Delta t}$ para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt . As linhas tracejadas representam os ajustes na forma $\sigma_{\Delta t} = (\gamma \Delta t)^H$ (Eq. (5.27)). Para o processo aleatorizado, $H = 0,57; 0,51$ e $0,52$, respectivamente, para o real, o franco suíço e a libra esterlina. Para o processo não aleatorizado, $H = 0,63; 0,58$ e $0,55$, respectivamente para essas mesmas moedas.

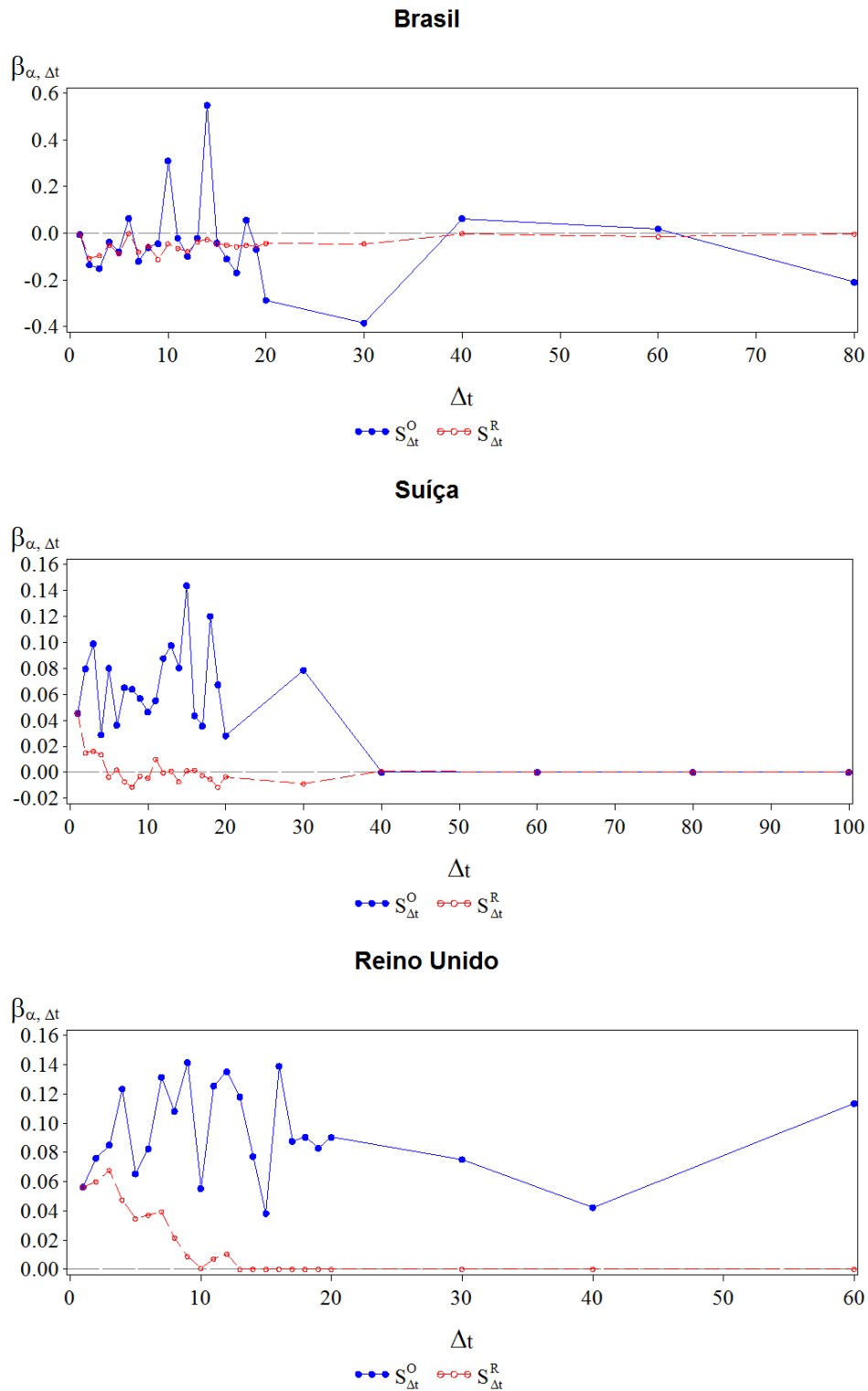


Figura 5.11: Estimativas de máxima verossimilhança para a assimetria efetiva $\beta_{\alpha, \Delta t} = \beta_{\Delta t} \tan \frac{\pi \alpha \Delta t}{2}$ para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt .

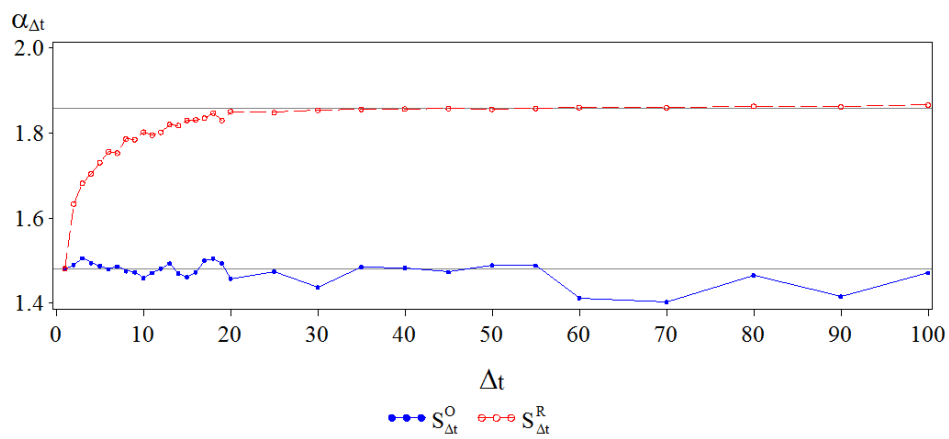


Figura 5.12: DJIA: estimativas de máxima verossimilhança de $\alpha_{\Delta t}$ para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt .

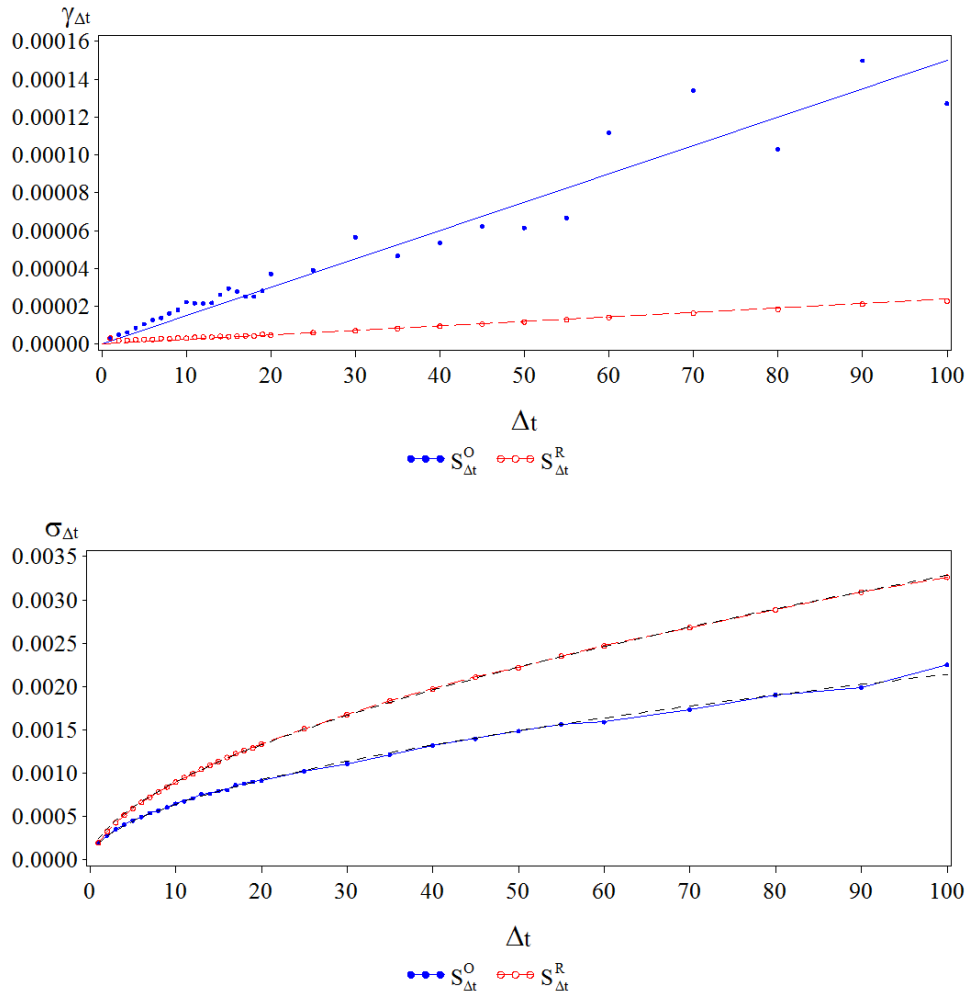


Figura 5.13: DJIA: estimativas de máxima verossimilhança de $\gamma_{\Delta t}$ (painel superior) e $\sigma_{\Delta t} = \gamma_{\Delta t}^{\frac{1}{\alpha}}$ (painel inferior) para as variáveis $S_{\Delta t}^R$ e $S_{\Delta t}^O$ segundo Δt . As curvas tracejadas representam os ajustes $\sigma_{\Delta t} \approx (0,0004\Delta t)^{0,54}$ (para $S_{\Delta t}^R$) e $\sigma_{\Delta t} \approx (0,0012\Delta t)^{0,52}$ (para $S_{\Delta t}^O$).

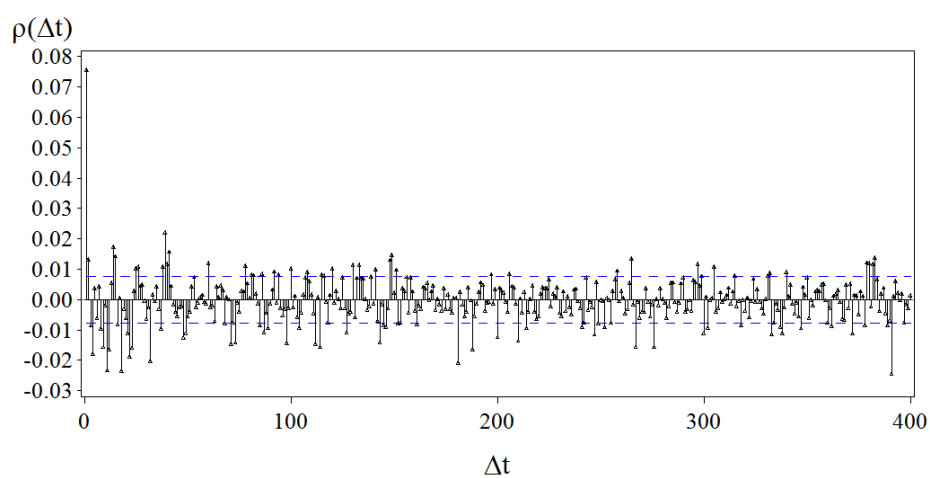


Figura 5.14: DJIA: gráfico da função de autocorrelação amostral dos retornos do índice DJIA, em que $\rho(\Delta t) = \text{Corr}(X_t, X_{t+\Delta t})$, $\Delta t \geq 1$. As autocorrelações são persistentes e há alternância de sinais positivos e negativos, o que indica que há dependência de longo alcance entre os retornos do índice DJIA.

Capítulo 6

Teste de independência

6.1 Introdução

Com base na sequência de estimativas de máxima verossimilhança dos índices de estabilidade $\{\alpha_{\Delta t}\}_{\Delta t \geq 1}$, o efeito da dependência serial na dinâmica da convergência do passeio aleatório $S_{\Delta t}$ [33, 35] foi discutido no capítulo anterior. Assim, é importante que se disponha de um bom instrumento para a estimação da dependência linear e não linear entre duas variáveis aleatórias (VA).

No caso linear, o coeficiente de correlação de Pearson (ρ) é uma medida natural da dependência entre duas VA X e Y conjuntamente gaussianas. Nesse caso, a função de densidade $f_{(X,Y)}(x, y)$ é especificada por uma matriz de covariância na forma

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix},$$

em que

$$\rho = \frac{\langle XY \rangle - \langle X \rangle \langle Y \rangle}{\sigma_X \sigma_Y},$$

$|\rho| \leq 1$, e $\sigma_X^2 > 0$ e $\sigma_Y^2 > 0$ são as variâncias de X e Y , respectivamente. Se $\rho = 0$, tem-se $f_{(X,Y)}(x, y) = f_X(x)f_Y(y)$, isto é, há independência entre X e Y . De modo análogo, se $\{X_t\}$ for um processo gaussiano estacionário, a função de autocorrelação $\rho(h)$ descreve a dependência serial entre X_t e $X_{t \pm h}$. O coeficiente ρ se relaciona com o conceito de regressão

linear. A média condicional $E[Y|X = x] = ax + b$ representa a reta de regressão de Y em x , em que $a = \rho\sigma_Y/\sigma_X$ é o coeficiente angular e b é o intercepto, e ρ^2 (coeficiente de determinação) representa a fração da variação total de Y explicada pela reta de regressão [89, 93].

No caso não linear, fora do ambiente gaussiano, $\rho = 0$ não necessariamente significa independência, pois é possível haver situações em que $\langle XY \rangle = \langle X \rangle \langle Y \rangle$, mesmo que haja dependência entre X e Y (Exemplo 6.2.1). Além disso, se pelo menos uma dessas variáveis for estável com $\alpha < 2$, então não se pode definir ρ por causa da inexistência do segundo momento.

Entre os poucos métodos estatísticos que possuem poder estatístico suficiente para detectar as diversas formas de dependência não linear [6] encontra-se o teste de HBKR (Hoeffding, Blum, Kiefer e Rosenblatt) [9, 54]. Esse teste é do tipo Cramér-von Mises [39] que considera um processo aleatório gerado pela diferença entre a distribuição conjunta empírica e o produto correspondente entre as marginais empíricas. Há testes mais recentes que envolvem outras caracterizações de processos aleatórios que são, essencialmente, equivalentes ao teste de HBKR para o caso bivariado [6, 7, 8, 40].

Considerando a necessidade de se avaliar a dependência entre duas VAs além do contexto gaussiano, este capítulo apresenta um novo teste estatístico assintótico com poder estatístico superior ao do teste de HBKR para a detecção de diferentes tipos de dependência não linear. Nosso procedimento também é do tipo Cramér-von Mises que se baseia na estatística χ^2 da razão de verossimilhança [89], o que equivale a minimizar a distância de Kullback-Leibler entre duas distribuições empíricas. O teste proposto não requer suposições acerca da distribuição dos dados (*distribution free*), sendo aplicável para situações com observações extremas.

O capítulo está organizado da seguinte forma. A Seção 6.2 trata da função característica multivariada e do conceito de independência. A Seção 6.3 apresenta o novo teste de independência, a estatística do teste e a função característica da sua distribuição amostral assintótica. Um experimento de Monte Carlo para a determinação dos valores críticos

assintóticos do teste é feito na Seção 6.4. A Seção 6.5 trata da validação e do poder do teste, considerando-se um exemplo de dependência não linear entre duas séries temporais não estacionárias de caudas pesadas. O poder do teste também é avaliado mediante dois outros exemplos, considerando-se dependência não linear entre séries temporais estacionárias [47] e entre duas séries temporais condicionalmente não estacionárias [6]. A aplicação do teste é ilustrada na Seção 6.6 com os dados de alta frequência da Bolsa de Valores de Nova Iorque (NYSE) e com as taxas de câmbio, comparando-se os resultados do nosso teste com aqueles produzidos pelo teste de HBKR; e a Seção 6.7 conclui este capítulo.

6.2 A FC multivariada e independência

Considere que $\mathbf{X} = (X_1, \dots, X_m)'$ seja um vetor aleatório m -dimensional com função de distribuição conjunta $F(\mathbf{x}; \boldsymbol{\theta})$, em que $\boldsymbol{\theta}$ representa o vetor de parâmetros e $\mathbf{x} = (x_1, \dots, x_m)' \in \mathbb{R}^m$. De modo análogo ao caso unidimensional, a FC de um vetor aleatório \mathbf{X} é definida como

$$\phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta}) = \langle e^{i\mathbf{q} \cdot \mathbf{X}} \rangle \quad (6.1)$$

$$= \int_{\mathbb{R}^m} e^{i\mathbf{q} \cdot \mathbf{x}} dF(\mathbf{x}; \boldsymbol{\theta}) \quad (6.2)$$

$$= \langle \cos(\mathbf{q} \cdot \mathbf{X}) \rangle + i \langle \text{sen}(\mathbf{q} \cdot \mathbf{X}) \rangle, \quad (6.3)$$

em que $\mathbf{q} = (q_1, \dots, q_m) \in \mathbb{R}^m$. Os aspectos básicos da FC multivariada são análogos aos do caso univariado apresentados no Cap. 2.

Em particular, se os elementos do vetor aleatório \mathbf{X} forem independentes, então

$$\phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta}) = \prod_{j=1}^m \phi_{X_j}(q_j; \boldsymbol{\theta}), \quad (6.4)$$

para qualquer $q_j \in \mathbb{R}$. Se $q_1 = \dots = q_m = q$, tem-se a FC da soma $S_m = \sum_{j=1}^m X_j$, pois

$$\phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta}) = \left\langle e^{iq \sum_{j=1}^m X_j} \right\rangle \quad (6.5)$$

$$= \phi_X^m(q; \boldsymbol{\theta}) \quad (6.6)$$

$$= \phi_{S_m}(q; \boldsymbol{\theta}). \quad (6.7)$$

Porém, $\phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta}) = \phi_X^m(q; \boldsymbol{\theta})$ não implica em independência entre os elementos de X , como mostra o próximo exemplo.

