



Anderson Alexander Gomes Cortines

**Fatores determinísticos e estocásticos das
grandezas observáveis financeiras**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Física do
Departamento de Física da PUC-Rio como requisito parcial para
obtenção do título de Doutor em Física

Orientadoras : Prof. Celia Beatriz Anteneodo
Prof. Rosane Riera Freire

Rio de Janeiro
março de 2009

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



Anderson Alexander Gomes Cortines

**Fatores determinísticos e estocásticos das
grandezas observáveis financeiras**

Tese apresentada ao Programa de Pós-graduação em Física do Departamento de Física do Centro Técnico Científico da PUC-Rio como requisito parcial para obtenção do título de Doutor em Física. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Celia Beatriz Anteneodo
Orientadora
Departamento de Física — PUC-Rio

Prof. Rosane Riera Freire
Co-Orientadora
Departamento de Física — PUC-Rio

Prof. Giovani Lopes Vasconcelos
Departamento de Física - UFPE

Prof. Jürgen Fritz Stilck
Instituto de Física - UFF

Prof. Luca Moriconi
Instituto de Física - UFRJ

Prof. Helio Cortes Vieira Lopes
Departamento de Matemática - PUC-Rio

Prof. José Eugênio Leal
Coordenador Setorial do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 20 de março de 2009

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Anderson Alexander Gomes Cortines

Graduou-se em Física na UFRRJ (Universidade Federal Rural do Rio de Janeiro) em 1999. Mestre em Física da Matéria Condensada Teórica com ênfase em Econofísica pela PUC-Rio em 2005. Pesquisa o comportamento do mercado financeiro, em especial a Bolsa de Valores de São Paulo.

Ficha Catalográfica

Cortines, Anderson Alexander Gomes

Fatores determinísticos e estocásticos das grandezas observáveis financeiras / Anderson Alexander Gomes Cortines; Orientadoras: Célia Beatriz Anteneodo e Rosane Riera Freire — Rio de Janeiro : PUC-Rio, Departamento de Física, 2009.

120 f.: il. (color.); 30 cm

1. Tese (Doutorado em Física) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Física, Rio de Janeiro, 2009.

Inclui referências bibliográficas.

1. Física – Tese. 2. Econofísica. 3. Equação de Fokker-Planck. 4. Equação de Langevin. 5. Índices de Mercado. 6. Volume de negociação. 7. Misturas Estatísticas. I. Anteneodo, Célia Beatriz. II. Freire, Rosane Riera. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Física. IV. Título.

CDD: 530

Agradecimentos

Às minhas orientadoras, Celia Anteneodo e Rosane Riera Freire, pela paciência, atenção, simpatia e estímulo para a realização deste trabalho;

À minha esposa Lílian, amor de minha vida, pelo apoio em todos os momentos;

À minha mãe Albina que me deu a sabedoria;

Às minhas filhas Rafaela e Mariana que me deram a inspiração;

À minha família pelo lazer salutar proporcionado nos finais de semana;

Ao CNPq e à PUC-Rio pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado;

Em especial à equipe de professores da PUC-Rio pelos ensinamentos, profissionalismo, integridade e atenção dispensada;

Aos professores que integram a banca de qualificação pela generosidade em participar desta defesa de tese;

Aos colegas de curso pela convivência e companheirismo, e que me fizeram adorar esse lugar, em especial aos residentes ou viajantes da sala 656 da última meia década: Vernek, Marcelo Apel, Diego, Petche, Felipe, Adriano, Laércio, Thiesen, Cintia, Wilson, Fábio, Thiago, Alexander, Mary, Mariana, Christian, Paulina, Jefferson, Lucas, Clara, Ricardo, Nei e Fernando Vilhena;

Ao pessoal do departamento de Física da PUC-Rio pelo constante apoio, em particular à Márcia, Giza, Majô e ao Julinho.

Resumo

Cortines, Anderson Alexander Gomes; Anteneodo, Celia Beatriz; Freire, Rosane Riera. **Fatores determinísticos e estocásticos das grandezas observáveis financeiras**. Rio de Janeiro, 2009. 120p. Tese de Doutorado — Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

As flutuações de preços e de outras grandezas observáveis nos mercados financeiros apresentam comportamentos não triviais, tais como longas correlações temporais, não gaussianidade ou leis de escala, cuja origem não é ainda bem compreendida. Neste trabalho investigamos possíveis mecanismos determinísticos e estocásticos responsáveis pelas distribuições de probabilidade anômalas observadas para os índices de mercado e para os volumes de ações comercializadas. No primeiro caso, consideramos a expansão de Kramers-Moyal como ponto de partida para descrever a evolução das densidades de probabilidade. Para a modelagem dos volumes negociados, consideramos misturas estatísticas que surgem das flutuações em escalas longas dos parâmetros internos que descrevem a dinâmica em escalas mais curtas. Este estudo provê uma demonstração consistente, a partir de análise empírica de séries temporais reais, de como funções de densidade de probabilidade com caudas em lei de potência podem emergir através de mecanismos diversos, tais como processos estocásticos com flutuações aditivo-multiplicativas, ou como resultado de misturas estatísticas.

Palavras-chave

Econofísica. Equação de Fokker-Planck. Equação de Langevin. Índices de Mercado. Volume de negociação. Misturas Estatísticas.

Abstract

Cortines, Anderson Alexander Gomes; Anteneodo, Celia Beatriz; Freire, Rosane Riera. **Deterministic and stochastic factors of financial observables**. Rio de Janeiro, 2009. 120p. Tese de Doutoramento — Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

The fluctuations of prices and other observables in financial markets have non-trivial behaviors, such as long temporal correlations, non-Gaussianity or scaling laws, whose origin is not well understood so far. In this work we have investigated possible deterministic and stochastic mechanisms responsible for the anomalous probability distributions observed for market indexes and volumes of traded shares. In the first case, we consider the Kramers-Moyal expansion as a starting point to describe the evolution of probability densities. For the modelling of trading volumes, we consider the mixed statistics that emerges from the long-scale fluctuations of inner parameters that describe the dynamics on shorter scales. This study provides a consistent demonstration, from empirical analysis of real time series, on how probability density functions with power laws tails may emerge through various mechanisms, such as stochastic processes with additive-multiplicative fluctuations or as a result of mixed statistics.

Keywords

Econophysics. Fokker Planck equation. Langevin equation. Market Index. Traded Volume. Statistical mixture.

Sumário

1	Introdução	16
1.1	Grandezas Observáveis Financeiras	17
1.2	Resultados na Literatura	21
1.3	Motivação e Objetivos	24
2	Referencial Teórico	27
2.1	Variáveis Aleatórias e Distribuições de Probabilidade	27
2.1.1	FDPs usuais	28
2.1.2	FDPs generalizadas	29
2.2	Processos Estocásticos	30
2.2.1	Classes de processos estocásticos	31
2.2.2	Processos markovianos	33
2.2.3	Expansão de Kramers-Moyal	34
2.2.4	Equação de Fokker-Planck	35
2.2.5	Equação de Itô-Langevin	36
2.2.6	Processos de Poisson	37
2.3	Misturas Estatísticas	38
3	Análise do Índice Acionário Brasileiro	39
3.1	Introdução	39
3.2	Análise de Kramers-Moyal	40
3.3	Dependência temporal dos parâmetros	44
3.4	Evolução temporal	47
3.5	Considerações finais	52
4	Análise Comparativa dos Índices de Mercados Mundiais	53
4.1	Introdução	53
4.2	Índices de Mercados Mundiais	53
4.3	Análise Preliminar	55
4.4	Análise de Kramers-Moyal	58
4.5	Análise Comparativa	61
4.6	Mais exemplos do cálculo dos coeficientes de KM para os mercados mundiais	66
4.7	Considerações finais	68
5	Análise do Volume de Ações Negociadas	69
5.1	Introdução	69
5.2	Estatística descritiva dos volumes negociados	69
5.3	Modelagem	72
5.4	Considerações finais	79
6	Reconstrução de Séries Financeiras	80
6.1	Introdução	80
6.2	Cálculo dos coeficientes de KM	80

6.3	Correção dos coeficientes de KM	82
6.4	Conclusões preliminares	87
7	Conclusões	89
7.1	Índice do Mercado Brasileiro	89
7.2	Índices de Mercados Mundiais	90
7.3	Volumes Negociados no Mercado Brasileiro	91
7.4	Conclusões Gerais	92
7.5	Perspectivas futuras	93
A	Apêndices	100
A.1	Considerações Teóricas Complementares	100
A.1.1	Sobre a solução invariante da EFP	100
A.1.2	Influência do reescalonamento no cálculo dos coeficientes de KM	102
A.1.3	Expansão de Itô-Taylor para coeficientes de KM[61]	103
A.1.4	Testes estatísticos de estacionaridade	107
A.2	Análises Complementares sobre o IBOVESPA	109
A.2.1	Limite para escalas curtas	109
A.2.2	Informações sobre as séries de dados utilizadas	110
A.2.3	Comparação dos resultados para retornos logarítmicos e incrementos	113
A.3	Séries Temporais dos Índices dos Mercados Mundiais	114
A.4	Análise Estatística da Quantidade de Transações	117
A.5	Lista de abreviaturas	120

Lista de figuras

- 1.1 Série do IBOVESPA entre 1991 e 2006 (período analisado nesta tese) com valores históricos e com preços deflacionados pelo IGP-DI conforme indicado em legenda. 19
- 1.2 Série dos retornos diários do IBOVESPA entre 1991 e 2006. 20
- 1.3 Série temporal dos volumes negociados na BOVESPA no período de 3/01/05 a 13/09/07. Os valores representam o volume negociado em intervalos $\Delta t = 30$ minutos, normalizados pelo valor médio em todo o período ($v = V/\langle V \rangle$). 21
- 1.4 Comparação dos histogramas de retornos normalizados para $\Delta t = 1$ minuto do IBOVESPA (símbolos), de novembro de 2002 a julho de 2006, com a distribuição gaussiana (linha cheia), em gráfico semi-logarítmico. 23
- 1.5 Distribuições empíricas (símbolos) e q -gaussianas (linhas cheias) para retornos normalizados das 10 empresas com maior volume de negociação na (a) NYSE e (b) Nasdaq em 2001. A linha pontilhada representa a distribuição gaussiana [9]. 24
- 1.6 Distribuições empíricas (pontos) e ajustes q -exponenciais (linhas) para volumes normalizados das 10 ações de maior volume na NYSE em 2001 [9]. 25

- 3.1 Representação das séries diária e intradiárias do IBOVESPA analisadas. Série intradiária de 15 minutos compreende período entre 21/01/98 e 31/03/03 (em vermelho) e série intradiária de 30 segundos relativa ao período entre 01/11/02 e 19/07/06 (em azul). Todas as séries estão deflacionadas pelo índice diário do IGP-DI. 40
- 3.2 Curvas de nível típicas da FDP conjunta para retornos obtida a partir de dados diários do IBOVESPA para $\Delta t_1 = 16$ dias e $\Delta t_2 = 8$ dias. Níveis das curvas correspondem a $\log_{10} = -1,5; -2,0; -2,5; -3,0; -3,5$ do centro até a borda. 41
- 3.3 Distribuição conjunta típica para retornos obtida para $\tau_2 = 6,88$ ($\Delta t_2 = 120$ min) e $\tau_1 = 6,18$ ($\Delta t_1 = 195$ min). 42
- 3.4 Corte de $P(r_2, \tau_2 | r_1, \tau_1)$ em $r_1 = 3\sigma_{\tau_2}$ obtido para $\tau_2 = 11,75$ ($\Delta t_2 = 4$ min), $\tau' = 10,75$ ($\Delta t' = 8$ min) e $\tau_1 = 9,75$ ($\Delta t_1 = 16$ min). Os símbolos cheios correspondem à FDP condicional calculada diretamente e os símbolos abertos à FDP integrada. Gráfico inserido: representação log-linear dos mesmos dados. 43
- 3.5 Coeficientes $\tilde{D}^{(1)}$ (a) e $\tilde{D}^{(2)}$ (b) em função de r , calculados em $\Delta t_2 = 109$ min e $\Delta t_1 = 125$ min (logo $\tau = 6,9$ e $\Delta\tau = 0,2$). As linhas cheias correspondem aos intervalos utilizados nos ajustes. 43
- 3.6 Coeficiente $\tilde{D}^{(4)}$ em função de r , calculado em $\tau = 8,35$ e diferentes valores de $\Delta\tau$, indicados na figura. Linhas sólidas correspondem a ajustes polinomiais de quarta ordem nos dados. 44

- 3.7 Dependência dos parâmetros a_1 e a_0 do coeficiente de tendência $D^{(1)}(r, \tau)$ com a escala temporal τ . A análise foi realizada com dados de séries diárias (círculos), de 15 minutos (quadrados) e de 30 segundos (triângulos). 45
- 3.8 Dependência dos parâmetros b_2, b_1 e b_0 do coeficiente de difusão $D^{(2)}(r, \tau)$ com a escala temporal τ . Símbolos como na Fig. 3.7. Em (c), incluímos para comparação $\sigma^2(\tau) = \sigma_\tau^2 / \sigma_{32}^2$ (símbolos pequenos) junto com seu comportamento assintótico (linhas finas) para escalas de tempo longas e curtas como previsto pela Eq. (3.14), com $\gamma = 1$ e 1, 17, respectivamente. 46
- 3.9 FDPs dos retornos normalizados. Comparação entre as soluções numéricas da EFP (3.9) (linhas cheias) e os dados empíricos (símbolos). As escalas temporais correspondem de $\tau = -2$ a $\tau \approx 15$. As FDPs estão deslocadas verticalmente, para melhor visualização. A condição inicial foi um ajuste gaussiano aos dados para $\tau = -2$ ($\Delta t = 128$ dias). As escalas τ apresentadas são informadas na tabela A.5 (apêndice). 47
- 3.10 FDPs empíricas dos retornos normalizados (símbolos), soluções numéricas da EFP (3.9) (linhas pretas) e ajustes q -gaussianos aos dados empíricos (linhas vermelhas). 50
- 3.11 Dependência do expoente μ em relação às escalas temporais τ . Símbolos correspondem ao ajuste da Eq. (3.12) aos histogramas empíricos, linhas pontilhadas representam os valores assintóticos dados pela Eq. (3.13). 51
- 4.1 Séries de retorno de preço dos índices mundiais apresentados em legenda. Linha vertical representa o início do ano de 1997. As séries foram deslocadas verticalmente para melhor visualização. 56
- 4.2 Séries de retornos diários dos índices de mercado mundiais (apresentados ao lado do respectivo índice) no período utilizado nesta tese. As séries foram deslocadas verticalmente para melhor visualização. 57
- 4.3 Cortes de nível através de $P(r_2, \tau_2 | r_1, \tau_1)$ em $r_1 = -2, 0, 0,5$ e $3,0$ (da esquerda para a direita) para $\tau_2 = 0, 24, \tau' = 0, 14$ e $\tau_1 = 0, 04$, para o índice de Cingapura STI. Os histogramas foram deslocados horizontalmente para melhor visualização. Os símbolos preenchidos correspondem às FDPs condicionais calculadas diretamente e os símbolos abertos às FDPs integradas de acordo com a Eq. (3.2). 57
- 4.4 Coeficientes $\tilde{D}^{(1)}$ (a), $\tilde{D}^{(2)}$ (b) e $\tilde{D}^{(4)}$ (c), em função de r , calculados em $\tau = 1, 2$ (escala temporal semanal) e $\Delta\tau = 0, 2$ para o índice alemão DAX. As linhas cheias correspondem às curvas ajustadas e as mais espessas aos intervalos utilizados para os ajustes. 58
- 4.5 Dependência dos parâmetros \tilde{a}_1 (a), \tilde{b}_2 (b) e \tilde{b}_0 (c) nas escalas temporais τ , para diferentes valores de $\Delta\tau$ indicados em (c), para o índice Mexicano IPC. As linhas cheias correspondem aos intervalos utilizados para a estimativa dos parâmetros a_1^s, b_2^s, A e γ , como explicado no texto, usando o resultado para $\Delta\tau = 0, 2$. 59

- 4.6 Parâmetro a_1^s medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos. O erro de a_1^s é de cerca de 10%. A linha cheia corresponde ao valor médio 0,57. 61
- 4.7 Parâmetro b_2^s medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos. 62
- 4.8 Inclinação γ (a) e nível A (b) do parâmetro b_0 medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos. 62
- 4.9 Representação dos mercados no plano de parâmetros (γ, A) , de acordo com os rótulos numéricos da tabela 4.1. Os símbolos mais escuros representam os mercados desenvolvidos. Os erros são de cerca de 5%. As linhas pontilhadas representam uma regressão linear como guia visual. 63
- 4.10 Representação dos mercados de acordo com a medida de $2H$ em [49] e γ (este trabalho). Símbolos são como os da Fig. 4.9. O erro de γ é de cerca de 5%. As fronteiras pontilhadas ($\gamma = 1$, $2H = 1$ e $\gamma = 2H$) estão desenhadas para comparação. 64
- 4.11 FDPs padronizadas para retornos diários para o índice argentino Merval (a) e índice americano Dow Jones (b). Os símbolos correspondem aos histogramas empíricos e as linhas sólidas aos ajustes de mínimos quadrados utilizando a forma padronizada da Eq. (2.10). 65
- 4.12 Mercados representados no plano $(\mu_{emp}, b_2^s/A)$. Símbolos são como os da Fig. 4.9. Mercados desenvolvidos representados com símbolos mais escuros. A linha pontilhada é a localização dos pontos dados pela Eq. (4.1). 66
- 4.13 Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jul.1997 a jul.2007, para os índices (a) BSE30 (Índia) e (b) CMA (Egito). 67
- 4.14 Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jan.1997 a jul.2007, para os índices (a) BEL20 (Bélgica) e (b) Nikkei (Japão). 67
- 4.15 Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jan.1997 a jul.2007, para os índices (a) FTSE100 (Reino Unido) e (b) DJIA (Estados Unidos). 67
- 5.1 Gráfico log-log da FDP dos volumes negociados a cada 30 minutos na BOVESPA no período entre 3/01/05 a 13/09/07. Os símbolos representam o histograma empírico e a linha sólida o ajuste por mínimos quadrados da Eq. (2.15) com parâmetros mostrados na Fig. 5.2. A linha pontilhada corresponde ao ajuste da Eq. (2.15) com $q = 1$, representado para comparação. Os gráficos inseridos correspondem às representações do gráfico principal em escalas linear-linear e log-linear. 70
- 5.2 Parâmetros de ajuste da distribuição q -Gama, β (a), q (b) e θ (c), para a FDP dos volumes normalizados entre 3/01/05 e 13/09/07, em função de Δt (triângulos). Os círculos correspondem aos valores teóricos obtidos das Eqs. (5.3). As linhas pontilhadas são um guia visual. As linhas cheias em (a) correspondem a um ajuste linear. 71

5.3	Função de autocorrelação para os dados de volume de negociação analisados nesta tese.	72
5.4	Número de ticks de 30 minutos em função do volume acumulado v_c na BOVESPA durante o período entre 3/01/05 e 13/09/07. Gráfico inserido: ampliação da curva.	74
5.5	Série temporal de σ em função de v_c para diferentes números de pontos por janela Δv_c (indicado na figura).	75
5.6	Histograma dos declives locais do gráfico mostrado na Fig. 5.4, calculados em janelas de 30 ticks (símbolos). Linhas sólidas correspondem ao melhor ajuste de mínimos quadrados de uma FDP Gama. Gráfico inserido: o mesmo gráfico em escala linear.	75
5.7	Ajustes da função q -Gama obtidos para as FDPs dos volumes de negociação para diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07: 60min(a), 90min(b), 120min(c), 180min(d). A q -Gama ajustada foi calculada com peso estatístico e desconsiderando volumes $\leq 10^{-1}$.	76
6.1	Coefficientes $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$ dos dados numéricos para $\tau = 1$. Linhas correspondem aos melhores ajustes, linear e quadrático respectivamente, por mínimos quadrados aos dados.	81
6.2	FDPs dos retornos normalizados. Símbolos correspondem aos histogramas empíricos. FDPs teóricas dadas pela Eq. (3.12) com $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ obtidas do cálculo dos coeficientes de tendência e difusão de acordo com as Eqs. (6.3) e (6.4) (linhas pontilhadas), e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ obtidos da Eqs. (6.15) e (6.16) (linhas cheias).	82
6.3	Séries temporais de retornos: empírica (a), artificiais geradas com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (b) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (c).	83
6.4	Função de autocorrelação linear $C(t)$ em função do tempo. Os símbolos correspondem aos dados empíricos, as linhas correspondem aos processos artificiais gerados com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (linhas pontilhadas) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (linhas cheias).	84
6.5	(a) Dependência de $a_1^{(n)}$ em função da ordem de aproximação n dada na Eq. (6.10), para diferentes valores de \tilde{a}_1 . (b) Ordem na qual o valor limite é atingido em função de \tilde{a}_1 , com erro de 5%.	86
6.6	FDPs conjuntas para retornos consecutivos, construídos a partir de: (a) dados empíricos, (b) processo artificial gerado com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$, (c) $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ e (d) dados empíricos misturados.	88
A.1	Parâmetro \tilde{a}_1 calculado para diversos $\Delta\tau$ a partir das séries de 15 minutos (a) e de 30 segundos (b) apresentadas nesta tese para o IBOVESPA.	109
A.2	Parâmetro \tilde{b}_0 calculado para diversos $\Delta\tau$ a partir das séries de 15 minutos (a) e de 30 segundos (b) apresentadas nesta tese para o IBOVESPA.	110

A.3	Coeficientes a_1 e b_2 (a) e b_0 (b) calculados a partir de séries de retornos e incrementos diários. Eles foram calculados em uma região de $\sigma = 1,75$ e com $\Delta\tau = 0,2$.	113
A.4	Séries de preço dos índices mundiais apresentados.	116
A.5	Séries temporais de volume e de quantidade de transações em um trecho (para melhor visualização do intervalo relativo a 1 dia) no ano de 2006 para a escala de 30 minutos.	117
A.6	Ajustes da função Gama obtidos para as FDPs da quantidade de transações para diferentes escalas para a BOVESPA no ano de 2006: 30min(a), 60min(b), 90min(c), 120min(d). A FDP Gama ajustada foi calculada com peso estatístico e para todo o intervalo de dados.	118
A.7	Número de ticks de 30 minutos em função da quantidade de transações acumulada Q_{T_c} na BOVESPA durante o ano de 2006.	118

Lista de tabelas

- 4.1 Índices mundiais e seus correspondentes rótulo, abreviatura, país e número de dias de negociação entre 1997 e 2007. 55
- 5.1 Parâmetros de ajuste da função q -Gama obtidos para as FDPs dos volumes de negociação em diferentes escalas de agregação para a BOVESPA no período entre 3.jan.05 e 13.set.2007. Incluímos também os valores médios $\langle V \rangle$ das séries originais. A $P(v)$ ajustada foi calculada com peso estatístico e descartando volumes ($v \leq 10^{-1}$). 72
- 5.2 Parâmetros obtidos para o ajuste da distribuição Gama para os histogramas das taxas em diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07. A $P(\sigma)$ foi ajustada utilizando toda a região dos dados. Para comparação, incluímos os valores teóricos dos parâmetros da FDP q -Gama, obtidos a partir da Eq. (5.3). 77
- A.1 Teste ADF para séries de preços, incrementos e retornos logarítmicos para os índices Merval, IBOVESPA e AOX e série de volumes para o BOVESPA, respectivamente. 108
- A.2 Teste PP para séries de preços, incrementos e retornos logarítmicos para os índices Merval, BOVESPA e AOX e série de volumes para o BOVESPA, respectivamente. 109
- A.3 Número médio de horas/dia de pregão da série de preços catalogadas a cada 15 minutos para o período entre 01/1998 e 06/2003 no IBOVESPA. 111
- A.4 Número médio de min/dia de pregão da série de preços catalogados a cada 30 segundos para o período entre 01/11/02 e 19/07/06 do IBOVESPA. 111
- A.5 Valores de τ segundo a Eq. (3.1) utilizados nas FDPs teóricas e empíricas apresentadas na Fig. 3.9 (de baixo para cima). As cores representam as séries de dados utilizadas para a construção do histograma empírico: preto - diária, vermelho - 15 minutos, azul - 30 segundos. 112
- A.6 Parâmetros das FDPs Gamas ajustadas para as distribuições de Q_T em diversas escalas. 119

"Como regra geral, o homem mais bem sucedido na vida é aquele que possui a melhor informação." Benjamin Disraeli

agentes possui um grande grau de incerteza devido à capacidade limitada de processamento das informações disponíveis.

Modelos matemáticos e estatísticos provenientes da descrição de sistemas físicos tem sido cada vez mais utilizados para realizar previsões sobre a dinâmica dos sistemas financeiros, obtendo leis para descrever seu comportamento esperado (médio) assim como as flutuações em torno deste comportamento padrão. Há um grande e crescente interesse econômico e científico em tentar prever a dinâmica do mercado e, em última análise, avaliar os riscos envolvidos nesta atividade econômica.

Um dos principais problemas na Física Estatística consiste no estudo das propriedades macroscópicas de sistemas complexos nos quais flutuações desempenham um papel central, por exemplo, na difusão e transições induzidas por ruído. Por essas características, as grandezas observáveis financeiras podem ser descritas com as ferramentas da Física Estatística.

De fato, algumas observáveis financeiras, como as flutuações de preços, tem sido recentemente estudadas em analogia com processos de difusão anômala [1], turbulência [2], intermitência [3] e outros fenômenos físicos. Reciprocamente, a análise da dinâmica de preços pode ajudar a esclarecer processos físicos mais gerais.

A descrição dos dados financeiros por processos estocásticos é de grande importância teórica, devido à dinâmica fortemente complexa das flutuações das séries temporais financeiras, que impõe novos desafios às tentativas de modelagem das leis dinâmicas responsáveis pelas propriedades estatísticas observadas.

1.1

Grandezas Observáveis Financeiras

Nesta tese, analisaremos duas grandezas observáveis financeiras importantes: o retorno de preços de índices de mercados e o volume de negociação. A nossa expectativa é que, ao analisar estas séries financeiras, possamos identificar os mecanismos determinísticos e estocásticos que governam a dinâmica do mercado financeiro.

As séries financeiras são coletadas e organizadas pelas Bolsas de Valores. Estas entidades financeiras divulgam um conjunto de informações sobre os negócios realizados ao longo do dia. Os principais indicadores referem-se a preços e volumes das ações negociadas. Para retratar o comportamento médio dos preços das ações e servir de parâmetro para identificar as tendências gerais do mercado de um país ou de segmentos específicos da economia, foram desenvolvidos os índices do mercado acionário, que hoje são considerados como

autênticos termômetros das expectativas dos investidores em relação ao futuro desempenho da economia.

Assim, os índices de mercado são formados por uma média dos preços das ações mais negociadas, ponderadas pelo volume da transação de cada uma. Dessa forma, podem ser efetivamente caracterizados como a medida da performance dos respectivos mercados. No Brasil, o índice que acompanha a evolução média das cotações das ações brasileiras, e que foi um dos índices analisados nesta tese, é o IBOVESPA. Além dele, existem diversos índices internacionais importantes. Abordaremos no nosso trabalho, índices que abrangem todas as regiões do mundo, bem como mercados desenvolvidos e emergentes.

Preços e índices relacionados à moeda possuem flutuações no tempo que são inerentes ao processo econômico. A unidade monetária analisada pode sofrer alterações devido a diversos fatores, como, por exemplo, crescimento ou recessão da economia, influência de outros mercados devido à globalização e inflação.

Entre as observáveis financeiras, as mais importantes são aquelas que determinam ganhos e perdas, e que podem ser estimadas a partir da estatística das flutuações das séries de preço em diferentes escalas de tempo Δt .

Há muitas formas de se definir a variável que avalia lucro e prejuízo. As mais utilizadas na literatura são os incrementos e os retornos logarítmicos, ou simplesmente retornos, de preços de índices.

A primeira e mais simples definição consiste em representar os incrementos de preço

$$\Delta x(t) = x(t + \Delta t) - x(t). \quad (1.1)$$

Contudo, variações no preço ao longo do tempo, como ajustes de moeda e inflação, afetam significativamente os incrementos. Essa observável é adequada apenas em períodos no qual a economia de determinado mercado se mantenha muito estável.

A segunda definição, mais utilizada na literatura, é o retorno logarítmico

$$r(t) = \log x(t + \Delta t) - \log x(t) = \log \frac{x(t + \Delta t)}{x(t)}. \quad (1.2)$$

Esta definição fornece uma medida em escala logarítmica da variação dos preços. Essa definição tem a vantagem de fornecer o lucro ou prejuízo percentual, informação mais adequada para os investidores, pois fornece o retorno do investimento relativo ao capital empregado.

Em economias com níveis de inflação históricos significantes (como a brasileira antes do Plano Real), é importante realizar um tratamento nas observáveis de preço visando descontar a inflação do período por um fator deflacionário $D(t)$. Com essa aproximação, os preços passam a refletir apenas a variação de seu valor intrínseco.

Todas as séries de preço do IBOVESPA utilizadas nesta tese foram deflacionadas pelo IGP-DI², tendo como referência o mês de agosto de 1994. O índice diário de deflação foi calculado a partir do índice mensal acumulado da data de referência até o dia em questão.