Exemplo 6.2.1. Considere o vetor aleatório $\mathbf{X} = (X_1, X_2)'$ cuja FD conjunta segue a forma [47]

$$f_{\mathbf{X}}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2) + g(x_1)h(x_2) - g(x_2)h(x_1), \quad (6.8)$$

em que $f_{X_k}(x_k)$ é uma FD univariada ($x_k \in \mathbb{R}$, $k = 1, 2$), e $g(x_k) \neq h(x_k)$ são funções integráveis ímpares. Por exemplo, suponha que X_1 e X_2 seguem uma distribuição exponencial dupla dada por $f_{X_k}(x_k) = \frac{1}{2}e^{-|x_k|}$, $g(x_k) = \frac{x_k}{2}e^{-\tau|x_k|}$ e $h(x_k) = \frac{x_k}{2}e^{-\xi|x_k|}$, em que $\tau \neq \xi > 1$ são tais que $f_{\mathbf{X}}(x_1, x_2) \geq 0$. Neste caso,

$$\begin{aligned} \phi_{\mathbf{X}}(q_1, q_2; \tau, \xi) &= \phi_{X_1}(q_1; \tau)\phi_{X_2}(q_2; \xi) + \\ &+ 4\tau\xi q_1 q_2 \left\{ \frac{1}{(\xi^2 + q_1^2)^2(\tau^2 + q_2^2)^2} - \frac{1}{(\tau^2 + q_1^2)^2(\xi^2 + q_2^2)^2} \right\}, \end{aligned} \quad (6.9)$$

em que $\phi_{X_1}(q_1; \tau) = \frac{1}{1+q_1^2}$ e $\phi_{X_2}(q_2; \xi) = \frac{1}{1+q_2^2}$. Apesar de X_1 e X_2 serem variáveis dependentes, observa-se que $\langle e^{iq(X_1+X_2)} \rangle = \langle e^{iqX_1} \rangle \langle e^{iqX_2} \rangle$ se $q_1 = q_2 = q$ em (6.9). Embora a correlação linear seja nula, X_1 e X_2 possuem uma estrutura de dependência (não linear). Como ilustração, as curvas de nível da FC da distribuição do vetor aleatório \mathbf{X} para $\tau = 1, 2$ e $\xi = 1, 9$ são mostradas no painel superior da Figura 6.2.1. Em contraste, as curvas da FC sob a hipótese de independência entre X_1 e X_2 são mostradas no painel inferior. Caso $q_1 = q_2 = q$, tem-se $\phi_{X_1+X_2}(q) = \phi_{X_1}(q)\phi_{X_2}(q)$.

□

Assim, considerando que $\langle X_1 X_2 \rangle = \langle X_1 \rangle \langle X_2 \rangle$ não necessariamente indica independência, e que os momentos da distribuição podem não existir, a próxima seção introduz um novo teste de independência para o caso bivariado ($m = 2$).

6.3 O teste de independência

Considere uma amostra aleatória $\mathbf{X}_1, \dots, \mathbf{X}_n$, em que $\mathbf{X}_i = (X_{i1}, X_{i2})'$ é um vetor bidimensional cujos elementos são VA absolutamente contínuas com função de distribuição

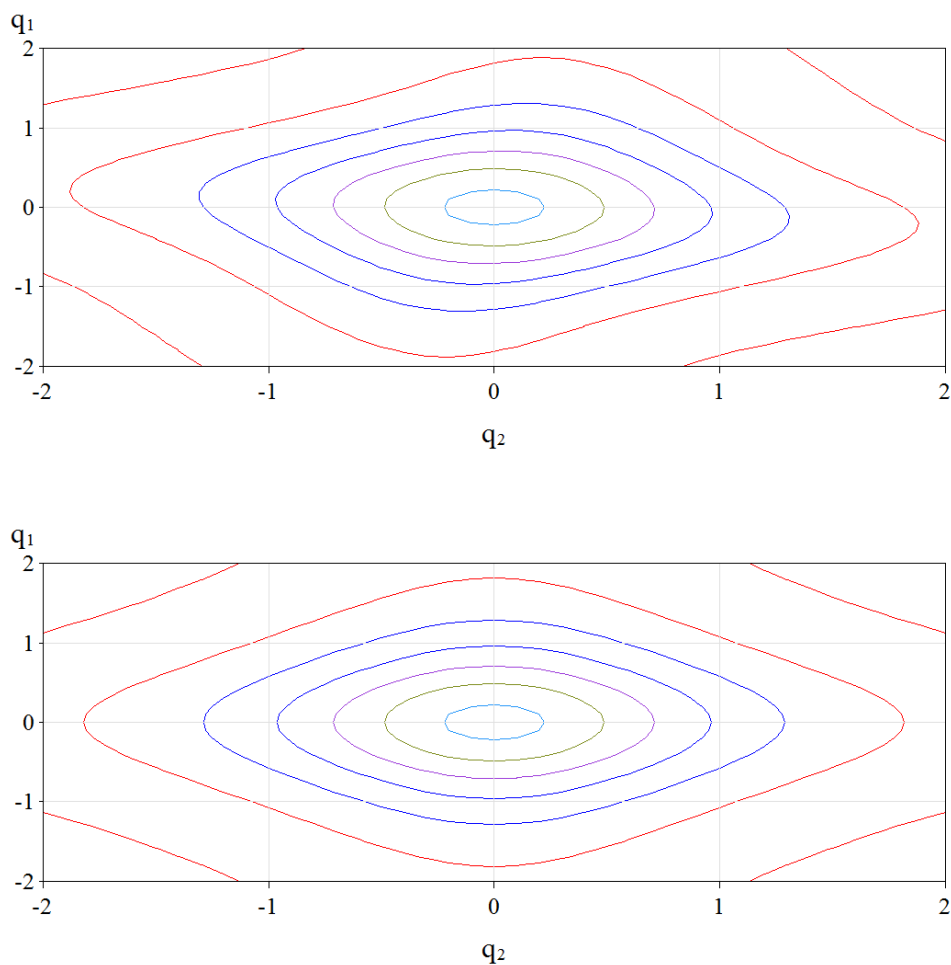


Figura 6.1: curvas de nível da FC da distribuição bivariada \mathbf{X} com $\tau = 1, 2$ e $\xi = 1, 9$ — painel superior — e as da FC sob hipótese de independência entre X_1 e X_2 — painel inferior —

acumulada (FDA) $F(\mathbf{x})$, em que $\mathbf{x} = (x_1, x_2)' \in \mathbb{R}^2$. O propósito é testar a hipótese nula de independência bivariada

$$F(\mathbf{x}) = F_1(x_1)F_2(x_2), \tag{6.10}$$

para todo x_j , em que $F_j(x_j)$ é a FDA marginal da VA X_j .

Para se testar a hipótese nula (6.10) pelo critério de Cramér-von Mises, calcula-se a distribuição conjunta empírica para cada ponto \mathbf{x} (Tab. 6.1), em que as realizações

(contagens) observadas para um dado vetor particular \mathbf{x} são

$$n_{k_1 k_2} = n_{k_1 k_2}(\mathbf{x}) = \sum_{i=1}^n \prod_{j=1}^2 \{1 - I(x_j - X_{ij})\}^{1-k_j} I^{k_j}(x_j - X_{ij}), \quad (6.11)$$

$$n_{k_1 \bullet} = n_{k_1 \bullet}(\mathbf{x}) = \sum_{i=1}^n \{1 - I(x_1 - X_{i1})\}^{1-k_1} I^{k_1}(x_1 - X_{i1}), \quad (6.12)$$

$$n_{\bullet k_2} = n_{\bullet k_2}(\mathbf{x}) = \sum_{i=1}^n \{1 - I(x_2 - X_{i2})\}^{1-k_2} I^{k_2}(x_2 - X_{i2}), \quad (6.13)$$

com $k_j = 0, 1$ e $I(y) = 1$, se $y \geq 0$; e $I(y) = 0$, se caso contrário.

Depois, considera-se a distribuição empírica sob a hipótese de independência (Tab. 6.2), em que

$$m_{k_1 k_2} = m_{k_1 k_2}(\mathbf{x}) = \frac{n_{k_1 \bullet}(\mathbf{x}) n_{\bullet k_2}(\mathbf{x})}{n} \quad (6.14)$$

são as realizações (contagens) esperadas para um dado ponto \mathbf{x} sob a hipótese (6.10).

Tabela 6.1: Distribuição bivariada empírica para um dado ponto $\mathbf{x} = (x_1, x_2)'$

eventos	$X_2 \leq x_2$	$X_2 > x_2$	total
$X_1 \leq x_1$	n_{11}	n_{10}	$n_{1\bullet}$
$X_1 > x_1$	n_{01}	n_{00}	$n_{0\bullet}$
total	$n_{\bullet 1}$	$n_{\bullet 0}$	n

Tabela 6.2: Distribuição esperada no ponto \mathbf{x} sob a hipótese de independência

eventos	$X_2 \leq x_2$	$X_2 > x_2$	total
$X_1 \leq x_1$	m_{11}	m_{10}	$n_{1\bullet}$
$X_1 > x_1$	m_{01}	m_{00}	$n_{0\bullet}$
total	$n_{\bullet 1}$	$n_{\bullet 0}$	n

Há diferentes maneiras para se medir a distância entre as Tabelas 6.1 e 6.2. Por exemplo, a estatística do teste de HBKR é dada por

$$B_{HBKR} = \int_{\mathbb{R}^2} Q^2(\mathbf{x}) d\hat{F}(\mathbf{x}), \quad (6.15)$$

em que $Q(\mathbf{x}) = (n_{11}(\mathbf{x}) - m_{11}(\mathbf{x}))/\sqrt{n}$ é um processo aleatório que mede a discrepância entre binomiais combinadas, e $\hat{F}(\mathbf{x}) = n_{11}/n$ denota a FDA empírica bivariada.

Como alternativa, sugerimos o processo aleatório

$$L^2(\mathbf{x}) = 2 \sum_{k_1=0}^1 \sum_{k_2=0}^1 n_{k_1 k_2} \ln \left\{ \frac{n_{k_1 k_2}}{m_{k_1 k_2}} \right\} \quad (6.16)$$

em lugar da distância $Q(\mathbf{x})$. A distância $L^2(\mathbf{x})$ é a estatística do teste da razão de verossimilhança generalizada para uma distribuição multinomial em uma tabela 2×2 . Essa estatística se relaciona com a distância de Kullback-Leibler entre as Tabelas 6.1 e 6.2. Por causa da equivalência assintótica com $L^2(\mathbf{x})$, também consideramos a estatística χ^2 de Pearson [1],

$$\chi^2(\mathbf{x}) = \sum_{k_1=0}^1 \sum_{k_2=0}^1 \frac{(n_{k_1 k_2} - m_{k_1 k_2})^2}{m_{k_1 k_2}}. \quad (6.17)$$

Enquanto a estatística $Q(x)$ considera apenas os eventos $[X_1 \leq x_1]$ e $[X_2 \leq x_2]$, as estatísticas $L^2(\mathbf{x})$ e $\chi^2(\mathbf{x})$ incluem os eventos complementares $[X_1 > x_1]$ e $[X_2 > x_2]$, além de possuírem propriedades ótimas relacionadas com o princípio da máxima verossimilhança [101].

Assintoticamente, a média [1] e a função de covariância de (6.17) são, respectivamente,

$$\lim_{n \rightarrow +\infty} E(\chi^2(\mathbf{x})) = 1 \quad (6.18)$$

e

$$\text{Cov}(\mathbf{s}_1, \mathbf{s}_2) = \frac{2(\min(u_1, u_2) - u_1 u_2)^2 (\min(v_1, v_2) - v_1 v_2)^2}{u_1 u_2 v_1 v_2 (1 - u_1)(1 - u_2)(1 - v_1)(1 - v_2)}, \quad (6.19)$$

em que $\mathbf{s}_j = (u_j, v_j) \in [0, 1] \times [0, 1]$, $\text{Cov}(\mathbf{s}_1, \mathbf{s}_2) = \lim_{n \rightarrow +\infty} E(\chi^2(\mathbf{x})\chi^2(\mathbf{y})) - E^2(\chi^2(\mathbf{x}))$. De modo semelhante ao do teste de HBKR [9, 59], o resultado (6.19) foi obtido com o auxílio de instrumentos computacionais.

Sob a hipótese (6.10), e considerando que (6.16) e (6.17) possuem a mesma distribuição limite, χ^2 com 1 grau de liberdade [1, 89], os resultados (6.18) e (6.19) também devem ser assintoticamente válidos para L^2 .

Desse modo, as estatísticas de Cramér-von Mises propostas são

$$B_{LR} = \int_{\mathbb{R}^2} L^2(\mathbf{x}) d\hat{F}(\mathbf{x}) \quad (6.20)$$

e

$$B_{\chi^2} = \int_{\mathbb{R}^2} \chi^2(\mathbf{x}) d\hat{F}(\mathbf{x}). \quad (6.21)$$

Como a estatística em (6.21) não depende da forma da FDA F , sem perda de generalidade, considera-se $\mathbf{s}_j = (u_j, v_j)$ como um ponto uniformemente distribuído no quadrado unitário $[0, 1] \times [0, 1]$. Assim, com base em (6.18), (6.19) e (6.21), conclui-se que

$$\mu_B = \langle B_{\chi^2} \rangle = \langle B_{LR} \rangle = 1 \quad (6.22)$$

e

$$\sigma_B^2 = \text{Var}(B_{\chi^2}) = \text{Var}(B_{LR}) = \frac{2(\pi^2 - 9)^2}{9} \approx 0.16805. \quad (6.23)$$

Para se obter a função característica da estatística no caso limite, $\phi_B(q) = \langle e^{iqB} \rangle$, em que $q \in \mathbb{R}$ e B representa a distribuição limite de B_{LR} e B_{χ^2} , consideramos a equação integral [9, 59]

$$\int_{[0,1] \times [0,1]} \text{Cov}(\mathbf{s}_1, \mathbf{s}_2) \gamma(\mathbf{s}_2) d\mathbf{s}_2 = \lambda \gamma(\mathbf{s}_1). \quad (6.24)$$

Com a determinação dos autovalores λ_j de (6.24), tem-se que $\phi_B(q) = \prod_{j,k} (1 - 2iq\lambda_j\lambda_k)^{-\frac{1}{2}}$, em que $\sum_{j,k} \lambda_j\lambda_k = 1$ para satisfazer (6.22) e (6.23). Por causa da fatorabilidade de $\text{Cov}(\mathbf{s}_1, \mathbf{s}_2)$, temos que $\text{Cov}(u_1, u_2) = \{\sqrt{2}(\min(u_1, u_2) - u_1u_2)^2\} / \{u_1v_1(1 - u_1)(1 - v_1)\}$, e assim, Eq. (6.24) torna-se equivalente ao problema de autovalores na forma diferencial

$$u(1 - u)\gamma''(u) - 2u\gamma'(u) + \frac{2\sqrt{2}}{\lambda}\gamma(u) = 0. \quad (6.25)$$

Considerando-se a Eq. (6.25) e a condição $\sum_{j,k} \lambda_j\lambda_k = 1$, os autovalores obtidos pelo método de Frobenius são $\lambda_j = 1/\{j(j + 1)\}$, em que $j = 1, 2, \dots$. Logo,

$$\phi_B(q) = \prod_{j,k} \left\{ 1 - \frac{2iq}{j(j + 1)k(k + 1)} \right\}^{-\frac{1}{2}}. \quad (6.26)$$

Os detalhamentos do desenvolvimento anterior se encontram no Apêndice F.

Com base na distribuição amostral caracterizada pela Eq. (6.26), a estatística (6.20) pode ser considerada para a detecção de dependência não linear. A seguir, a distribuição amostral e alguns pontos críticos serão obtidos computacionalmente.

6.4 Valores críticos assintóticos

Com respeito à distribuição amostral assintótica, a Eq. (6.26) sugere que a estatística B é uma soma de VA independentes gama com parâmetros de forma iguais a $1/2$, mas com diferentes parâmetros escala iguais a $2/\{j(j+1)k(k+1)\}$, $j, k \geq 1$. Para efetuar a soma numericamente, consideramos uma soma truncada em que $1 \leq j, k \leq 200$. Em seguida, foram geradas 100.000 replicações dessas somas truncadas pelo método de Monte Carlo. Como $\sum_{j=1}^{200} \sum_{k=1}^{200} \lambda_j \lambda_k = 40000/40401 \approx 0.99$, há um vício decorrente do truncamento. Esse vício foi corrigido adicionando-se o valor $401/40401$ para cada realização B . A Tab. 6.3 mostra que a média e a variância das realizações simuladas se aproximam dos valores teóricos assintóticos correspondentes (6.22) e (6.23).

Tabela 6.3: Estatística B simulada: média e variância empíricas, e seus valores teóricos correspondentes

distribuição	média	variância
empírica	0,998	0,1671
teórica	1,000	0,1680

Com base na distribuição empírica, a hipótese nula (6.10) é rejeitada se $B > b$, em que b é o valor crítico relativo ao nível de significância desejado α . Considerando-se $\alpha = 0, 1\%, 1\%, 2\%, 2, 5\%, 5\%$, e 10% , os respectivos valores críticos b após a correção do vício são apresentados na Tab. 6.4.

Tabela 6.4: Estatística B : níveis de significância ns e os valores críticos correspondentes b para se testar a hipótese de independência

$ns(\%)$	0,1	1,0	2,0	2,5	5,0	10,0
b	3,527	2,468	2,164	2,072	1,778	1,507

6.5 Validação e o poder do teste

Outro experimento de Monte Carlo foi realizado para a validação do teste e também para comparar seu poder estatístico com os do teste de correlação de Pearson ρ e do teste de HBKR. Para a validação foram geradas replicações de amostras aleatórias sob a hipótese nula (6.10). Em seguida, os percentuais empíricos de resultados significativos a $ns = 1\%, 2, 5\%$ e 5% foram comparados com os respectivos valores esperados (Tab. 6.4). O poder também foi estimado com base nos percentuais de resultados significativos com $ns = 1\%, 2, 5\%$ and 5% a partir de amostras geradas sob diferentes hipóteses alternativas, mediante os exemplos que se seguem.

No Exemplo 6.5.1, consideramos um caso de dependência não linear entre duas distribuições de caudas pesadas. Nesse exemplo, os parâmetros de escala das variáveis envolvidas são mutuamente dependentes, de modo que a série temporal resultante não é estacionária. O Exemplo 6.5.2 contempla o caso de dependência não linear entre duas séries estacionárias, e finalmente o Exemplo 6.5.3 considera o caso de dependência não linear entre duas séries condicionalmente não estacionárias.

Exemplo 6.5.1. Considere que $X|U = \gamma_1$ e $Y|V = \gamma_2$, condicionalmente, são distribuições de Cauchy cujas FD são

$$f_{X|U=\gamma_1}(x) = \frac{\gamma_1}{\pi(\gamma_1^2 + x^2)} \quad (6.27)$$

e

$$f_{Y|V=\gamma_2}(y) = \frac{\gamma_2}{\pi(\gamma_2^2 + y^2)}, \quad (6.28)$$

em que γ_1 e γ_2 são, respectivamente, as realizações das variáveis aleatórias $U = 30Z_1^2$ e $V = (1/U)^\xi(30Z_2^2)^{1-\xi}$; Z_1 e Z_2 são distribuições normais padrão; e $0 \leq \xi \leq 1$. Se $\xi = 0$, as marginais não condicionais X e Y são independentes; caso contrário, possuem dependência não linear. Assim, 800 replicações $\{(X_i, Y_i)\}_{i=1, \dots, 1000}$ foram geradas. Embora B_{χ^2} e B_{LR} sejam assintoticamente equivalentes, os resultados para B_{χ^2} foram particularmente desinteressantes, pois a estatística χ^2 não é apropriada para baixas contagens esperadas [1]. O processo aleatório $\chi^2(\mathbf{x})$ se relaciona com o teste de HBKR como $\chi^2(\mathbf{x}) = w(\mathbf{x})Q^2(\mathbf{x})$, em que $w^{-1}(\mathbf{x}) = (1 - \hat{F}_1(x))\hat{F}_1(x)(1 - \hat{F}_2(y))\hat{F}_2(y)$ é uma função peso proporcional às variâncias estimadas das contagens $n\hat{F}_1(x)$ e $n\hat{F}_2(y)$. Logo, pequenos valores $Q^2(\mathbf{x})$ podem resultar em grandes valores $\chi^2(\mathbf{x})$ dependendo do peso $w(\mathbf{x})$.

Entretanto, resultados interessantes foram produzidos com base na estatística B_{LR} . A Tab. 6.5 apresenta a frequência percentual de rejeição da hipótese de independência, com base nos valores críticos fornecidos pela Tab. 6.4. Para o caso de independência ($\xi = 0$), há concordância entre os níveis de significância empíricos e teóricos, exceto o teste ρ , que se mostrou excessivamente conservador. Com $\xi = 1$, a estrutura não linear foi totalmente detectada tanto pelo teste de HBKR como pelo teste sugerido, enquanto ρ falhou completamente.

Tabela 6.5: Validação ($\xi = 0$) e poder empírico, %

	$\xi = 1$			$\xi = 0.3$			$\xi = 0$		
ns	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}
5,0	0,0	100	100	0,8	34,0	85,1	1,9	4,4	4,6
2,5	0,0	100	100	0,5	13,8	65,1	1,8	2,7	2,6
1,0	0,0	100	100	0,3	4,0	31,8	0,6	0,9	1,2

Mesmo que X e Y sejam geradas independentemente, com $\xi = 1$ ambas estão relacionadas pelo parâmetro de escala e, como resultado, uma estrutura não linear surge inevitavelmente [35]. Como um caso intermediário, com $\xi = 0.3$, nosso teste apresentou poder superior ao do teste de HBKR.

□

Exemplo 6.5.2. Considere novamente a densidade apresentada no Ex. 6.2.1 [47]

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) + g(x)h(y) - g(y)h(x), \quad (6.29)$$

em que $f_X(x)$ e $f_Y(y)$ são FD univariadas, e $g(x) \neq h(x)$ são funções integráveis ímpares. Por exemplo, considere a distribuição exponencial dupla $f_X(x) = \frac{1}{2}e^{-|x|}$, $f_Y(y) = \frac{1}{2}e^{-|y|}$, e $g(x) = \frac{x}{2}e^{-\tau|x|}$, $h(x) = \frac{x}{2}e^{-\xi|x|}$, em que $\tau \neq \xi > 1$ deve ser tal que $f_{X,Y}(x, y) \geq 0$. Obviamente, X e Y são dependentes, apesar de $E(e^{iq(X+Y)}) = E(e^{iqX})E(e^{iqY})$, para todo $q \in \mathbb{R}$. Para efetuarmos a simulação de Monte Carlo, consideramos a FD condicional

$$F_{X|Y=y}(x) = \begin{cases} 1 - \frac{1}{2}e^{-|x|} + J(x, y), & \text{se } x > 0; \\ \frac{1}{2}e^{-|x|} + J(x, y), & \text{caso contrário,} \end{cases} \quad (6.30)$$

em que

$$J(x, y) = \frac{y}{2} \left\{ \frac{e^{-(\xi-1)|y|}}{\tau^2} \left(\Gamma(|x|, 2, 1/\tau) - 1 \right) - \frac{e^{-(\tau-1)|y|}}{\xi^2} \left(\Gamma(|x|, 2, 1/\xi) - 1 \right) \right\},$$

e $\Gamma(|x|, 2, 1/\tau) = \int_0^{|x|} qe^{-\tau q} dq$ é a função gamma incompleta inferior. A simulação de Monte Carlo foi efetuada seguindo-se os passos: (1) uma realização y é gerada, e (2) dado o valor y gerado no passo 1, uma realização x é gerada com base na Eq. (6.30) e no método da transformação inversa acoplado com um algoritmo de Newton-Raphson para determinação de raízes.

Sob a hipótese alternativa na Eq. (6.29), 1.000 replicações de amostras aleatórias $\{(X_i, Y_i)\}_{i=1, \dots, 1000}$ foram geradas para alguns valores adequados de τ e ξ . Os resultados são mostrados na Tab. 6.6. Observa-se que B_{LR} se mostrou mais poderoso que ρ e B_{HBKR} .

□

Exemplo 6.5.3. Considere agora o caso do anel bivariado [6], definindo-se $X|U = \theta \sim N(\cos \theta, \sigma^2)$ e $Y|U = \theta \sim N(\sin \theta, \sigma^2)$, em que $U \sim \text{Uniform}(0, 2\pi)$. Novamente, 1.000 replicações de amostras aleatórias $\{(X_i, Y_i)\}_{i=1, \dots, 1000}$ foram geradas usando-se $\sigma = 0, 1, 0, 5$ e $0,8$. Os resultados são mostrados na Tab. 6.7. Observa-se que o teste proposto é mais poderoso do que o teste de HBKR no caso $\sigma = 0, 5$.

Tabela 6.6: Poder empírico no Exemplo 6.5.2, %

	$\tau = 1,9$ e $\xi = 1,2$			$\tau = 1,6$ e $\xi = 1,2$			$\tau = 3,0$ e $\xi = 1,1$		
ns	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}
5,0	4,6	54,6	93,2	5,7	44,7	93,9	5,3	41,6	69,7
2,5	1,8	23,1	77,3	2,7	19,7	76,7	2,8	17,9	41,9
1,0	0,6	10,2	58,3	1,7	8,8	58,3	1,1	8,0	25,3

□

Tabela 6.7: Poder empírico no Exemplo 6.5.3, %

	$\sigma = 0,1$			$\sigma = 0,5$			$\sigma = 0,8$		
ns	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}	ρ	B_{HBKR}	B_{LR}
5,0	0,8	100	100	3,0	25,4	74,9	5,4	5,7	7,0
2,5	0,2	100	100	1,2	7,2	46,4	2,3	2,9	3,4
1,0	0,0	100	100	0,3	2,9	22,9	1,1	1,4	1,6

6.6 Ilustrações

6.6.1 Ações da bolsa de Nova Iorque

Como ilustração, consideramos as séries temporais *intraday* de preços de algumas ações negociadas na NYSE descritas no Cap. 1. Seja X_t o retorno de uma ação no instante t . Neste exemplo, o interesse é avaliar a dependência contemporânea entre os retornos do par de ações, $(X_t, Y_t)'$. Para se testar a hipótese de independência na Eq. (6.10), nós primeiros consideramos a estatística ρ . Como se espera, as correlações lineares entre essas séries são estatisticamente significativas [91, 62, 65], variando de 0,11 a 0,63 com *P-values* inferiores a 0,0001 (Tab. 6.8). Assim, para separar a dependência linear da não linear,

consideramos a matriz de correlação

$$\Omega = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad (6.31)$$

e a transformação linear

$$(X_t^*, Y_t^*)' = \Omega^{-\frac{1}{2}} \left(\frac{X_t - \langle X_t \rangle}{\sigma_X}, \frac{Y_t - \langle Y_t \rangle}{\sigma_Y} \right)'. \quad (6.32)$$

Os retornos transformados X_t^* e Y_t^* não são correlacionados [76]. Considerando $(X_t^*, Y_t^*)'$ em nosso teste, encontramos grandes valores de B com pequenos P -values ($<.0001$), o que nos força a rejeitar a hipótese de independência (Tab. 6.9). Assim, concluímos que a dependência não linear é um aspecto comum entre esses retornos. A Fig. 6.2 apresenta diagramas de dispersão entre alguns retornos transformados. Visualmente nenhum padrão é perceptível, embora os resultados na Tab. 6.9 sugere fortemente que X_t^* e Y_t^* são dependentes.

Tabela 6.8: Resultados para o teste ρ (P -values $<.0001$)

Stock	RLJ	ESGR	PFE	ABT	PGN	C	WPO	MU	CSCO	CORE
JAZZ	0,23	0,19	0,37	0,41	0,36	0,44	0,13	0,35	0,39	0,23
RLJ		0,11	0,22	0,23	0,22	0,26	0,13	0,21	0,21	0,20
ESGR			0,14	0,19	0,15	0,22	0,15	0,16	0,13	0,29
PFE				0,63	0,50	0,51	0,17	0,40	0,47	0,19
ABT					0,58	0,55	0,19	0,41	0,49	0,22
PGN						0,49	0,13	0,35	0,44	0,22
C							0,16	0,45	0,52	0,25
WPO								0,16	0,19	0,15
MU									0,43	0,21
CSCO										0,20

6.6.2 Taxas de câmbio

Com respeito às taxas de câmbio, consideramos as autocorrelações de primeira ordem, ou seja, $\rho = \text{Corr}(X_t, X_{t-1})$ [20]. Os resultados se encontram na Tab. 6.10. Para algumas moedas, como o Dólar australiano e o Real, apesar de as autocorrelações de primeira ordem não serem estatisticamente significativas (com P -values superiores a 1%), o teste proposto e o de HBKR indicam associações significativas entre X_t e X_{t-1}), havendo, portanto,

Tabela 6.9: Resultados do teste de independência considerando-se as variáveis transformadas (P -values $< 0,0001$)

Stock	RLJ	ESGR	PFE	ABT	PGN	C	WPO	MU	CSCO	CORE
JAZZ	351,6	41,1	153,4	194,1	293,2	272,0	27,4	412,9	162,5	81,3
RLJ		54,8	428,1	454,9	527,7	495,0	147,2	513,2	409,6	129,6
ESGR			39,5	50,7	50,5	65,5	8,8	81,6	27,8	39,3
PFE				197,7	219,9	132,2	90,0	339,3	129,0	76,6
ABT					334,5	178,6	74,0	440,4	164,0	90,3
PGN						307,6	99,2	503,9	334,3	137,0
C							39,0	451,1	198,6	119,2
WPO								152,3	100,1	27,3
MU									386,6	170,7
CSCO										92,1

dependência não linear. Apenas o Euro não apresentou fortes indícios de dependência não linear (P -value = 4%). Para algumas moedas, a Fig. 6.3 mostra as estimativas das curvas de nível referentes às densidades bivariadas obtidas pelo método do Kernel. No Cap. 7, discutiremos acerca desses padrões, como perspectivas para estudos futuros.

Tabela 6.10: Resultados da aplicação do teste para dependência entre X_t e X_{t-1} (P -values aproximados entre parênteses).

país	moeda	B_{LR}	B_{HBKR}	ρ
África do Sul	Rand	90,767 (<0,01)	2,946 (<0,01)	-0,0291 (<0,01)
Austrália	Dólar australiano	42,712 (<0,01)	1,243 (<0,01)	-0,0103 (0,30)
Brasil	Real	26,990 (<0,01)	0,685 (<0,01)	0,0116 (0,44)
Canadá	Dólar canadense	29,393 (<0,01)	0,622 (<0,01)	0,0188 (0,06)
Coréia do Sul	Won	71,519 (<0,01)	1,838 (<0,01)	0,0910 (<0,01)
Dinamarca	Coroa dinamarquesa	8,828 (<0,01)	0,263 (<0,01)	-0,0302 (<0,01)
Índia	Rúpia	59,511 (<0,01)	1,901 (<0,01)	-0,1068 (<0,01)
Japão	Yen	17,726 (<0,01)	0,523 (<0,01)	0,0202 (0,04)
México	Peso	15,536 (<0,01)	0,303 (<0,01)	-0,1153 (<0,01)
Nova Zelândia	Dólar neozelandês	32,082 (<0,01)	0,822 (<0,01)	-0,0106 (0,28)
Noruega	Coroa norueguesa	14,611 (<0,01)	0,350 (<0,01)	-0,0120 (0,22)
Reino Unido	Libra Esterlina	16,290 (<0,01)	0,419 (<0,01)	0,0453 (<0,01)
Singapura	Dólar de Singapura	15,771 (<0,01)	0,355 (<0,01)	-0,0576 (<0,01)
Suécia	Coroa Sueca	16,265 (<0,01)	0,405 (<0,01)	-0,0161 (0,10)
Suíça	Franco Suíço	7,695 (<0,01)	0,200 (<0,01)	0,0141 (0,15)
Tailândia	Baht	64,648 (<0,01)	1,681 (<0,01)	-0,0776 (<0,01)
Taiwan	Dólar de Taiwan	48,469 (<0,01)	1,508 (<0,01)	-0,0279 (0,21)
Zona do Euro	Euro	2,1445 (0,04)	0,060 (0,04)	0,0132 (0,45)

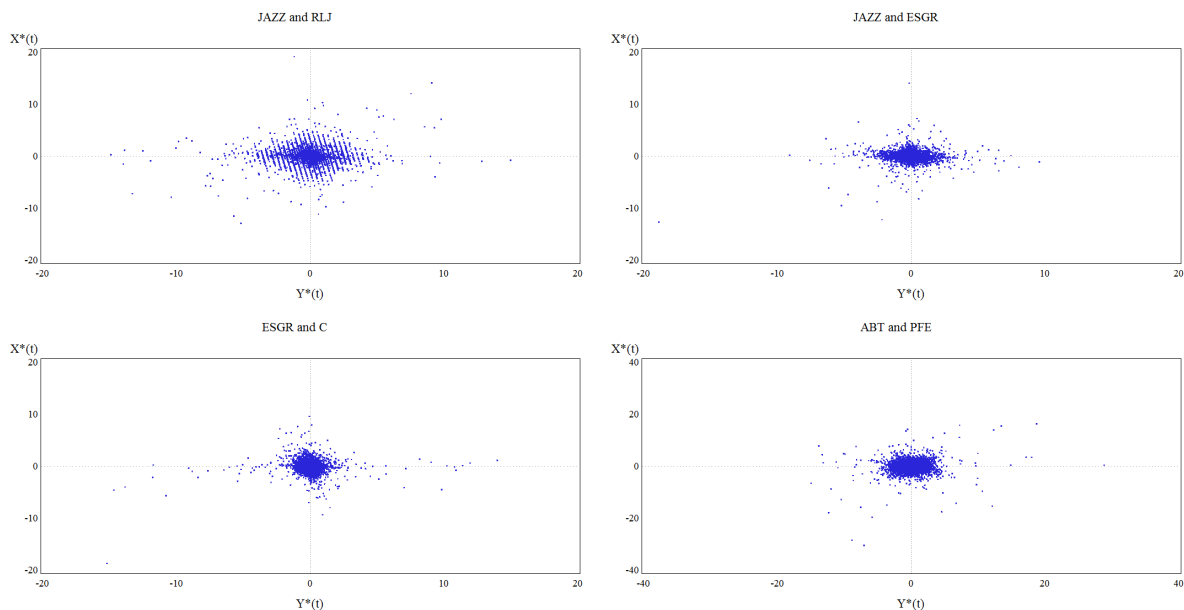


Figura 6.2: Diagramas de dispersão entre os retornos transformados de algumas ações. Embora nenhuma associação seja aparente, o que é esperado de retornos transformados, as variáveis X_t^* e Y_t^* são dependentes (Tab. 6.9).

6.7 Discussão

Este capítulo apresentou um novo teste assintótico para avaliação da independência entre duas VA contínuas. Nosso teste é do tipo Cramér-von-Mises cujo processo aleatório se baseia na estatística χ^2 da razão de verossimilhança. A equivalência assintótica entre $L^2(\mathbf{x})$ e $\chi^2(\mathbf{x})$ foi considerada para a obtenção dos momentos da estatística B do teste proposto. Desenvolvemos a forma assintótica da função característica teórica da estatística B e, a partir dela, determinamos computacionalmente os valores críticos aproximados mediante 100.000 replicações de Monte Carlo, considerando-se os níveis de significância de 0,1%, 1%, 2%, 2,5%, 5%, e 10%.