A seguir, na Fig. 1.1, apresentamos a série temporal diária do IBOVESPA nominal e a deflacionada pelo IGP-DI, que foi a utilizada nesta tese.

É importante ressaltar que a análise apresentada neste estudo não é designada para capturar valores extremos e seus efeitos, como *booms* ou *crashes* econômicos.



Figura 1.1: Série do IBOVESPA entre 1991 e 2006 (período analisado nesta tese) com valores históricos e com preços deflacionados pelo IGP-DI conforme indicado em legenda.

Apresentamos, como exemplo, na Fig. 1.2, a série de retorno diário do IBOVESPA deflacionado no período indicado na figura. Podemos perceber

²Índice Geral de Preços. É calculado mensalmente pela Fundação Getúlio Vargas. O IGP-DI/FGV foi instituído em 1944 com a finalidade de medir o comportamento de preços em geral da economia brasileira. É uma média aritmética, ponderada dos índices: Índice de Preços no Atacado (60%), Índice de Preços ao Consumidor (30%) e Índice Nacional da Construção Civil (10%). DI ou Disponibilidade Interna é a consideração das variações de preços que afetam diretamente as atividades econômicas localizadas no território brasileiro.

que após a implantação do Plano Real em 1994, a amplitude dos retornos gradualmente diminuiu em relação ao período anterior, com exceção de algumas épocas de crises internacionais, como por exemplo as crises dos Tigres Asiáticos em 1997 e do Brasil em 1999.

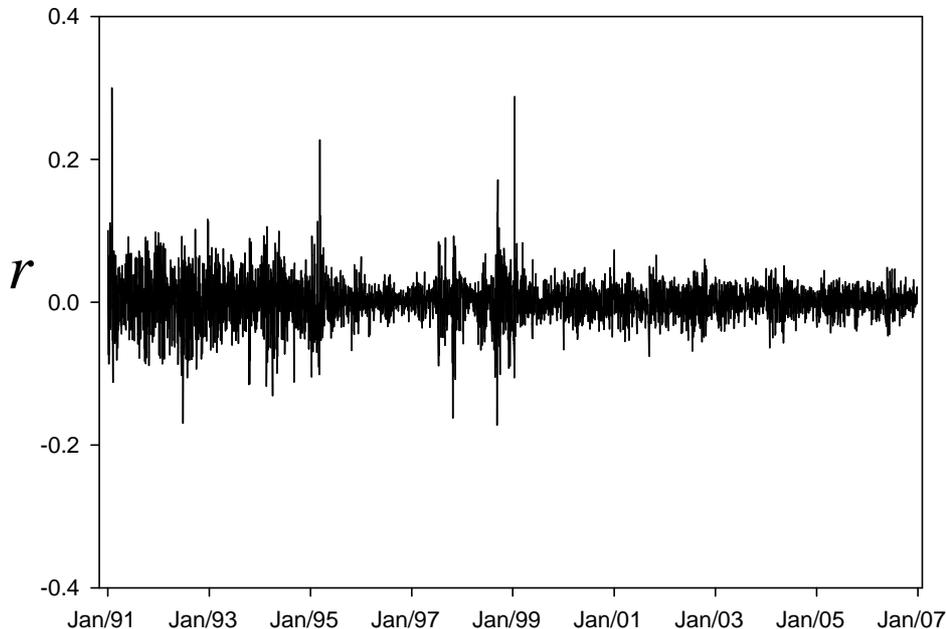


Figura 1.2: Série dos retornos diários do IBOVESPA entre 1991 e 2006.

A segunda observável que será estudada nesta tese é o volume de negociação, que pode ser definido como o número de ações ou então, alternativamente, de contratos negociados entre vendedores e compradores durante um determinado período de tempo. Nesse estudo, utilizamos a primeira definição. O volume é um indicador importante na análise do mercado pois ele mede o valor de uma jogada de negociação, e serve como medida da atividade financeira. Se os mercados fizerem uma forte pressão para os preços moverem-se para cima ou para baixo, a percepção dessa força depende do volume para esse período. Quanto maior o volume de negociação que gerou o movimento do preço, mais significativo é este movimento.

O volume de negociação nas Bolsas de Valores é uma medida da liquidez das ações e também pode fornecer informações úteis para a realização de estratégias de negociação [4].

Tomamos como exemplo representativo nesta tese, o volume negociado intradiário da BOVESPA, de 3 de janeiro de 2005 a 13 de setembro de 2007. Os dados foram obtidos em intervalos de 30 minutos. A série original de volume de negociação, $V(t)$, foi reescalada pelo valor empírico médio $\langle V \rangle$ e é

apresentada na Fig. 1.3, para o período estudado nesta tese. Podemos perceber o padrão intermitente da atividade de negociação.

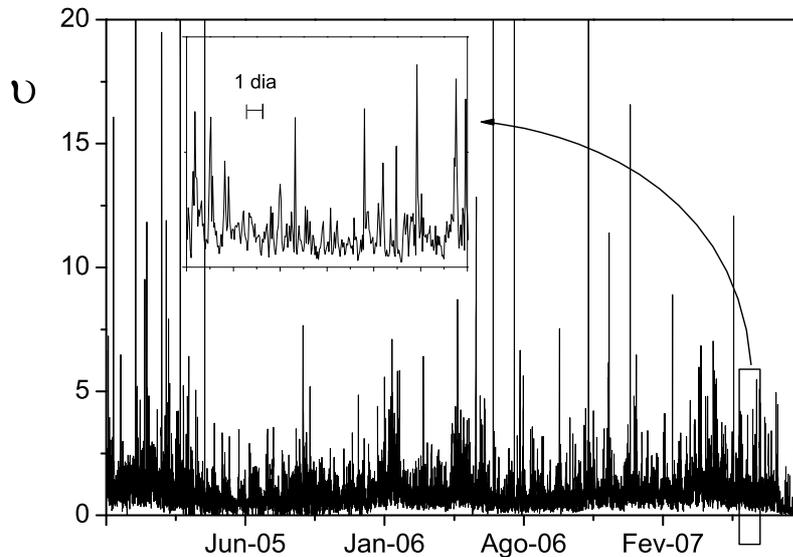


Figura 1.3: Série temporal dos volumes negociados na BOVESPA no período de 3/01/05 a 13/09/07. Os valores representam o volume negociado em intervalos $\Delta t = 30$ minutos, normalizados pelo valor médio em todo o período ($v = V/\langle V \rangle$).

1.2

Resultados na Literatura

A maioria dos modelos matemáticos utilizados em Econofísica tem sua origem na tese de doutorado de Louis Bachelier sobre a Teoria da Especulação na Bolsa de Paris em 1900, na qual ele propôs um modelo para o processo estocástico dos preços, descrito por uma caminhada aleatória, ou *random walk*, com flutuações regidas por uma distribuição gaussiana [5].

Mais recentemente, uma explicação importante da hipótese de random walk foi apresentada por Fama e Samuelson [6] na década de 70, que consiste na formulação da Hipótese do Mercado Eficiente, e que se tornou um paradigma utilizado até hoje.

A Hipótese do Mercado Eficiente estabelece que o preço atual da ação reflete todas as informações e expectativas dos participantes do mercado. Conseqüentemente, nenhum lucro pode ser obtido a partir de negociações baseadas em informações do mercado, pois estas já foram absorvidas no preço. Assim, o retorno esperado de investimento futuro é nulo. Baseado nesta hipótese, o movimento da cotação de uma ação seria completamente imprevisível a partir de informações tais como preço e volume de transação passados.

Assim, segundo essa teoria, a propriedade de ajustamento instantâneo dos preços das ações devido às informações públicas implica na independência entre variações sucessivas de preços na seqüência das negociações. Um mercado que apresenta tal comportamento é, por definição, um mercado random walk, onde a série de mudanças de preços não tem memória, isto é, a história passada da série não pode ser usada para prever o futuro. Neste mercado informacionalmente eficiente, onde todos os participantes detêm o mesmo nível de informação e agem racionalmente, ajustando os preços imediatamente, não haveria espaço para que os fundamentalistas obtivessem lucros acima da média do mercado. Da mesma forma, os grafistas também não poderiam antecipar nenhuma nova tendência ou movimento de preços em formação. Neste caso, somente restariam os fatos imprevisíveis a comandar as quebras de expectativas.

No entanto, no mercado real, a atualização dos preços não é instantânea e as negociações tem custo. O grau de ineficiência do mercado é o que permite oportunidades de lucro que compensam os investidores dos custos das transações e da obtenção de informação. O mercado real segue portanto apenas aproximadamente o mercado eficiente, pois caso contrário, qualquer esforço de previsão baseada em análise da série histórica seria inútil. Apesar da Hipótese do Mercado Eficiente ser uma idealização economicamente irrealizável, ele serve como referência para outras modelagens do processo estocástico dos preços.

Nessa mesma época, na década de 60, Mandelbrot [7], ao analisar o mercado americano de algodão, verificou que as distribuições de retornos de preços se comportavam segundo as chamadas distribuições de Lévy, com caudas mais longas que as distribuições gaussianas. A partir de então, numerosas observações nos mercados por todo o mundo mostraram que as caudas das distribuições de diversos ativos, tais como *commodities*, taxas de câmbio e retorno de preços de ações, decaíam mais devagar do que a distribuição gaussiana predizia. Muitos modelos foram propostos na literatura para caracterizar a distribuição de retornos de preços nestes mercados, para várias escalas de tempo e com especial interesse para as distribuições leptocúrticas³ - assim chamadas as distribuições com caudas mais acentuadas do que a da distribuição normal.

A Fig. 1.4 apresenta a distribuição empírica para os retornos normalizados do IBOVESPA na escala de tempo de $\Delta t=1$ minuto e a sua comparação com a distribuição gaussiana. O período de observação considerado neste exemplo foi de 4 anos, entre Novembro de 2002 e Julho de 2006. Verifica-

³Distribuições leptocúrticas são aquelas que possuem curtose maior do que a da distribuição normal, igual a 3. Comparando-se duas distribuições com mesmo desvio padrão, a distribuição com maior curtose possui caudas mais longas.

se que a distribuição gaussiana é inadequada para descrever a distribuição empírica.

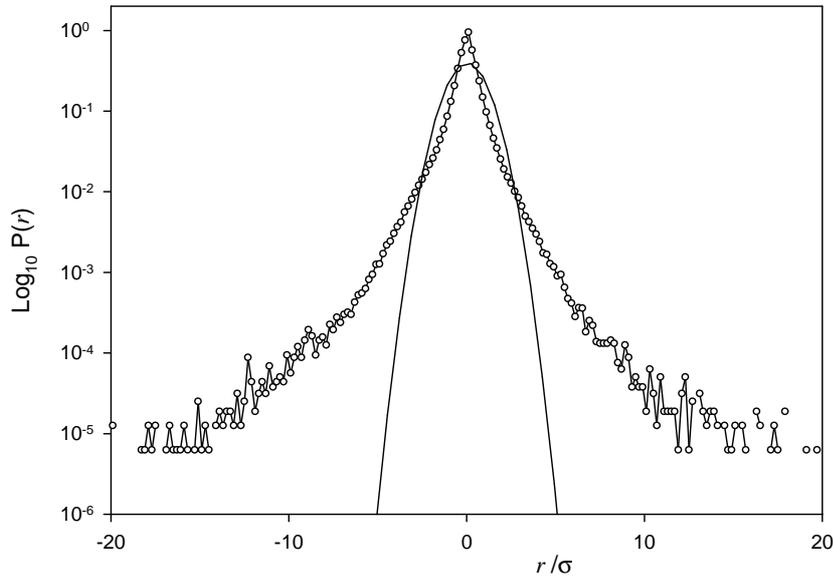


Figura 1.4: Comparação dos histogramas de retornos normalizados para $\Delta t = 1$ minuto do IBOVESPA (símbolos), de novembro de 2002 a julho de 2006, com a distribuição gaussiana (linha cheia), em gráfico semi-logarítmico.

A análise de vários tipos de títulos financeiros em diferentes mercados tem fornecido distribuições com caudas em lei de potência com expoente de decaimento tipicamente na faixa de $2 < \alpha < 5$, fora do domínio das distribuições de Lévy. Distribuições com caudas exponenciais também tem sido reportadas para alguns mercados mundiais [8, 9].

Recentemente foi apresentada uma modelagem para os dados de alta frequência do mercado americano utilizando distribuições q -gaussianas [10, 11], uma generalização da distribuição gaussiana, que será descrita com mais detalhes no Capítulo 2. A Fig. 1.5 mostra respectivamente as distribuições empíricas para o NYSE e para a Bolsa eletrônica NASDAQ em 2001. As distribuições são bem descritas por q -gaussianas com parâmetro ótimo $q \cong 1.4$, como verificado na Fig. 1.5 [12].

Por outro lado, sabe-se que a dinâmica de preço em escalas de tempo longas (regime mensal), é regida por distribuições gaussianas. Os dois regimes de escalas longas e curtas tem sido modelados na literatura como dois regimes separados. Um dos objetivos desta tese é o de prover um caminho possível para conectar estes dois regimes através de uma descrição única.

O volume de negociações financeiras foi objeto de análise estatística apenas recentemente. Mostrou-se primeiramente que a função de densidade

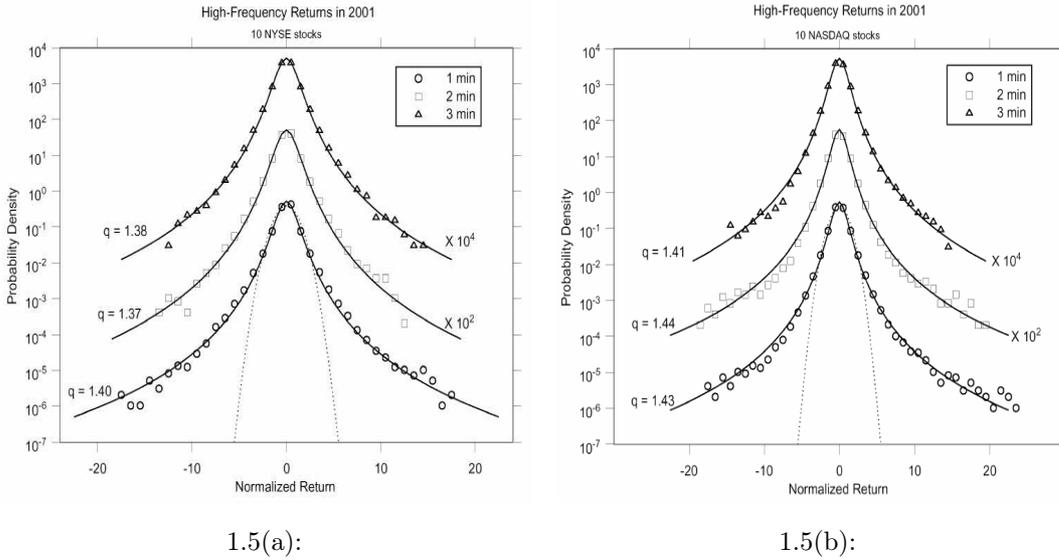


Figura 1.5: Distribuições empíricas (símbolos) e q -gaussianas (linhas cheias) para retornos normalizados das 10 empresas com maior volume de negociação na (a) NYSE e (b) Nasdaq em 2001. A linha pontilhada representa a distribuição gaussiana [9].

de probabilidade (FDP) dos volumes de ações individuais, negociadas em alta-frequência, tem decaimento em lei de potência [13, 14]. A forma da distribuição q -Gama [15], uma generalização da função Gama, que será descrita com mais detalhes no Capítulo 2, foi proposta para a descrição da FDP dos volumes de negociação de alta frequência em diversos mercados: para as 10 ações de maior volume no NYSE e no Nasdaq em 2001, com *ticks*⁴ de 1, 2 e 3 minutos [12, 15]; para as 30 ações do Dow Jones Industrial Average em 2004, com ticks de 1 minuto [16, 17, 18]; e para as ações do Citigroup no NYSE em 2004 [19], com ticks de 1 minuto, conforme apresentado na Fig. 1.6 [12].

1.3 Motivação e Objetivos

Uma característica típica dos sistemas complexos é a existência de estruturas não-triviais em diferentes escalas de tempo. No mercado financeiro, em particular, as flutuações de preço em escalas longas e curtas, são caracterizadas por FDPs gaussianas e com cauda longa, respectivamente.

Para a descrição de ambos os regimes, assim como da transição entre eles, vamos adotar o procedimento proposto por Friedrich e colaboradores [20], que explora uma correspondência entre a dinâmica do mercado financeiro e a turbulência hidrodinâmica [2], assumindo a existência de um fluxo de informação em cascata em uma hierarquia de escalas temporais longas e curtas.

⁴Marcadores de registro de dados financeiros.

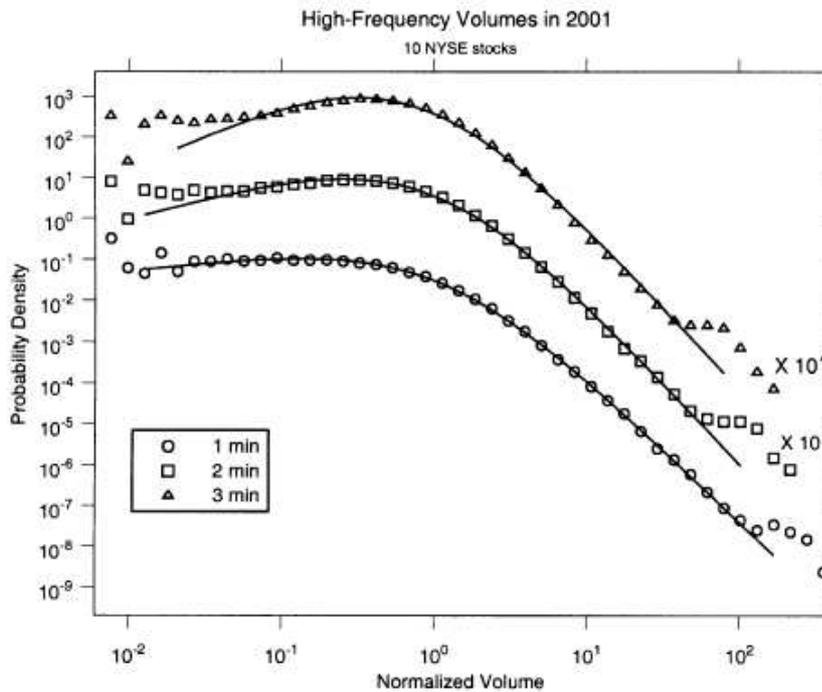


Figura 1.6: Distribuições empíricas (pontos) e ajustes q -exponenciais (linhas) para volumes normalizados das 10 ações de maior volume na NYSE em 2001 [9].

Para tal, computamos diretamente das séries temporais empíricas de retorno, os coeficientes da expansão de Kramers-Moyal (KM) que governam a evolução temporal das FDPs ao longo de diversas escalas de tempo. Reproduzimos a plena evolução dos histogramas empíricos dos retornos do IBOVESPA, abraçando a transição de FDPs gaussianas às com caudas longas, quando se passa de escalas de tempo longas para curtas.

Apresentamos também soluções analíticas consistentes com a família das FDPs que emergem da equação de evolução temporal. Mostramos que estas distribuições pertencem à classe de distribuições q -gaussianas, que compreendem as famílias de FDPs invariantes e não-estacionárias observadas em ambos os limites temporais assintóticos.

Em seguida, estendemos nossa análise a diversos mercados ao redor do mundo. Por um lado, a detecção de padrões universais entre esses mercados é importante para identificar os mecanismos fundamentais que governam a dinâmica dos preços. Por outro lado, é também importante quantificar e distinguir os mercados, de acordo com algumas propriedades, tais como eficiência, resistência a pressões externas, à especulação internacional, ou o tempo de relaxação para que os efeitos provocados por choques macroeconômicos se dissipem.

Neste cenário, a obtenção dos coeficientes de KM diretamente das

séries temporais empíricas, pode ser considerada uma técnica alternativa para comparar as dinâmicas dos índices de mercado, uma vez que o conhecimento destes coeficientes permite acessar quantitativamente as leis determinísticas e estocásticas que governam a dinâmica dos preços.

Para este fim, foi realizada uma análise extensiva, obtendo e comparando os coeficientes de 24 índices diários de mercados ao redor do mundo, englobando desde mercados desenvolvidos até os considerados emergentes.

Nossa análise revela a existência de padrões universais dos coeficientes de KM, e que estes são descritos em termos de um pequeno número de parâmetros microscópicos.

Esses parâmetros permitem quantificar características, tais como desvios da gaussianidade ou da eficiência, fornecendo uma ferramenta útil para discriminar os diversos mercados. Esta abordagem permite ainda ranquear ou segregar os mercados, de acordo com os parâmetros microscópicos que caracterizam os coeficientes de KM.

No caso dos volumes financeiros, também podemos identificar processos caracterizados por diferentes escalas temporais. Por um lado, o volume de ações negociadas ao longo do pregão é o resultado de um processo de negociação caracterizado por várias etapas: a chegada de ordens de compra, de venda e de ajuste entre essas ordens. Este processo determina uma escala microscópica de formação de preços e de volume negociados. Por outro lado, as flutuações no ambiente do mercado tais como fluxo não estacionário de notícias e rumores, afetam o volume global de negociação, implicando na existência de inomogeneidades na atividade do mercado. A flutuação desta atividade, de acordo com a época de observação, por sua vez possui uma escala temporal característica mais longa do que os tempos característicos de formação do preço.

Para a análise dos volumes, vamos caracterizar sua FDP tendo como base processos de Poisson microscópicos de diferentes escalas misturadas, no qual a atividade do mercado é também um parâmetro flutuante.

Na literatura, as tentativas de modelagem, tanto da dinâmica de preços como a de volumes, são geralmente construídas fenomenologicamente a partir de modelos propostos *a priori*.

Um dos objetivos desta tese é desvendar a dinâmica estocástica subjacente das flutuações dos preços e dos volumes, através de medidas diretas das séries temporais, identificando os fatores determinísticos e estocásticos.

2

Referencial Teórico

Neste capítulo, apresentaremos de forma resumida, a base matemática fundamental para os capítulos seguintes e as principais ferramentas teóricas utilizadas no desenvolvimento de nosso trabalho [21, 22, 23, 24].

2.1

Variáveis Aleatórias e Distribuições de Probabilidade

Uma *variável aleatória* unidimensional X é uma função $S \mapsto S_X \subset \mathbb{R}$ que associa a cada elemento s de um espaço amostral S (conjunto dos resultados possíveis (numéricos ou não) de um experimento aleatório), um número real $X(s)$. A variável X pode ser discreta (S_X finito ou infinito enumerável) ou contínua (S_X infinito não-enumerável).

A *distribuição de probabilidade* de uma variável aleatória X cujo contradomínio é $S_X = \{x_1, x_2, \dots\}$ (caso discreto) é dada pelo conjunto de pares (x_i, p_i) , $i = 1, 2, \dots$, onde $p_i \equiv P(X = x_i)$, probabilidade de x_i , devem satisfazer

- I- $p_i \geq 0$ (não-negatividade);
- II- $\sum_{i \geq 1} p_i = 1$ (normalização).

No caso contínuo, a distribuição de probabilidade de uma variável aleatória X é dada pela função f_X , chamada *função densidade de probabilidade* (FDP), tal que $f_X(x)dx$ representa a probabilidade $P(x \leq X \leq x+dx)$. Assim, a FDP permite calcular a probabilidade de que X se encontre dentro de um intervalo real $[a, b]$:

$$P(a \leq X \leq b) = \int_a^b dx f_X(x). \quad (2.1)$$

A FDP *conjunta* de duas (ou mais) variáveis aleatórias X, Y (facilmente generalizável para N) é $f_{X,Y}(x, y)$, tal que $f_{X,Y}(x, y)dxdy = P(x \leq X \leq x + dx, y \leq Y \leq y + dy)$. O par (X, Y) representa também uma variável aleatória bidimensional.

A FDP *marginal*, por exemplo, da variável X é dada por

$$f_X(x) = \int_{-\infty}^{\infty} dy f_{X,Y}(x, y), \quad (2.2)$$

enquanto a FDP *condicionada* de X , dado um certo valor de $Y = y$, é

$$f_{X|Y}(x, y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad \text{com } f_Y > 0, \quad (2.3)$$

onde f_Y é a FDP marginal de Y .

Uma FDP é caracterizada pelos *momentos*. O momento de ordem n da variável X é dado por:

$$\langle X^n \rangle = \int_{-\infty}^{\infty} dx x^n f_X(x). \quad (2.4)$$

O momento *centrado* de ordem n da variável X é:

$$\langle [X - \mu_X]^n \rangle = \int_{-\infty}^{\infty} dx (x - \mu_X)^n f_X(x), \quad (2.5)$$

onde $\mu_X \equiv \langle X \rangle$ é a *média* (aritmética) ou *valor esperado*. A *variância* ou *desvio quadrático* é o momento centrado de ordem 2, $\langle (X - \mu_X)^2 \rangle$, e o *desvio padrão* a sua raiz quadrada:

$$\sigma_X = \sqrt{\langle (X - \mu_X)^2 \rangle}. \quad (2.6)$$

2.1.1 FDPs usuais

É comum utilizar como primeira aproximação ou primeira tentativa de modelagem das distribuições empíricas, uma distribuição gaussiana ou normal

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_0)^2}{2\sigma^2}}, \quad (2.7)$$

com média μ_0 e desvio padrão σ .

Esta FDP aparece em diversos fenômenos na natureza, e em particular em processos relacionados à difusão normal ou Browniana. A razão da presença praticamente universal da distribuição gaussiana se deve a ela emergir naturalmente como uma distribuição limite para processos aleatórios, como consequência do Teorema Central do Limite (TCL): a soma de variáveis aleatórias independentes e identicamente distribuídas cujo segundo momento é finito é normalmente distribuída. Se as variáveis não possuem a mesma distribuição, deve ser satisfeita adicionalmente a condição de Lindeberg [21] para a validade do TCL.

Sendo distribuição limite por agregação de variáveis aleatórias minimamente bem comportadas, a gaussiana é estável. Diz-se que uma família de variáveis independentes e identicamente distribuídas possui distribuição estável se a soma delas possuir a mesma distribuição das variáveis individualmente.

Uma outra distribuição usual para variáveis aleatórias definidas positivas é a distribuição Gama ou erlangiana [25]

$$f(x) = \frac{x^{\beta-1} e^{-x/\theta}}{\theta^\beta \Gamma(\beta)}, \quad (2.8)$$

sendo $\beta, \theta > 0$ e onde $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ é a função Gama.

Esta FDP decai exponencialmente e inclui como casos particulares a FDP exponencial (se $\beta = 1$) e a χ -quadrado (se $\beta = n/2$, com n inteiro e $\theta = 2$).

Embora a parte central de uma distribuição geralmente é a região de maior probabilidade de ocorrência, é nas suas caudas que obtemos informações relativas aos valores extremos. Assim, em qualquer modelagem de distribuição de retornos de preços ou outras variáveis financeiras é fundamental a análise das caudas das FDPs, pois permite estimar lucros e prejuízos relevantes para o mercado financeiro.

Em geral, a frequência de ocorrência de valores extremos nas séries financeiras apresenta desvios do comportamento exponencial, sendo ordens de grandeza maiores do que os previstos pelas FDPs usuais, como ilustrado na Fig. 1.4. Diz-se que esses histogramas empíricos possuem caudas gordas ou pesadas. A seguir, vamos analisar as propriedades de algumas distribuições que foram utilizadas na modelagem de variáveis financeiras nesta tese.

2.1.2

FDPs generalizadas

Em problemas tradicionais da mecânica estatística de equilíbrio a energia e a entropia são quantidades extensivas. Para que estes resultados sejam válidos é necessário que diferentes regiões do sistema sejam independentes. Existem sistemas com interação de longo alcance, no entanto, para os quais não é possível assumir esta independência como, por exemplo, estrelas interagindo sob a influência da interação gravitacional. Da mesma forma, sistemas complexos cujo estado fundamental é altamente degenerado ou que possuem memória microscópica de longo alcance apresentam tempo de relaxação ao equilíbrio muito longo e não podem ser na prática descritos pela mecânica estatística de Boltzmann-Gibbs. Recentemente, foi proposto por Tsallis [10] um formalismo para a análise desses sistemas baseado em estatísticas não-extensivas. Deste formalismo, emergem naturalmente FDPs em que a função exponencial é generalizada segundo

$$\exp_q(x) = [1 + (1 - q)x]^{\frac{1}{1-q}}, \quad (2.9)$$

onde $q \in \mathbb{R}$.