A validação e o desempenho do poder do teste foram avaliados também com base em experimento de Monte Carlo, considerando-se um exemplo de estrutura de dependência bivariada não linear entre distribuições com caudas pesadas, além de dois outros exemplos de dependência não linear para os casos de séries temporais estacionárias e condicional-

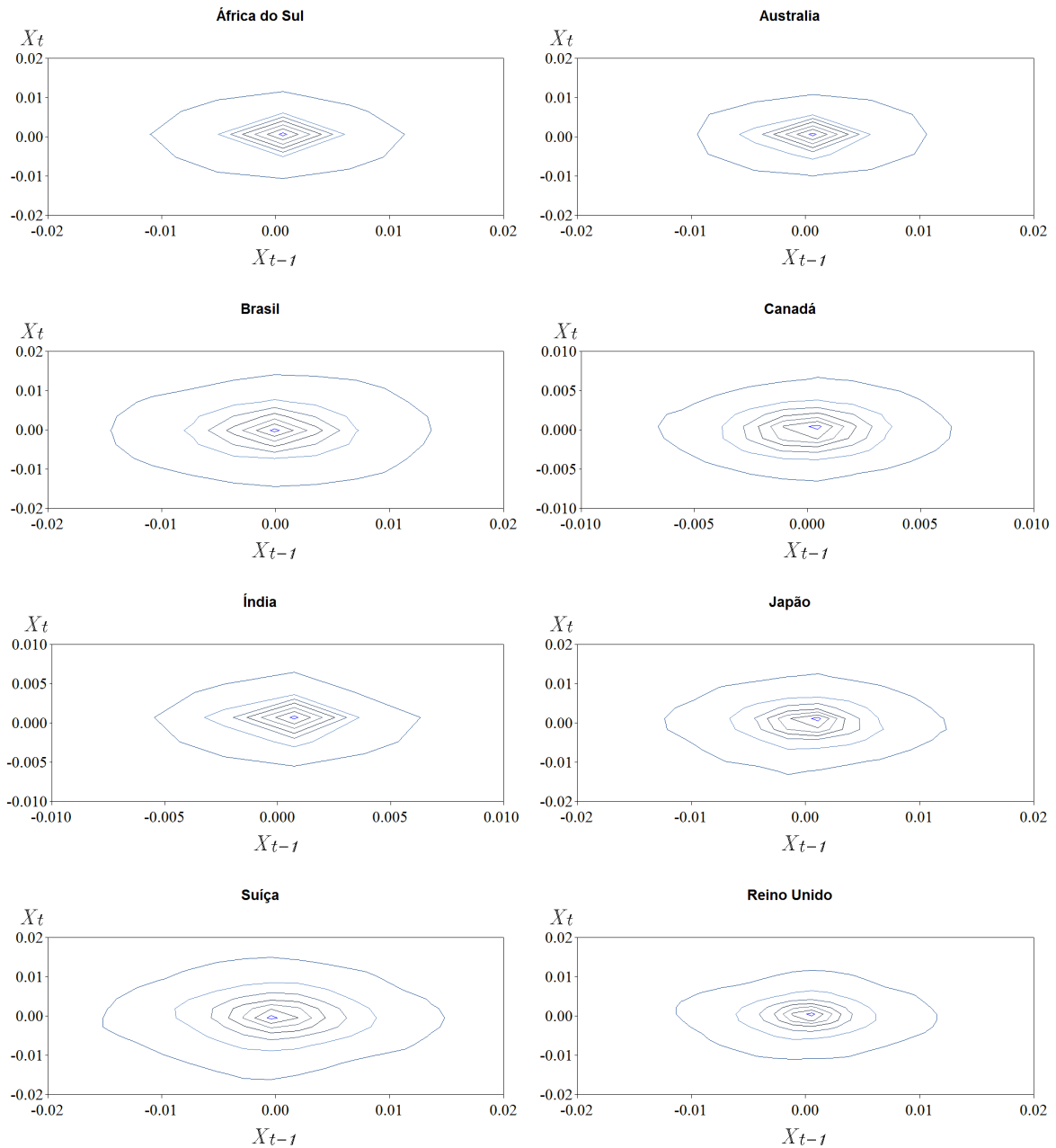


Figura 6.3: Diagramas de contorno das estimativas das densidades bivariadas obtidas pelo método do Kernel. As curvas de nível (de dentro para fora) representam os percentis 1, 5, 10, 50, 90, 95, 99 e 100%

mente não estacionárias.

Os resultados confirmam que nosso teste é consistente e poderoso na presença de estrutura de dependência não linear em um contexto bivariado não gaussiano. Para ilustrar, o teste foi aplicado em dados financeiros de alta frequência que exibiram extrema

volatilidade “*flash crashes*” e em taxas de câmbio, revelando que tanto as dependências linear como a não linear são aspectos comuns a esses dados [80].

Capítulo 7

Considerações finais

Neste trabalho apresentamos uma expansão da função característica (FC) que se mostrou útil para a estimação por máxima verossimilhança (MV) dos parâmetros de distribuições sob a hipótese de estabilidade. Essas estimativas permitem descrever a dinâmica da convergência de processos para a gaussiana. Os resultados sugerem que a lentidão dessa convergência se deve à dependência serial de curto e de longo alcance. Também observamos que a função característica empírica (FCE) sofre truncamento natural devido à finitude do tamanho amostral (n). Para um valor n fixo, o *scaling* da FCE apresenta uma quebra sempre no mesmo patamar $\phi_{0,n}$, independentemente da forma da distribuição dos dados. Finalmente, introduzimos um novo método assintótico que permite testar a hipótese de independência entre dois conjuntos de dados. Nosso teste é do tipo Cramér-von Mises, em que o processo empírico é obtido com base na divergência de Kullback-Leibler, e se mostrou estatisticamente poderoso para detectar dependência não linear fora do ambiente gaussiano [80].

A expansão proposta para a FC pode ser escrita como (Cap. 2)

$$\phi(q; \boldsymbol{\theta}) = \Phi(D_q; \boldsymbol{\theta})\delta(q),$$

em que $D_q = \frac{d}{dq}$ é o operador diferencial, $\delta(q)$ é a função delta de Dirac,

$$\Phi(D_q; \boldsymbol{\theta}) = \sum_{k=0}^{+\infty} c_k D_q^k,$$

e

$$c_k = \frac{2\pi f^{(k)}(-i)^k}{k!}.$$

Essa expansão é aplicável para qualquer distribuição cuja função de densidade possua representação na forma de série de Maclaurin, o que contempla, por exemplo, as distribuições estáveis. Essa forma não foi encontrada nas obras clássicas que tratam sobre as funções características [47, 66, 120]. O operador $\Phi(D_q; \boldsymbol{\theta})$ foi denominado *polinômio característico*, uma vez que seus coeficientes $\{c_k\}$ permitem caracterizar a forma da distribuição. Por exemplo, para as distribuições estáveis com $\alpha \neq 1$ (Cap. 3), temos

$$c_{2j+1} = i \frac{2\Gamma\left(\frac{2(j+1)}{\alpha}\right)}{\Gamma(2(j+1))\alpha\gamma_*^{\frac{2(j+1)}{\alpha}}} \cdot A_{1,j}$$

e

$$c_{2j} = \frac{2\Gamma\left(\frac{2j+1}{\alpha}\right)}{\Gamma(2j+1)\alpha\gamma_*^{\frac{2j+1}{\alpha}}} \cdot A_{2,j},$$

em que $\gamma_* = \gamma \cdot \sqrt{1 + \beta_\alpha^2}$, $\beta_\alpha = \beta \tan \frac{\pi\alpha}{2}$ representa a assimetria efetiva, e as componentes $A_{1,k}$ e $A_{2,k}$, que não dependem do parâmetro de escala γ , são dadas por

$$A_{1,j} = \text{sen} \left\{ \frac{2j+2}{\alpha} \arctan(\beta_\alpha) \right\},$$

e

$$A_{2,j} = \text{cos} \left\{ \frac{2j+1}{\alpha} \arctan(\beta_\alpha) \right\}.$$

Caso a distribuição seja simétrica, tem-se $A_{1,j} = 0$ e $A_{2,j} = 1$ para $\forall j \geq 0$. Mas para o caso assimétrico, essas funções não são constantes e um estudo acerca do comportamento dessas funções foi apresentado no Cap. 3. Embora a sequência de coeficientes $\{c_k\}$ possa divergir, dependendo do parâmetro de escala γ , é possível obter uma sequência convergente equivalente mediante transformação de escala.

No Capítulo 2 mostramos que a minimização da divergência de Kullback-Leibler entre duas distribuições X e Y , com respeito ao parâmetro $\theta_j \in \boldsymbol{\theta}$ da distribuição X , produz uma equação na forma

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \left\{ \phi_Y(q; \boldsymbol{\theta}') - \phi_X(q; \boldsymbol{\theta}) \right\} dq = 0,$$

em que

$$\omega_j(q; \boldsymbol{\theta}) = \{\overline{\Phi(D_q; \boldsymbol{\theta})}\}^{-1} \overline{h_j(q; \boldsymbol{\theta})}$$

e

$$h_j(q; \boldsymbol{\theta}) = \frac{d}{d\theta_j} \phi_X(q; \boldsymbol{\theta}).$$

Em particular, se Y e X representam, respectivamente, uma distribuição empírica e um modelo hipotético de distribuição de probabilidade, e considerando que $\hat{\phi}(q)$ representa a FC empírica obtida com base em uma amostra aleatória de tamanho n , tem-se que a equação

$$\int_{-\infty}^{+\infty} \omega_j(q; \boldsymbol{\theta}) \{ \hat{\phi}(q) - \phi_X(q; \boldsymbol{\theta}) \} dq = 0$$

compõe o sistema de equações de verossimilhança para grandes amostras ($n \rightarrow \infty$). A solução $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_p)'$ desse sistema de equações ($j = 1, \dots, p$) representa a estimativa de máxima verossimilhança do vetor de parâmetros $\boldsymbol{\theta} \in \mathbb{R}^p$.

Entretanto, na prática, n é finito e, conseqüentemente, $\hat{\phi}(q)$ é a FC de uma distribuição uniforme discreta, enquanto $\phi_X(q; \boldsymbol{\theta})$ é a de uma distribuição absolutamente contínua. Esse contraste gera um truncamento natural da FCE, conforme prevê o teorema de Glivenko-Cantelli (Cap. 4). Com base na distribuição amostral da FCE e na teoria de testes de hipóteses [101], estabelecemos um critério para o truncamento da FC na forma

$$\hat{\phi}_*(q) = |\hat{\phi}(q)| I(|\hat{\phi}(q)| \geq \phi_{0,n}),$$

em que $I(|\hat{\phi}(q)| \geq \phi_{0,n}) = 1$ se $|\hat{\phi}(q)| \geq \phi_{0,n}$, e $I(|\hat{\phi}(q)| \geq \phi_{0,n}) = 0$ se $|\hat{\phi}(q)| < \phi_{0,n}$, $\phi_{0,n} = \sqrt{z_2^2/2n}$ e z_2^2 é o quantil desejado da distribuição χ^2 com dois graus de liberdade.

Portanto, para fins de estimação, a equação de verossimilhança deve se restringir a $q \in Q = [-q_n, q_n]$, em que Q foi denominado *intervalo de operação* da FCE para o qual $\hat{\phi}_*(q) > 0$. Em particular, para uma distribuição estável assimétrica X com vetor de parâmetros $\boldsymbol{\theta} = (\alpha, \gamma, \beta)'$, vale lembrar que $|\phi_X(q; \boldsymbol{\theta})|$ representa a FC da distribuição simetrizada $\boldsymbol{\theta}_s = (\alpha, \gamma, 0)'$ (Cap. 3). Assim, mesmo que haja assimetria, os parâmetros α e γ podem ser estimados independentemente do parâmetro de assimetria β com base

em $|\hat{\phi}(q)|$. Considera-se que $\hat{\phi}_2(q)$ (parte real da FCE) seja estatisticamente significativa para um nível de significância ns se

$$|\sqrt{2n}\hat{\phi}_2(q)| > z_{ns}, \quad (7.1)$$

em que $ns = P(|Z| > z_{ns})$ e $Z \sim N(0, 1)$.

Experimentos de Monte Carlo foram realizados para simular passeios aleatórios na forma $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$, em que $\{X_t\}$ é uma sequência IID de VAs estáveis. Com base nessas simulações, observamos que o padrão esperado $\ln(-\ln |\hat{\phi}_{S_{\Delta t}}(q)|)$ versus $\ln(q)$ sofre uma quebra de *scaling* sempre no mesmo patamar, independentemente do valor Δt . Esse fato permite explicar as quebras abruptas de *scaling* de outras características distribucionais, como o momentos absolutos das distribuições quase estáveis [84] e os truncamentos não abruptos da FD.

Com base no método de estimação por máxima verossimilhança (EMV) via funções características apresentado no Cap. 5, foi possível avaliar a estabilidade ou quase estabilidade do passeio aleatório $S_{\Delta t} = X_1 + \dots + X_{\Delta t}$ para cada $\Delta t \geq 1$. Se, por exemplo, $\{X_t\}_{1 \leq t \leq \Delta t}$ for uma sequência IID de VA estáveis com parâmetros α e γ , então $\alpha_{S_{\Delta t}} = \alpha$ e $\gamma_{S_{\Delta t}} = \Delta t \gamma$. Porém, se essas VA IID não forem estáveis, então $\alpha_{S_{\Delta t}} \rightarrow 2$ à medida que Δt aumenta e $\gamma_{S_{\Delta t}} = \Delta t \gamma$. E, se não houver independência nem estabilidade, então é possível que $\alpha_{S_{\Delta t}} \approx \alpha$ enquanto $\Delta t_0 \leq \Delta t \leq \Delta t_1$, e depois $\alpha_{S_{\Delta t}} \rightarrow 2$. Mas nesse caso, $\gamma_{S_{\Delta t}} \neq \Delta t \gamma$ por causa da autocorrelação serial de curto ou de longo alcance.

O Cap. 6 apresentou um novo teste estatístico para a detecção de dependência não linear. Essa ferramenta é útil, pois a dependência não linear também proporciona lentidão na convergência do processo $S_{\Delta t}$ para a gaussiana [33, 35]. O teste proposto é do tipo Cramer-von Mises e considera a estatística qui-quadrado da razão de verossimilhança, o que equivale a minimizar a divergência de Kullback-Leibler entre duas distribuições empíricas. Nosso teste não requer suposições acerca da distribuição dos dados (*distribution free*), e se mostrou mais poderoso do que o clássico teste de HBKR para a detecção da dependência não linear em situações variadas.

7.1 Perspectivas para trabalhos futuros

7.1.1 Representação em séries

Nesta seção, discutiremos acerca da representação de uma distribuição estável simétrica X como uma soma convergente de produtos entre variáveis aleatórias independentes [103]. Uma dessas variáveis se relaciona com o tempo de chegada de uma partícula em um processo de Poisson.

Considere que $N(\tau)$ representa o número registrado de partículas no intervalo de tempo $[0, \tau]$, que evolui segundo um processo de Poisson com taxa de chegada igual a $\nu\tau$. Se G_k representa o instante em que se registra a chegada da k -ésima partícula no processo, então sabe-se que a distribuição do intervalo de tempo entre duas partículas consecutivas, $G_{k+1} - G_k$, é exponencial com média ν^{-1} . Desse modo, pela hipótese de independência entre os intervalos de tempo, a distribuição do tempo de chegada $G_k = G_1 + G_2 + \dots + G_k$ segue uma distribuição gama cujo parâmetro de forma é igual a k e o de escala é igual a ν . Outro aspecto conhecido é que a distribuição de Poisson é estável. Se $N_1(\tau)$ e $N_2(\tau)$ são dois processos de Poisson independentes com taxas respectivamente iguais a ν_1 e ν_2 , então $N_1(\tau) + N_2(\tau)$ segue um processo de Poisson com taxa igual a $\nu_1 + \nu_2$. E se o processo de Poisson com taxa ν se ramifica em A e B , de modo que uma partícula segue para o ramo A com probabilidade $0 \leq p \leq 1$ e para o B com probabilidade $(1 - p)$, então as distribuições dos números de partículas em A e B também seguem distribuições de Poisson com taxas $p\nu$ e $(1 - p)\nu$, respectivamente [89].

Sejam $\{G_k\}_{k \geq 1}$ uma sequência de tempos aleatórios independentes, em que $G_k \sim \text{gamma}(k, 1)$, e $\{\epsilon_k\}$ uma sequência independente de $\{G_k\}$ formada por ruídos aleatórios IID, e considere a soma

$$X = \sum_{k=1}^{\infty} G_k^{-\frac{1}{\alpha}} \epsilon_k, \quad (7.2)$$

em que $0 < \alpha \leq 2$. Nesse caso, se X_1 e X_2 são duas cópias independentes de X , tem-se

que

$$X = \gamma_1 X_1 + \gamma_2 X_2 \tag{7.3}$$

$$= \sum_{i=1}^{\infty} (\gamma_1^{-\alpha} G_{1,i})^{-\frac{1}{\alpha}} \epsilon_{1,i} + \sum_{j=1}^{\infty} (\gamma_2^{-\alpha} G_{2,j})^{-\frac{1}{\alpha}} \epsilon_{2,j}, \tag{7.4}$$

em que $\gamma_1 > 0$, $\gamma_2 > 0$ e $\gamma_1^\alpha + \gamma_2^\alpha = 1$. A relação acima pode ser verificada com base nas propriedades da distribuição de Poisson descritas no início desta subseção. Observe que $\gamma_1^{-\alpha} G_{1,k}$ e $\gamma_2^{-\alpha} G_{2,k}$ seguem distribuições gama com parâmetros de escala, respectivamente, iguais a γ_1^α e γ_2^α . Conseqüentemente, no sentido contrário da ramificação (sobreposição), o processo resultante é de Poisson com taxa igual a $\gamma_1^\alpha + \gamma_2^\alpha = 1$. Como G_k é um passeio aleatório com incrementos são positivos, para uma dada realização, a sobreposição dos tempos observados nos processos X_1 e X_2 produz uma sequência de tempos equivalente a de um processo de Poisson com taxa unitária, como ilustra a Fig. 7.1. Assim, $(\gamma_1^{-\alpha} G_{1,i})^{-\frac{1}{\alpha}} + (\gamma_2^{-\alpha} G_{2,j})^{-\frac{1}{\alpha}} = G_k^{-\frac{1}{\alpha}}$.

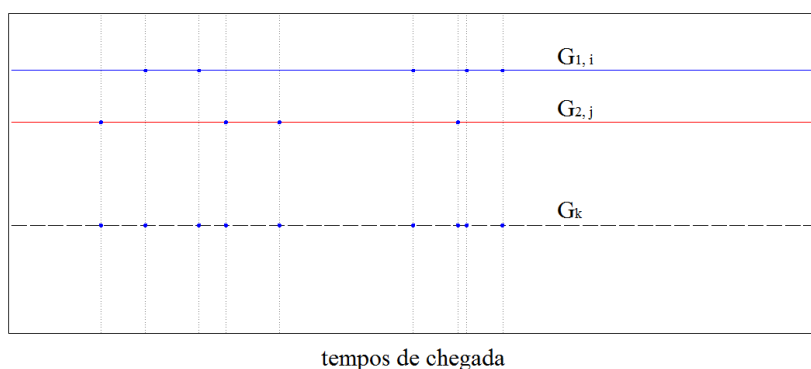


Figura 7.1: Realizações de tempos de chegada $G_{1,i} \sim \text{gamma}(i, \gamma_1^\alpha)$ e $G_{2,j} \sim \text{gamma}(j, \gamma_2^\alpha)$, em que $i = 1, \dots, 5$ e $j = 1, \dots, 4$, $\gamma_1 > 0$, $\gamma_2 > 0$ e $\gamma_1^\alpha + \gamma_2^\alpha = 1$. Esses tempos foram projetados na linha inferior, representando as realizações de tempos de chegada $G_k \sim \text{gamma}(k, 1)$, $k = 1, \dots, 9$ de um processo de Poisson sobreposto.

Assim, por (7.3), tem-se que a soma X definida em é estável. A variável $G_1^{-\frac{1}{\alpha}}$ segue a distribuição de Fréchet, um caso particular da distribuição generalizada de valores extremos. As demais variáveis, $G_k^{-\frac{1}{\alpha}}$, em que $k \geq 2$, seguem distribuições do tipo gama

inversa. Ao contrário da primeira variável, estas possuem caudas menos pesadas.

Assim, posteriormente, podemos estudar o comportamento da distribuição X no caso em que o sistema se limita um número finito (N) de partículas, i.e.,

$$X = \sum_{k=1}^N G_k^{-\frac{1}{\alpha}} \epsilon_k; \quad (7.5)$$

ou, no caso em que desconsidera a VA de Fréchet da soma, ou seja,

$$X = \sum_{k=2}^{\infty} G_k^{-\frac{1}{\alpha}} \epsilon_k. \quad (7.6)$$

7.1.2 Estudo da origem do agrupamento de volatilidades e das correlações de longo alcance

Ao mesmo tempo que o truncamento da distribuição de Lévy não permite descrever agrupamentos de volatilidades, os modelos de séries temporais para a volatilidade (os modelos GARCH) não são consistentes com as propriedades de quase-estabilidade [74]. Em um processo linear estacionário, os agrupamentos de volatilidades surgem quando há entradas aleatórias de grandes magnitudes. Essas grandes entradas momentaneamente produzem desequilíbrios no sistema estocástico linear, que se manifestam por meio de agrupamentos de volatilidades. Como ilustração, a Fig. 7.2 mostra a evolução temporal de um processo linear [11] na forma

$$Y_t = -0.7X_{t-1} + X_t - 0.7 * X_{t-1}, \quad (7.7)$$

em que Y_t representa um retorno no instante t , X_t representa um ruído aleatório com distribuição assimétrica estável com $\alpha = 1.6$, $\gamma = 10$ e $\beta = 0.5$. No painel inferior da Fig. 7.2, os agrupamentos de volatilidades na série $\{Y_t\}$ são produzidos pelos choques de grandes magnitudes que se destacam no painel superior da Fig. 7.2. Embora a estrutura do processo Y_t seja de curto alcance, sua função de autocorrelação amostral (Fig. 7.3) inesperadamente mostra uma estrutura de longo alcance.

Por isso, outro assunto a ser investigado é a relação entre os agrupamentos de volatilidades, os processos lineares com erros estáveis ou quase estáveis e os processos de

memória longa. Evidentemente, outros modelos de séries temporais, como os não lineares [118, 38, 83] e os determinísticos [113]), podem ser considerados nessa investigação.

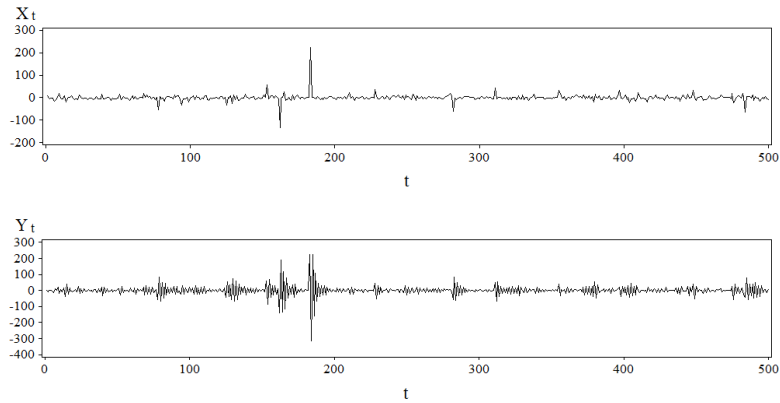


Figura 7.2: Simulação de uma série temporal $\{Y_t\}$ (painel inferior), segundo um processo ARMA(1,1) com ruídos assimétricos de cauda pesada na forma $Y_t = -0.7X_{t-1} + X_t + -0.7 * X_{t-1}$, em que $t = 1, \dots, 500$; e $\{X_t\}$ (painel superior) são ruídos IID conforme uma distribuição assimétrica estável com $\alpha = 1.6$, $\gamma = 10$ e $\beta = 0.5$. Notam-se os pequenos agrupamentos de volatilidade na série $\{Y_t\}$.

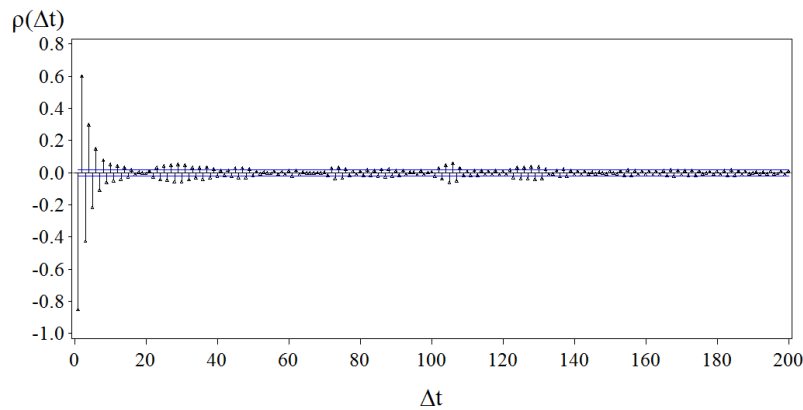


Figura 7.3: Função de autocorrelação (FAC) amostral da série temporal simulada $\{Y_t\}$, $\rho(\Delta t) = \text{Corr}(Y_t, Y_{t-\Delta t})$. O comportamento da FAC sugere memória de longo alcance, embora seu processo gerador seja de curto alcance.

7.1.3 A FCE multivariada e outras medidas de independência

Seja $\mathbf{X}_1, \dots, \mathbf{X}_n$ uma amostra aleatória simples de n vetores de dimensão m . A FCE associada a essa amostra é definida como

$$\hat{\phi}(\mathbf{q}; \boldsymbol{\theta}) = \frac{1}{n} \sum_{j=1}^n e^{i\mathbf{q} \cdot \mathbf{X}_j} \quad (7.8)$$

$$= \int_{\mathbb{R}^n} e^{i\mathbf{q} \cdot \mathbf{x}} d\hat{F}(\mathbf{x}; \boldsymbol{\theta}) \quad (7.9)$$

$$= \frac{1}{n} \sum_{j=1}^n \cos(\mathbf{q} \cdot \mathbf{X}_j) + i \frac{1}{n} \sum_{j=1}^n \text{sen}(\mathbf{q} \cdot \mathbf{X}_j), \quad (7.10)$$

em que $\mathbf{q} = (q_1, \dots, q_m)' \in \mathbb{R}^m$. Se os elementos dos vetores aleatórios \mathbf{X}_j forem independentes, então

$$\hat{\phi}(\mathbf{q}; \boldsymbol{\theta}) = \prod_{j=1}^n \hat{\phi}(q_j; \boldsymbol{\theta}), \quad (7.11)$$

para qualquer $q_j \in \mathbb{R}$. A avaliação da significância da FCE multivariada também segue o mesmo procedimento descrito no Cap. 4, já que a FCE multivariada pode ser tratada como a FCE de uma soma ponderada. Observe que

$$\mathbf{q} \cdot \mathbf{X} = q_1 X_1 + q_2 X_2 + \dots + q_m X_m \quad (7.12)$$

$$= q(q'_1 X_1 + q'_2 X_2 + \dots + q'_m X_m), \quad (7.13)$$

em que $q'_j = \frac{q_j}{q}$, com $q'_j \equiv 0$ se $q = 0$. Assim, dado um vetor \mathbf{q} , a FCE multivariada é equivalente à FCE da variável aleatória $S_m = q'_1 X_1 + q'_2 X_2 + \dots + q'_m X_m$.

Em particular, em caso de independência, tem-se

$$\hat{\phi}(\mathbf{q}; \boldsymbol{\theta}) = \hat{\phi}^m(q; \boldsymbol{\theta}), \quad (7.14)$$

em que $\mathbf{q} = (q, q, \dots, q)'$, o que é equivalente à FCE de uma soma de m variáveis aleatórias IID, $S_m = X_1 + \dots + X_m$. Porém, fora do contexto gaussiano, $\hat{\phi}(\mathbf{q}; \boldsymbol{\theta}) = \hat{\phi}^m(q; \boldsymbol{\theta})$ não necessariamente representa independência mútua entre os elementos do vetor aleatório.

Com respeito à distância L_2 entre as distribuições de dois vetores aleatórios \mathbf{X} e \mathbf{Y} , com base na Eq. (2.36), a extensão para o caso multivariado [120] é

$$\int_{\mathbb{R}^m} (f_{\mathbf{Y}}(\mathbf{x}; \boldsymbol{\theta}') - f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\theta}))^2 d\mathbf{x} = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} |\phi_{\mathbf{Y}}(\mathbf{q}; \boldsymbol{\theta}') - \phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta})|^2 d\mathbf{q}. \quad (7.15)$$

Com base em (7.15), pode-se definir

$$\varsigma(\mathbf{X}) = \int_{\mathbb{R}^m} |\phi_{\mathbf{X}}(\mathbf{q}; \boldsymbol{\theta}) - \prod_{j=1}^m \phi_{X_j}(q_j; \boldsymbol{\theta})|^2 d\mathbf{q} \quad (7.16)$$

como uma medida de dependência generalizada entre os elementos de \mathbf{X} .

Exemplo 7.1.1. Considerando a FD dada no exemplo 6.2.1, enquanto o coeficiente de correlação não permite medir a dependência entre X_1 e X_2 , a medida $\varsigma(X_1, X_2)$ para esse caso é dada por

$$\varsigma(X_1, X_2) = \int \int |4\tau\xi q_1 q_2 \left\{ \frac{1}{(\xi^2 + q_1^2)^2 (\tau^2 + q_2^2)^2} - \frac{1}{(\tau^2 + q_1^2)^2 (\xi^2 + q_2^2)^2} \right\}|^2 dq_1 dq_2.$$

Por exemplo, se $\tau = 1, 2$ e $\xi = 1, 9$, tem-se que $\varsigma(X_1, X_2) = 0, 0914$.

□

Do ponto de vista empírico, futuramente podemos estudar as propriedades da estatística

$$\hat{\varsigma}(\mathbf{X}) = \int_{\mathbb{R}^m} |\hat{\phi}_{\mathbf{X}}(\mathbf{q}) - \prod_{j=1}^m \hat{\phi}_{X_j}(q_j)|^2 d\mathbf{q} \quad (7.17)$$

para a avaliação da dependência contemporânea e serial.

7.1.4 Extensões do teste de independência

No Cap. 6 introduzimos um novo teste de independência para o caso bivariado ($m = 2$). Para o caso multivariado ($m > 2$), considera-se uma amostra aleatória $\mathbf{X}_1, \dots, \mathbf{X}_n$, em que $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{im})'$ é um vetor aleatório m -dimensional definida por uma função de distribuição acumulada (FDA) $F(\mathbf{x})$, com $\mathbf{x} = (x_1, \dots, x_m)' \in \mathbb{R}^m$. O propósito é avaliar a hipótese de independência conjunta

$$F(\mathbf{x}) = \prod_{j=1}^m F_j(x_j), \quad (7.18)$$

para todo x_j , em que $F_j(x_j)$ representa a FDA marginal da VA X_j . Os processos aleatórios para a construção do teste são

$$L^2(\mathbf{x}) = 2 \sum_{k_1=0}^1 \cdots \sum_{k_m=0}^1 n_{k_1 \dots k_m} \ln \left\{ \frac{n_{k_1 \dots k_m}}{\langle n \rangle_{k_1 \dots k_m}} \right\} \quad (7.19)$$

e

$$\chi^2(\mathbf{x}) = \sum_{k_1=0}^1 \cdots \sum_{k_m=0}^1 \frac{(n_{k_1 \dots k_m} - \langle n \rangle_{k_1 \dots k_m})^2}{\langle n \rangle_{k_1 \dots k_m}}, \quad (7.20)$$

em que

$$n_{k_1 \dots k_m} = n_{k_1 \dots k_m}(\mathbf{x}) = \sum_{i=1}^n \prod_{j=1}^m \{1 - I(x_j - X_{ij})\}^{1-k_j} I^{k_j}(x_j - X_{ij}), \quad (7.21)$$

e

$$\langle n \rangle_{k_1 \dots k_m} = \langle n \rangle_{k_1 \dots k_m}(\mathbf{x}) = n \prod_{j=1}^p \{1 - \hat{F}_j(x_j)\}^{1-k_j} \hat{F}_j^{k_j}(x_j) \quad (7.22)$$

representam, respectivamente, as contagens observadas e as esperadas para um dado vetor \mathbf{x} . Com base nos processos $L^2(\mathbf{x})$ e $\chi^2(\mathbf{x})$ discutidos no Cap. 6, as estatísticas do teste para o caso $m > 2$ são

$$B_{LR} = \int_{\mathbb{R}^m} L^2(\mathbf{x}) d\hat{F}(\mathbf{x}) \quad (7.23)$$

e

$$B_{\chi^2} = \int_{\mathbb{R}^m} \chi^2(\mathbf{x}) d\hat{F}(\mathbf{x}). \quad (7.24)$$

Para cada ponto \mathbf{x} , os processos $L^2(\mathbf{x})$ e $\chi^2(\mathbf{x})$ seguem uma distribuição χ^2 com $2^m - m - 1$ graus de liberdade, de modo que

$$\langle B_{LR} \rangle = \langle B_{\chi^2} \rangle = \langle B \rangle = 2^m - m - 1. \quad (7.25)$$

Caso não seja possível determinar a função de covariância para $m > 2$, a distribuição amostral da estatística B poderá ser estudada empiricamente mediante simulações de Monte Carlo [123].

A estatística B possivelmente pode ser aplicada para se testar a aderência dos dados (*goodness of fit*) com respeito a determinado modelo hipotético [1]. Assim, o teste poderia ser modificado para se avaliar, por exemplo, a normalidade multivariada dos dados ou a presença de uma estrutura particular de dependência não linear.



Apêndice A

Addendum matemático

Este apêndice lista alguns resultados matemáticos elementares úteis [49], contemplando integrais, propriedades das funções gama e delta e as dos coeficientes binomiais.

A.1 Integrais

$$\int_{-\infty}^{+\infty} e^{\pm iqu} du = 2\pi\delta(q). \quad (\text{A.1})$$

$$\int_{-\infty}^{+\infty} ue^{iqu} du = -2\pi i\delta^{(1)}(q). \quad (\text{A.2})$$

$$\int_{-\infty}^{+\infty} u^2 e^{iqu} du = -2\pi\delta^{(2)}(q). \quad (\text{A.3})$$

$$\int_{-\infty}^{+\infty} u^3 e^{iqu} du = 2\pi i\delta^{(3)}(q). \quad (\text{A.4})$$

$$\int_{-\infty}^{+\infty} u^k e^{iqu} du = 2\pi(-i)^k \delta^{(k)}(q). \quad (\text{A.5})$$

$$\int_0^{+\infty} (e^{iu} - 1)u^{-1-\alpha} du = \Gamma(-\alpha)e^{-i\pi\alpha/2}, 0 < \alpha < 1. \quad (\text{A.6})$$

$$\int_0^{+\infty} (e^{iu} - 1 - iu)u^{-1-\alpha} du = \Gamma(-\alpha)e^{-i\pi\alpha/2}, 1 < \alpha < 2. \quad (\text{A.7})$$

$$\int_0^{+\infty} (e^{iqu} - 1 - iquI_{(0,1]}(u))u^{-2} du = -\frac{\pi q}{2} - iq \ln q + icq, \quad (\text{A.8})$$

em que $q > 0$, $c = \int_1^{+\infty} u^{-2} \text{sen} u du + \int_0^1 u^{-2} (\text{sen} u - u) du$ e $I_{(0,1]}(u) = 1$ se $0 < u \leq 1$ e $I_{(0,1]}(u) = 0$ se caso contrário.

$$\int_0^{+\infty} u^{\alpha-1} e^{-\gamma u} \cos(\beta u) du = \frac{\Gamma(\alpha)}{(\gamma^2 + \beta^2)^{\alpha/2}} \cos\left(\alpha \arctan \frac{\beta}{\gamma}\right), \alpha > 0, \gamma > 0 \quad (\text{A.9})$$

$$\int_0^{+\infty} u^{\alpha-1} e^{-\gamma u} \text{sen}(\beta u) du = \frac{\Gamma(\alpha)}{(\gamma^2 + \beta^2)^{\alpha/2}} \text{sen}\left(\alpha \arctan \frac{\beta}{\gamma}\right), \alpha > 0, \gamma > 0 \quad (\text{A.10})$$

$$\int_0^{+\infty} \frac{1 - \cos u}{u^2} du = \frac{\pi}{2}. \quad (\text{A.11})$$

A.2 Função gama

Define-se a função gama como

$$\Gamma(a) = \int_0^{+\infty} u^{a-1} e^{-u} du, \quad (\text{A.12})$$

em que $a \in \mathbb{R} - \{0, -1, -2, -3, \dots\}$.

$$\Gamma(a+1) = a\Gamma(a). \quad (\text{A.13})$$

$$\Gamma(n+1) = n!, n \in \mathbb{N}. \quad (\text{A.14})$$

$$\Gamma(n+1/2) = \frac{(2n)!\sqrt{\pi}}{4^n n!} = \frac{(2n-1)!!\sqrt{\pi}}{2^n}. \quad (\text{A.15})$$

$$\Gamma(a)\Gamma(a + 1/2) = 2^{1-2a}\sqrt{\pi}\Gamma(2a). \quad (\text{A.16})$$

$$\Gamma(a)\Gamma(a + 1/n)\Gamma(a + 2/n)\cdots\Gamma(a + (n - 1)/n) = (2\pi)^{\frac{n-1}{2}} n^{\frac{1}{2}-na}\Gamma(na). \quad (\text{A.17})$$

Para a grande,

$$\ln \Gamma(a) \approx (a - \frac{1}{2}) \ln a - a + \frac{1}{2} \ln(2\pi). \quad (\text{A.18})$$

A.3 Derivadas da função delta

$$2\pi(-i)^k \delta^{(k)}(q) = \int_{-\infty}^{+\infty} u^k e^{iqu} du. \quad (\text{A.19})$$

$$2\pi(i)^k \delta^{(k)}(q) = \int_{-\infty}^{+\infty} u^k e^{-iqu} du. \quad (\text{A.20})$$

$$q^n \delta^{(n)}(q) = (-1)^n n! \delta(q). \quad (\text{A.21})$$

$$q \delta^{(n)}(q) = -n \delta^{(n-1)}(q). \quad (\text{A.22})$$

A.4 Coeficientes binomiais

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}. \quad (\text{A.23})$$

$$\sum_{k=0}^n \binom{n}{k}^2 = \sum_{k=0}^n \binom{n}{k} \binom{n}{n-k} = \binom{2n}{n}. \quad (\text{A.24})$$

$$\sum_{j=l}^{\infty} \frac{1}{\binom{j}{k}} = \frac{k}{(k-1)\binom{l-1}{k-1}}, \text{ para } k \geq 2. \quad (\text{A.25})$$

Apêndice B

Uma relação binomial da função escore

Proposição. Considere uma função de densidade $f(x; \theta) \in C^k$, em que $\theta \in \mathbb{R}$, tal que $g(x; \theta) = \frac{d}{d\theta} f(x; \theta) \in C^k$. Nessa situação,

$$g^{(k)}(x; \theta) = \sum_{l=0}^k \binom{k}{l} s^{(k-l)}(x; \theta) f^{(l)}(x; \theta), \quad (\text{B.1})$$

em que $s(x; \theta) = \frac{d}{d\theta} \ln f(x; \theta)$ é a função escore.