Esta função, conhecida como q -exponencial, permite generalizar várias FDPs usuais, tal como a gaussiana, segundo a expressão

$$G_q(x) = \mathcal{N}_q \exp_q(-\beta(x - \mu_0)^2) = \mathcal{N}_q [1 - \beta(1 - q)(x - \mu_0)^2]^{\frac{1}{1-q}}, \quad (2.10)$$

onde $\beta > 0$ é um parâmetro de escala e \mathcal{N}_q é a constante de normalização, que para $1 < q < 3$ é dada por:

$$\mathcal{N}_q = \sqrt{\left(\frac{\beta(q-1)}{\pi}\right)} \frac{\Gamma\left[\frac{1}{(q-1)}\right]}{\Gamma\left[\frac{(3-q)}{2(q-1)}\right]}. \quad (2.11)$$

Esta FDP, conhecida como q -gaussiana, possui como limite do parâmetro $q \rightarrow 1$ a gaussiana usual. Assintoticamente, a q -gaussiana tem comportamento em lei de potência para $q > 1$. Com efeito, de (2.10), com $\beta^{-1/2}|x| \gg 1$:

$$G_q(x) \approx 1/|x|^{\frac{2}{(q-1)}}. \quad (2.12)$$

O parâmetro β controla a largura, e conseqüentemente, pelo vínculo de normalização, a altura da FDP em $x = 0$. A variância é finita somente para $q < 5/3$. Neste caso é dada por:

$$\sigma^2 = \langle (x - \mu_0)^2 \rangle = \int (x - \mu_0)^2 G_q(x) dx = \frac{1}{\beta(5 - 3q)} \quad (2.13)$$

e o parâmetro de escala β é inversamente proporcional à variância. De acordo com a Eq. (2.13), podemos reescrever a Eq. (2.10) como:

$$G_q(x) = \mathcal{N}_q \left[1 + \frac{(q-1)}{(5-3q)} \frac{(x - \mu_0)^2}{\sigma^2} \right]^{\frac{1}{(1-q)}}. \quad (2.14)$$

A distribuição Gama, dada pela Eq. (2.8), também pode ser generalizada de forma a descrever distribuições com caudas pesadas, através de

$$\Gamma_q(x) = \mathcal{N}_q x^{\beta-1} \left[1 + (q-1) \frac{x}{\theta} \right]^{\frac{1}{(1-q)}}, \quad (2.15)$$

onde

$$\mathcal{N}_q = \left(\frac{q-1}{\theta}\right)^\beta \frac{\Gamma\left[\frac{1}{(q-1)}\right]}{\Gamma[\beta] \Gamma\left[\frac{1}{q-1} - \beta\right]}. \quad (2.16)$$

Esta distribuição, conhecida como q -Gama, no limite $q \rightarrow 1$ tende à distribuição Gama usual e para $q > 1$ possui cauda em lei de potência.

2.2

Processos Estocásticos

Uma função aleatória $\mathcal{P}_X(t)$ é uma aplicação que atribui a cada valor do índice $t \in \mathbb{T}$ uma variável aleatória X_t . Um processo estocástico (PE) é uma função aleatória do tempo [24]. O processo pode ser de tempo discreto

ou contínuo dependendo da natureza do conjunto \mathbb{T} . Atribuindo-se um valor particular $x(t)$ à variável aleatória X_t (que alternativamente denotaremos $X(t)$), para cada t , temos uma realização do processo, uma função ordinária do tempo. Um PE $\mathcal{P}_X(t)$ também pode ser visto como o conjunto ou *ensemble* de todas as suas realizações $\{x(t), t \in \mathbb{T}\}$ ¹.

2.2.1

Classes de processos estocásticos

As características que permitem distinguir entre PEs são a natureza do espaço de estados (espaço amostral de cada $X(t)$, com $t \in \mathbb{T}$), a natureza do conjunto de índices \mathbb{T} e as relações de dependência entre as variáveis aleatórias $X(t)$, que se derivam das distribuições conjuntas.

Conhecendo a hierarquia infinita de FDPs conjuntas correspondentes a n valores arbitrários da variável temporal $f_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) \equiv f_n(X(t_1) = x_1, \dots; X(t_n) = x_n)$, com $n = 1, 2, \dots$, o PE (seja discreto ou contínuo) fica completamente definido [23]. As FDPs f_j da hierarquia com $1 \leq j < n$ são obtidas por integração de f_n . Dadas estas funções podemos calcular valores médios. Por exemplo, a função de autocorrelação de dois tempos é dada por

$$\langle X(t_1)X(t_2) \rangle = \int \int dx_1 dx_2 x_1 x_2 f_2(x_1, t_1; x_2, t_2). \quad (2.17)$$

Se as f_j não mudam ao substituir t_i por $t_i + \tau$ (com τ arbitrário), então trata-se de um *processo estacionário*, ou seja, não são afetadas por um deslocamento no tempo:

$$f_n(x_1, t_1 + \Delta t; \dots; t_n + \Delta t) = f_n(x_1, t_1; \dots; x_n, t_n), \quad (2.18)$$

para todo n, τ e t_1, \dots, t_n . Em particular os momentos resultam ser estacionários no tempo. O PE é dito estacionário em sentido amplo, se possui segundos momentos finitos e se a covariância² $\text{cov}(X(t_i), X(t_i + \tau))$ depende somente de $\tau, \forall t_i \in \mathbb{T}$.

Uma possível classificação dos processos estocásticos é a seguinte [24]:

Processos puramente aleatórios:

A FDP condicional $f_{1|n-1}$, com $n > 1$, independe dos valores em instantes anteriores, ou seja

$$f_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = f_1(x_n, t_n). \quad (2.19)$$

¹Veja também a interessante descrição alternativa de van Kampen [23].

²Se X e Y são duas variáveis aleatórias contínuas, a covariância é dada por $\int (x - \mu_X)(y - \mu_Y) dx dy$.

Um exemplo deste tipo de processo é dado pela sequência de resultados ao jogar muitas vezes um dado. Nesse caso particular as variáveis X_t são identicamente distribuídas além de serem independentes.

Processos markovianos:

A FDP condicional $f_{1|n-1}$, com $n > 1$, depende somente do valor no instante anterior, ou seja, se o valor presente é conhecido exatamente, o conhecimento futuro não é alterado por informação adicional sobre o passado. Mais formalmente, para qualquer conjunto de tempos sucessivos $t_1 < t_2 < \dots < t_n$:

$$f_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = f_{1|1}(x_n, t_n | x_{n-1}, t_{n-1}). \quad (2.20)$$

Um exemplo de processo markoviano é dado pelo caminhante aleatório. A probabilidade de transição da posição i' ao tempo t para a posição i ao tempo $t + 1$ é

$$P(i, t + 1 | i', t) = p\delta_{i-1, i'} + (1 - p)\delta_{i+1, i'}, \quad (2.21)$$

que depende somente da informação ao tempo imediatamente anterior.

Processos mais gerais:

Uma possível generalização dos PEs descritos acima corresponde ao caso em que as FDPs condicionais dependem de dois ou mais tempos.

Seja, por exemplo, um caminhante que apresenta tendência a persistir ou mudar de direção: ele continua na mesma direção com probabilidade p e muda de direção com probabilidade $q = 1 - p$. Neste caso a variável X_t que descreve a posição do caminhante ao tempo t não é mais markoviana, já que a probabilidade condicional depende não somente de X_{t-1} mas também de X_{t-2} .

Porém, isto pode ser remediado introduzindo uma variável de duas componentes (X_t, X_{t-1}) [23]. Esta nova variável resulta ser markoviana. Com efeito, sua probabilidade de transição é dada por

$$P(i_1, i_2, t + 1 | i'_1, i'_2, t) = \delta_{i_2 i'_1} [p\delta_{i_1 - i_2, i'_1 - i'_2} + (1 - p)\delta_{i_1, i'_2}]. \quad (2.22)$$

Como consequência geral, nota-se que um processo pode ser markoviano ou não dependendo das variáveis utilizadas para descrevê-lo. Se a memória da caminhada envolvesse mais tempos, seriam necessárias mais variáveis para poder tratar o processo como markoviano. Porém, se a memória se estende sobre *todos* os passos anteriores, esse procedimento não pode ser aplicado e

o processo é irremediavelmente não-markoviano. Um exemplo é o dos “self-avoiding random walks” ou polímeros infinitos com volume excluído.

Dentre os diversos tipos de processos estocásticos, aprofundaremos nas propriedades dos processos markovianos.

2.2.2

Processos markovianos

Um processo de Markov é um processo estocástico que não possui memória de “tempos” anteriores. Isto corresponde a casos experimentais em que a “janela” de tempo entre medidas Δt excede o tempo de correlação $\Delta t_C \ll \Delta t$ do processo. Assim, a probabilidade condicional deste processo perde a memória das situações anteriores à última medida, verificando a Eq. (2.20). Apenas o conhecimento sobre o ponto (x_{n-1}, t_{n-1}) é necessário para sabermos o que acontece em (x_n, t_n) .

A sua hierarquia de FDPs f_n fica totalmente caracterizada pelas densidades $f_{1|1}$ e f_1 . Por exemplo, para $t_1 < t_2 < t_3$:

$$\begin{aligned} f_3(x_1, t_1; x_2, t_2; x_3, t_3) &= f_2(x_1, t_1; x_2, t_2) f_{1|1}(x_3, t_3 | x_2, t_2; x_1, t_1) \\ &= f_1(x_1, t_1) f_{1|1}(x_2, t_2 | x_1, t_1) f_{1|1}(x_3, t_3 | x_2, t_2). \end{aligned} \quad (2.23)$$

Podemos proceder da mesma maneira para todas as FDPs f_n dessa hierarquia. Assim podemos expressar f_n por um produto de probabilidades condicionais e por f_1 . Se soubermos as formas de $f_{1|1}(x_2, t_2 | x_1, t_1)$ e $f_1(x_1, t_1)$, podemos calcular todas as quantidades de interesse para o processo. Devido a essas relações, as probabilidades condicionais também são conhecidas como *probabilidades de transição*.

No processo markoviano só há memória do valor da variável aleatória para o último tempo, onde medimos x . O intervalo de tempo $t_2 - t_1$ da FDP condicional $f(x_2, t_2 | x_1, t_1)$ de um processo de Markov é arbitrário. Se a diferença é grande, a dependência de f em x_1 será pequena (ou seja, a memória do valor da variável aleatória é quase inexistente). Se a diferença de tempo é infinitesimal (porém maior que o tempo de correlação), a probabilidade condicional terá idealmente valor exato em x_1 , ou seja, $\lim_{t_2 \rightarrow t_1} f(x_2, t_2 | x_1, t_1) = \delta(x_1 - x_2)$.

Equação de Chapman-Kolmogorov

Integrando a Eq. (2.23) em x_2 (para $t_1 \leq t_2 \leq t_3$) obtemos:

$$f_2(x_1, t_1; x_3, t_3) = f_1(x_1, t_1) \int dx_2 f_{1|1}(x_3, t_3 | x_2, t_2) f_{1|1}(x_2, t_2 | x_1, t_1), \quad (2.24)$$

logo

$$f_{1|1}(x_3, t_3|x_1, t_1) = \int dx_2 f_{1|1}(x_3, t_3|x_2, t_2) f_{1|1}(x_2, t_2|x_1, t_1). \quad (2.25)$$

A Eq. (2.25) é chamada de *equação de Chapman-Kolmogorov* (CK). Ela é uma identidade que deve ser obedecida por todo processo de Markov. É essencial que $t_1 \leq t_2 \leq t_3$. Ou seja, a integração é feita sobre um estado intermediário x_2 . A FDP f_1 também deve satisfazer

$$f_1(x_2, t_2) = \int dx_1 f_{1|1}(x_2, t_2|x_1, t_1) f_1(x_1, t_1). \quad (2.26)$$

A Eq. (2.25) pode ser interpretada como a probabilidade de transição de x_1 no instante de tempo t_1 para x_3 no instante t_3 é a mesma que a probabilidade de transição de x_1 em t_1 para x_2 em t_2 multiplicado pela probabilidade de transição de x_2 em t_2 para x_3 em t_3 para todos os possíveis valores de x_2 . Para um processo markoviano estacionário, a probabilidade $f_{1|1}(x_2, t_2|x_1, t_1)$ depende apenas do intervalo $\Delta t = t_2 - t_1$. No que segue, simplificaremos a notação por $f_1(x, t) \equiv f(x, t)$.

2.2.3

Expansão de Kramers-Moyal

Da definição de probabilidades condicionais podemos verificar que a densidade de probabilidade $f(x, t + \Delta t)$ em um instante $t + \Delta t$ e a densidade de probabilidade $f(x, t)$ em um instante t estão conectadas (para $\Delta t \geq 0$) segundo

$$f(x, t + \Delta t) = \int f(x, t + \Delta t|x', t) f(x', t) dx'. \quad (2.27)$$

Para obter o diferencial $\partial f(x, t)/\partial t$, precisamos conhecer as probabilidades de transição $f(x, t + \Delta t|x', t)$ para Δt pequeno. Por outro lado, assumimos que conhecemos todos os momentos M_k , (para $k \geq 1$):

$$M_k(x', t, \Delta t) = \langle [X(t + \Delta t) - X(t)]^k \rangle_{X(t)=x'} = \int (x - x')^k f(x, t + \Delta t|x', t) dx. \quad (2.28)$$

Existem diversas formas de obter uma expansão geral para a probabilidade de transição. Um dos métodos mais diretos é utilizar a expansão de Taylor para a FDP e para a probabilidade de transição. Definindo $\Delta x = x - x'$, podemos expandir o integrando da Eq. (2.27) em séries de Taylor:

$$\begin{aligned} f(x, t + \Delta t|x', t) f(x', t) &= f(x - \Delta x + \Delta x, t + \Delta t|x - \Delta x, t) f(x - \Delta x, t) \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (\Delta x)^k \frac{\partial^k}{\partial x^k} f(x + \Delta x, t + \Delta t|x, t) f(x, t). \end{aligned} \quad (2.29)$$

Inserindo a Eq. (2.29) na Eq. (2.27) e integrando em relação a Δx ,

obtemos diretamente:

$$f(x, t + \Delta t) = f(x, t) + \sum_{k=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^k \frac{M_k(x, t, \Delta t)}{k!} f(x + \Delta x, t + \Delta t | x, t) f(x, t). \quad (2.30)$$

Assumindo que os momentos M_k podem ser expandidos em séries de Taylor em relação a Δt ($k \geq 1$),

$$\frac{M_k(x, t, \Delta t)}{k!} = D^{(k)}(x, t) \Delta t + \mathcal{O}(\Delta t^2) \quad (2.31)$$

e levando em conta apenas os termos lineares em Δt , temos:

$$\frac{\partial f(x, t)}{\partial t} = \sum_{k=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^k D^{(k)}(x, t) f(x, t). \quad (2.32)$$

A Eq. (2.32) é conhecida como Expansão de Kramers-Moyal.

Para processos markovianos, $D^{(k)}$ não depende dos valores de $X(t')$ em instantes de tempo anteriores. Dessa forma, a Eq. (2.32) pode ser entendida como uma equação diferencial de primeira ordem e a densidade de probabilidade $f(x, t)$ pode ser totalmente determinada por integração com distribuição inicial $f(x, t_0)$ ($t > t_0$) e com condições de contorno adequadas.

A probabilidade de transição $f(x, t | x', t')$ é a FDP $f(x, t)$ para a condição inicial especial $f(x, t') = \delta(x - x')$. Assim, a probabilidade condicional também deve seguir a Eq. (2.32).

2.2.4

Equação de Fokker-Planck

Se a expansão de KM pode ser truncada após o segundo termo, a equação resultante é chamada de Equação de Fokker-Planck (EFP). Ela se escreve de forma geral:

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} [D^{(1)} f(x, t)] + \frac{\partial^2}{\partial x^2} [D^{(2)} f(x, t)]. \quad (2.33)$$

Neste formalismo $D^{(2)} > 0$ é chamado de coeficiente de difusão local e $D^{(1)}$ é um campo de força externo, conhecido como coeficiente de tendência. Esses coeficientes podem ou não depender do tempo. Matematicamente a Eq. (2.33) é um equação diferencial parcial de segunda ordem linear parabólica. Na literatura matemática, a Eq. (2.33) também é conhecida como *Equação de Kolmogorov*.

Um importante teorema relacionado à EFP é o teorema de Pawula [24]. De acordo com este teorema, se qualquer coeficiente $D^{(2s)} = 0$ para $s \geq 2$, todos os coeficientes $D^{(n)}$ com $n \geq 3$ devem ser zero, reduzindo naturalmente uma expansão de KM a uma equação EFP.

Os coeficientes de KM, $D^{(k)}(x, t)$ são definidos, a partir da Eq. (2.31) como:

$$D^{(k)}(x, t) = \lim_{\Delta t \rightarrow 0} \tilde{D}^{(k)}(x, t, \Delta t), \quad (2.34)$$

com

$$\tilde{D}^{(k)}(x, t, \Delta t) = \frac{M^{(k)}(x, t, \Delta t)}{k! \Delta t}, \quad (2.35)$$

sendo $M^{(k)}$ os momentos das FDPs, obtidos pela Eq. (2.28).

A EFP proporciona uma poderosa ferramenta para lidar com os problemas levantados até aqui e tem sido utilizada em muitos campos diferentes em ciências naturais, incluindo os de estado sólido e física de plasma, ótica quântica, reação cinética química e nuclear, biologia molecular e dinâmica populacional [24].

Esta abordagem também tem sido aplicada para muitas observáveis econômicas tais como taxas de câmbio [20, 26, 27] e preço de petróleo [28], reproduzindo com sucesso, e a evolução completa dos histogramas empíricos de retorno ao longo de diversas escalas temporais [20, 26, 29].

Quando a EFP não admite solução analítica, ela pode ser integrada numericamente. Uma forma de realizar essa integração numérica é utilizar variáveis discretas l , que podem ser definidas como $x_i = r l_i$, com tempos discretizados $t_i = l_i \Delta t$. Dessa forma, se os diferenciais forem aproximados por diferenças consistentes, resolver a EFP se reduz a iterações de uma equação de diferenças. Essas diferenças devem ser estáveis a ponto da probabilidade dos erros não crescerem.

2.2.5

Equação de Itô-Langevin

Muitos processos estocásticos podem ser representados mediante equações diferenciais estocásticas, do tipo

$$\dot{x} = D^{(1)} + \sqrt{2D^{(2)}}\eta(t), \quad (2.36)$$

onde η é um ruído branco, com $\langle \eta(t) \rangle = 0$ e $\langle \eta(t)\eta(t') \rangle = \Gamma\delta(t - t')$.

Esta equação, que tem a forma de uma equação de Langevin generalizada, tem sido utilizada há muitos anos para descrever problemas relacionados a movimentos Brownianos. Ela leva em consideração, além da contribuição determinística para a evolução da variável x , uma outra, que representa as forças flutuantes ou aleatórias.

É fácil mostrar que a EFP associada à equação estocástica (2.36), com a convenção de Itô (que será adotada nesta tese), é a Eq. (2.33) [24].

2.2.6

Processos de Poisson

Um tipo de processo estocástico muito frequente em processos físicos e que também será considerado nesta tese é o processo de Poisson. Seja um processo $\mathcal{P}_X(t)$ de tempo contínuo em que a variável aleatória X_t conta o número de vezes que ocorre um dado "evento" durante o intervalo $[0, t)$ e $p_n(t) = P(X_t = n)$, com $n = 0, 1, 2, \dots$. Um processo de Poisson é caracterizado pelas seguintes propriedades:

- as variáveis que representam o número de ocorrências em intervalos disjuntos são variáveis aleatórias independentes;
- se Y_t representa o número de ocorrências durante $[t_o, t_o + t)$, então para qualquer $t_o > 0$, X_t e Y_t tem a mesma FDP;
- $p_1(\Delta t) \simeq \lambda \Delta t$, onde λ é uma constante positiva, se $\Delta t > 0$ for suficientemente pequeno;
- $\sum_{k \geq 2} p_k(\Delta t) \simeq 0$, ou seja, a probabilidade de duas ou mais ocorrências durante um intervalo suficientemente pequeno é desprezível;
- $X_0 = 0$, ou seja, $p_0(0) = 1$.

Estas hipóteses permitem deduzir uma expressão para $p_n(t)$ chegando-se a $p_n(t) = e^{-\lambda t} (\lambda t)^n / n!$ [21].

Para um processo de Poisson a distribuição de intervalos de tempo T entre chegadas é a exponencial

$$f_T(t) = \frac{1}{\tau} e^{-t/\tau}, \quad (2.37)$$

onde $\tau = 1/\lambda$.

Consideremos agora eventos cuja ocorrência depende da realização de várias etapas (um número β de etapas) independentes e ocultas, cada uma resultante de um processo de Poisson. Então a distribuição de intervalos de tempo entre os eventos observáveis será a distribuição da soma de variáveis exponenciais. Por convolução temos

$$f_T(t) = \frac{1}{\tau \Gamma(\beta)} (t/\tau)^{\beta-1} e^{-t/\tau} \equiv \Gamma_{\beta, \tau}(t). \quad (2.38)$$

definida para $t \geq 0$. No caso particular $\beta = 1$, recupera-se a FDP exponencial.

2.3

Misturas Estatísticas

Analisando as interações de hádrons em cascatas de raios cósmicos [30], Wilk e Wlodarczyk observaram desvios da distribuição de momentos transversos em relação à distribuição exponencial esperada. Eles explicaram os desvios mediante possíveis flutuações da seção transversa σ e ainda determinaram que a grandeza

$$\omega = \frac{\langle \sigma^2 \rangle - \langle \sigma \rangle^2}{\langle \sigma \rangle^2}, \quad (2.39)$$

permitia medir o grau de afastamento da lei exponencial.

Eles obtiveram a lei empírica, supondo que a FDP observada podia ser derivada como

$$f(x) \propto \int_0^\infty d\sigma e^{-\sigma x} g(\sigma), \quad (2.40)$$

onde σ é o parâmetro flutuante com FDP $g(\sigma)$. Ou seja, a lei exponencial $e^{-\sigma x}$ pode ser vista como uma FDP condicionada dado o valor de σ ($f(x|\sigma) \propto e^{-\sigma x}$), de tal forma que

$$f(x) = \int_0^\infty d\sigma f(x|\sigma) g(\sigma). \quad (2.41)$$

Este tipo de composição de probabilidades é conhecida como *mistura estatística*.

A partir do trabalho de Wilk e Wlodarczyk [30], Cohen e Beck [31] consideraram sistemas de não-equilíbrio com dinâmica complexa em estados estacionários com grandes flutuações de quantidades intensivas (e.g. temperatura, potencial químico). Dependendo das propriedades estatísticas das flutuações, eles reproduziram uma descrição de mecânica estatística efetiva. Esta proposta generaliza a apresentada por Tsallis [10, 32], onde, em particular, o parâmetro entrópico q surge como $q = 1 + \omega$.

Desta perspectiva, a FDP estacionária de certos sistemas fora do equilíbrio surge do fator de Boltzmann efetivo

$$B(E) = \int_0^\infty d\beta f(\beta) e^{-\beta E}, \quad (2.42)$$

onde $f(\beta)$ é a distribuição da temperatura inversa β . Por isso esta abordagem é denominada *superestatística*, já que $B(E)$ representa a estatística da estatística $e^{-\beta E}$.

Cabe notar que muito antes desta interpretação de mecânica estatística, misturas estatísticas eram consideradas nas ciências econômicas e afins [33, 34].

3

Análise do Índice Acionário Brasileiro

3.1

Introdução

No Brasil, a primeira Bolsa de Valores foi fundada em 1890 com o nome de Bolsa Livre. Essa instituição pioneira passou por diversas reestruturações ao longo da história, modificando algumas vezes seu nome. Em 1967 passa a chamar-se Bolsa de Valores de São Paulo, ou simplesmente BOVESPA. Nela se negociam títulos emitidos por empresas privadas ou estatais no Brasil. Em 2007, passa a chamar-se BOVESPA Holding S.A. e, por fim, em 2008, há uma integração com a Bolsa de Mercadorias & Futuros, BM&F S.A., transformando-a na atual Bolsa de Valores, Mercadorias e Futuros, BM&FBOVESPA S.A., atualmente a terceira maior Bolsa mundial em valor de mercado, a US\$ 23,34 bilhões (nos preços de 2 de maio), atrás apenas da alemã Deutsche Börse (US\$ 31,41 bilhões) e da americana CME (Chicago Mercantile Exchange), com valor de US\$ 25,98 bilhões.

O Índice BOVESPA, ou IBOVESPA, é o índice que acompanha a evolução média das cotações das ações brasileiras. Sua composição é basicamente formulada da mesma forma que a maioria dos índices mundiais. É um valor atual, em moeda corrente, de uma carteira teórica de ações. Para que sua representatividade se mantenha ao longo do tempo, é feita uma reavaliação periódica, alterando-se a composição e peso da carteira. Considerando-se seu rigor metodológico e o fato de que a BM&FBOVESPA S.A. concentra mais de 90% do total de negócios do país, e é responsável por 100% das ações transacionadas no mercado brasileiro, trata-se do mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro disponível, permitindo tanto avaliações de curtíssimo prazo, como observações de expressivas séries de tempo.

Para a análise do IBOVESPA, selecionamos três conjuntos de dados: 3960 preços de fechamento, no período entre 02 de janeiro de 1991 e 28 de dezembro de 2006, 37984 valores do índice gravados a cada 15 minutos entre 21 de janeiro de 1998 e 31 de março de 2003 e 794310 valores do índice catalogados

a cada 30 segundos entre 01 de novembro de 2002 e 19 de julho de 2006. Todas as séries estão deflacionadas pelo índice diário do IGP-DI. O tempo foi contado em dias de pregão, descartando finais de semana, feriados e períodos sem atividade. Apresentamos abaixo na Fig. 3.1 as séries de preço diária e intradiárias utilizadas nesta tese.

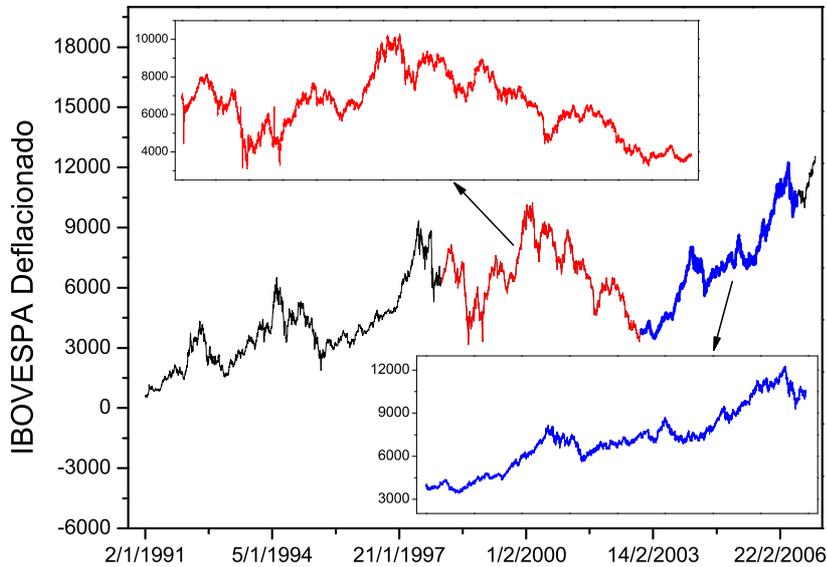


Figura 3.1: Representação das séries diária e intradiárias do IBOVESPA analisadas. Série intradiária de 15 minutos compreende período entre 21/01/98 e 31/03/03 (em vermelho) e série intradiária de 30 segundos relativa ao período entre 01/11/02 e 19/07/06 (em azul). Todas as séries estão deflacionadas pelo índice diário do IGP-DI.

3.2

Análise de Kramers-Moyal

Neste capítulo da tese, vamos investigar a evolução "temporal" das FDPs dos incrementos logarítmicos de preço ou retornos r , definidos na Eq. (1.2).

Consideramos, como equação de evolução geral para as FDPs dos retornos, a expansão de KM, válida para processos de Markov, apresentada na Eq. (2.32).

Em analogia com o tratamento de fenômenos turbulentos [20], sobre modelos de eventos que ocorrem em escalas hierárquicas, definimos, como escala temporal, uma escala logarítmica reversa, dada por

$$\tau = \log_2(\Delta t_0/\Delta t), \quad (3.1)$$

onde Δt_0 é uma escala arbitrária, definida aqui como 32 dias de negociação (ver também Apêndice A.2.2).

A utilização da escala logarítmica nos permite avançar mais rapidamente no tempo t . Notem que quanto maior a escala de tempo Δt , menor a escala hierárquica τ . A escala reversa surge do pressuposto de que a informação flui através de cascatas de energia a partir de escalas de tempo mais longas em direção a escalas mais finas. [20]

Os retornos medidos foram dados, por conveniência, em unidades de desvio padrão σ_{32} da respectiva série de dados em uma escala temporal (arbitrária) de $\Delta t_0 = 32$ dias.

De acordo com a Eq. (2.3), as FDPs condicionais, dadas por $P(r_2, \tau_2 | r_1, \tau_1) = P(r_2, \tau_2; r_1, \tau_1) / P(r_1, \tau_1)$ com $\tau_2 > \tau_1$, foram obtidas das séries reais construindo os histogramas para a probabilidade conjunta $P(r_2, \tau_2; r_1, \tau_1)$, computando os pares de retornos r_i que incidem no mesmo instante inicial.

A forma típica da FDP conjunta do IBOVESPA é apresentada na Fig. 3.2 bidimensionalmente para retornos diários e na Fig. 3.3, tridimensionalmente para retornos intradiários. A FDP conjunta descreve a dependência estatística dos incrementos em escalas temporais próximas.

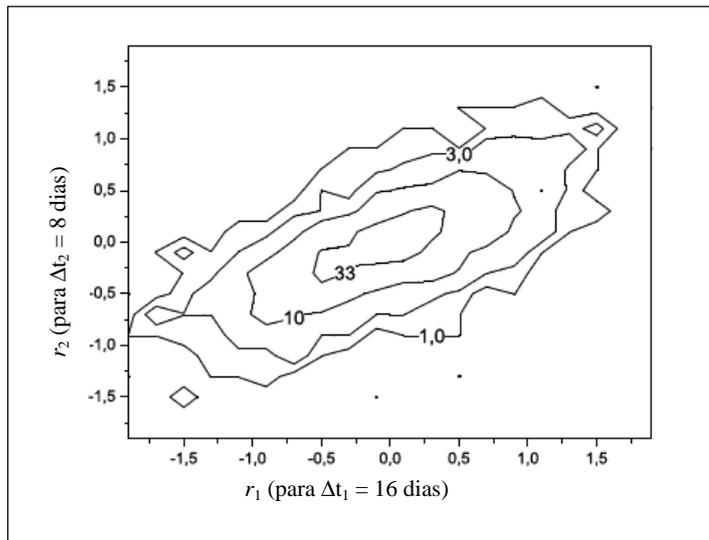


Figura 3.2: Curvas de nível típicas da FDP conjunta para retornos obtida a partir de dados diários do IBOVESPA para $\Delta t_1 = 16$ dias e $\Delta t_2 = 8$ dias. Níveis das curvas correspondem a $\log_{10} = -1,5; -2,0; -2,5; -3,0; -3,5$ do centro até a borda.

As propriedades markovianas do processo estocástico podem ser investigadas mediante a análise das probabilidades condicionais. Para tal, calculamos a equação de CK, a partir da Eq. (2.25):

$$P(r_2, \tau_2 | r_1, \tau_1) = \int dr' P(r_2, \tau_2 | r', \tau') P(r', \tau' | r_1, \tau_1), \quad (3.2)$$

onde $\tau_2 > \tau' > \tau_1$. A validade dessa equação foi examinada para diferentes tripletos (τ_2, τ', τ_1) . Como um exemplo ilustrativo, na Fig. 3.4, comparamos

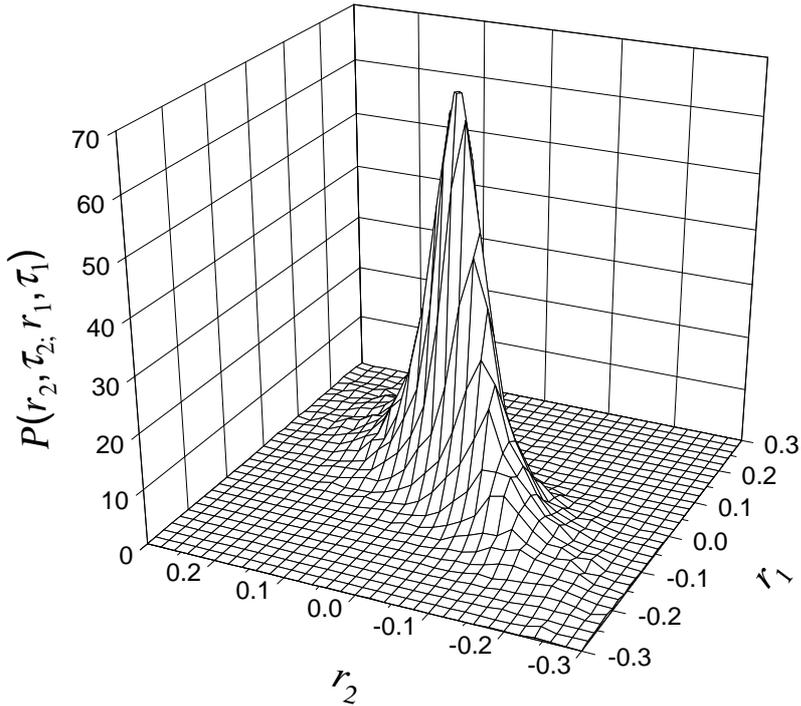


Figura 3.3: Distribuição conjunta típica para retornos obtida para $\tau_2 = 6,88$ ($\Delta t_2 = 120$ min) e $\tau_1 = 6,18$ ($\Delta t_1 = 195$ min).

as FDPs condicionais integradas e as obtidas diretamente (o lado direito e o esquerdo da Eq. (3.2), respectivamente). A boa concordância entre as FDPs condicionais calculadas diretamente e as integradas sustenta a markovianidade, validando dessa forma a nossa abordagem.

Os primeiros coeficientes $\tilde{D}^{(k)}$ foram computados, para $\tau = \frac{\tau_1 + \tau_2}{2}$, com $\Delta\tau = \tau_2 - \tau_1$, diretamente a partir dos dados, utilizando suas definições estatísticas, dadas pela Eq. (2.35). Para cada par de valores $(\tau, \Delta\tau)$, encontramos que $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$ como função de r , seguem, em muito boa aproximação, leis lineares e quadráticas, respectivamente, como ilustrado na Fig. 3.5. A saber,

$$\begin{aligned}\tilde{D}^{(1)} &= -\tilde{a}_1 r + \tilde{a}_0, \\ \tilde{D}^{(2)} &= \tilde{b}_2 [r]^2 + \tilde{b}_1 r + \tilde{b}_0.\end{aligned}\tag{3.3}$$

Mediante ajustes linear e quadrático para $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$, respectivamente, obtivemos os parâmetros $\{\tilde{a}_i, \tilde{b}_j\}$ para cada par $(\tau, \Delta\tau)$. Os valores limites $\{a_i, b_j\}$ para $\Delta\tau \rightarrow 0$ determinam a dependência em r dos coeficientes de tendência $D^{(1)}$ e de difusão $D^{(2)}$. Comportamentos similares para $D^{(1)}$ e $D^{(2)}$ foram observados para incrementos logarítmicos e lineares de índices de preço e para taxas de câmbio envolvendo mercados nos E.U.A., Alemanha e

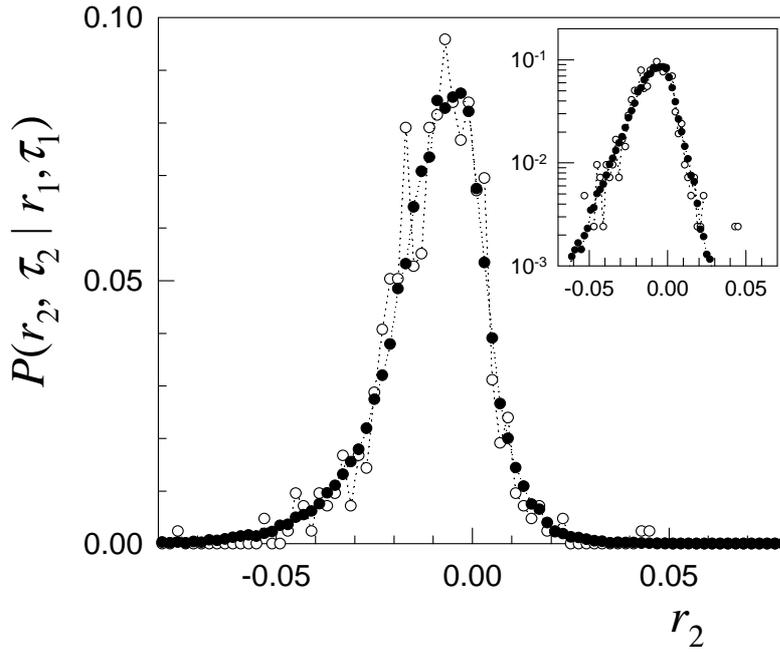


Figura 3.4: Corte de $P(r_2, \tau_2 | r_1, \tau_1)$ em $r_1 = 3\sigma_{\tau_2}$ obtido para $\tau_2 = 11,75$ ($\Delta t_2 = 4$ min), $\tau' = 10,75$ ($\Delta t' = 8$ min) e $\tau_1 = 9,75$ ($\Delta t_1 = 16$ min). Os símbolos cheios correspondem à FDP condicional calculada diretamente e os símbolos abertos à FDP integrada. Gráfico inserido: representação log-linear dos mesmos dados.

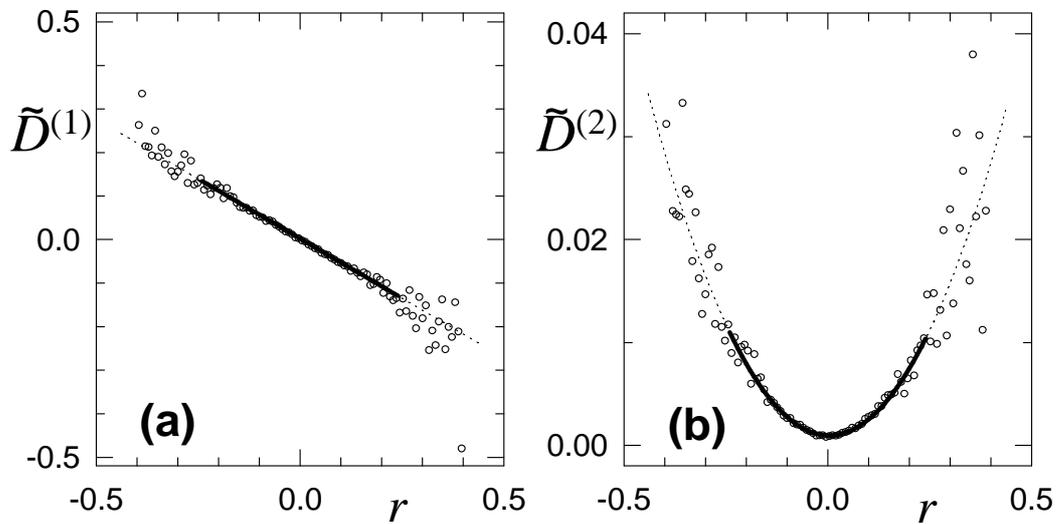


Figura 3.5: Coeficientes $\tilde{D}^{(1)}$ (a) e $\tilde{D}^{(2)}$ (b) em função de r , calculados em $\Delta t_2 = 109$ min e $\Delta t_1 = 125$ min (logo $\tau = 6,9$ e $\Delta\tau = 0,2$). As linhas cheias correspondem aos intervalos utilizados nos ajustes.

Japão [20, 35, 36].

Obtivemos também os coeficientes de quarta ordem $\tilde{D}^{(4)}$ em função de r , e realizamos um ajuste polinomial de quarta ordem, para cada par $(\tau, \Delta\tau)$. Observamos que ao diminuir $\Delta\tau$, com τ fixo, as curvas de $\tilde{D}^{(4)}(\Delta x, \tau, \Delta\tau)$ se

achata, como apresentado na Fig. 3.6. Supondo o limite $D^{(4)}(\Delta x, \tau)$ em (2.34) desprezivelmente pequeno, de acordo com o teorema de Pawula [24], a expansão de KM (2.32) pode ser truncada após a segunda ordem, reduzindo-a dessa forma a uma EFP.

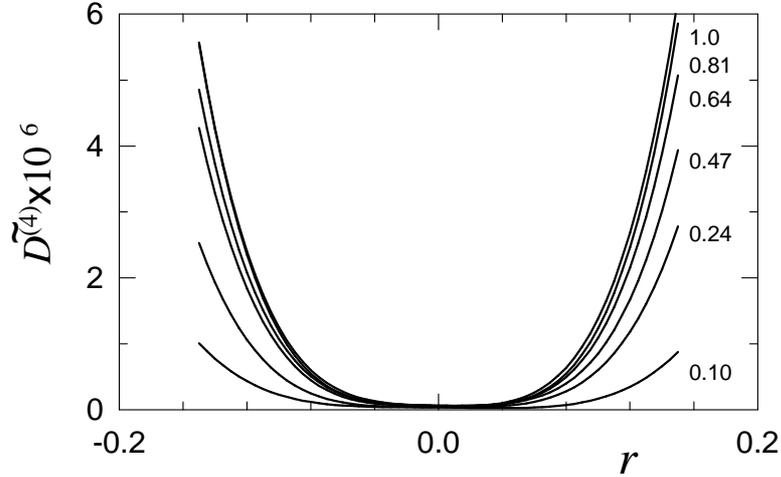


Figura 3.6: Coeficiente $\tilde{D}^{(4)}$ em função de r , calculado em $\tau = 8,35$ e diferentes valores de $\Delta\tau$, indicados na figura. Linhas sólidas correspondem a ajustes polinomiais de quarta ordem nos dados.

É importante mencionar que as escalas de tempo das amostras impõem limites naturais ao valor mínimo de $\Delta\tau$ acessível empiricamente. Por exemplo, a escala diária dos dados impõe o limite de 1 dia para a variação entre as escalas Δt_1 e Δt_2 consideradas na Eq.(2.34). Assim, $\Delta\tau = 0,1$, ($\Delta t_1/\Delta t_2 \simeq 1,07$) só pode ser atingido para $\tau \leq 1,09$ ($\Delta t \geq 15$ dias), enquanto $\Delta\tau = 0,2$ ($\Delta t_1/\Delta t_2 \simeq 1,15$) só é acessível quando $\tau \leq 2$ ($\Delta t \geq 8$ dias).

Levando-se em conta estas restrições, observa-se que dentro das flutuações estatísticas, o comportamento limite dos parâmetros pode ser estimado considerando os resultados obtidos para $\Delta\tau = 0,2$ como aqueles que definem $\{a_i, b_j\}$ para cada τ . A dependência em τ de \tilde{a}_i e \tilde{b}_i é exemplificada no Apêndice A.2.1, para diferentes valores de $\Delta\tau$.

3.3

Dependência temporal dos parâmetros

A dependência dos parâmetros limites $\{a_i, b_j\}$ de acordo com a escala τ é apresentada nas Figs. 3.7-3.8. Apresentamos essa dependência nas Eqs. (3.4)-(3.8):

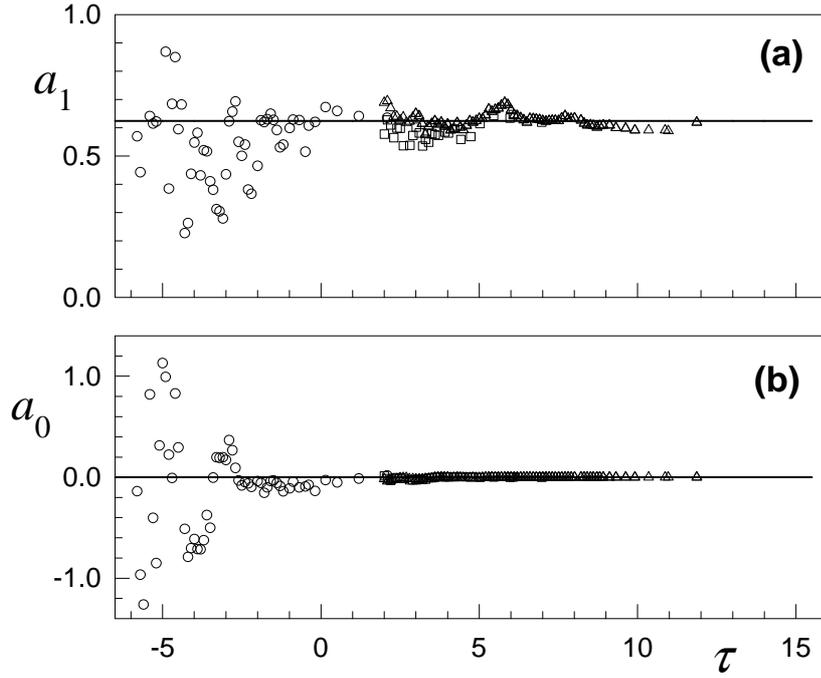


Figura 3.7: Dependência dos parâmetros a_1 e a_0 do coeficiente de tendência $D^{(1)}(r, \tau)$ com a escala temporal τ . A análise foi realizada com dados de séries diárias (círculos), de 15 minutos (quadrados) e de 30 segundos (triângulos).

$$\begin{aligned} a_0 &= 0, \\ a_1 &= -0,642, \end{aligned} \quad (3.4)$$

$$b_1 = 0,5 \left(1 - (1 + \exp(-\tau + 3,5/0,2))^{-1} (1 + \exp(\tau + 2,5/1,7))^{-1} - 1 \right), \quad (3.5)$$

$$b_2 = 0,1364 \left(1 - (1 + \exp(\tau + 2,5/0,49))^{-1} \right), \quad (3.6)$$

$$b_0 = 0,2/2^\tau, \quad (3.7)$$

$$b_0 = 0,3/2^{1,17\tau}. \quad (3.8)$$

Eq. (3.7) foi usada para $\tau < 6,6$ e Eq. (3.8) para $\tau \geq 6,6$.

Os parâmetros de ajuste a_0 e a_1 permanecem aproximadamente constantes ao longo das escalas diária e intradiárias. As flutuações observadas nas figuras para τ pequeno são devidas à menor significância estatística dos resultados.

O parâmetro de difusão b_2 aumenta com τ desde próximo de zero, em escalas mensais, até um valor limite, no regime de alta frequência,

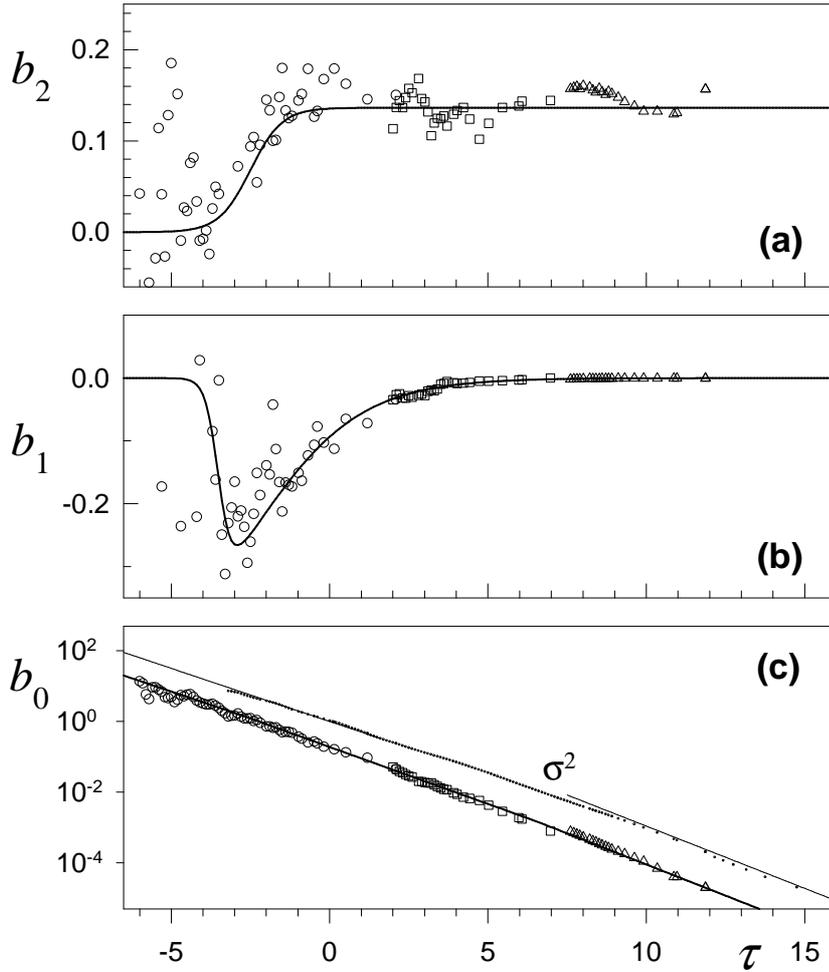


Figura 3.8: Dependência dos parâmetros b_2, b_1 e b_0 do coeficiente de difusão $D^{(2)}(r, \tau)$ com a escala temporal τ . Símbolos como na Fig. 3.7. Em (c), incluímos para comparação $\sigma^2(\tau) = \sigma_\tau^2 / \sigma_{32}^2$ (símbolos pequenos) junto com seu comportamento assintótico (linhas finas) para escalas de tempo longas e curtas como previsto pela Eq. (3.14), com $\gamma = 1$ e $1,17$, respectivamente.

sinalizando uma transição na dinâmica subjacente dos retornos. Por outro lado, o parâmetro b_1 apresenta valores não nulos apenas em escalas semanal/mensal, indicando desvios de simetria das FDPs.

Entretanto, o parâmetro de difusão b_0 apresenta um decaimento exponencial persistente à medida que τ aumenta, caracterizado por $b_0(\tau) \approx 2^{-\gamma\tau}$. Comportamento semelhante de b_0 foi relatado para taxas de câmbio [20], embora, para um menor alcance de escalas temporais. É importante ressaltar que, apesar do decaimento exponencial de b_0 , ele não pode ser genericamente negligenciado para grandes escalas em τ pois ele representa a amplitude da componente aditiva do ruído, que proporciona estabilidade ao processo estocástico para pequenos r . De fato, a variância σ_τ^2 também segue um decaimento exponencial do tipo $2^{-\gamma\tau}$, como mostrado na Fig. 3.8(c), definindo assim o nível de referência para b_0 . Dois regimes associados a decaimentos um

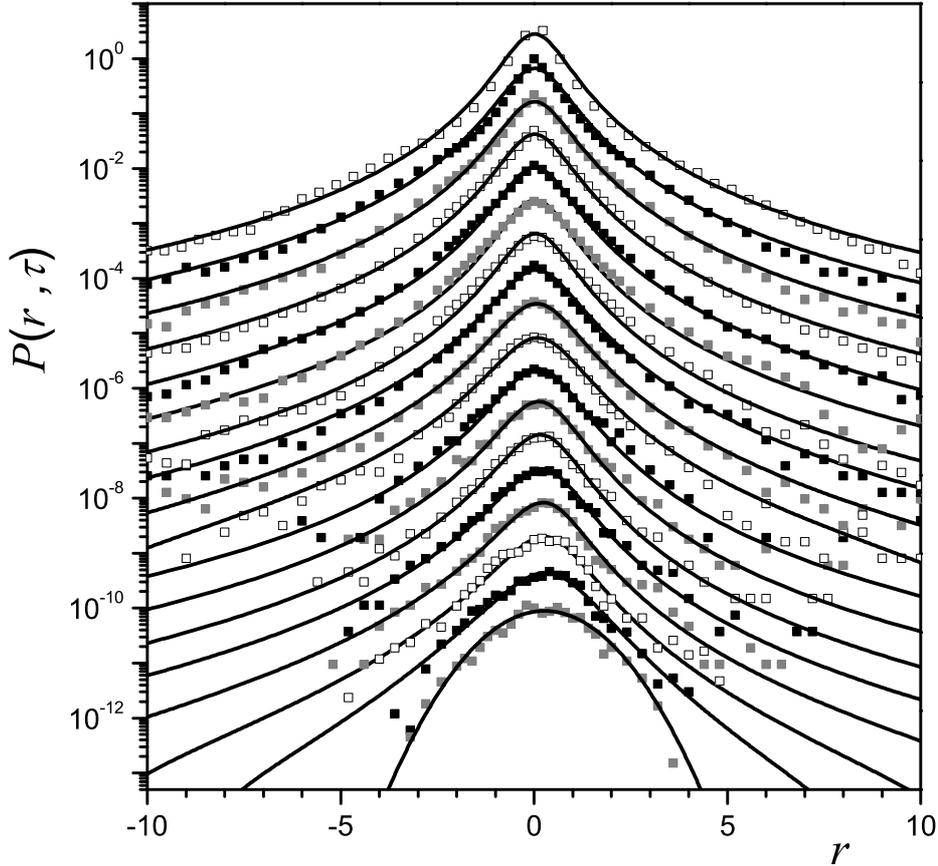


Figura 3.9: FDPs dos retornos normalizados. Comparação entre as soluções numéricas da EFP (3.9) (linhas cheias) e os dados empíricos (símbolos). As escalas temporais correspondem de $\tau = -2$ a $\tau \approx 15$. As FDPs estão deslocadas verticalmente, para melhor visualização. A condição inicial foi um ajuste gaussiano aos dados para $\tau = -2$ ($\Delta t=128$ dias). As escalas τ apresentadas são informadas na tabela A.5 (apêndice).

pouco diferentes são observados para as duas quantidades (ver Fig. 3.8(c)), sugerindo que $b_0(\tau)$ esteja relacionado à σ_τ^2 . No limite de τ pequeno ambas as quantidades são caracterizadas por $\gamma = 1$, o que corresponde ao regime difusivo normal, enquanto que em escalas de alta frequência, $\gamma > 1$, correspondente a um regime superdifusivo. Para $\tau > 11$, a impossibilidade de obtenção de resultados para $\Delta\tau$ suficientemente pequenos impedem a estimativa dos parâmetros, segundo a Eq. (2.34).

3.4 Evolução temporal

A partir da EFP (2.33), mapeando $x \rightarrow r$ e $t \rightarrow \tau$, a equação de evolução resultante se escreve explicitamente

$$\partial_\tau P = \partial_r(a_1 P) + \partial_r^2((b_0 + b_1 r + b_2 r^2)P), \quad (3.9)$$

onde a dependência em τ dos parâmetros foi suavizada através dos *ansätze* representados nas Figs. 3.7-3.8. A Eq. (3.9) foi integrada numericamente por meio de um esquema FTCS (do inglês *Forward-Time Centred-Space*) [37].

Um ajuste gaussiano no histograma empírico em $\Delta t=128$ dias ($\tau = -2$) foi utilizado como condição inicial. A evolução foi conduzida em direção à escala de 30 segundos, a mais alta resolução temporal dos nossos dados. Para $\tau > 11$, o prosseguimento da evolução da EFP, foi realizado pela extrapolação da dependência em τ dos coeficientes a partir dos *ansätze* utilizados. Na Fig. 3.9, mostramos as FDPs de retornos normalizados (reescalados por σ_τ) gerados pela EFP, juntamente com os dados empíricos. Sua adequação é notavelmente boa em toda a região de dados, validando fortemente a nossa estimativa dos coeficientes de KM.

Dentro da dinâmica de Langevin, $D^{(1)}$ e $D^{(2)}$ estão relacionados com as forças deterministas e aleatórias, respectivamente [24]. Apesar do amplo leque de escalas de tempo analisadas, a intensidade da força harmônica restauradora, dada por $D^{(1)}(r, \tau) \simeq -a_1(\tau)r$, manteve-se praticamente constante para $\tau > -2$. Isto significa que os mecanismos de relaxação dos mercados, regidos entre outros fatores, por limite de capital e aversão ao risco, são semelhantes em diversas escalas temporais.

O coeficiente de difusão no entanto, apresenta características diferenciadas de acordo com a escala temporal. Para a maioria das escalas temporais, $D^{(2)}(r, \tau) \simeq b_0(\tau) + b_2(\tau) r^2$ é dominado pelo termo independente e pelo termo quadrático, associados respectivamente, às componentes aditiva e multiplicativa do ruído. Devido ao caráter acumulativo das flutuações, a componente aditiva b_0 aumenta com Δt . Entretanto, a mudança de b_2 para um patamar mais elevado para pequenos intervalos Δt indica um grande efeito multiplicativo nessa região temporal. Isto significa que o comportamento endógeno do mercado, que cria espontaneamente a amplificação dos mecanismos de resposta a flutuações de preços, apresenta diferentes níveis para as escalas micro e macroscópicas. A presença de ruído multiplicativo é um conhecido mecanismo gerador de FDPs estacionárias com caudas gordas [38, 39, 40]. Verificou-se que para um grande conjunto de parâmetros de controle, caudas em lei de potência prevalecem, cujo expoente depende da razão a_1/b_2 , sendo independente de b_0 .

À luz desses últimos resultados, os platôs observados para a_1 e b_2 nas Figs. 3.7 e 3.8 sugerem que a FDP seja invariante com relação à τ em ambos os limites assintóticos. De fato, a EFP (3.9) admite soluções τ -invariantes como será discutido a seguir.

De acordo com nossos resultados empíricos, para escalas de tempo longas

e curtas, assumindo valores estacionários para a_1 e b_2 e uma lei exponencial para b_0 , temos:

$$b_0(\tau) = A2^{-\gamma\tau}, \quad (3.10)$$

e negligenciando b_1 , a Eq. (3.9) é reescrita como:

$$\partial_\tau P = \partial_r(a_1 P) + \partial_r^2((b_0(\tau) + b_2 r^2)P). \quad (3.11)$$

A solução invariante da EFP (3.11) é (ver Apêndice A.1.1):

$$P(r, \tau) \sim 1/(b_0(\tau) + b_2 r^2)^{(\mu+1)/2}, \text{ com} \quad (3.12)$$

$$\mu = 1 + (a_1 - B/2)/b_2, \quad (3.13)$$

onde $B \equiv -b'_0/b_0 \equiv \gamma \ln 2$ não depende do tempo.

Comparando-se com a Eq. (2.10), a solução (3.12) é uma q -gaussiana com $q = 1 + 2/(\mu + 1)$ (ver Apêndice A.1.1). No caso particular $b_2 \rightarrow 0$, as soluções se reduzem à forma gaussiana ($q=1$).

A variância reescalada $\sigma^2(\tau) \equiv \sigma_\tau^2/\sigma_{32}^2$ das soluções invariantes é (ver no Apêndice A.1.1):

$$\sigma^2(\tau) = b_0(\tau)/(a_1 - B/2 - b_2). \quad (3.14)$$

No limite de grandes escalas de tempo Δt , o coeficiente de difusão é dominado pelo termo estacionário independente, b_0 , que obedece a Eq. (3.10), com $A \equiv b_0(0)$. A substituição dos valores numéricos de $b_0(0)$ e a_1 , na Eq. (3.14) retorna, em muito boa aproximação¹, $\sigma^2(\tau) = 2^{-\tau}$ (difusão normal na escala linear de tempo) de acordo com os resultados numéricos (mostrados na Fig. 3.8(c)). Como consequência, nesse limite, a equação de evolução recupera a gaussianidade, governada por um equilíbrio entre a força harmônica determinística e o ruído aditivo dependente do tempo.