Demonstração:

Como

$$s(x; \theta) = g(x; \theta) / f(x; \theta), \quad (\text{B.2})$$

tem-se imediatamente para $k = 0$ que $g(x; \theta) = f(x; \theta)s(x; \theta)$. Agora, por simplicidade, considere que $f = f(x; \theta)$, $g = g(x; \theta)$ e $s = s(x; \theta)$.

Para $k = 1$, derivando (B.2) se obtém $s^{(1)} = \frac{g^{(1)}}{f} - s \frac{f^{(1)}}{f}$ de modo que

$$g^{(1)} = s^{(1)} f + s f^{(1)}.$$

Para $k = 2$,

$$\begin{aligned} s^{(2)} &= \frac{g^{(2)}}{f} - \frac{g^{(1)} f^{(1)}}{f} - s^{(1)} \frac{f^{(1)}}{f} - s \left\{ \frac{f^{(2)}}{f} - \left(\frac{f^{(1)}}{f} \right)^2 \right\} \\ &= \frac{g^{(2)}}{f} - \left\{ s^{(1)} + s \frac{f^{(1)}}{f} \right\} \frac{f^{(1)}}{f} - s^{(1)} \frac{f^{(1)}}{f} - s \frac{f^{(2)}}{f} + s \left(\frac{f^{(1)}}{f} \right)^2 \\ &= \frac{g^{(2)}}{f} - 2s^{(1)} \frac{f^{(1)}}{f} - s \frac{f^{(2)}}{f}. \end{aligned}$$

Assim,

$$g^{(2)} = s^{(2)} f + s^{(1)} f^{(1)} + s f^{(2)}.$$

Para $k = 3$,

$$\begin{aligned} s^{(3)} &= \frac{g^{(3)}}{f} - \frac{g^{(2)} f^{(1)}}{f} - 2s^{(2)} \frac{f^{(1)}}{f} - 2s^{(1)} \left\{ \frac{f^{(2)}}{f} - \left(\frac{f^{(1)}}{f} \right)^2 \right\} - s^{(1)} \frac{f^{(2)}}{f} - s \left\{ \frac{f^{(3)}}{f} - \frac{f^{(2)} g^{(1)}}{f} \right\} \\ &= \frac{g^{(3)}}{f} - 3s^{(2)} \frac{f^{(1)}}{f} - 3s^{(1)} \frac{f^{(2)}}{f} - s \frac{f^{(3)}}{f}, \end{aligned}$$

de modo que

$$g^{(3)} = s^{(3)} f + 3s^{(2)} f^{(1)} + 3s^{(1)} f^{(2)} + s f^{(3)}.$$

Repetindo-se o mesmo procedimento para $k = 4$, obtém-se

$$s^{(4)} = \frac{g^{(4)}}{f} - \sum_{l=1}^4 \binom{4}{l} s^{(4-l)} \frac{f^{(l)}}{f},$$

ou seja,

$$g^{(4)}(x; \boldsymbol{\theta}) = \sum_{l=0}^4 \binom{4}{l} s^{(4-l)} f^{(l)}.$$

Para $n \geq 1$, suponha que

$$s^{(n)} = \frac{g^{(n)}}{f} - \sum_{l=1}^n \binom{n}{l} s^{(n-l)} \frac{f^{(l)}}{f}$$

ou

$$g^{(n)}(x; \boldsymbol{\theta}) = \sum_{l=0}^n \binom{n}{l} s^{(n-l)} f^{(l)}.$$

Agora resta mostrar que a relação vale para $n + 1$, isto é,

$$s^{(n+1)} = \frac{g^{(n+1)}}{f} - \sum_{l=1}^{n+1} \binom{n+1}{l} s^{(n+1-l)} \frac{f^{(l)}}{f}.$$

De fato,

$$\begin{aligned}
 s^{(n+1)} &= \frac{g^{(n+1)}}{f} - \frac{g^{(n)} f^{(1)}}{f} - \sum_{l=1}^n \binom{n}{l} \left\{ s^{(n+1-l)} \frac{f^{(l)}}{f} + s^{(n-l)} \left(\frac{f^{(l+1)}}{f} - \frac{f^{(l)} f^{(1)}}{f} \right) \right\} \\
 &= \frac{g^{(n+1)}}{f} - s^{(n)} \frac{f^{(1)}}{f} - \sum_{l=1}^n \binom{n}{l} \left\{ s^{(n+1-l)} \frac{f^{(l)}}{f} + s^{(n-l)} \frac{f^{(l+1)}}{f} \right\} \\
 &= \frac{g^{(n+1)}}{f} - s^{(n)} \frac{f^{(1)}}{f} - \sum_{l=1}^n \binom{n}{l} s^{(n+1-l)} \frac{f^{(l)}}{f} - \sum_{j=2}^{n+1} \binom{n}{j-1} s^{(n+1-j)} \frac{f^{(j)}}{f} \\
 &= \frac{g^{(n+1)}}{f} - \sum_{l=1}^n \left\{ \binom{n}{l} + \binom{n}{l-1} \right\} s^{(n+1-l)} \frac{f^{(l)}}{f} - \binom{n+1}{n+1} s \frac{f^{(n+1)}}{f} \\
 &= \frac{g^{(n+1)}}{f} - \sum_{l=1}^{n+1} \binom{n+1}{l} s^{(n+1-l)} \frac{f^{(l)}}{f}.
 \end{aligned}$$

◇

Apêndice C

A fórmula de Lévy-Khinchine

Considere que X seja uma variável aleatória (VA) infinitamente divisível com função de distribuição $F(x)$ e função característica (FC) $\phi_X(q)$. Logo, $\forall n \in \mathbb{N}^+$, existe uma VA X_n tal que $\phi_X(q) = \phi_{X_n}^n(q)$, em que $\phi_{X_n}(q)$ é FC de X_n . Para $\phi_X(q) \neq 0$ e $n \gg 1$, tem-se

$$\begin{aligned} n\{\phi_{X_n}(q) - 1\} &= n\{\phi_X^{1/n}(q) - 1\} \\ &= n\{e^{\frac{1}{n} \ln \phi_X(q)} - 1\} \\ &= n\{1 + \frac{1}{n} \ln \phi_X(q) + O(1/n) - 1\} \approx \ln \phi_X(q). \end{aligned}$$

Assim, para n grande,

$$n\{\phi_{X_n}(q) - 1\} = n \int_{-\infty}^{+\infty} (e^{iqx} - 1) dF_{X_n}(x) \approx \ln \phi_X(q).$$

Agora, considere as integrais [47]

$$G_n(u) = n \int_{-\infty}^u \frac{x^2}{1+x^2} dF_{X_n}(x)$$

e

$$I_n(q) = n \int_{-\infty}^{+\infty} (e^{iqu} - 1) \frac{1+u^2}{u^2} dG_n(u),$$

de modo que, pelas propriedades da integral de Lebesgue,

$$I_n(q) = n \int_{-\infty}^{+\infty} (e^{iqu} - 1) \frac{1+u^2}{u^2} \frac{u^2}{1+u^2} dF_{X_n}(x) \approx \ln \phi_X(q).$$

Assim, pode-se concluir que

$$\Re(I_n(q)) \approx \ln |\phi_X(q)| \quad (\text{C.1})$$

(pois se $z = re^{ix} \in \mathbb{C}$, então $\ln z = \ln r + ix$; isto é, $\Re(\ln z) = \ln r = \ln |z|$). Agora, devemos verificar que $G_n(+\infty)$ é limitada. Para isso, considere as integrais

$$A_n = \int_{|u| \leq 1} dG_n(u)$$

e

$$B_n = \int_{|u| > 1} dG_n(u),$$

de modo que $A_n + B_n = \int dG_n(u)$. Considerando $0 \leq q \leq 2$, com base em (C.1) podemos escrever

$$-\ln \phi_X(q) \approx \int \left(1 - \cos(qu)\right) \frac{1+u^2}{u^2} dG_n(u).$$

Dado $\epsilon > 0$, temos

$$-\ln |\phi_X(q)| + \epsilon \geq \int_{|u| \leq 1} \left(1 - \cos(qu)\right) \frac{1+u^2}{u^2} dG_n(u) \quad (\text{C.2})$$

e

$$-\ln |\phi_X(q)| + \epsilon \geq \int_{|u| > 1} \left(1 - \cos(qu)\right) \frac{1+u^2}{u^2} dG_n(u). \quad (\text{C.3})$$

Como $\cos u \approx 1 - \frac{u^2}{2} + \frac{u^4}{4!}$, temos $\frac{1-\cos u}{u^2} = \frac{1}{2} - \frac{u^2}{4!} + r$. Se $|u| \leq 1$, então $r > 0$, já que $u^{2k+2} < u^{2k}$ e, além disso, $-u^2 \geq 1$. Logo, $\frac{1-\cos u}{u^2} > \frac{1}{2} - \frac{u^2}{4!} > \frac{1}{2} - \frac{1}{24} > \frac{1}{3}$. Daí, considerando $\frac{1-\cos u}{u^2} > \frac{1}{3}$ e $q = 1$ na Eq. (C.2), temos

$$\begin{aligned} -\ln |\phi_X(1)| + \epsilon &\geq \int_{|u| \leq 1} \frac{1+u^2}{3} dG_n(u) \\ &= \frac{1}{3} \int_{|u| \leq 1} dG_n(u) + \frac{1}{3} \int_{|u| \leq 1} u^2 dG_n(u) > \frac{A_n}{3}, \end{aligned}$$

desde que $\int_{|u| \leq 1} u^2 dG_n(u) > 0$. Agora, integrando a Eq. (C.3) para $0 \leq q \leq 2$, obtemos

$$\begin{aligned} -\int_0^2 \left(\ln |\phi_X(q)| - \epsilon \right) dq &\geq \int_0^2 \int_{|u| > 1} \left(1 - \cos(qu)\right) \frac{1+u^2}{u^2} dG_n(u) dq \\ -\frac{1}{2} \int_0^2 \ln |\phi_X(q)| dq + \epsilon &\geq \int_{|u| > 1} \left(1 - \frac{\text{sen}(2u)}{2u}\right) \frac{1+u^2}{u^2} dG_n(u) \\ &> \int_{|u| > 1} \left(1 - \frac{\text{sen}(2u)}{2u}\right) dG_n(u) > \frac{B_n}{2} \end{aligned}$$

(como $|u| > 1$, $\text{sen}2u < 1$, ou seja, $\text{sen}2u < |u|$; assim, $1 - \frac{\text{sen}2u}{|2u|} > \frac{1}{2}$ e, como $\text{sen}(u)$ é função ímpar, podemos escrever $1 - \frac{\text{sen}2u}{2u} > \frac{1}{2}$). Logo, como $\ln |\phi_X(1)|$ e $\frac{1}{2} \int_0^2 \ln |\phi_X(q)| dq$ são finitos, conclui-se que $G_n(+\infty)$ é limitada.

Agora, devemos mostrar que

$$\lim_{U \rightarrow \infty} \int_{|u| > U} dG_n(u) = 0.$$

Dado $\epsilon > 0$ e n suficientemente grande, temos

$$-\ln |\phi_X(q)| + \epsilon \geq \int_{|u| > U} \left(1 - \cos(qu)\right) \frac{1+u^2}{u^2} dG_n(u) \geq \int_{|u| > U} \left(1 - \cos(qu)\right) dG_n(u).$$

Integrando a expressão anterior para $0 \leq q \leq 2/U$, $U \geq 1$, obtemos

$$\begin{aligned} -\int_0^{2/U} \ln |\phi_X(q)| dq + \frac{2}{U} \epsilon &\geq \int_0^{2/U} \int_{|u| > U} \left(1 - \cos(qu)\right) dG_n(u) dq \\ -\frac{U}{2} \int_0^{2/U} \ln |\phi_X(q)| dq + \epsilon &= \int_{|u| > U} \left(1 - \frac{U \text{sen}(2u/U)}{2u}\right) dG_n(u). \end{aligned}$$

Como $|u| \geq U$, temos que $1 - \frac{\text{sen}(2u/U)}{2u/U} \geq \frac{1}{2}$ (pois $|u|/U \geq 1$, ou seja, $|u/U| \geq 1$). Daí,

$$-\frac{U}{2} \int_0^{2/U} \ln |\phi_X(q)| dq + \epsilon \geq \frac{1}{2} \int_{|u| > U} dG_n(u),$$

ou seja,

$$\int_{|u| > U} dG_n(u) \leq -U \int_0^{2/U} \ln |\phi_X(q)| dq + 2\epsilon.$$

Por outro lado, pelo teorema do valor médio,

$$\left| \int_0^{2/U} \ln |\phi_X(q)| dq \right| = \frac{2}{U} \left| \ln |\phi_X(\xi)| \right|,$$

em que $\xi \in [0, 2/U]$. Portanto, $\frac{U}{2} \left| \int_0^{2/U} \ln |\phi_X(q)| dq \right| = \left| \ln |\phi_X(\xi)| \right|$ e, assim,

$$\frac{U}{2} \left| \int_0^{2/U} \ln |\phi_X(q)| dq \right| \leq \max_{0 \leq q \leq 2/U} \left| \ln |\phi_X(\xi)| \right| < \epsilon,$$

o que permite concluir que $\int_{|u| > U} dG_n(u) \leq 4\epsilon$. Seja

$$\mu = \int \frac{1}{u} dG_n(u) = n \int \frac{u^2}{(1+u^2)u} dF_{X_n}(u) = n \int \frac{u}{1+u^2} dF_{X_n}(u).$$

Finalmente, temos [47]

$$\begin{aligned}\ln \phi_X(q) &\approx I_n(q) = \int \left(e^{iqu} - 1 \right) \frac{1+u^2}{u^2} dG_n(u) \\ &= \int \left(e^{iqu} - 1 \right) \frac{1+u^2}{u^2} dG_n(u) + i\mu q - i\mu q \\ &= \int \left(e^{iqu} - 1 \right) \frac{1+u^2}{u^2} dG_n(u) + i\mu q - \int \frac{iq}{u} dG_n(q) \\ &= \int \left(e^{iqu} - 1 - \frac{iqu}{1+u^2} \right) \frac{1+u^2}{u^2} dG_n(u) + i\mu q.\end{aligned}\tag{C.4}$$

◇

Apêndice D

A informação de Fisher

Considere uma distribuição X descrita por uma função de densidade $f(x; \theta)$, em que $\theta \in \mathbb{R}$ é um parâmetro desconhecido. A função escore é definida como

$$s(x; \theta) = \frac{d}{d\theta} \ln f(x; \theta).$$

O escore $s(x; \theta)$ indica a sensibilidade relativa de $f(x; \theta)$ a variações do parâmetro θ , e seu valor esperado é nulo, pois

$$\begin{aligned} \langle s(x; \theta) \rangle &= \int s(x; \theta) f(x; \theta) dx \\ &= \frac{d}{d\theta} \int f(x; \theta) dx = 0. \end{aligned}$$

Define-se a informação de Fisher como a variância de $s(x; \theta)$, ou seja,

$$\text{IF}(\theta) = \langle s^2(x; \theta) \rangle. \quad (\text{D.1})$$

Considere que $\hat{\theta}(\mathbf{x})$ é um estimador não viciado de θ , ou seja, $b(\theta) = \langle \hat{\theta}(\mathbf{x}) - \theta \rangle = 0$.