No limite oposto para grandes τ , b_2 possui um valor estacionário não nulo, enquanto $b_0(\tau)$ evolui de acordo com a Eq. (3.10), com $\gamma > 1$. Nesse caso, a Eq. (3.14) também prevê um comportamento assintótico para $\sigma^2(\tau)$ em acordo com os valores empíricos, conforme mostrado na Fig. 3.8(c). Como a evolução dos demais parâmetros, com exceção de b_0 , é lenta, também investigamos se as FDPs poderiam ser efetivamente descritas pelo *ansatz* (3.12) em um regime temporal estendido.

Os resultados dos ajustes (por mínimos quadrados) estão em boa concordância com às FDPs empíricas, para praticamente todos as escalas τ ,

¹Veja que $\sigma^2(0) = 1$ por construção.

como podemos verificar na Fig. 3.10. Para $\tau \approx 0$, os valores não nulos de b_1 , impõem uma correção à forma q -gaussiana, devido à presença de assimetria.

Os valores de μ resultantes dos ajustes para as q -gaussianas são apresentados na Fig. 3.11, juntamente com os valores assintóticos obtidos a partir da Eq. (3.13). Para pequenos valores de τ , os valores crescentes dos μ efetivos apontam para a latência do regime gaussiano. Para grandes valores de τ , os expoentes de cauda empíricos tendem para um valor estacionário em bom acordo com os valores teóricos.

Caudas em lei de potência são frequentemente citadas na literatura para descrever ativos financeiros em regimes de alta frequência [1, 41, 42]. O *ansatz* da forma q -gaussiana já havia sido proposto para dados de alta-frequência [1, 12, 43] e para retornos logarítmicos diários [44], através de uma abordagem fenomenológica. No entanto, em nossa análise, elas surgem naturalmente da equação de evolução obtida empiricamente através do cálculo dos coeficientes de KM ao longo das escalas temporais.

Lembremos que os resultados da evolução das FDPs para $\tau > 11$ ($\Delta t < 4$ minutos), mostrados na Fig. 3.9, foram gerados pela extrapolação da

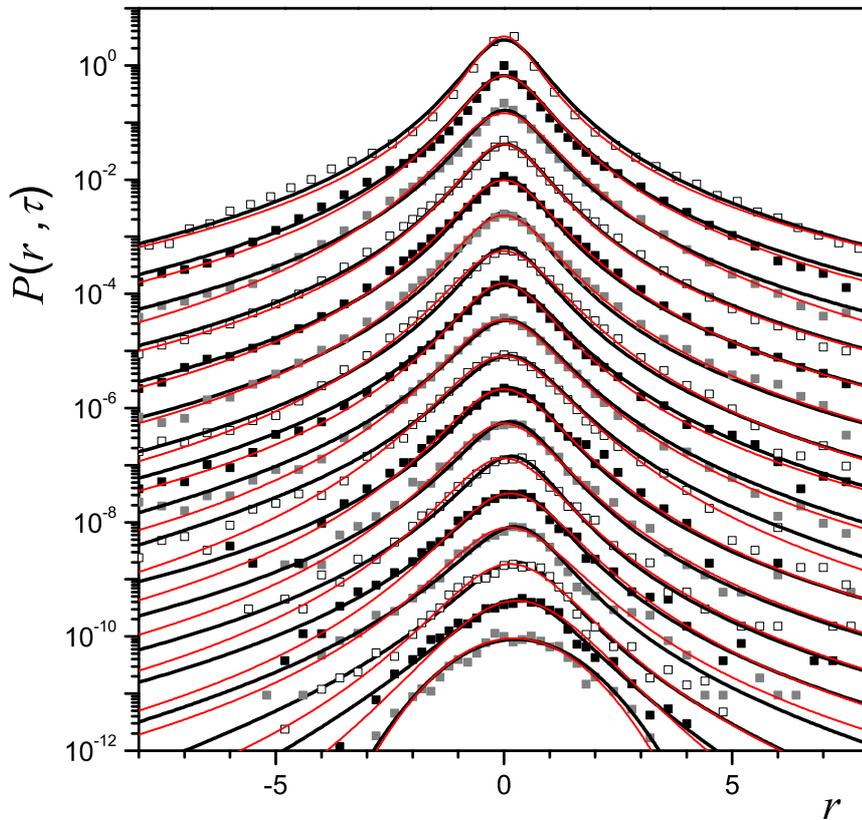


Figura 3.10: FDPs empíricas dos retornos normalizados (símbolos), soluções numéricas da EFP (3.9) (linhas pretas) e ajustes q -gaussianos aos dados empíricos (linhas vermelhas).

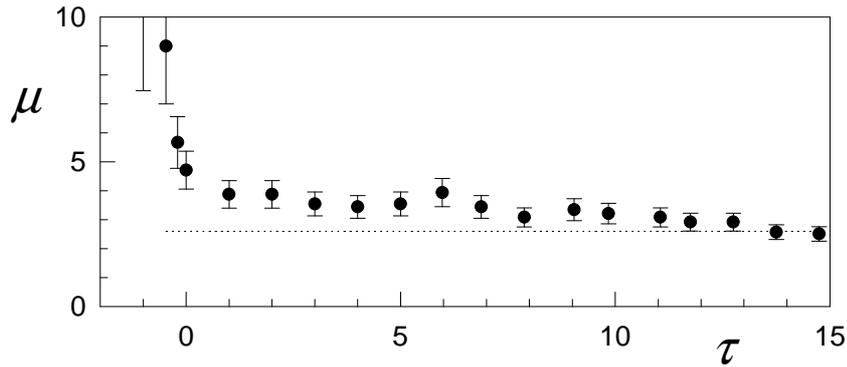


Figura 3.11: Dependência do expoente μ em relação às escalas temporais τ . Símbolos correspondem ao ajuste da Eq. (3.12) aos histogramas empíricos, linhas pontilhadas representam os valores assintóticos dados pela Eq. (3.13).

dependência em τ dos coeficientes obtidos, que no entanto, forneceram uma boa previsão dos histogramas empíricos até $\tau = 14,75$ ($\Delta t = 30$ segundos). Se por um lado, isso sugere a confiabilidade dos parâmetros estimados, por outro, evidencia a previsibilidade das estatísticas intradiárias, devido à existência de efeitos de memória na dinâmica dos retornos de preços no regime de alta frequência. Vale a pena mencionar que o desvio observado do histograma empírico para as menores escalas de tempo analisadas (30 segundos) exprime o início de um regime não-markoviano [45].

3.5

Considerações finais

Do conhecimento dos coeficientes de KM, é possível acessar quantitativamente as leis determinísticas e estocásticas que governam a dinâmica dos preços de mercado.

Esta abordagem tem sido aplicada para muitas grandezas econômicas tais como índices de mercado [35, 44], taxas de câmbio [20, 26, 27] e preço de petróleo [28], reproduzindo com sucesso a evolução completa dos histogramas empíricos de retorno ao longo de diversas escalas temporais [20, 26].

Nosso estudo revela padrões universais na dependência funcional dos coeficientes de KM, que podem ser descritos por um pequeno número de parâmetros microscópicos.

Neste capítulo do trabalho, também apresentamos soluções assintóticas consistentes para a equação de evolução temporal que emerge a partir da análise empírica. Dessas soluções surgem novas relações matemáticas conectando características das FDPs, tais como o expoente das caudas, aos parâmetros dos coeficientes de KM.

Os resultados revelam um caminho dinâmico que conduz de estatísticas gaussianas a caudas em lei de potência, oferecendo perspectivas de descrição de outros sistemas complexos onde transições semelhantes são observadas.

O conteúdo deste capítulo pode também ser encontrado na Ref. [29].

4

Análise Comparativa dos Índices de Mercados Mundiais

4.1

Introdução

A detecção de padrões universais dentre os diferentes mercados financeiros é importante pois revela os mecanismos fundamentais que governam a dinâmica dos preços. Por outro lado, é também importante identificar as características próprias de cada mercado, tais como eficiência, resistência a pressões externas e a especulação internacional, ou o tempo característico de dissipação dos efeitos provocados por choques macroeconômicos, para efetuar comparações entre eles.

Estudos comparativos já foram realizados utilizando diversas quantidades como volatilidade [46, 47, 48, 49], expoentes de Hurst [46, 50, 51, 52, 53, 54] e outras propriedades estatísticas dos mercados [55, 56]. Neste capítulo temos como objetivo mostrar que a obtenção dos coeficientes de KM diretamente das séries temporais empíricas é uma técnica alternativa para comparar as flutuações dinâmicas dos índices de mercado.

Com efeito, após termos verificado a robustez da expansão de KM para o índice do mercado brasileiro, estendemos nossa análise para outros mercados mundiais. Realizamos uma análise comparativa de 24 índices diários de Bolsas ao redor do mundo, englobando mercados desenvolvidos e emergentes.

Nesta parte da tese, visamos elucidar características comuns e distintas da dinâmica dos índices de mercado e, potencialmente, ranquear ou segregar os mercados, de acordo com parâmetros microscópicos que caracterizem os coeficientes de KM.

4.2

Índices de Mercados Mundiais

Existem diversos índices importantes no mundo. Abordaremos no nosso trabalho, 24 índices ao redor do mundo (listados na tabela 4.1). Este conjunto engloba todos os continentes e inclui tanto mercados desenvolvidos como

emergentes ¹.

Como exemplo, podemos citar os mais abordados na literatura científica:

- New York Stock Exchange Composite Index (NYSE) - É o índice da Bolsa de Valores de Nova York, a maior, mais importante e tradicional Bolsa de Valores do mundo em volume de ações negociadas, com U\$21,79 trilhões de dólares negociados em 2006. Localizada na Wall Street, Nova York, ali são negociados títulos e ações das principais empresas dos Estados Unidos e do mundo.
- S&P 500 - Índice calculado pela consultoria americana Standard & Poor's que reflete o desempenho de 500 maiores ações atualmente negociadas nos Estados Unidos.
- Dow Jones Industrial Average (DJIA) - Índice americano baseado na cotação das 30 maiores e mais importantes empresas industriais do país.
- National Association of Security Dealers Automated Quotation System (Nasdaq) - É o índice da primeira Bolsa do mundo a negociar exclusivamente ações de empresas de Internet, informática e de alta tecnologia. Como as operações de compra e venda são feitas por meio de computadores no mercado de balcão, a Nasdaq ficou conhecida como a "Bolsa Eletrônica". Ela é atualmente composta por mais de 3000 ações diferentes.
- Nikkei 225 - Índice da Bolsa de Valores de Tóquio, que reflete o preço das 225 ações mais negociadas no mercado japonês.
- DAX - Principal índice alemão, baseado no preço das 30 ações mais negociadas na Bolsa de Valores de Frankfurt.

As séries dos índices de preço e dos respectivos retornos diários dos mercados considerados são apresentadas na Fig. 4.1. Todas as séries de preço foram coletadas do website da Yahoo finance².

Para compararmos os diversos mercados ao redor do mundo, é necessário considerar um período próximo ou, se possível, um mesmo período de tempo. Somente há informação sobre um grande número de mercados mundiais a partir de 1997. Dessa forma, consideramos o período de 1997 a 2007 neste estudo. Os índices BSE30, CMA, JSXC, KOSPI e TWI foram considerados, devido a disponibilidade dos dados, a partir em julho de 1997 e os demais a partir de janeiro de 1997. Os retornos diários para o período analisado na tese são apresentados na Fig. 4.2.

¹Usamos as classificações de mercados utilizadas em <http://globaledge.msu.edu>. Outras fontes, tais como as <http://www.ftse.com> ou <http://www.mscibarra.com>, coincidem, com exceção de Hong-Kong e Cingapura que são classificadas como desenvolvidos.

²yahoo.finance.com

Rótulo	índice	país	# dias
1	AEX	Holanda	2678
2	ATX	Áustria	2597
3	BEL 20	Bélgica	2669
4	CAC 40	França	2669
5	DAX	Alemanha	2662
6	FTSE 100	Reino Unido	2652
7	SMI	Suíça	2647
8	BSE 30	Índia	2477
9	HSI	Hong Kong	2599
10	JSXC	Indonésia	2428
11	KLSEC	Malásia	2591
12	KOSPI	Coréia do Sul	2490
13	Nikkei 225	Japão	2581
14	STI	Cingapura	2643
15	TWI	Taiwan	2462
16	DJIA	EUA	2666
17	Nasdaq	EUA	2652
18	NYSE	EUA	2652
19	S&P 500	EUA	2645
20	IBOVESPA	Brasil	2596
21	IPC	México	2631
22	Merval	Argentina	2597
23	CMA	Egito	1964
24	AOX	Austrália	2666

Tabela 4.1: Índices mundiais e seus correspondentes rótulo, abreviatura, país e número de dias de negociação entre 1997 e 2007.

4.3

Análise Preliminar

Nesta tese, o tempo foi contado em dias de pregão, descartando finais de semana e feriados. Os retornos medidos foram dados em unidades de desvio padrão σ_{32} da respectiva série de dados na escala temporal de 32 dias.

A estacionariedade dos retornos foi verificada através de dois métodos usualmente utilizados em testes econométricos: Augmented Dickey-Fuller(ADF) e Philip-Perron(PP) [57] (ver Apêndice A.1.4). Embora algumas séries apresentem focos repentinos de volatilidade, ambos os testes forneceram diagnósticos positivos para estacionariedade para todos os mercados examinados. Isso significa que as séries de dados permitem considerar, pelo menos em um bom grau de aproximação, medidas estacionárias. Em qualquer caso, os resultados podem ser adotados como médios ou efetivos para cada série completa.

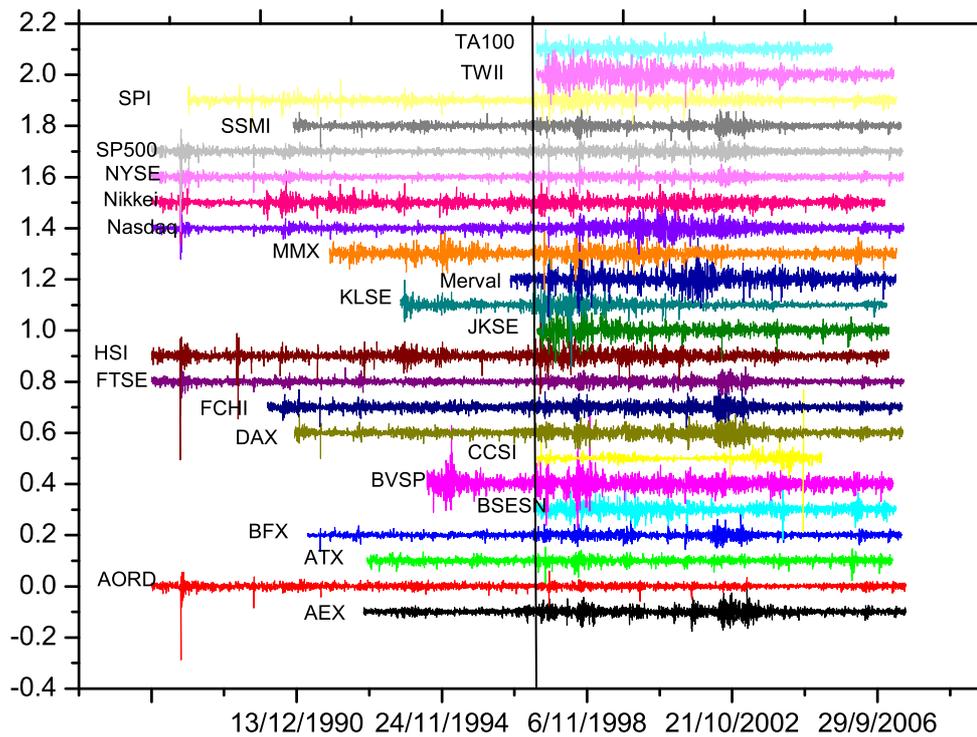


Figura 4.1: Séries de retorno de preço dos índices mundiais apresentados em legenda. Linha vertical representa o início do ano de 1997. As séries foram deslocados verticalmente para melhor visualização.

Com o objetivo de investigar as propriedades markovianas dos processos estocásticos, calculamos a equação de CK, apresentada na Eq. (2.25).

Como no caso do mercado brasileiro, a validade da equação de CK foi examinada para diferentes tripletos (τ_2, τ', τ_1) , ao comparar FDPs condicionais calculadas diretamente a partir de $P(r_2, \tau_2 | r_1, \tau_1)$ com as FDPs integradas calculadas com o membro direito da Eq. (3.2). Como um exemplo ilustrativo, na Fig. 4.3, comparamos cortes para valores escolhidos de r_1 para duas FDPs diferentes computadas para o índice de Cingapura STI. Dentro dos erros estatísticos, os histogramas coincidem, providenciando evidências para a validade da Eq. (3.2). Resultados similares foram encontrados para todos os mercados analisados.

A verificação direta da equação de CK pode também ser utilizada para estimar a escala temporal mínima na qual as séries de retornos podem ser aproximadas por um processo markoviano [28]. Escolhendo diferentes r_1 , realizamos o teste para vários valores fixos de τ_1 enquanto diminuimos $\tau_2 - \tau'$. Não verificamos desvios significativos até a escala de resolução mais curta de nossos dados.

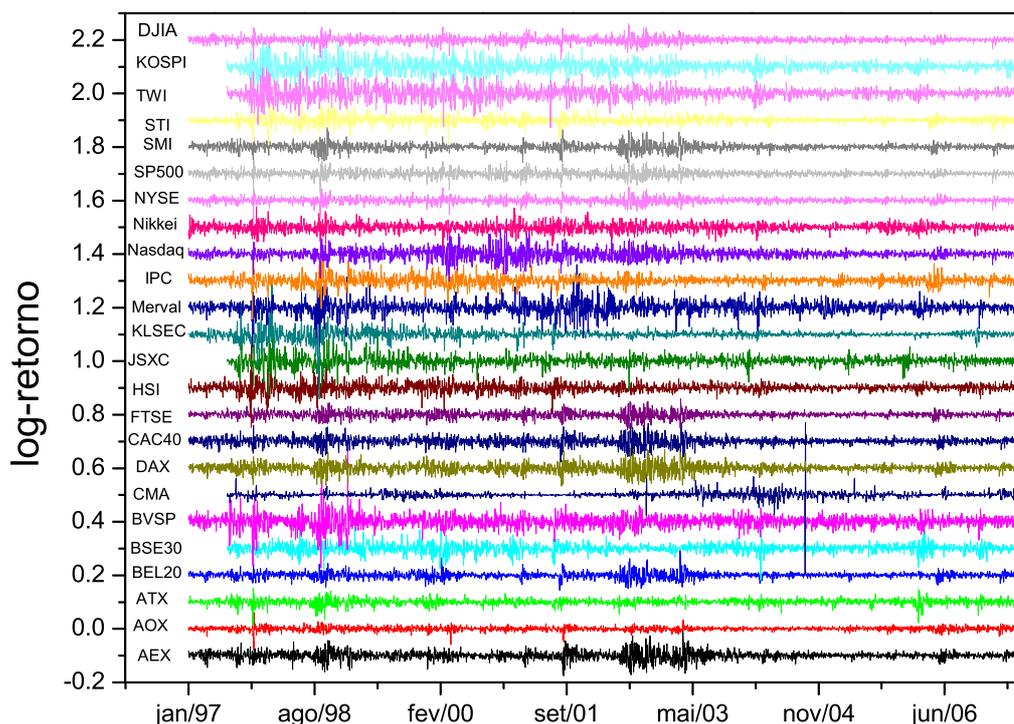


Figura 4.2: Séries de retornos diários dos índices de mercado mundiais (apresentados ao lado do respectivo índice) no período utilizado nesta tese. As séries foram deslocadas verticalmente para melhor visualização.

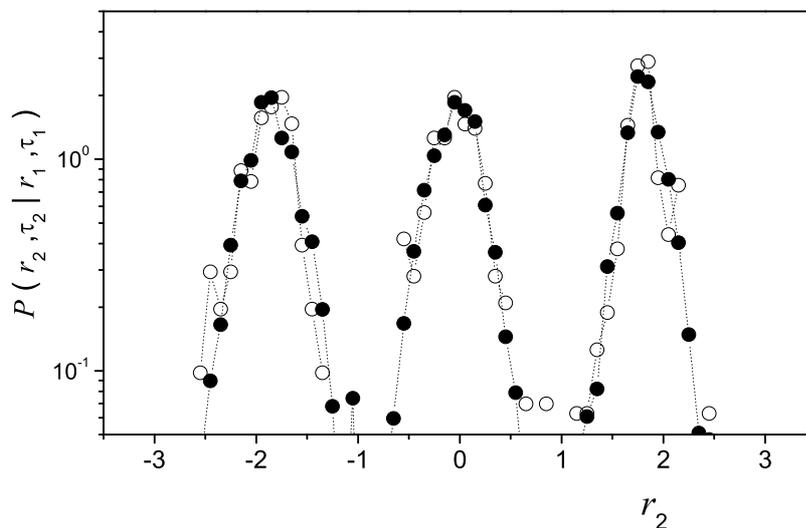


Figura 4.3: Cortes de nível através de $P(r_2, \tau_2 | r_1, \tau_1)$ em $r_1 = -2,0, 0,5$ e $3,0$ (da esquerda para a direita) para $\tau_2 = 0,24$, $\tau' = 0,14$ e $\tau_1 = 0,04$, para o índice de Cingapura STI. Os histogramas foram deslocados horizontalmente para melhor visualização. Os símbolos preenchidos correspondem às FDPs condicionais calculadas diretamente e os símbolos abertos às FDPs integradas de acordo com a Eq. (3.2).

4.4 Análise de Kramers-Moyal

Uma vez atestada a validade da expansão de KM para nossos dados, passamos ao cálculo de seus coeficientes. Encontramos que as séries de dados analisadas apresentam, para todo $-4 < \tau < 4$ (512 dias $< \Delta t < 2$ dias) e $\Delta\tau$ suficientemente pequeno, um comportamento comum dos coeficientes $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$ definidos na Eq. (2.35). Nas Figs. 4.4.(a) e 4.4.(b), ilustramos as formas típicas desses coeficientes em função de r .

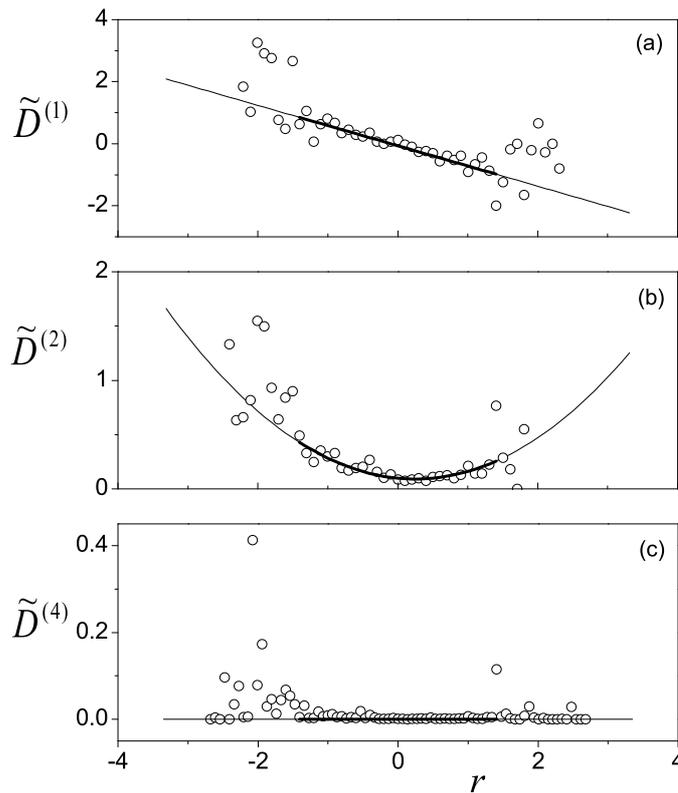


Figura 4.4: Coeficientes $\tilde{D}^{(1)}$ (a), $\tilde{D}^{(2)}$ (b) e $\tilde{D}^{(4)}$ (c), em função de r , calculados em $\tau = 1, 2$ (escala temporal semanal) e $\Delta\tau = 0, 2$ para o índice alemão DAX. As linhas cheias correspondem às curvas ajustadas e as mais espessas aos intervalos utilizados para os ajustes.

De acordo com a Eq. (2.34), os valores de $\{\tilde{a}_i, \tilde{b}_j\}$, definidos nas Eqs. (3.3), determinam, no limite $\Delta\tau \rightarrow 0$, para um dado τ , os parâmetros $\{a_i, b_j\}$ que descrevem a dependência em r dos coeficientes $D^{(1)}$ e $D^{(2)}$. Deste modo, em concordância com os resultados apresentados na literatura para um grupo restrito de mercados [20, 26, 29, 35, 44], obtivemos que a dependência em r desses coeficientes é bem descrita por formas lineares e quadráticas respectivamente.

O coeficiente de quarta ordem $D^{(4)}$ é consistente com um valor próximo de zero para todos os mercados analisados, como sugerido na Fig. 4.4.(c) para o índice alemão DAX. Assim, a Eq. (2.32) pode ser truncada na segunda ordem da expansão, reduzindo-se a EFP dada pela Eq. (3.11). Portanto, esta EFP representa uma evolução universal para as FDPs dos retornos, totalmente caracterizada por poucos parâmetros dependentes de τ .

A dependência em τ dos parâmetros \tilde{a}_i e \tilde{b}_i é exemplificada na Fig. 4.5, utilizando o índice mexicano IPC, para diferentes valores de $\Delta\tau$.

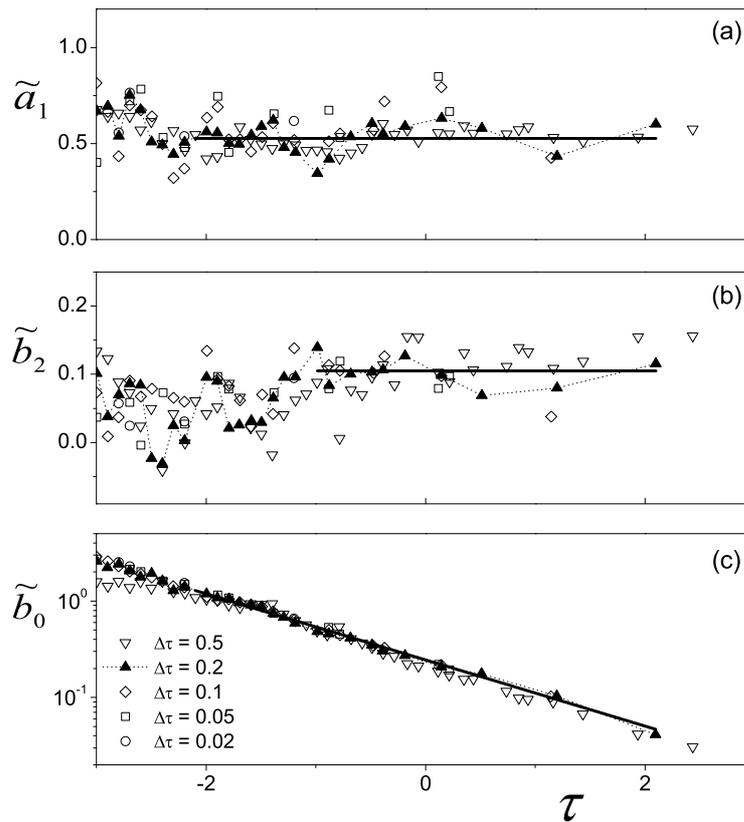


Figura 4.5: Dependência dos parâmetros \tilde{a}_1 (a), \tilde{b}_2 (b) e \tilde{b}_0 (c) nas escalas temporais τ , para diferentes valores de $\Delta\tau$ indicados em (c), para o índice Mexicano IPC. As linhas cheias correspondem aos intervalos utilizados para a estimativa dos parâmetros a_1^s, b_2^s, A e γ , como explicado no texto, usando o resultado para $\Delta\tau = 0, 2$.

Em nossa análise, descartamos os resultados para $(\tau < -2)$, devido à estatística pobre desta região temporal. Levando-se em conta as restrições de acessibilidade empírica do limite de $\Delta\tau \rightarrow 0$, apresentados na seção (3.2) (que é diferente para cada τ), observa-se que dentro das flutuações estatísticas, o comportamento limite dos parâmetros para toda a região $(\tau \geq -2)$ pode ser estimado a partir dos resultados para $\Delta\tau = 0, 2$.

Assim, nossa análise comparativa é feita considerando os resultados obtidos para $\Delta\tau = 0, 2$ como aqueles que definem $\{a_i, b_j\}$ para cada τ .

Como resultados gerais para a dependência em τ dos parâmetros dos coeficientes de KM, encontramos para todos os mercados, os comportamentos ilustrados na Fig. 4.5:

- (i) O parâmetro a_1 apresenta comportamento estacionário dentro das escalas temporais $-2 \leq \tau \leq 2$ (Fig. 4.5.(a)). A média de a_1 nesse intervalo foi definida como o valor estacionário para a_1^s .
- (ii) O parâmetro b_2 apresenta um aumento em função de τ de um valor próximo do zero até um nível estacionário, b_2^s (Fig. 4.5.(b)). Considerando que em $\tau = -1$ todos os mercados atingiram um nível estacionário, estimamos b_2^s como a média de b_2 no intervalo $-1 \leq \tau \leq 2$.
- (iii) O parâmetro b_0 cai exponencialmente com o tempo segundo a Eq. (3.10)(Fig. 4.5.(c)), analogamente ao obtido anteriormente para o IBOVESPA. Ajustamos uma lei exponencial única no intervalo $-2 \leq \tau \leq 2$, obtendo os valores efetivos para A e γ para cada mercado.

Os parâmetros a_0 e b_1 (não mostrados) não aparecem como medidas significantes para comparações entre os mercados. Analogamente ao obtido para o IBOVESPA, encontramos valores do parâmetro a_0 desprezíveis, enquanto o parâmetro b_1 é notavelmente negativo para todos os mercados. Este resultado credencia b_1 como um parâmetro relevante que regula as assimetrias observadas nas distribuições empíricas de retornos em escalas temporais semanais/mensais. No entanto, b_1 não exhibe valores estacionários nessas escalas de tempo, tornando difícil a sua utilização para fins comparativos. Além disso, ele tende a desaparecer para grandes τ , concomitantemente com a extinção da assimetrias nas escalas intradiárias.

Em conclusão, nosso estudo mostra que a equação de evolução das FDPs dos retornos se reduz a EFP (3.11) para todos os mercados analisados. Cada coeficiente de KM ($D^{(1)}$ ou $D^{(2)}$) tem a mesma dependência funcional de r para todos os mercados, e a dependência em τ dos seus parâmetros também apresenta características universais para as escalas semanais/mensais consideradas. Nossa análise preliminar, forneceu o seguinte conjunto de quantidades relevantes, computadas como definido anteriormente: os valores estacionários a_1^s e b_2^s , e os parâmetros A e γ que caracterizam a dependência em τ de b_0 .

4.5 Análise Comparativa

Ao longo desta seção, usaremos os rótulos numéricos mostrados na tabela 4.1 para identificar os mercados. Depois de terem sido detectadas propriedades dinâmicas qualitativamente compartilhadas por todos esses mercados, agora, vamos realizar comparações quantitativas das medidas que surgiram a partir da análise preliminar.

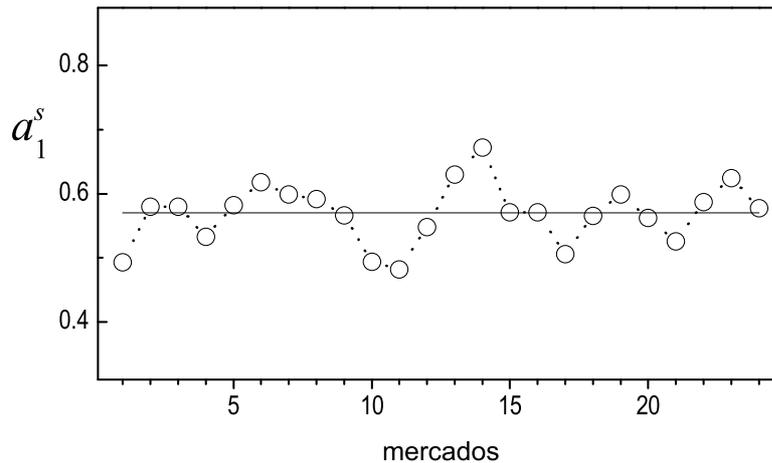


Figura 4.6: Parâmetro a_1^s medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos. O erro de a_1^s é de cerca de 10%. A linha cheia corresponde ao valor médio 0,57.

O parâmetro a_1^s estima a força de restauração do mercado contra flutuações de preço. Seu valor para cada mercado é representado na Fig. 4.6. Encontramos um comportamento universal entre os mercados, com valor médio 0,57 e desvio padrão 0,05 (menor que o erro de a_1^s). Embora estes resultados descartem a_1^s como uma medida de classificação, é um resultado relevante uma vez que manifesta que as leis determinísticas que governam a dinâmica dos mercados são universais não só qualitativamente, mas também quantitativamente, nas escalas de tempo analisadas.

A seguir, apresentamos nas Figs. 4.7 e 4.8 os valores numéricos para os parâmetros do coeficiente de difusão para os mercados mundiais analisados.

Enquanto A descreve a amplitude de b_0 em $\tau = 0$, γ é o inverso do tempo de decaimento. Observa-se que A apresenta uma grande dispersão entre mercados (um fator de cerca de três entre os valores extremos).

Na Fig. 4.9 apresentamos os mercados (com símbolos mais escuros para os mercados desenvolvidos) no plano (γ, A) dos dois parâmetros livres da componente aditiva do ruído b_0 .

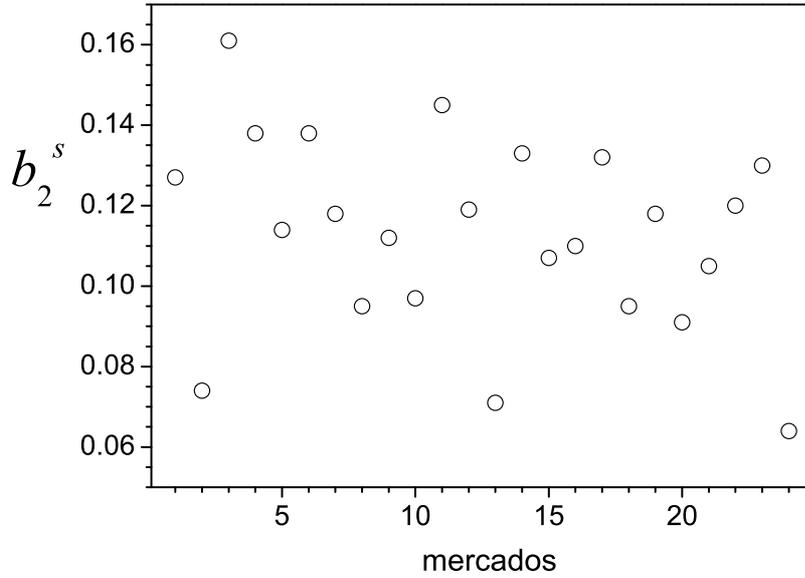


Figura 4.7: Parâmetro b_2^s medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos.

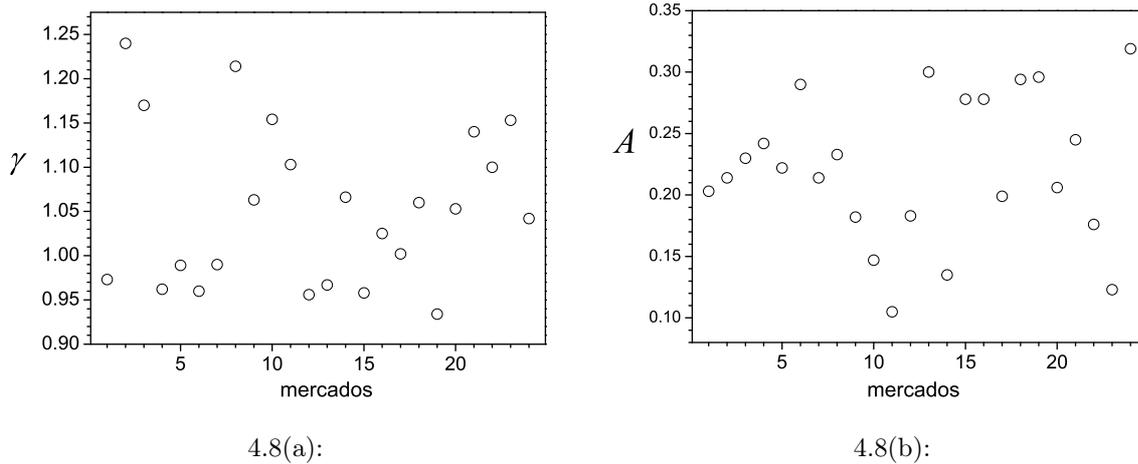


Figura 4.8: Inclinação γ (a) e nível A (b) do parâmetro b_0 medido para todos os mercados listados na tabela 4.1, de acordo com seus rótulos.

A Fig. 4.9 também mostra que mercados desenvolvidos, exceto os da Áustria e Bélgica (rótulos 2 e 3, respectivamente), se aglutinam na região de A grande e γ pequeno. Esta descoberta sugere que esse conjunto de parâmetros tem capacidade em potencial para classificar os mercados. Note que valores grandes de A e pequenos de γ implicam que a componente aditiva da flutuação dos índices é grande e possui um lento e estável decaimento ao longo das escalas temporais, sugerindo mercados mais eficientes.

De fato, a solução invariante (3.12) é consistente com $b_0(\tau) \propto \sigma^2(\tau)$, conforme mostrado no Apêndice A.1.1. De (3.10), a componente aditiva do ruído escala com o intervalo de tempo Δt segundo $b_0 = \alpha (\Delta t)^\gamma$. Logo, desvios

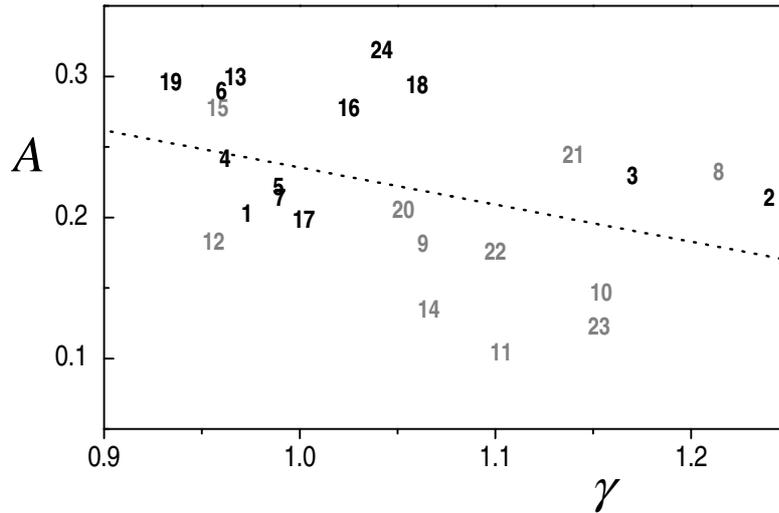


Figura 4.9: Representação dos mercados no plano de parâmetros (γ, A) , de acordo com os rótulos numéricos da tabela 4.1. Os símbolos mais escuros representam os mercados desenvolvidos. Os erros são de cerca de 5%. As linhas pontilhadas representam uma regressão linear como guia visual.

da linearidade ($\gamma \neq 1$) indicam difusão anômala, sugerindo que γ possa ter um papel similar ao do expoente de Hurst H (usualmente associado com o escalonamento da variância com o tempo $\sigma^2(\Delta t) \propto \Delta t^{2H}$), fornecendo uma medida de eficiência do mercado. Índices de mercado tem sido ranqueados na literatura de acordo com seus expoentes de Hurst através de diferentes abordagens [50, 51, 52, 53, 54]. É importante mencionar que embora alguns trabalhos tenham reportado expoentes de Hurst não estacionários [52, 53, 54, 58], somente os expoentes globais ou médios fornecem uma ferramenta para comparação.

Na Fig. 4.10, exibimos os valores de γ , avaliados neste trabalho, e os valores de H , relatados na literatura. Consideramos a Ref. [50], devido ao grande número de mercados comuns em ambas as amostras, incluindo desenvolvidos e emergentes, bem como o alinhamento dos períodos analisados. A Fig. 4.10 mostra claramente uma correlação positiva entre ambos os parâmetros. Além disso, há uma tendência para os mercados mais maduros serem ranqueados como os mais eficientes, com valores de γ (e $2H$) menores daqueles dos mercados emergentes. A fronteira de eficiência ($\gamma = 1$ e $2H = 1$) também mostra a boa concordância entre as segregações geradas por cada medida, qualificando γ como um parâmetro significativo de classificação da eficiência dos mercados.

Os valores de b_2^s , que descrevem a amplitude da componente multiplicativa randômica no mecanismo de flutuação de preço, apresenta uma variação relativa grande entre os mercados (em torno de um fator 3 entre os

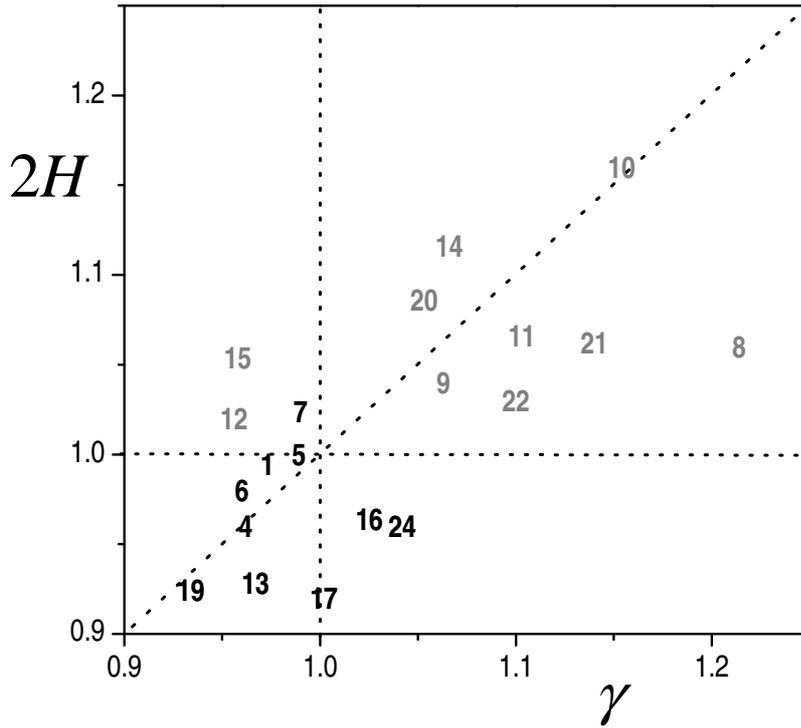


Figura 4.10: Representação dos mercados de acordo com a medida de $2H$ em [49] e γ (este trabalho). Símbolos são como os da Fig. 4.9. O erro de γ é de cerca de 5%. As fronteiras pontilhadas ($\gamma = 1$, $2H = 1$ e $\gamma = 2H$) estão desenhadas para comparação.

valores detetados mínimos e máximos), (ver Apêndice ??). Contudo, a simples visualização dos valores de b_2^s não manifesta uma ordem econômica clara de classificação dos mercados. De fato, uma quantidade mais significativa para caracterizar a dinâmica dos mercados é a força relativa das componentes de ruído multiplicativo e aditivo. Esse fato será discutido a seguir em conexão com as caudas das FDPs de retorno.

Assumindo, de acordo com nossas observações, dependências exponenciais em τ para b_0 , valores estacionários para a_1 e b_2 , assim como valores desprezíveis para a_0 e valor assintoticamente desprezível para b_1 , mostramos a EFP resultante (3.11) admite uma solução assintótica da forma (3.12) com expoente μ dado por (3.13) e variância reescalada $\sigma^2(\tau)$ dada por (3.14).

Devido ao escalonamento dos retornos com σ_{32} , $\sigma^2(\tau)$ é unitária em $\tau = 0$. Então, a partir de (3.14), segue-se que

$$\mu = 2 + A/b_2. \tag{4.1}$$

Por consistência da solução da EFP (3.11) apresentada no Apêndice A.1.1, os parâmetros de KM não são independentes. De acordo com as Eqs. (3.10) e (3.14), os parâmetros estão relacionados através da fórmula $A = a_1 - b_2 - \gamma \ln 2/2$. Esta relação vale somente aproximadamente,

uma vez que o limite assintótico intradiário, $\tau \gg 1$, não foi atingido nas escalas analisadas. Da mesma forma, o valor teórico de μ , dado pela Eq. (4.1), determina o expoente das caudas das FDP somente assintoticamente para grandes τ (pequenos intervalos de tempo). Ainda assim, este limite é alcançado somente se a dependência exponencial em τ do b_0 , juntamente com os níveis estacionários a_1 e b_2 persistirem para muitas hierarquias de escalas temporais.

Com o objetivo de explicitar possíveis conexões entre as medidas microscópicas e as propriedades das caudas das FDPs, obtivemos os valores empíricos de μ , μ_{emp} , na escala diária ($\tau = 5$), a maior escala temporal dos nossos dados. Eles foram computados efetuando-se um ajuste de mínimos quadrados com a forma padronizada da Eq. (2.10):

$$P(x) = \frac{\Gamma((\mu + 1)/2)}{\Gamma(\mu/2)\sqrt{\pi(\mu - 2)}} \left(1 + \frac{x^2}{\mu - 2}\right)^{-(\mu+1)/2}, \quad (4.2)$$

com somente um parâmetro de ajuste, μ , como ilustrado na Fig. 4.11 para os índices argentino Merval e americano Dow Jones.

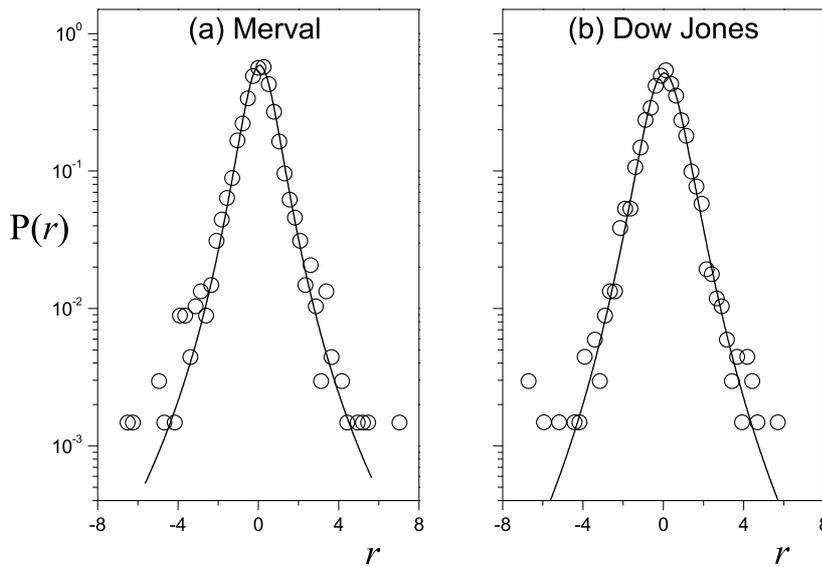


Figura 4.11: FDPs padronizadas para retornos diários para o índice argentino Merval (a) e índice americano Dow Jones (b). Os símbolos correspondem aos histogramas empíricos e as linhas sólidas aos ajustes de mínimos quadrados utilizando a forma padronizada da Eq. (2.10).

Na Fig. 4.12, os mercados estão representados no plano do expoente μ_{emp} e da razão b_2^s/A . Pode-se observar uma correlação notável entre ambas as quantidades que seguem aproximadamente a relação teórica dada pela Eq. (4.1). Isto sugere que um conteúdo de informação relevante foi capturado pela razão das amplitudes dos ruídos multiplicativo/aditivo. No processo aditivo-multiplicativo, o papel do último ruído para gerar caudas em lei de

potência é bem conhecido [39, 40, 59, 60]. Em concordância, uma razão b_2^s/A maior leva a valores menores de μ_{emp} , provocando desvios mais fortes da gaussianidade, que é recuperada no limite $\mu \rightarrow \infty$. Pode-se também observar a tendência seguida pelos mercados desenvolvidos, assim como pelos grandes mercados emergentes, de ter menor razão b_2^s/A . Este resultado indica que a medida composta b_2^s/A tem um potencial papel classificador. Assim, independentemente de γ , ela representa uma perspectiva alternativa para examinar a eficiência como afastamento da gaussianidade.

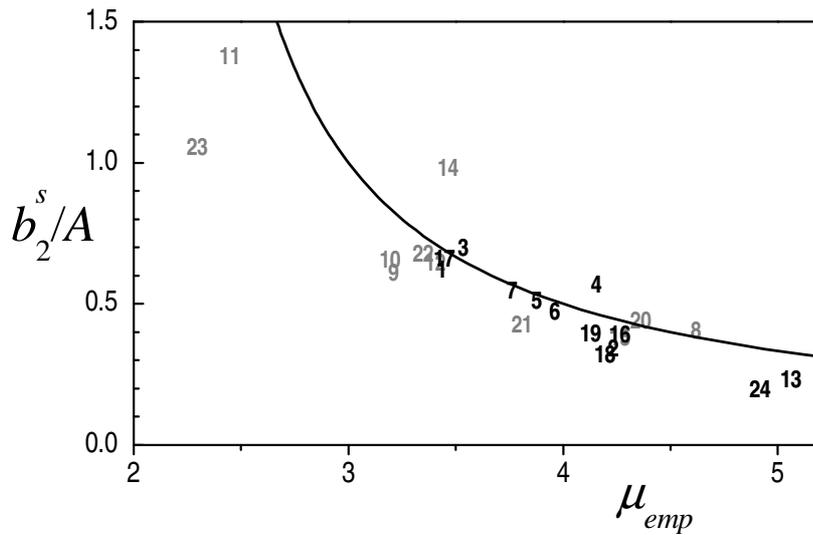
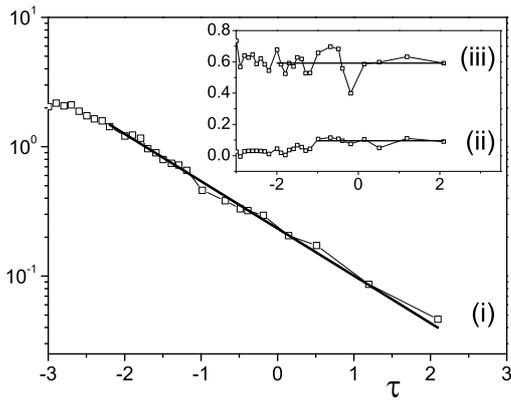


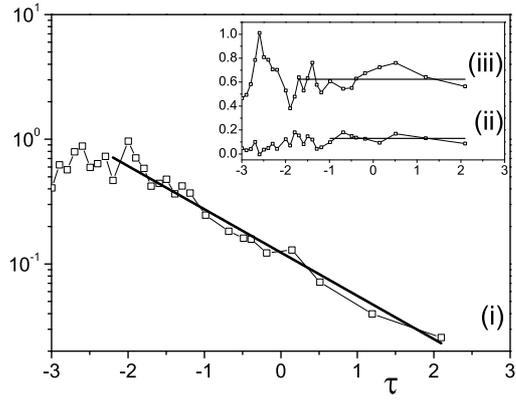
Figura 4.12: Mercados representados no plano $(\mu_{emp}, b_2^s/A)$. Símbolos são como os da Fig. 4.9. Mercados desenvolvidos representados com símbolos mais escuros. A linha pontilhada é a localização dos pontos dados pela Eq. (4.1).

4.6 Mais exemplos do cálculo dos coeficientes de KM para os mercados mundiais

Como ilustração do procedimento utilizado para as determinações dos parâmetros dos coeficientes de KM obtidos a partir das séries de retorno, apresentamos mais alguns exemplos para os mercados mundiais. Os demais mercados não apresentados aqui, possuem comportamentos semelhantes.

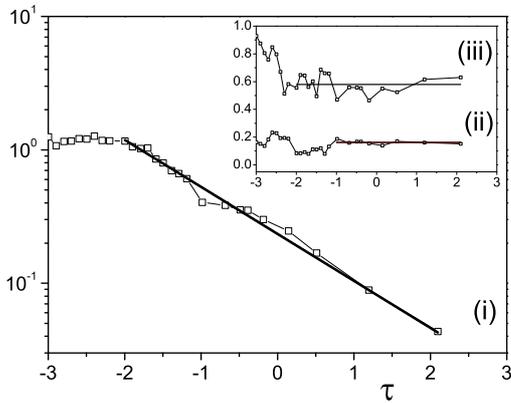


4.13(a):

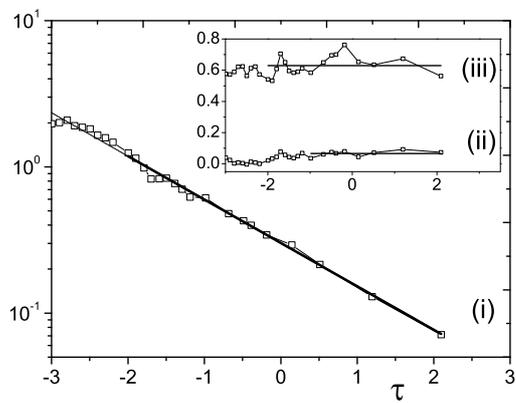


4.13(b):

Figura 4.13: Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jul.1997 a jul.2007, para os índices (a) BSE30 (Índia) e (b) CMA (Egito).

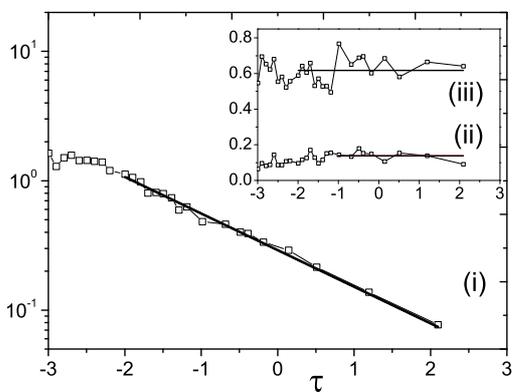


4.14(a):

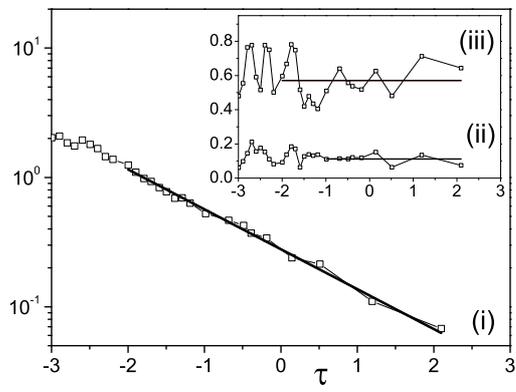


4.14(b):

Figura 4.14: Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jan.1997 a jul.2007, para os índices (a) BEL20 (Bélgica) e (b) Nikkei (Japão).



4.15(a):



4.15(b):

Figura 4.15: Parâmetros (i) \tilde{b}_0 , (ii) \tilde{b}_2 e (iii) \tilde{a}_1 calculados para o período de jan.1997 a jul.2007, para os índices (a) FTSE100 (Reino Unido) e (b) DJIA (Estados Unidos).

4.7

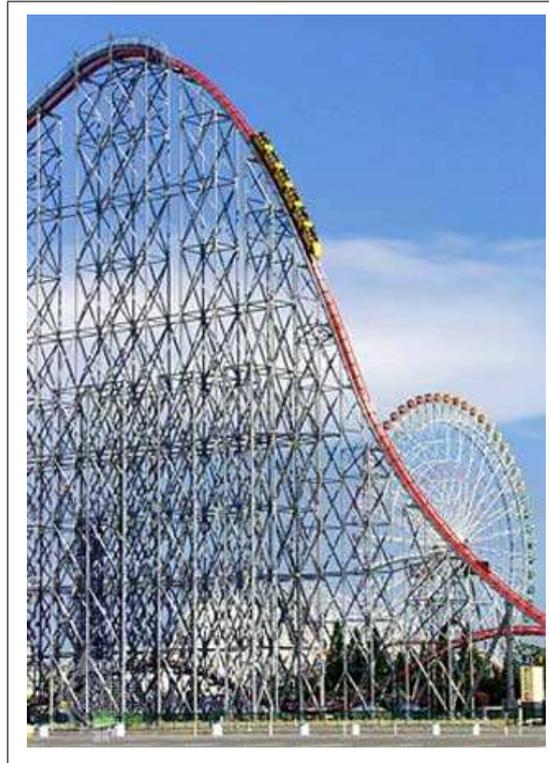
Considerações finais

Nosso estudo revela, por um lado, padrões universais nos coeficientes de KM dos diferentes mercados mundiais, envolvendo tanto fatores determinísticos quanto estocásticos da dinâmica dos retornos de índices de preços. Por outro lado, esta técnica possibilitou a detecção de diferenças quantitativas de alguns parâmetros, fornecendo uma ferramenta alternativa para discriminar dinâmicas de mercado. A partir desses resultados, é possível quantificar algumas características observadas, tais como desvios da gaussianidade das FDPs de retornos e a eficiência dos mercados.

O conteúdo deste capítulo pode também ser encontrado na Ref. [61].

5

Análise do Volume de Ações Negociadas



5.1

Introdução

Neste capítulo investigamos a estatística dos volumes negociados em mercados financeiros. Mostramos que processos estocásticos de volumes negociados podem ser entendidos tendo como base processos de Poisson misturados em uma escala mesoscópica.

5.2

Estatística descritiva dos volumes negociados

Utilizamos como exemplo representativo o volume negociado intradiário da BOVESPA, de 3 de janeiro de 2005 a 13 de setembro de 2007. Os dados foram obtidos em intervalos de 30 minutos, consistindo de 9970 pontos. O tempo é contado apenas em dias de negócio, descartando finais de semana e feriados. A série original de volume negociado, $V(t)$, foi reescalada pelo valor

empírico médio $\langle V \rangle$. A série de $v = V/\langle V \rangle$ está representada na Fig. 1.3, e mostra o padrão intermitente da atividade de negociação.

Verificamos que nossos histogramas de volumes negociados empíricos em diferentes escalas temporais de agregação Δt , podem ser bem aproximados pela FDP q -Gama dada na Eq. (2.15), como ilustrado na Fig. 5.1 (linhas cheias).