Derivando-se $b(\theta)$ com respeito a θ , obtemos

$$\begin{aligned} \frac{d}{d\theta} \left\{ \int_{\mathbf{x} \in \mathbb{R}^n} (\hat{\theta}(\mathbf{x}) - \theta) f(\mathbf{x}; \theta) d\mathbf{x} \right\} &= \\ \int (\hat{\theta}(\mathbf{x}) - \theta) \frac{d}{d\theta} f(\mathbf{x}; \theta) d\mathbf{x} - \int f(\mathbf{x}; \theta) d\mathbf{x} &= \\ \int (\hat{\theta}(\mathbf{x}) - \theta) \frac{d}{d\theta} \{ \ln f(\mathbf{x}; \theta) \} f(\mathbf{x}; \theta) d\mathbf{x} - 1 &= \\ \int (\hat{\theta}(\mathbf{x}) - \theta) \sqrt{f(\mathbf{x}; \theta)} \frac{d}{d\theta} \{ \ln f(\mathbf{x}; \theta) \} \sqrt{f(\mathbf{x}; \theta)} d\mathbf{x} - 1 &= 0. \end{aligned}$$

Aplicando a desigualdade de Cauchy-Schwarz, e considerando que as variáveis aleatórias são i.i.d., obtemos

$$\int (\hat{\theta}(\mathbf{x}) - \theta)^2 f(\mathbf{x}; \theta) d\mathbf{x} \int \left\{ \frac{d}{d\theta} \ln f(\mathbf{x}; \theta) \right\}^2 f(\mathbf{x}; \theta) d\mathbf{x} \geq 1.$$

Desenvolvendo a desigualdade anterior, obtemos

$$\begin{aligned} \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle \left\langle \left\{ \frac{d}{d\theta} \sum_{j=1}^n \ln f(x_j; \theta) \right\}^2 \right\rangle &= \\ \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle \left\langle \left(\frac{d}{d\theta} \sum_{j=1}^n \ln f(x_j; \theta) \right) \left(\frac{d}{d\theta} \sum_{k=1}^n \ln f(x_k; \theta) \right) \right\rangle &= \\ \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle \sum_{j=1}^n \sum_{k=1}^n \left\langle \left(\frac{d}{d\theta} \ln f(x_j; \theta) \right) \left(\frac{d}{d\theta} \ln f(x_k; \theta) \right) \right\rangle &= \\ \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle \sum_{j=1}^n \left\langle \left(\frac{d}{d\theta} \ln f(x_j; \theta) \right)^2 \right\rangle &= \\ \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle n \langle s^2(x; \theta) \rangle &= \\ \langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle n I(\theta) &\geq 1, \end{aligned}$$

de modo que

$$\langle (\hat{\theta}(\mathbf{x}) - \theta)^2 \rangle \geq \frac{1}{n \text{IF}(\theta)}, \quad (\text{D.2})$$

o que significa que o erro quadrático médio de um estimador não viciado $\hat{\theta}(X)$ será pelo menos igual a $\frac{1}{n \text{IF}(\theta)}$. Esse resultado é chamado desigualdade de Cramér-Rao, e o termo $\frac{1}{n \text{IF}(\theta)}$ é chamado de limite inferior de Cramér-Rao. A igualdade ocorre se $\hat{\theta}(\mathbf{x}) - \theta$ for proporcional a $\frac{d}{d\theta} \ln f(\mathbf{x}; \theta)$ com probabilidade 1, isto é, se existir uma função $c(\theta; n)$ tal que

$$\frac{d}{d\theta} \ln f(\mathbf{x}; \theta) = c(\theta; n)(\hat{\theta}(\mathbf{x}) - \theta). \quad (\text{D.3})$$

Logo, se $\hat{\theta}(\mathbf{x})$ for um estimador não viciado de m.v., conclui-se que seu erro quadrático médio é igual a $\frac{1}{n \text{IF}(\theta)}$, pois $\frac{d}{d\theta} \ln f(\mathbf{x}; \theta) = 0$.

Apêndice E

Simulação de uma VA estável

Este apêndice apresenta o método de Janicki-Weron para a simulação de realizações de variáveis aleatórias estáveis [58, 122]. Seja X uma VA estável com parâmetros $0 < \alpha \leq 2$, $\gamma = 1$, $|\beta| \leq 1$ e $\mu = 0$. Uma realização X pode ser obtida da seguinte maneira:

- gerar uma variável aleatória uniforme (U) no intervalo $(-\frac{\pi}{2}, \frac{\pi}{2})$;
- gerar uma variável aleatória exponencial (W) com média 1, independentemente de U ;
- para $\alpha \neq 1$, calcular

$$X = (1 + \beta_\alpha^2)^{\frac{1}{2\alpha}} \frac{\text{sen}(\alpha U + \arctan \beta_\alpha)}{\cos^{\frac{1}{\alpha}} U} \left(\frac{\cos[(1 - \alpha)U - \arctan \beta_\alpha]}{W} \right)^{\frac{1-\alpha}{\alpha}}, \quad (\text{E.1})$$

em que $\beta_\alpha = \beta \tan \frac{\pi\alpha}{2}$;

- para $\alpha = 1$,

$$X = \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta U \right) \tan U - \beta \ln \left(\frac{\frac{\pi}{2} W \cos U}{\frac{\pi}{2} + \beta U} \right) \right]. \quad (\text{E.2})$$

Se Y for uma VA estável com parâmetros $0 < \alpha \leq 2$, $\gamma > 0$, $|\beta| \leq 1$ e $\mu \in \mathbb{R}$, então uma realização Y pode ser obtida mediante a transformação

$$Y = \begin{cases} \gamma^{\frac{1}{\alpha}} X + \mu, & \text{se } \alpha \neq 1, \\ \gamma X + \mu + \frac{2\beta\gamma}{\pi} \ln \gamma, & \text{se } \alpha = 1. \end{cases} \quad (\text{E.3})$$

◇

Apêndice F

Detalhamento do Cap. 6

Este apêndice apresenta os detalhes relativos ao desenvolvimento da estatística do teste de independência proposto no Cap. 6. A Seção F.1 mostra que a estatística B é uma convolução de variáveis aleatórias independentes que seguem distribuições gama com o mesmo parâmetro de forma, mas com os de escala distintos. Na Seção F.2 esboçamos a solução do problema de autovalores, Eqs. (6.24) ou (6.25).

F.1 Forma geral da função característica de B

Considere o problema de autovalores na forma integral

$$\int_0^1 \text{Cov}(u, v)\gamma(v)dv = \lambda\gamma(u), \quad (\text{F.1})$$

em que $\text{Cov}(u, v)$ é a função de covariância de um processo gaussiano $L(u)$ com média nula, cuja solução são os autovalores $\lambda_1, \lambda_2, \dots$ e as autofunções correspondentes são $\gamma_1(u), \gamma_2(u), \dots$. Agora, considere [9, 59]

$$Y(u) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} \gamma_j(u) Z_j, \quad (\text{F.2})$$

em que Z_1, Z_2, \dots são variáveis gaussianas independentes com média 0 e variância 1. O processo $Y(u)$ é gaussiano com média nula e função de covariância

$$\begin{aligned}
 \langle Y(u)Y(v) \rangle &= \sum_{j,k \geq 1} \sqrt{\lambda_j \lambda_k} \gamma_j(u) \gamma_k(v) \langle Z_j Z_k \rangle \\
 &= \sum_{j \geq 1} \lambda_j \gamma_j(u) \gamma_j(v) \\
 &= \sum_{j \geq 1} \int_0^1 \text{Cov}(u, v') \gamma_j(v') dv' \gamma_j(v) \\
 &= \int_0^1 \text{Cov}(u, v') \sum_{j \geq 1} \gamma_j(v') \gamma_j(v) dv' \\
 &= \int_0^1 \text{Cov}(u, v') \delta(v' - v) dv' \\
 &= \text{Cov}(u, v).
 \end{aligned} \tag{F.3}$$

Portanto, os processos gaussianos $L(u)$ e $Y(u)$ são idênticos e, assim,

$$\begin{aligned}
 \int_0^1 L^2(u) du &= \int_0^1 Y^2(u) du \\
 &= \int_0^1 \sum_{j,k \geq 1} \sqrt{\lambda_j \lambda_k} \gamma_j(u) \gamma_k(u) Z_j Z_k du \\
 &= \int_0^1 \sum_{j \geq 1} \lambda_j Z_j^2 \gamma_j^2(u) du \\
 &= \sum_{j \geq 1} \lambda_j Z_j^2 \int_0^1 \gamma_j^2(u) du \\
 &= \sum_{j \geq 1} \lambda_j Z_j^2,
 \end{aligned} \tag{F.4}$$

em que $\{Z_j^2\}$ é uma sequência de variáveis aleatórias independentes χ^2 com 1 grau de liberdade. Logo, a função característica do processo $\int_0^1 L^2(u) du$ é

$$\left\langle \exp \left(iq \int_0^1 L^2(u) du \right) \right\rangle = \prod_{j=1}^{\infty} (1 - 2iq \lambda_j)^{-\frac{1}{2}}. \tag{F.5}$$

Com base em (F.4), temos

$$\begin{aligned}
 \left\langle \int_0^1 L^2(u) du \right\rangle &= \sum_{j \geq 1} \lambda_j \langle Z_j^2 \rangle \\
 &= \sum_{j \geq 1} \lambda_j,
 \end{aligned} \tag{F.6}$$

e

$$\begin{aligned}
 \text{Var}\left(\int_0^1 L^2(u)du\right) &= \text{Var}\left(\sum_{j \geq 1} \lambda_j Z_j^2\right) \\
 &= \sum_{j \geq 1} \lambda_j^2 \text{Var}(Z_j^2) \\
 &= 2 \sum_{j \geq 1} \lambda_j^2.
 \end{aligned} \tag{F.7}$$

F.2 Solução do problema de autovalores

Pela fatorabilidade da função de covariância (Eq. (6.19)), temos

$$\text{Cov}(u, v) = \frac{\sqrt{2}(\min\{u, v\} - uv)^2}{uv(1-u)(1-v)},$$

em que $(u, v) \in [0, 1] \times [0, 1]$. Ou seja,

$$\text{Cov}(u, v) = \begin{cases} \frac{\sqrt{2}(u-uv)^2}{uv(1-u)(1-v)} = \frac{\sqrt{2}u(1-v)}{v(1-u)}, & \text{se } u < v, \\ \frac{\sqrt{2}(v-uv)^2}{uv(1-u)(1-v)} = \frac{\sqrt{2}v(1-u)}{u(1-v)}, & \text{se } u > v. \end{cases} \tag{F.8}$$

Assim, para $u < v$,

$$\begin{aligned}
 \frac{\partial \text{Cov}(u, v)}{\partial u} &= \frac{\sqrt{2}(1-v)}{v(1-u)^2} = \frac{\sqrt{2}u(1-v)}{v(1-u)} \cdot \frac{1}{u(1-u)} \\
 &= \frac{\text{Cov}(u, v)}{u(1-u)},
 \end{aligned} \tag{F.9}$$

enquanto, para $u > v$,

$$\begin{aligned}
 \frac{\partial \text{Cov}(u, v)}{\partial u} &= -\frac{\sqrt{2}v}{u^2(1-v)} = -\frac{\sqrt{2}v(1-u)}{u(1-v)} \cdot \frac{1}{u(1-u)} \\
 &= -\frac{\text{Cov}(u, v)}{u(1-u)}.
 \end{aligned} \tag{F.10}$$

Considere novamente o problema de autovalores na forma integral

$$\int_0^1 \text{Cov}(u, v)\gamma(v)dv = \lambda\gamma(u). \tag{F.11}$$

Derivando-a com respeito a u , obtém-se

$$\begin{aligned}\lambda\gamma'(u) &= \frac{\partial}{\partial u} \int_0^1 \text{Cov}(u, v)\gamma(v)dv = \int_0^1 \frac{\partial}{\partial u} \text{Cov}(u, v)\gamma(v)dv \\ &= \int_{0 < v < u} \frac{\partial}{\partial u} \text{Cov}(u, v)\gamma(v)dv + \int_{u < v < 1} \frac{\partial}{\partial u} \text{Cov}(u, v)\gamma(v)dv \\ &= - \int_0^u \frac{\text{Cov}(u, v)}{u(1-u)}\gamma(v)dv + \int_u^1 \frac{\text{Cov}(u, v)}{u(1-u)}\gamma(v)dv,\end{aligned}$$

de modo que,

$$\begin{aligned}\lambda u(1-u)\gamma'(u) &= - \int_0^u \text{Cov}(u, v)\gamma(v)dv + \int_u^1 \text{Cov}(u, v)\gamma(v)dv \\ &= - \int_0^u \text{Cov}(u, v)\gamma(v)dv + \int_0^1 \text{Cov}(u, v)\gamma(v)dv - \int_0^u \text{Cov}(u, v)\gamma(v)dv \\ &= -2 \int_0^u \text{Cov}(u, v)\gamma(v)dv + \int_0^1 \text{Cov}(u, v)\gamma(v)dv.\end{aligned}$$

Derivando novamente a expressão acima com respeito a u , e considerando que $\text{Cov}(u, u) = \sqrt{2}$, temos

$$\begin{aligned}\lambda(1-2u)\gamma'(u) + \lambda u(1-u)\gamma''(u) &= -2\text{Cov}(u, u)\gamma(u) + \lambda\gamma'(u) \\ &= -2\sqrt{2}\gamma(u) + \lambda\gamma'(u),\end{aligned}$$

ou seja, o problema na forma integral (F.2) é equivalente ao problema de autovalores na forma diferencial

$$u(1-u)\gamma''(u) - 2u\gamma'(u) + \frac{2\sqrt{2}}{\lambda}\gamma(u) = 0. \quad (\text{F.12})$$

A solução da Eq. (F.12) pode ser determinada pelo método de Frobenius. Considerando a série (de potências) de Frobenius ($j \geq 0$)

$$\gamma(u) = \sum_{k \geq 0} c_k u^{j+k}, \quad (\text{F.13})$$

temos as séries

$$\gamma'(u) = \sum_{k \geq 0} c_k (j+k) u^{j+k-1} \quad (\text{F.14})$$

$$\gamma''(u) = \sum_{k \geq 0} c_k (j+k)(j+k-1) u^{j+k-2}. \quad (\text{F.15})$$

Substituindo-se as séries (F.13), (F.14) e (F.15) em (F.12), obtemos

$$\begin{aligned}
 & u(1-u) \sum_{k \geq 0} c_k(j+k)(j+k-1)u^{j+k-2} - 2u \sum_{k \geq 0} c_k(j+k)u^{j+k-1} \\
 & \qquad \qquad \qquad + \frac{2\sqrt{2}}{\lambda} \sum_{k \geq 0} c_k u^{j+k} = \\
 & \sum_{k \geq 0} c_k(j+k)(j+k-1)u^{j+k-1} - \sum_{k \geq 0} c_k(j+k)(j+k-1)u^{j+k} \\
 & \qquad \qquad \qquad - 2 \sum_{k \geq 0} c_k(j+k)u^{j+k} + \frac{2\sqrt{2}}{\lambda} \sum_{k \geq 0} c_k u^{j+k} = \\
 & \sum_{k \geq 0} c_k(j+k)(j+k-1)u^{j+k-1} + \sum_{k \geq 0} c_k \left\{ \frac{2\sqrt{2}}{\lambda} - (j+k)(j+k+1) \right\} u^{j+k} = \\
 & \sum_{k^* \geq -1} c_{k^*+1}(j+k^*+1)(j+k^*)u^{j+k^*} + \sum_{k \geq 0} c_k \left\{ \frac{2\sqrt{2}}{\lambda} - (j+k)(j+k+1) \right\} u^{j+k} = \\
 & \sum_{k \geq -1} c_{k+1}(j+k+1)(j+k)u^{j+k} + \sum_{k \geq 0} c_k \left\{ \frac{2\sqrt{2}}{\lambda} - (j+k)(j+k+1) \right\} u^{j+k} = \\
 & c_0 j(j-1)u^{j-1} + \sum_{k \geq 0} \left\{ c_{k+1}(j+k+1)(j+k) + c_k \left[\frac{2\sqrt{2}}{\lambda} - (j+k)(j+k+1) \right] \right\} u^{j+k} = 0.
 \end{aligned} \tag{F.16}$$

A equação indicial se origina da menor potência u^{j-1} . Assim, para que a equação acima se anule, primeiramente é necessário que $c_0 j(j-1) = 0$; ou seja, $j = 0$ ou $j = 1$. Em seguida, os demais termos devem se anular, i.e.,

$$c_{k+1}(j+k+1)(j+k) + c_k \left[\frac{2\sqrt{2}}{\lambda} - (j+k)(j+k+1) \right] = 0,$$

ou, na forma recursiva,

$$c_{k+1} = \frac{(j+k)(j+k+1) - \frac{2\sqrt{2}}{\lambda}}{(j+k+1)(j+k)} \cdot c_k.$$

Assim, para $j = 0$ e $k \geq 1$,

$$c_{k+1} = \frac{k(k+1) - \frac{2\sqrt{2}}{\lambda}}{(k+1)k} \cdot c_k,$$

enquanto para $j = 1$ e $k \geq 0$,

$$c_{k+1} = \frac{(k+1)(k+2) - \frac{2\sqrt{2}}{\lambda}}{(k+1)(k+2)} \cdot c_k.$$

Pelo teste da razão, se $c_k \neq 0$, a série diverge, pois $\limsup_{k \rightarrow \infty} \left| \frac{c_{k+1}}{c_k} \right| = 1$. Por outro lado, pela relação de recorrência, se $c_k = 0$ para algum k , então $c_{k^*} = 0$ para todo $k^* > k$.

Assim, conclui-se que

$$k(k+1) - \frac{2\sqrt{2}}{\lambda} = 0,$$

ou seja, para $k \geq 1$, os autovalores são

$$\lambda_k = \frac{2\sqrt{2}}{k(k+1)}\eta, \quad (\text{F.17})$$

em que η é uma constante de normalização. Para determiná-la, considera-se a restrição $\sum_{k \geq 1} \lambda_k = \langle B_{LR} \rangle$ e a propriedade (A.25) dos coeficientes binomiais.

Para o caso bivariado ($m = 2$),

$$\begin{aligned} \sum_{j,k \geq 1} \lambda_{j,k} &= \langle B_{LR} \rangle = 1 = (2\sqrt{2})^2 \eta^2 \sum_{j \geq 1} \frac{1}{j(j+1)} \sum_{k \geq 1} \frac{1}{k(k+1)} \\ &= 8\eta^2 \sum_{j \geq 2} \frac{1}{j(j-1)} \sum_{k \geq 2} \frac{1}{k(k-1)} \\ &= 8\eta^2 \sum_{j \geq 2} \frac{1}{2 \binom{j}{2}} \sum_{k \geq 2} \frac{1}{2 \binom{k}{2}} \\ &= 2\eta^2 \sum_{j \geq 2} \frac{1}{\binom{j}{2}} \sum_{k \geq 2} \frac{1}{\binom{k}{2}} = 8\eta^2. \end{aligned}$$

Logo, para este caso, $\eta = \frac{1}{2\sqrt{2}}$, de modo que $\lambda_{j,k} = \frac{(2\sqrt{2})^2 \eta^2}{j(j+1)k(k+1)} = \frac{1}{j(j+1)k(k+1)}$.