Os ajustes dos histogramas dos volumes de negociação usando a FDP Gama dada por

$$P(v) = \frac{\eta}{\Gamma(\beta)} (\eta v)^{\beta-1} \exp(-\eta v), \quad (5.1)$$

são bastante bons para valores intermediários, como ilustrado na Fig. 5.1 (linhas pontilhadas). Contudo, eles manifestam desvios para volumes pequenos e grandes.

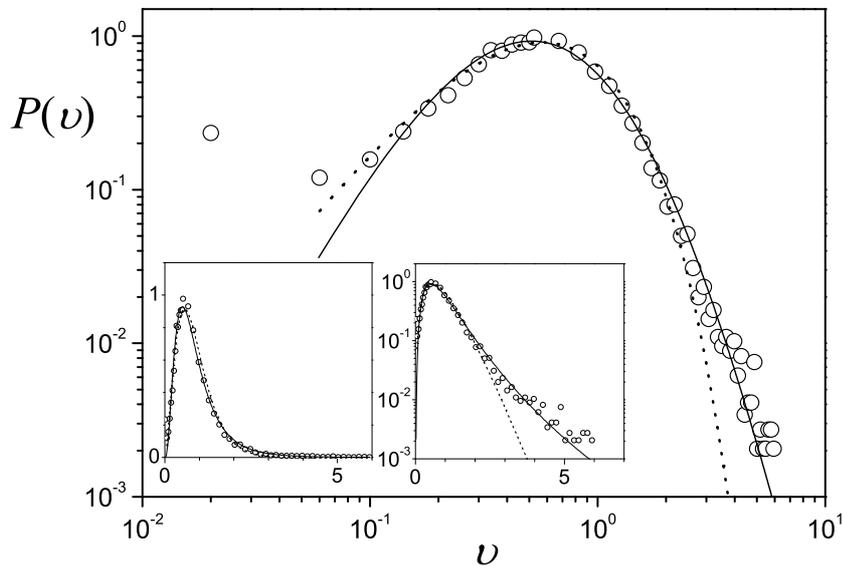


Figura 5.1: Gráfico log-log da FDP dos volumes negociados a cada 30 minutos na BOVESPA no período entre 3/01/05 a 13/09/07. Os símbolos representam o histograma empírico e a linha sólida o ajuste por mínimos quadrados da Eq. (2.15) com parâmetros mostrados na Fig. 5.2. A linha pontilhada corresponde ao ajuste da Eq. (2.15) com $q = 1$, representado para comparação. Os gráficos inseridos correspondem às representações do gráfico principal em escalas linear-linear e log-linear.

O ajuste da q -Gama apresenta um desvio notável somente para volumes muito pequenos (ver também Refs. [12, 15, 16]). O excesso na frequência empírica pode ser devido a um patamar mínimo para transações nos mercados, cujas estatísticas não fazem parte do escopo da tese. Todavia, há uma concordância muito boa para toda a região de ajuste englobando a maioria dos volumes significantes. Um ajuste de mínimos quadrados na Eq. (2.15) foi realizado ao longo da região dos dados da amostra, descartando a região de menor volume (tipicamente abaixo de $v \simeq 10^{-1}$). Neste ajuste, foi considerado

o peso estatístico $1/P(v)$ com o objetivo de ponderar volumes extremos. Os valores resultantes do ajuste dos parâmetros são apresentados na Fig. 5.2 e na tabela 5.1. De forma consistente com um volume reescalonado médio unitário, enquanto os três parâmetros variam, a média da distribuição q -Gama, $\beta\theta/((1-q)\beta+2-q)$, permanece próxima a 1 (com a discrepância de no máximo 6%).

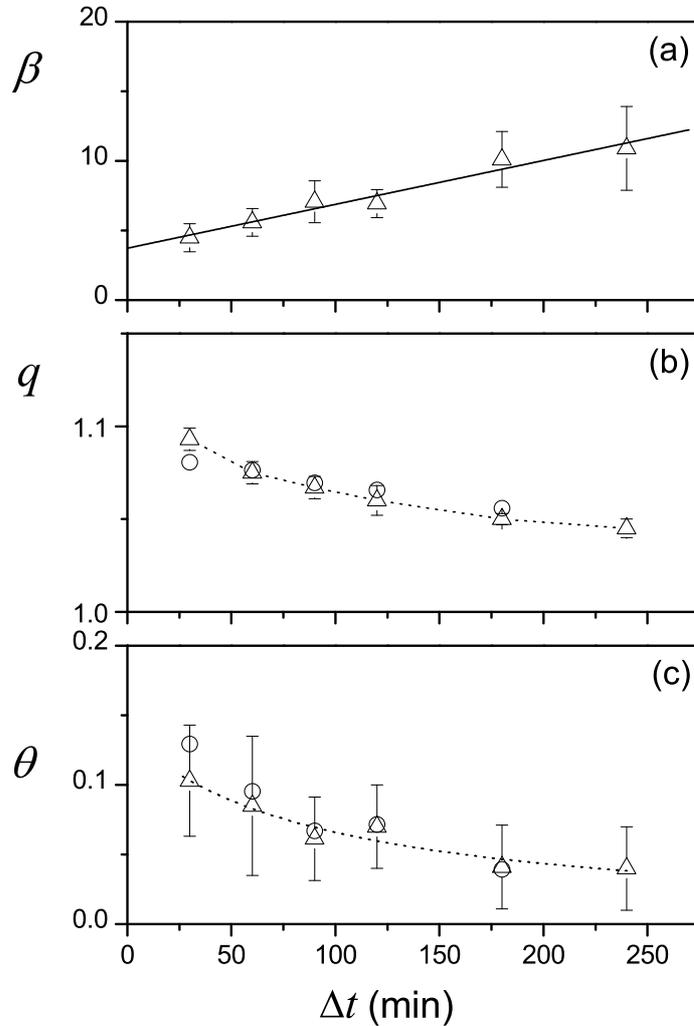


Figura 5.2: Parâmetros de ajuste da distribuição q -Gama, β (a), q (b) e θ (c), para a FDP dos volumes normalizados entre 3/01/05 e 13/09/07, em função de Δt (triângulos). Os círculos correspondem aos valores teóricos obtidos das Eqs. (5.3). As linhas pontilhadas são um guia visual. As linhas cheias em (a) correspondem a um ajuste linear.

Para esclarecer o surgimento das FDPs q -Gama neste cenário, é importante verificar a independência de sucessivos volumes. Dessa forma, calculamos as funções de autocorrelação de primeira e segunda ordem, e não encontramos níveis importantes de correlação para nenhuma janela de tamanho maior do que 30 minutos. Na Fig. 5.3 apresentamos o função de autocorrelação de primeira ordem do volume C_v em função do intervalo de tempo . Ela

Escala	β	θ	q	$10^{-9}\langle V \rangle$
30min	4,48	0,103	1,093	1,61
60min	5,58	0,085	1,075	3,21
90min	7,06	0,061	1,067	4,81
120min	6,93	0,070	1,060	6,42
180min	10,1	0,041	1,050	9,63
240min	10,9	0,040	1,045	12,8

Tabela 5.1: Parâmetros de ajuste da função q -Gama obtidos para as FDPs dos volumes de negociação em diferentes escalas de agregação para a BOVESPA no período entre 3.jan.05 e 13.set.2007. Incluímos também os valores médios $\langle V \rangle$ das séries originais. A $P(v)$ ajustada foi calculada com peso estatístico e descartando volumes ($v \leq 10^{-1}$).

mostra que os efeitos de memória em sucessivos ticks de volumes neste nível de agregação de dados é fraco.

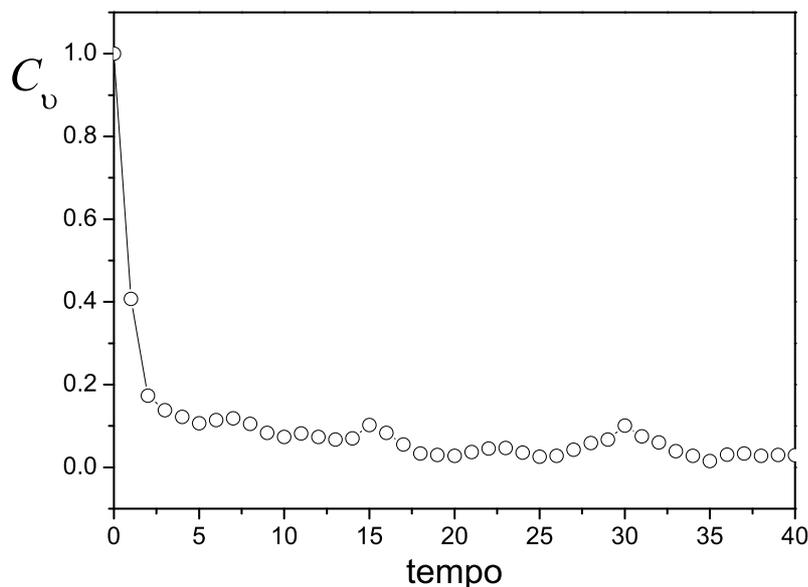


Figura 5.3: Função de autocorrelação para os dados de volume de negociação analisados nesta tese.

5.3 Modelagem

Nesta seção, modelamos o processo que gera as distribuições empíricas dos volumes negociados. Nossa argumentação se dará em termos de processos de ponto ou de chegada.

Em um determinado nível de agregação de dados, cada intervalo de largura Δt contém o volume total negociado naquele intervalo. Em Δt , ocorre um número de negociações, cada uma envolvendo um dado volume de ações.

Uma vez que nossa base de dados não fornece nenhuma informação sobre os volumes individuais em cada negociação, pensamos, alternativamente, na chegada de intervalos através do fluxo de volumes cumulativos. O intervalo de volume acumulado entre sucessivos ticks é $\Delta v_c = v$. Então, deduzimos a estatística de v em analogia àquela dos intervalos de tempo (ou tempos de espera) em processos de chegada usuais.

Nesse cenário, a estatística do processo de ponto se constitui, em essência, em um processo de chegada poissoniano. De fato, é uma aproximação válida considerar:

- (i) o número de eventos (ticks) em incrementos de volumes sem sobreposição como independentes e estacionários (nas escalas de tempo intradiárias analisadas);
- (ii) que para δv_c suficientemente pequeno, a probabilidade de acontecimento de um tick é proporcional a δv_c ;
- (iii) que a probabilidade de mais de um tick chegar em δv_c é desprezível.

Nesse caso, se a taxa de ocorrência (número médio de ticks por volume acumulado) é η , então a FDP dos intervalos v seguirá uma lei exponencial simples $P(v) = \eta \exp(-\eta v)$, se assemelhando aos tempos de espera entre ocorrências sucessivas em processos de chegada poissonianos [21]. Contudo, as transações no mercado são atividades complexas envolvendo muitos passos para tomar uma decisão no nível microscópico de eventos individuais, e os volumes em qualquer escala maior são construídos de volumes na escala desses eventos. Então, é razoável assumir que, ao invés de ser um processo poissoniano simples, o processo poissoniano dos volumes é descrito por β estágios ocultos, cada um associado a uma estatística exponencial simples, com taxas similares. Como vimos no Capítulo 2 (2.38), para estágios independentes, ocorrendo a uma taxa η , a distribuição dos volumes será dada pela Eq. (5.1), também conhecida como a distribuição Gama, $\Gamma_{\beta,1/\eta}$.

Nesse caso, o parâmetro de forma β pode ser interpretado como um número efetivo de estágios do processo combinado. Lembramos também que a Eq. (5.1) é equivalente à Eq. (2.15) quando $q \rightarrow 1$.

Entretanto, o comportamento ligeiramente não exponencial na cauda pode ter uma origem mais fundamental. A análise das correlações indica que os desvios não são devidos à presença de fortes correlações. Então, passamos a investigar outras fontes das caudas longas observadas. Por um lado, a negociação não é homogênea. Por outro, como vimos no Capítulo 2, foi mostrado por Wilk e Wlodarczyk [30] que as distribuições com cauda em lei de potência emergem de forma natural de flutuações lentas de um parâmetro

relacionado ao inverso de uma medida de temperatura aparecendo em uma distribuição originalmente exponencial. Assim, observamos cuidadosamente o possível caráter estocástico da taxa η na Eq. (5.1), que quantifica o número médio de eventos (ticks) por volume acumulado v_c de cada estágio. Para tal, construímos um gráfico do número de ticks em função do v_c , que é apresentado na Fig. 5.4.

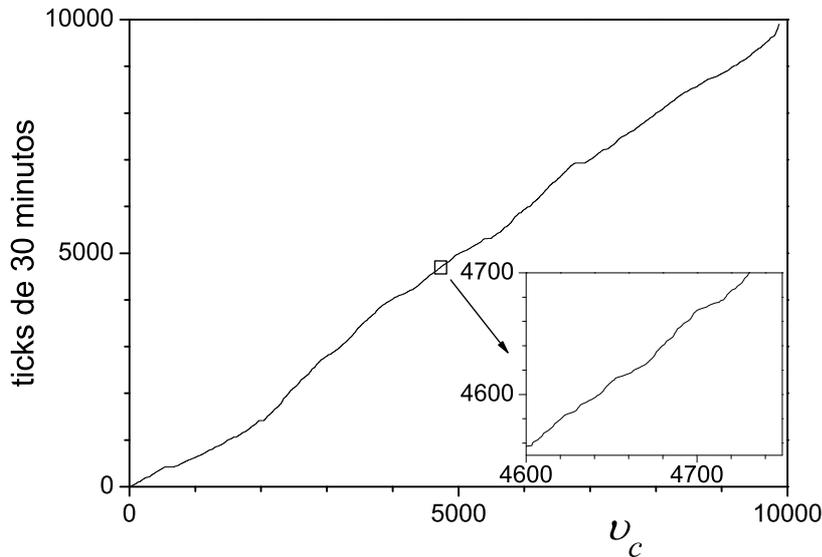


Figura 5.4: Número de ticks de 30 minutos em função do volume acumulado v_c na BOVESPA durante o período entre 3/01/05 e 13/09/07. Gráfico inserido: ampliação da curva.

Pode-se observar que o declive local efetivo, o qual chamaremos de σ , flutua em diferentes escalas, como mostrado na Fig. 5.4. Podemos notar que σ^{-1} é aproximadamente uma medida local de $\langle v \rangle$, que, de acordo com a Eq. (5.1), satisfaz $\langle v \rangle = \beta \eta^{-1}$. A taxa local é então $\eta = \beta \sigma$ correspondendo à taxa de cada estágio (escondido) do processo. De fato, a natureza composta do processo de volume implica que a taxa efetiva observada é diminuída quando comparada com a taxa real de estágio simples.

Com o objetivo de quantificar as flutuações da taxa, os declives locais da Fig. 5.4 foram calculados para janelas de igual comprimento Δv_c . Pequenas janelas são muito ruidosas enquanto janelas mais largas escondem a estrutura das flutuações. Portanto, consideramos janelas de tamanho intermediário correspondentes tipicamente a intervalos de 30 ticks (cerca de 2 dias de negociação). Na Fig. 5.5 apresentamos a série temporal de σ em função do número de pontos por janela.

O histograma de declives locais é apresentado na Fig. 5.6. Esta figura mostra que a estatística dos declives σ (logo, aquela das taxas locais η) segue, em boa aproximação, uma distribuição Gama. O teste de ajuste de

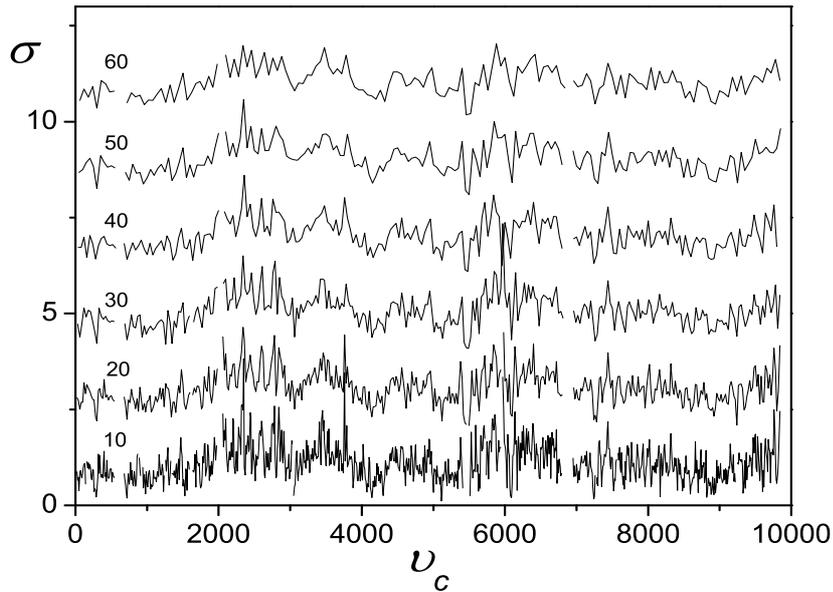


Figura 5.5: Série temporal de σ em função de ν_c para diferentes números de pontos por janela $\Delta \nu_c$ (indicado na figura).

Kolmogorov-Smirnoff (teste KS) [37], mostra como resultado uma adequação forte à hipótese da distribuição Gama para σ . De forma contrária, outras distribuições padrões de dois parâmetros foram rejeitadas, por exemplo a Gama inversa, com valor de referência (p-valor) menor que 0,035.

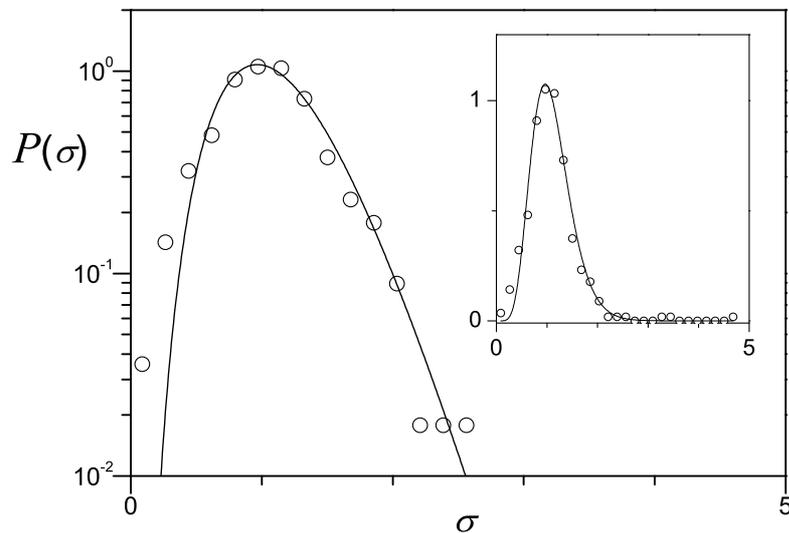


Figura 5.6: Histograma dos declives locais do gráfico mostrado na Fig. 5.4, calculados em janelas de 30 ticks (símbolos). Linhas sólidas correspondem ao melhor ajuste de mínimos quadrados de uma FDP Gama. Gráfico inserido: o mesmo gráfico em escala linear.

Os parâmetros da FDP $\Gamma_{\alpha,1/\kappa}$ associada aos histogramas de η foram

obtidos do ajuste da FDP $\Gamma_{\alpha,1/\beta\kappa}$ ao histograma de $\sigma = \eta/\beta$. Valores para diferentes escalas de agregação analisadas podem ser encontrados na tabela 5.2. Espera-se que a distribuição dos declives dependa somente fracamente do intervalo de ticks Δt . De fato, os valores α e $\kappa\beta$ não são afetados significativamente pela escolha de Δt . Além disso, os resultados empíricos estão de acordo com a média teórica dos declives, $\langle\sigma\rangle = \alpha/(\kappa\beta)$, que é unitária por construção ($\langle\sigma\rangle = \langle v\rangle \equiv 1$).

A partir de séries de volumes de 30 minutos, podemos gerar novas séries em diferentes escalas, agregando os dados, obtendo assim, séries de volumes acumulados a cada 60 minutos, 120 minutos, etc.

Observamos que nas diferentes escalas de agregação a distribuição dos volumes segue a forma q -Gama. Apresentamos a seguir, na Fig. 5.7, as q -Gamas obtidas para as escalas utilizadas na tese e apresentadas na tabela 5.1.

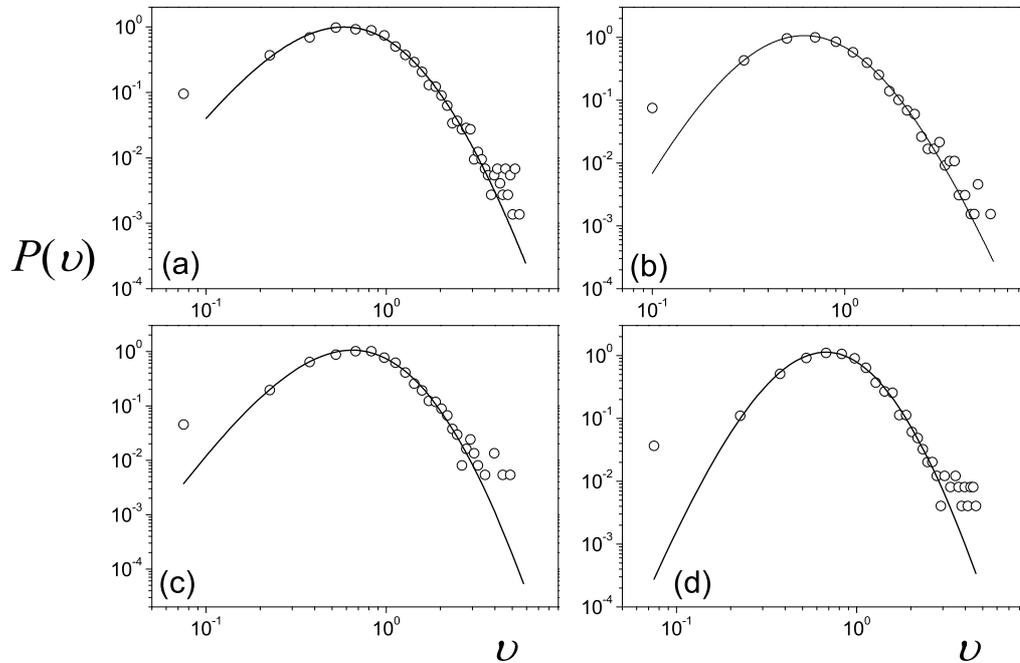


Figura 5.7: Ajustes da função q -Gama obtidos para as FDPs dos volumes de negociação para diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07: 60min(a), 90min(b), 120min(c), 180min(d). A q -Gama ajustada foi calculada com peso estatístico e desconsiderando volumes $\leq 10^{-1}$.

O fato da taxa de ocorrência η não ser uma constante, e sim uma quantidade flutuante, pode explicar o comportamento das caudas das distribuições dos volumes negociados. Se η flutua com uma escala característica maior que o processo estocástico subjacente, podemos interpretar a PF $\Gamma_{\beta,1/\eta}$ em (5.1) como a probabilidade condicional $f_{v|\eta}$ da variável v dado η , onde η é uma variável estocástica. Além disso, vamos considerar que η possui

Escala	α	κ	$1/\beta\kappa$	q_{th}	θ_{th}	θ_{th}/θ	q_{th}/q
30min	7,93	1,605	0,139	1,081	0,129	1,25	0,99
60min	7,50	1,248	0,144	1,076	0,095	1,12	1,00
90min	7,33	0,966	0,147	1,069	0,067	1,09	1,00
120min	8,30	1,090	0,132	1,066	0,072	1,02	1,00
180min	7,78	0,704	0,141	1,056	0,039	0,96	1,00

Tabela 5.2: Parâmetros obtidos para o ajuste da distribuição Gama para os histogramas das taxas em diferentes escalas para a BOVESPA no período de 3/01/05 a 13/09/07. A $P(\sigma)$ foi ajustada utilizando toda a região dos dados. Para comparação, incluímos os valores teóricos dos parâmetros da FDP q -Gama, obtidos a partir da Eq. (5.3).

uma distribuição Gama, i.e., $f_\eta(x) = \Gamma_{\alpha,\kappa}(x)$. Embora, essa seja apenas uma estimativa da distribuição das taxas locais, ela leva em conta suas características principais, com pequenos desvios nas caudas, como observado na Fig. 5.6. Com esses pressupostos, a PF marginal dos volumes passa a ter a forma:

$$P(v) = \int dx f_\eta(x) f_{v|\gamma}(v, x) = \mathcal{N} v^{\beta-1} e_q^{-\eta v}, \quad (5.2)$$

onde \mathcal{N} é um fator de normalização e a função q -exponencial (e_q) foi definida no Capítulo 2.

Assim, vemos que para a variável aleatória η , regida pela distribuição Gama de parâmetros α e κ , a FDP incondicional (ou misturada) dos volumes se torna uma q -Gama.

Os valores teóricos dos parâmetros da FDP q -Gama resultante generalizada (2.15) são:

$$\beta_{th} = \beta, \quad q_{th} = 1 + \frac{1}{\alpha + \beta} \quad \text{e} \quad \theta_{th} = \frac{\kappa}{\alpha + \beta}. \quad (5.3)$$

Note-se que, enquanto β controla o aumento da FDP para volumes menores, α controla a cauda: lei de potência com expoente $\alpha + 1$ para α finito, e lei exponencial para $\alpha \rightarrow \infty$.

Obtivemos uma boa concordância entre os valores q_{th} e θ_{th} e os resultados dos ajustes da distribuição empírica dos volumes, nas diferentes escalas de agregação. Esta comparação é apresentada na tabela 5.2.

Em resumo, nossa análise leva à FDP q -Gama que descreve, em muito boa aproximação os histogramas empíricos (veja Fig. 5.1), em todas as escalas de agregação.

Na Fig. 5.2(a), pode-se observar que β aumenta linearmente com Δt a partir de um valor $\beta_0 > 1$, nas diferentes escalas de agregação. Isso implica a existência de um número invariante efetivo de estágios β_0 associado com o

nível de tempo mais fundamental da descrição das observáveis financeiras. A linearidade caracteriza as propriedades de agregação do processo do volume. Na Fig. 5.2(b) pode-se observar que q decai enquanto Δt aumenta. Isso implica que inomogeneidades se reduzem para grandes intervalos de acumulação Δt . É notável que, para as 10 ações de maior volume na Nasdaq e NYSE para intervalos de 1, 2 e 3 minutos, são observados [12, 15] comportamentos similares.

Das Eqs. (5.3), q é determinado de uma maneira única a partir do número efetivo $\alpha + \beta$ de estágios ou graus de liberdade independentes que especificam ambas as distribuições Gama. Então, por um lado, q reflete o grau de granularidade na descrição do processo de negociação (através de β). Por outro, isso também reflete a presença de inomogeneidades (através de α), decorrentes das flutuações no ambiente do mercado, tais como o fluxo não estacionário de notícias e rumores. Ambos os efeitos se dissipam para grandes intervalos de acumulação Δt .

Investigamos também uma outra grandeza observável financeira, a série da quantidade de transações realizadas, ou ações negociadas, para diferentes escalas de agregação. Ao contrário da FDP dos volumes, a FDP obtida segue uma distribuição Gama, e não foram observadas flutuações no gráfico da taxa local do gráfico do número de ticks em função da grandeza acumulada. Apresentamos este resultado no Apêndice A.4.

5.4

Considerações finais

Em resumo, investigamos um banco de dados real dos volumes negociados em uma escala temporal mesoscópica onde a natureza granular das transações e propriedades de agregação estão presentes. Além de fornecer novas perspectivas para a compreensão da dinâmica dos volumes, o estudo presente provê uma demonstração consistente com dados reais sobre como as FDPs com leis de potência emergem de efeitos de flutuação observáveis, de acordo com a visão superestatística.

O conteúdo deste capítulo pode também ser encontrado na Ref. [62].

6

Reconstrução de Séries Financeiras

6.1

Introdução

Muitos fenômenos físicos que envolvem grandezas flutuantes podem ser descritos por uma equação diferencial estocástica do tipo de Itô-Langevin com a forma apresentada na Eq. (2.36). Sabe-se que a FDP associada é governada pela EFP (2.33) correspondente [24].

Para uma série temporal ideal, regida pela Eq. (2.36) e colecionada por um período suficientemente longo e com alta resolução, a dinâmica do processo pode ser reconstruída diretamente obtendo $D^{(1)}$ e $D^{(2)}$, como foi feito nos Capítulos 3 e 4, usando sua definição estatística, dada pela Eq. (2.34). Nesta seção da tese reconstruiremos mediante esse método e utilizando um algoritmo numérico, uma série de retornos intradiários do IBOVESPA.

6.2

Cálculo dos coeficientes de KM

Utilizamos como exemplo de reconstrução de uma série temporal, a série de retornos normalizados do IBOVESPA para a escala temporal linear de quatro minutos. Eles foram calculados a partir de uma série de preços catalogados a cada 30 segundos, sem sobreposição dos dados. O período adotado foi de 01 de novembro de 2002 a 19 de julho de 2006, totalizando 794.310 dados. Desta forma, a série de retornos de quatro minutos possui 99.288 dados.

Definimos como unidade de tempo a escala de 4 minutos, para a qual, portanto, $\tau \equiv \Delta t = 1$.