Analogamente, para $m = 3$,

$$\begin{aligned} \sum_{j,k,l \geq 1} \lambda_{j,k,l} &= 4 \\ &= (2\sqrt{2})^3 \eta^3 \sum_{j \geq 1} \frac{1}{j(j+1)} \sum_{k \geq 1} \frac{1}{k(k+1)} \sum_{l \geq 1} \frac{1}{l(l+1)} \\ &= (2\sqrt{2})^3 \eta^3. \end{aligned}$$

Logo, $\lambda_{j,k,l} = \frac{(2\sqrt{2})^3 \eta^3}{j(j+1)k(k+1)l(l+1)} = \frac{4}{j(j+1)k(k+1)l(l+1)}$. Assim, temos para o caso multidimensional em geral,

$$\lambda_{j_1, \dots, j_m} = \frac{\langle B_{LR} \rangle}{\prod_{i=1}^m j_i(j_i+1)}. \quad (\text{F.18})$$

◇

Referências Bibliográficas

- [1] A. Agresti. *An Introduction to Categorical Data Analysis*. John Wiley & Sons, 2007.
- [2] D. Applebaum. Lévy processes — from probability to finance and quantum groups. *Notices of the AMS*, 51(11):1336–1347, 2004.
- [3] D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2004.
- [4] R. W. Arad. Parameter estimation for symmetric stable distribution. *International Economic Review*, 21(1):209–220, 1980.
- [5] M. Ausloos and K. Ivanova. Introducing False EUR and False EUR exchange rates. *Physica A*, 286:353–366, 2000.
- [6] N. K. Bakirov, M. L. Rizzo, and G. J. Székely. A multivariate nonparametric test of independence. *Journal of Multivariate Analysis*, 97:1742–1756, 2006.
- [7] R. Beran, M. Bilodeau, and P. L. de Micheaux. Nonparametric tests of independence between random vectors. *Journal of Multivariate Analysis*, 98:1805–1824, 2007.
- [8] M. Bilodeau and P. L. de Micheaux. A multivariate empirical characteristic function test of independence with Normal marginals. *Journal of Multivariate Analysis*, 95:345–369, 2005.
- [9] J. R. Blum, J. Kiefer, and M. Rosenblatt. Distribution free tests of independence based on the sample distribution function. *The Annals of Mathematical Statistics*, 32:485–498, 1961.

- [10] O. Brandouy, J.-P. Delahaye, L. Ma, and H. Zenil. Algorithmic complexity of financial motions. Technical report, Department of Economics, University of Trento, 2012.
- [11] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer, 2nd edition, 2006.
- [12] A. Carbone, G. Castelli, and H. E. Stanley. Time-dependent Hurst exponent in financial time series. *Physica A*, 344:267–271, 2004.
- [13] S. Csörgő. Limit behaviour of the empirical characteristic function. *The Annals of Probability*, 9(1):130–144, 1981.
- [14] B. M. Tabak D. O. Cajueiro. Possible causes of long-range dependence in the Brazilian stock market. *The Annals of Mathematical Statistics*, 345:635–645, 2005.
- [15] S. Da Silva. Chaos and exchange rates. *Estudos Empresariais*, 6(1):9–15, 2001.
- [16] S. Da Silva, R. Matsushita, I. Gléria, and A. Figueiredo. Hurst exponents, power laws and efficiency in the Brazilian foreign exchange market. *Economics Bulletin*, 7(1):1–11, 2007.
- [17] S. Da Silva, R. Matsushita, I. Gléria, A. Figueiredo, and P. Rathie. International finance, Lévy distributions, and the econophysics of exchange rates. *Communications in Nonlinear Science and Numerical Simulation*, 10:365–393, 2005.
- [18] B. De Finetti. *Theory of Probability*, volume 1. John Wiley & Sons, 1974.
- [19] B. De Finetti. *Theory of Probability*, volume 2. John Wiley & Sons, 1975.
- [20] M. A. Delgado. Testing serial independence using the sample distribution function. *Journal of Time Series Analysis*, 17:271–286, 1996.
- [21] C. Dorea, C. Otiniano, R. Matsushita, and P. Rathie. Lévy flight approximations for scaled transformations of random walks. *Computational Statistics & Data Analysis*, 51:6343–6354, 2007.

- [22] W. H. DuMouchel. Stable distributions in statistical inference: 2. information from stably distributed samples. *Journal of the American Statistical Association*, 70:386–393, 1975.
- [23] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall, 1993.
- [24] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. John Wiley & Sons, 2nd edition, 1971.
- [25] A. Feuerverger. A consistent test for bivariate dependence. *International Statistical Review*, 61(3):419–433, 1993.
- [26] A. Feuerverger and P. McDunnough. On some Fourier methods for inference. *Journal of the American Statistical Association*, 76(374):379–387, 1981.
- [27] A. Feuerverger and P. McDunnough. On the efficiency of empirical characteristic function procedures. *Journal of Royal Statistical Society. Series B (Methodological)*, 43(1):20–27, 1981.
- [28] A. Feuerverger and R. A. Mureika. The empirical characteristic function and its applications. *The Annals of Statistics*, 5(1):88–97, 1977.
- [29] A. Figueiredo, M. T. de Castro, S. Da Silva, and I. Gleria. Jump diffusion models and the evolution of financial prices. *Physics Letters A*, 375(34):3055–3061, 2011.
- [30] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. Autocorrelation as a source of truncated Lévy flights in foreign exchange rates. *Physica A*, 323:601–625, 2003.
- [31] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. On the origins of truncated Lévy flights. *Physics Letters A*, 315:51–60, 2003.
- [32] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. Autocorrelation and the sum of stochastic variables. *Physics Letters A*, 326:166–170, 2004.

- [33] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. Lévy flights, autocorrelation, and slow convergence. *Physica A*, 337:369–383, 2004.
- [34] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. Financial volatility and independent and identically distributed variables. *Physica A*, 346:484–498, 2005.
- [35] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. Nonidentically distributed variables and nonlinear autocorrelation. *Physica A*, 363:171–180, 2006.
- [36] A. Figueiredo, I. Gléria, R. Matsushita, and S. Da Silva. The Lévy sections theorem revisited. *Journal of Physics A*, 40:5783–5794, 2007.
- [37] A. Figueiredo, R. Matsushita, S. Da Silva, M. Serva, G. M. Viswanathan, C. Nascimento, and I. Gléria. The Lévy sections theorem: An application to econophysics. *Physica A*, 386:756–759, 2007.
- [38] P. H. Franses and D. van Dijk. *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press, 2000.
- [39] C. Genest, J. F. Quessy, and B. Rémillard. Local efficiency of a Cramér-von Mises test of independence. *Journal of Multivariate Analysis*, 97:274–294, 2006.
- [40] K. Ghoudi, R. J. Kulperger, and B. Rémillard. A nonparametric test of serial independence for time series and residuals. *Journal of Multivariate Analysis*, 79:191–218, 2001.
- [41] R. Giglio, S. Da Silva, I. Gléria, A. Ranciaro, R. Matsushita, and A. Figueiredo. Efficiency of financial markets and algorithmic complexity. *Journal of Physics: Conference Series*, 246:012032, 2010.
- [42] R. Giglio, R. Matsushita, and S. Da Silva. The relative efficiency of stockmarkets. *Economics Bulletin*, 7(6):1–12, 2008.

- [43] R. Giglio, R. Matsushita, A. Figueiredo, I. Gléria, and S. Da Silva. Algorithmic complexity theory and the relative efficiency of financial markets. *Europhysics Letters*, 84:48005, 2008.
- [44] C. G. Gilmore. An examination of nonlinear dependence in exchange rates, using recent methods from Chaos Theory. *Global Finance Journal*, 12:139–151, 2001.
- [45] I. Gléria, A. Figueiredo, R. Matsushita, P. Rathie, and S. Da Silva. Exponentially damped Lévy flights, multiscaling and slow convergence in stockmarkets. *Physica A*, 342:200–206, 2004.
- [46] I. Gléria, R. Matsushita, and S. Da Silva. Scaling power laws in the Sao Paulo Stock Exchange. *Economics Bulletin*, 7(3):1–12, 2002.
- [47] B. V. Gnedenko. *The Theory of Probability*. Mir Publishers, 1973.
- [48] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, 1954.
- [49] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier, 7th edition, 2007.
- [50] H. M. Gupta and J. R. Campanha. The gradually truncated Lévy flight for systems with power-law distributions. *Physica A*, 268:231–239, 1999.
- [51] H. M. Gupta and J. R. Campanha. The gradually truncated Lévy flight: stochastic process for complex systems. *Physica A*, 275:531–543, 2000.
- [52] P. Hall and A. H. Welsh. A test for normality based on the empirical characteristic function. *Biometrika*, 70(2):485–489, 1983.
- [53] W. Hardle, J. S. Marron, and M. P. Wand. Bandwidth choice for density derivatives. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):223–232, 1990.

- [54] W. Hoeffding. A nonparametric test of independence. *The Annals of Mathematical Statistics*, 19:546–557, 1948.
- [55] Y. Hong. Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *Journal of the American Statistical Association*, 94(448):1201–1220, 1999.
- [56] M. Hušková and S. G. Meintanis. Testing procedures based on the empirical characteristic functions I: goodness-of-fit, testing for symmetry and independence. *Tatra Mountains Mathematical Publications*, 39:225–233, 2008.
- [57] I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff Publishing, 1971.
- [58] A. Janicki and A. Weron. *Simulation and Chaotic Behavior of α -Stable Stochastic Processes*. Marcel Dekker, 1994.
- [59] M. Kac. On some connections between probability theory and differential and integral equations. *Proceedings of the Second Berkeley Symposium of Mathematical Statistics and Probability*, pages 180–215, 1951.
- [60] B. Klar and S. G. Meintanis. Tests for Normal mixtures based on the empirical characteristic function. *Computational Statistics & Data Analysis*, 49:227–242, 2005.
- [61] I. Koponen. Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physical Review E*, 52(1):1197–1199, 1995.
- [62] V. Kulkarni and N. Deo. Correlation and volatility in an Indian stock market: a random matrix approach. *The European Physical Journal B*, 60:101–109, 2007.
- [63] S. Kullback. *Information Theory and Statistics*. Dover Publications, 1968.
- [64] P. Lévy. Eléments aléatoires. In D. Dugué, editor, *Oeuvres de Paul Lévy*, volume 3. Gauthier-Villars, 1976.

- [65] L. Liu and J. Wan. A study of correlations between crude oil spot and futures markets: a rolling sample test. *Physica A*, 390:3754–3766, 2011.
- [66] E. Lukacs. *Characteristic Functions*. Charles Griffin & Co, 2nd edition, 1970.
- [67] E. Lukacs and O. Szász. On analytic characteristic functions. *Pacific Journal of Mathematics*, 2(4):615–625, 1952.
- [68] S. Luo, Z. Wang, and Q. Zhang. An inequality for characteristic functions and its applications to uncertainty relations and the quantum Zeno effect. *Journal of Physics A*, 35(28):5935–5941, 2002.
- [69] B. B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.
- [70] B. B. Mandelbrot. The variation of some other speculative prices. *Journal of Business*, 40:393–413, 1967.
- [71] B. B. Mandelbrot. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk (Selecta Volume E)*. Springer-Verlag, 1997.
- [72] B. B. Mandelbrot. Heavy tails in finance for independent or multifractal price increments. In S. T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*, volume 3, pages 4–34. Elsevier, 2003.
- [73] R. N. Mantegna and H. E. Stanley. Scaling behavior in the dynamics of an economic index. *Nature*, 376:46–49, 1995.
- [74] R. N. Mantegna and H. E. Stanley. Modeling of financial data: comparison of the truncated Lévy flight and the ARCH(1) and GARCH(1,1) processes. *Physica A*, 254:77–84, 1998.
- [75] R. N. Mantegna and H. E. Stanley. *An Introduction to Econophysics*. Cambridge University Press, 2000.

- [76] K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, 1979.
- [77] M. Markatou, J. L. Horowitz, and R. V. Lenth. Robust scale estimation based on the empirical characteristic function. *Statistics & Probability Letters*, 25:185–192, 1995.
- [78] R. Matsushita and S. Da Silva. A log-periodic fit for the flash crash of May 6, 2010. *Economics Bulletin*, 31(2):1772–1779, 2011.
- [79] R. Matsushita, S. Da Silva, A. Figueiredo, and I. Gléria. Log-periodic crashes revisited. *Physica A*, 364(3):331–335, 2006.
- [80] R. Matsushita, A. Figueiredo, and S. Da Silva. A suggested statistical test for measuring bivariate nonlinear dependence. *Physica A*, 391:4891–4898, 2012.
- [81] R. Matsushita, I. Gléria, A. Figueiredo, and S. Da Silva. Fractal structure in the Chinese yuan/us dollar rate. *Economics Bulletin*, 7(2):1–13, 2003.
- [82] R. Matsushita, I. Gléria, A. Figueiredo, and S. Da Silva. Are pound and euro the same currency? *Physics Letters A*, 368:173–180, 2007.
- [83] R. Matsushita, I. Gléria, A. Figueiredo, and S. Da Silva. The Chinese chaos game. *Physica A*, 378:427–442, 2007.
- [84] R. Matsushita, I. Gléria, A. Figueiredo, P. Rathie, and S. Da Silva. Exponentially damped Lévy flights, multiscaling and exchange rates. *Physica A*, 333:353–369, 2004.
- [85] R. Matsushita, P. Rathie, and S. Da Silva. Exponentially damped Lévy flights. *Physica A*, 326:544–555, 2003.
- [86] J. H. McCulloch. Numerical approximation of the symmetric stable distribution and density. Technical report, Department of Economics, Ohio State University, 1994.

- [87] J. H. McCulloch. Financial applications of stable distributions. In G. S. Maddala and C. R. Rao, editors, *Handbook of Statistics*, volume 14, pages 393–425. Elsevier, 1996.
- [88] L. C. Miranda and R. Riera. Truncated Lévy walks and an emerging market economic index. *Physica A*, 297:509–520, 2001.
- [89] A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 3rd edition, 1987.
- [90] U. Müller, M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morge-negg. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking and Finance*, 14:1189–1208, 2001.
- [91] T. Nakamura and M. Small. Correlation structures in short-term variabilities of stock indices and exchange rates. *Physica A*, 383:96–101, 2007.
- [92] H. Nakao. Multi-scaling properties of truncated Lévy flights. *Physics Letters A*, 266:282–289, 2000.
- [93] R. B. Nelsen. Correlation, regression lines, and moments of inertia. *The American Statistician*, 52(4):343–345, 1998.
- [94] J. P. Nolan. Maximum likelihood estimation and diagnostics for stable distributions. In S.I. Resnick O.E. Barndorff-Nielsen, T. Mikosch, editor, *Lévy Processes: Theory and Applications*, pages 379–400. Birkhäuser, 2001.
- [95] A. S. Paulson, E. W. Holcomb, and R. A. Leitch. The estimation of the parameters of the stable laws. *Biometrika (1975)*, 62, 1, p. 163, 62(1):163–170, 1975.
- [96] A. M. M. Polito, A. Figueiredo, T. M. da Rocha Filho, F. V. Prudente, and L. S. Costa. The characteristic function method applied to molecular dynamics of inelas-tic granular gases. *Physica A*, 373:392–416, 2007.

- [97] S. J. Press. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, 67(340):842–846, 1972.
- [98] C. R. Rao. *Linear Statistical Inference and Its Applications*. John Wiley & Sons, 2nd edition, 2002.
- [99] P. Rathie, C. Dorea, and R. Matsushita. Lévy distribution, H-function and applications to currency data. *Proceeding of the Seventh International Conference of the Society for Special Functions and their Applications (SSFA)*, 7:17–26, 2006.
- [100] B. Rémillard and R. Theodorescu. Estimation based on the empirical characteristic function. In N. Balakrishnan, I. A. Ibragimov, and V. B. Nevzorov, editors, *Asymptotic Methods in Probability and Statistics with Applications*, pages 435–450. Birkhäuser, 2001.
- [101] G. G. Roussas. *A Course in Mathematical Statistics*. Academic Press, 2nd edition, 1997.
- [102] R. Y. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley & Sons, 1981.
- [103] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall/CRC, 1994.
- [104] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.
- [105] W. Schoutens. *Lévy Processes in Finance*. John Wiley & Sons, 2003.
- [106] J. Shao and D. Tu. *The Jackknife and Bootstrap*. Springer, 1995.
- [107] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.
- [108] J. S. Simonoff. *Smoothing Methods in Statistics*. Springer, 1996.

- [109] J. A. Skjeltorp. Scaling in the Norwegian stock market. *Physica A*, 283:486–525, 2001.
- [110] D. Sornette and A. Johansen. Significance of log-periodic precursors to financial crashes. *Quantitative Finance*, 1:452–471, 2001.
- [111] D. Sornette and C. Vanneste. Dynamics and memory effects in rupture of thermal fuse networks. *Physical Review Letters*, 68:612–615, 1992.
- [112] D. Sornette and W.X. Zhou. The US 2000 - 2002 market descent: how much longer and deeper? *Quantitative Finance*, 2:468–481, 2002.
- [113] J. C. Sprott. *Chaos and Time-Series Analysis*. Oxford University Press, 2004.
- [114] H. E. Stanley, L. A. N. Amaral, X. Gabaix, P. Gopikrishnan, and V. Plerou. Similarities and differences between physics and economics. *Physica A*, 299:1–15, 2001.
- [115] T. M. Stoker. Smoothing bias in density derivative estimation. *Journal of the American Statistical Association*, 88(423):855–863, 1993.
- [116] B. W. Stuck. Distinguishing stable probability measures. part I: discrete time. *Bell System Technical Journal*, 55:1125–1182, 1976.
- [117] E. Taufer and N. Leonenko. Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes. *Journal of Statistical Planning and Inference*, 139:3050–3063, 2009.
- [118] H. Tong. *Non-linear Time Series: a Dynamical System Approach*. Oxford Science Publications, 1999.
- [119] R. S. Tsay. *Analysis of Financial Time Series*. John Wiley & Sons, 2nd edition, 2005.
- [120] N. G. Ushakov. *Selected Topics in Characteristic Functions*. VSP, 1999.

- [121] N. G. van Kampen. An equation for the characteristic function of a Markov process and its application to a Langevin process. *Physics Letters A*, 76:104–106, 1980.
- [122] A. Weron and R. Weron. Computer simulation of Lévy α -stable variables and processes. In P. Garbaczewski, M. Wolf, and A. Weron, editors, *Chaos — The Interplay Between Stochastic and Deterministic Behaviour*, volume 457, pages 379–392. Springer, 1995.
- [123] G. E. Wilding and G. S. Mudholkar. Empirical approximations for Hoeffding’s test of bivariate independence using two Weibull extensions. *Statistical Methodology*, 5:160–170, 2008.
- [124] L. Xu, P. C. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley. Quantifying signals with power-law correlations: a comparative study of detrended fluctuation analysis and detrended moving average techniques. *Physical Review E*, 71(051101):1–14, 2005.
- [125] J. Yu. Empirical characteristic function estimation and its applications. *Econometric Reviews*, 23(2):93–123, 2004.