Da obtenção dos momentos condicionais, calculamos as funções $\tilde{D}^{(1)}(x, \tau)$ e $\tilde{D}^{(2)}(x, \tau)$ na Eq. (2.35), obtendo formas lineares e quadráticas em x respectivamente, como apresentado na Fig. 6.1. Para $\tau = 1$, obtivemos $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85 \pm 0, 05; 0, 45 \pm 0, 03; 0, 40 \pm 0, 02)$.

Para calcular os coeficientes de KM necessários para a integração da Eq. (2.36), consideramos como primeira aproximação, as funções $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$,

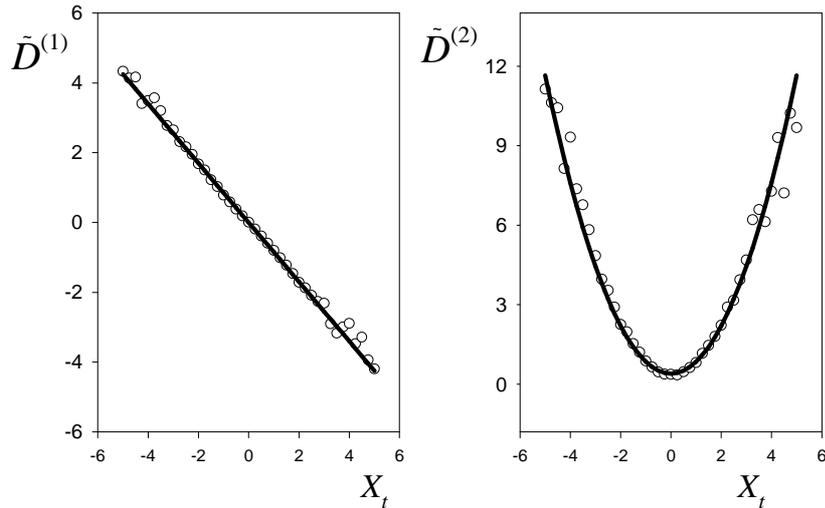


Figura 6.1: Coeficientes $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$ dos dados numéricos para $\tau = 1$. Linhas correspondem aos melhores ajustes, linear e quadrático respectivamente, por mínimos quadrados aos dados.

obtidas para a menor escala disponível, ou seja, $\tau = 1$. Dado que $D^{(1)}(x) = -a_1x$ e $D^{(2)}(x) = b_0 + b_2x^2$, a solução estacionária da Eq. (2.33) é

$$P(x) = \frac{P_0}{\left[1 + \frac{b_2}{b_0}x^2\right]^{1+\frac{a_1}{b_2}}}, \quad (6.1)$$

que possui forma q -gaussiana e constante de normalização P_0 dada por $\sqrt{\frac{b_2}{\pi b_0} \frac{\Gamma[\frac{a_1}{b_2}+1]}{\Gamma[\frac{a_1}{b_2}+\frac{1}{2}]}}$. A condição

$$b_0 = a_1 - b_2 \quad (6.2)$$

surge do fato de que, uma vez que os dados estão normalizados, a variância $\sigma^2 = b_0/(a_1 - b_2)$, deve ser unitária.

Os valores dos parâmetros dos coeficientes de KM calculados com $\tau = 1$ produzem uma FDP estacionária, Eq. (6.1), em boa concordância com o histograma empírico, conforme mostrado na Fig. 6.2.

A Eq. (2.36) foi integrada numericamente através do método de Euler [63]. Para comparação, utilizamos também os métodos de Milstein e Runge-Kutta [63], e os resultados não foram significativamente diferentes. O passo de integração foi $dt = 10^{-3}$, e os dados foram gravados a cada $1/dt$ passos, consistentemente com nossa escolha de $\tau = 1$. Apresentamos a série empírica na Fig. 6.3(a) e a gerada numericamente na Fig. 6.3(b). O histograma da série artificial coincide, naturalmente, com a distribuição estacionária dada pela Eq. (6.1).

Além de obter os histogramas, em concordância com os valores teóricos, também monitoramos as correlações. Na Fig. 6.4, observamos que a função de

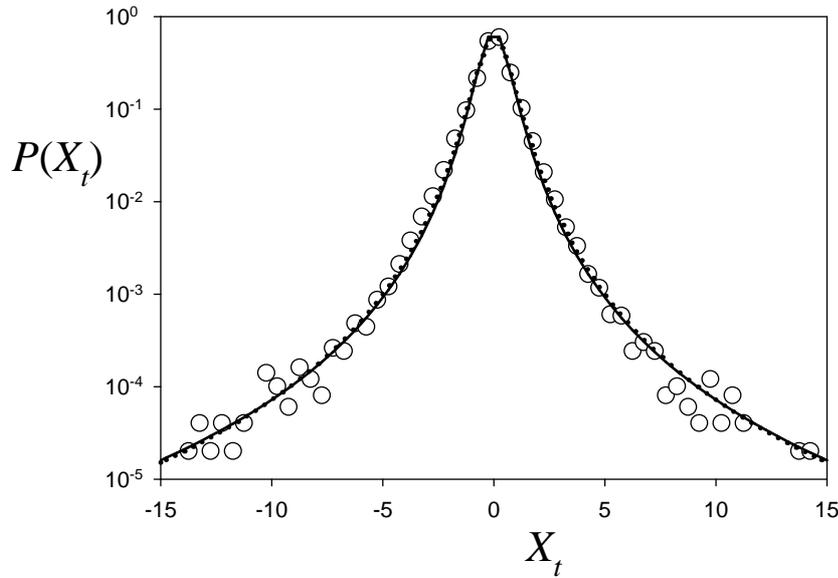


Figura 6.2: FDPs dos retornos normalizados. Símbolos correspondem aos histogramas empíricos. FDPs teóricas dadas pela Eq. (3.12) com $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ obtidas do cálculo dos coeficientes de tendência e difusão de acordo com as Eqs. (6.3) e (6.4) (linhas pontilhadas), e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ obtidos da Eqs. (6.15) e (6.16) (linhas cheias).

autocorrelação da série obtida a partir dos coeficientes para τ -finito possuía um tempo de correlação maior que o da série empírica, apesar das FDPs serem muito semelhantes.

6.3

Correção dos coeficientes de KM

O resultado apresentado acima nos leva a algumas considerações. Em alguns casos, estimativas numéricas não podem ser sempre extrapoladas em direção ao limite $\tau \rightarrow 0$, devido à taxa limitada de aquisição dos dados reais. Dessa forma, como vimos, temos acesso apenas às estimações finitas em τ dos parâmetros $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$, onde τ nem sempre é pequeno em comparação com os tempos característicos do processo. Assim, para algumas séries temporais reais, como as séries financeiras, os métodos de reconstrução precisam ser revisados.

O erro devido ao uso de τ -finito pode ser estimado da expansão da Eq. (2.36). Recentemente, Sura e Barsugli [64] obtiveram as correções de primeira ordem para os coeficientes dados pela Eq. (2.35). Para um dado incremento τ , as aproximações em diferenças finitas apresentadas na Eq. (2.35) podem ser reescritas como:

$$\tilde{D}^{(1)}(x, \tau) = \frac{1}{\tau} \langle X(t + \tau) - X(t) \rangle |_{X(t)=x}, \quad (6.3)$$

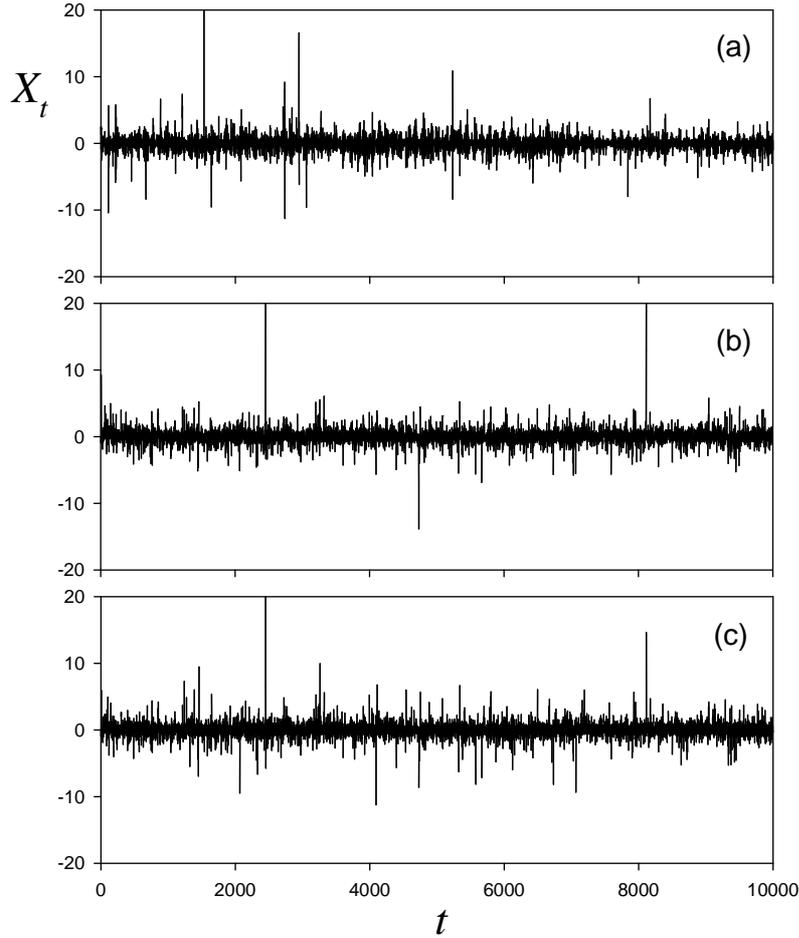


Figura 6.3: Séries temporais de retornos: empírica (a), artificiais geradas com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (b) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (c).

$$2\tilde{D}^{(2)}(x, \tau) = \frac{1}{\tau} \langle (X(t+\tau) - X(t))^2 \rangle |_{X(t)=x}. \quad (6.4)$$

Uma das formas para estimarmos as correções introduzidas quando se adota um incremento de tempo τ -finito nas Eqs. (6.3) e (6.4), é a expansão em série de Itô-Taylor [63] de $X(t+\tau)$. Para obtê-la, vamos considerar a Eq. (2.36) em sua forma integral:

$$X_{t+\tau} = X_t + \int_t^{t+\tau} D^{(1)}(X_s) ds + \int_t^{t+\tau} \sqrt{2D^{(2)}(X_s)} dW_s \quad (6.5)$$

Na Eq. (6.5), os valores de $D^{(1)}$ e $D^{(2)}$ em instantes intermediários são obtidos utilizando-se a fórmula de Itô [63] para funções arbitrárias da variável estocástica X_t (ver Apêndice A.1.3):

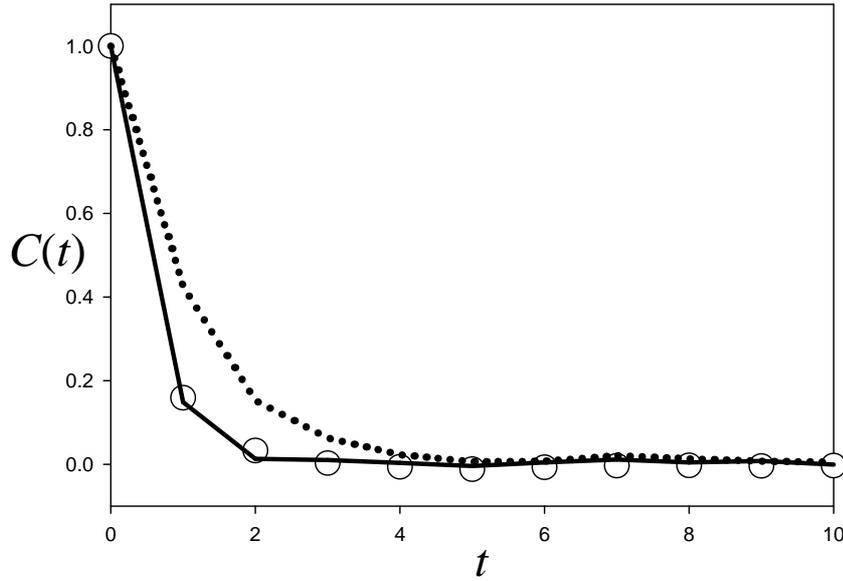


Figura 6.4: Função de autocorrelação linear $C(t)$ em função do tempo. Os símbolos correspondem aos dados empíricos, as linhas correspondem aos processos artificiais gerados com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$ (linhas pontilhadas) e $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ (linhas cheias).

$$dF(X_t, t) = (\partial_t F + D^{(1)} \partial_{X_t} F + D^{(2)} \partial_{X_t X_t} F) dt + (\sqrt{2D^{(2)}} \partial_{X_t} F) dW \equiv L^0 F dt + L^1 F dW. \quad (6.6)$$

A primeira aplicação da fórmula de Itô à Eq. (6.5) fornece (ver Apêndice A.1.3):

$$X_{t+\tau} = X_t + \int_t^{t+\tau} \left(D^{(1)}(X_t) + \int_t^s L_0(D^{(1)}(X_{s'})) ds' + \int_t^s L_1(D^{(1)}(X_{s'})) dW_{s'} \right) ds + \int_t^{t+\tau} \left(\sqrt{2D^{(2)}}(X_t) + \int_t^s L_0 \sqrt{2D^{(2)}}(X_{s'}) ds' + \int_t^s L_1 \sqrt{2D^{(2)}}(X_{s'}) dW_{s'} \right) dW_s. \quad (6.7)$$

Após aplicações sucessivas, podem ser obtidas correções de ordem superior.

Considerando apenas os termos da expansão até a primeira ordem nas Eqs. (6.3) e (6.4), obtemos (ver Apêndice A.1.3):

$$\tilde{D}^{(1)}(x, \tau) = D^{(1)} + \left(\frac{D^{(1)}(D^{(1)})'}{2} + \frac{D^{(2)}(D^{(1)})''}{4} \right) \tau + \mathcal{O}(\tau^2) \quad (6.8)$$

e

$$\begin{aligned} \tilde{D}^{(2)}(x, \tau) = & D^{(2)} + \left[\frac{1}{2}(D^{(1)})^2 + D^{(2)}(D^{(1)})' + D^{(1)}\sqrt{D^{(2)}} \left(\sqrt{D^{(2)}} \right)' \right. \\ & \left. + \{D^{(2)} \left(\sqrt{D^{(2)}} \right)'^2 + \left(\sqrt{D^{(2)}} \right)^3 \left(\sqrt{D^{(2)}} \right)'' \} \right] \tau + \mathcal{O}(\tau^2). \end{aligned} \quad (6.9)$$

As Eqs. (6.8) e (6.9) nos fornecem uma alternativa direta para corrigir os erros devido ao τ -finito. Apesar disso, precisamos especificar as formas de $D^{(1)}$, $(D^{(1)})'$, $D^{(2)}$ e $(D^{(2)})'$. Outras maneiras de se determinar o erro em coeficientes de KM são referenciadas em [64].

Embora as correções de primeira ordem geralmente sejam suficientes, há situações onde esta correção não é apropriada. Neste caso, consideraremos uma correção exata desenvolvida recentemente para os dois primeiros coeficientes de KM calculados com τ -finitos [65]:

$$\tilde{a}_1 = a_1 \sum_{j \geq 0} \frac{[-a_1]^j}{(j+1)!} \tau^j, \quad (6.10)$$

$$\tilde{b}_0 = b_0 \sum_{j \geq 0} \frac{[-2(a_1 - b_2)]^j}{(j+1)!} \tau^j, \quad (6.11)$$

$$\tilde{b}_2 = b_2 \sum_{j \geq 0} \frac{\frac{1}{2}[2(a_1 - b_2)]^{j+1} - [a_1]^{j+1}}{(j+1)!} \tau^j. \quad (6.12)$$

Se a Eq. (6.2) é válida, obtemos a relação $\tilde{b}_0 = \tilde{a}_1 - \tilde{b}_2$, válida também para qualquer ordem de truncamento das séries acima.

Somando as séries infinitas e levando em conta a condição dada pela Eq. (6.2), obtém-se

$$\tilde{a}_1 = \frac{1 - \exp(-a_1\tau)}{\tau}, \quad (6.13)$$

$$\tilde{b}_0 = \frac{1 - \exp(-2b_0\tau)}{2\tau}. \quad (6.14)$$

As Eqs. (6.13) e (6.14) podem ser invertidas para se extrair os parâmetros exatos em função dos estimados:

$$a_1 = \frac{\ln(1 - \tilde{a}_1\tau)}{-\tau}, \quad (6.15)$$

$$b_0 = \frac{\ln(1 - 2\tilde{b}_0\tau)}{-2\tau}. \quad (6.16)$$

Também podemos extrair valores aproximados dos parâmetros

verdadeiros a partir das séries truncadas apresentadas nas Eqs. (6.10), (6.11) e (6.12). Chamando de $a_1^{(n)}$ o valor obtido pela inversão da Eq. (6.10) truncada em uma ordem n para um dado \tilde{a}_1 , podemos analisar a convergência desse parâmetro. Na Fig. 6.5, observamos que a convergência, dentro de uma margem de erro de 5%, depende do valor de \tilde{a}_1 . É mais rápida para valores pequenos de \tilde{a}_1 , ou seja, tão logo $1/\tilde{a}_1$ torna-se grande em comparação com a escala $\tau = 1$. Contudo, este não parece ser o caso dos dados empíricos das séries financeiras analisadas nesta tese: para $\tilde{a}_1 \approx 0,85$ a convergência é atingida somente na sexta ordem.

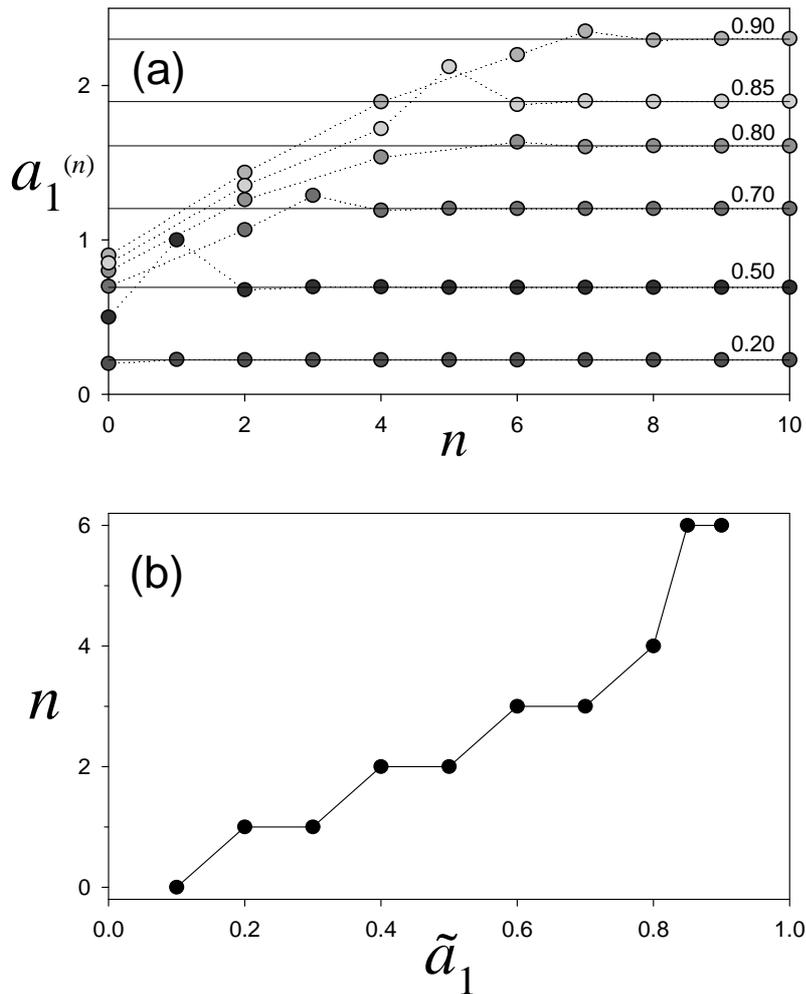


Figura 6.5: (a) Dependência de $a_1^{(n)}$ em função da ordem de aproximação n dada na Eq. (6.10), para diferentes valores de \tilde{a}_1 . (b) Ordem na qual o valor limite é atingido em função de \tilde{a}_1 , com erro de 5%.

A partir das Eqs. (6.15), (6.16) e (6.2), obtivemos os parâmetros exatos $(a_1, b_0 \text{ e } b_2) = (1, 90; 1, 10; 0, 81)$.

Também comparamos os resultados teóricos obtidos para $\tau = 1$ com a análise de séries temporais geradas numericamente usando as estimativas

teóricas de $D^{(1)}$ e $D^{(2)}$. A Eq. (2.36) também foi resolvida mediante o método de Euler, com o mesmo procedimento adotado para gerar a série artificial de retornos a partir de $\tilde{D}^{(1)}$ e $\tilde{D}^{(2)}$. Como podemos verificar na Fig. 6.3(c), a série artificial também possui boa concordância com a série empírica da Fig. 6.3(a).

Uma vez que $a_1/b_2 \approx \tilde{a}_1/\tilde{b}_2$, ambos os conjuntos de valores, os exatos e os para τ -finito produzem uma FDP teórica em muito boa aproximação com os histogramas empíricos, como ilustrado na Fig. 6.2.

A principal diferença está no cálculo das correlações. Na Fig. 6.4, observamos a boa aproximação entre a função de autocorrelação gerada a partir das séries artificiais com os parâmetros limites e com os empíricos, enquanto que os coeficientes obtidos com τ -finito geravam um tempo de correlação maior, apesar das FDPs serem muito semelhantes.

Finalmente, se o processo é estacionário e markoviano, verificado através da equação de CK e dos testes de estacionariedade PP e ADF, como mostrado nos capítulos anteriores, então a probabilidade conjunta $P(x_t, x_{t+\tau})$ contém toda a informação do processo estocástico. Na Fig. 6.6, apresentamos as curvas de nível dos FDPs conjuntas obtidas das séries temporais empíricas e artificiais para $\tau = 1$.

Vemos que o processo artificial, gerado levando-se em conta as correções às estimativas de τ -finito dos coeficientes de KM, fornece comportamento mais próximo aos dados empíricos, quando comparado ao processo artificial sem correção. Em particular, a probabilidade conjunta na Fig. 6.6(c) mostra uma dependência mais fraca entre retornos sucessivos de acordo com o obtido na Fig. 6.4.

6.4

Conclusões preliminares

Neste capítulo da tese, reconstruímos artificialmente uma série empírica de retornos intradiários. Além de reproduzir os histogramas empíricos, a série artificial possui níveis de correlação linear realistas. Para fazer esta reconstrução, vimos a necessidade de realizar correções de ordens superiores aos cálculos dos coeficientes de KM para τ -finito. Para tal, utilizamos uma expansão teórica de ordem infinita apresentada em [65]. O conteúdo deste capítulo pode também ser encontrado na Ref. [66].

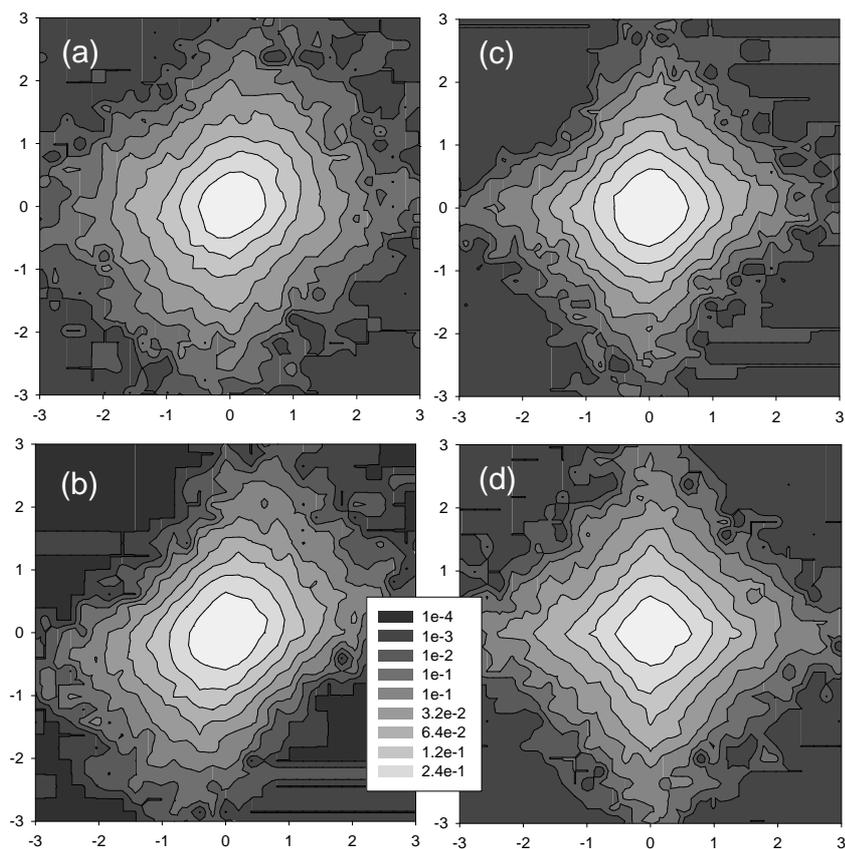


Figura 6.6: FDPs conjuntas para retornos consecutivos, construídos a partir de: (a) dados empíricos, (b) processo artificial gerado com parâmetros $(\tilde{a}_1, \tilde{b}_2, \tilde{b}_0) = (0, 85; 0, 45; 0, 40)$, (c) $(a_1, b_2, b_0) = (1, 9; 1, 10; 0, 81)$ e (d) dados empíricos misturados.

7

Conclusões

Esta tese teve como objetivo descrever a dinâmica subjacente das grandezas observáveis financeiras, tais como retorno de preços e volume de ações negociadas, a partir da análise empírica das respectivas séries temporais.

Vários resultados importantes foram obtidos, que não estão restritos à complexidade dos dados financeiros, mas podem fornecer subsídios para diversos sistemas complexos em geral.

7.1

Índice do Mercado Brasileiro

Obtivemos uma descrição consistente da evolução das FDPs empíricas dos retornos em uma vasta região de escalas temporais, governadas desde gaussianas até distribuições com estatísticas de caudas gordas. Mediante uma técnica única de análise, o cálculo dos coeficientes de KM, pudemos encontrar as características comuns e distintas que emergem de ambos os regimes.

Especificamente, nossa investigação revelou que os coeficientes de tendência e difusão tem a mesma dependência funcional em r , e que a dependência em τ dos seus parâmetros é a que distingue os regimes.

A forma q -gaussiana foi recentemente proposta *ad hoc* para descrever histogramas com cauda em lei de potência para retornos de preços de alta-frequência, e até mesmo diários. No nosso caso, essas distribuições surgem naturalmente da equação de evolução obtida do cálculo dos coeficientes de KM em diferentes escalas de tempo.

Também mostramos que os limites para escalas de tempo pequenas e grandes englobam PFs assintoticamente invariantes em forma. Dessas soluções, surgem novas relações conectando as características das PFs (como μ e σ) aos parâmetros de tendência e difusão. Além disso, somado a uma interpretação dos mecanismos relevantes que governam a dinâmica subjacente do mercado, nossas descobertas podem prover idéias úteis para um cenário mais geral no qual a evolução observada é q -gaussiana.

7.2

Índices de Mercados Mundiais

Analisamos os coeficientes de KM que governam a dinâmica dos preços em escalas temporais macroscópicas para um conjunto representativo de mercados ao redor do mundo. Como resultado geral, encontramos que a equação de evolução para as FDPs dos retornos pode ser representada adequadamente por uma EFP universal, plenamente caracterizada por poucos parâmetros dependentes de τ . Nossa investigação também revelou que os coeficientes de tendência e difusão tem a mesma dependência funcional em r para os diferentes mercados, e que a dependência em τ dos seus parâmetros também apresenta características universais nas escalas temporais mensais/semanais estudadas. Parâmetros dotados com informação significativa da dinâmica dos mercados foram analisados para realizar possíveis medidas classificatórias.

Os valores estacionários a_1^s , descrevendo a intensidade da força restauradora representada por $D^{(1)}$, não apresenta diferenças significativas entre os mercados (Fig. 4.6). Este resultado descarta a_1^s como medida potencial de ranqueamento. Contudo, isso expressa um achado importante: as leis determinísticas que restringem a dinâmica da maioria dos mercados nas escalas mensais/semanais agem com similar intensidade. Portanto, mecanismos de flexibilização são universais tanto qualitativa como quantitativamente, apesar da diversidade de fluxo de informação, bem como a aversão ao risco, de mercado para mercado.

Em contrapartida, nosso estudo demonstrou que, embora qualitativamente semelhantes entre os mercados analisados, ambos os componentes estocásticos aditivo e multiplicativo, traçados a partir de $D^{(2)}$, são quantitativamente distintos para cada mercado, fornecendo parâmetros importantes para a distinção entre os mercados.

Em particular, a amplitude da componente aditiva do ruído, b_0 , decai exponencialmente (com escala temporal logarítmica reversa, τ) com expoente γ , que, à semelhança do expoente de Hurst, surge como uma medida classificatória da eficiência do mercado (Fig. 4.10). Além disso, A , a amplitude característica de b_0 em escalas mensais, apresenta uma grande dispersão de mercado para mercado, permitindo a separação entre mercados maduros e emergentes dentro do plano de parâmetros (γ, A) (Fig. 4.9).

A amplitude da componente multiplicativa das flutuações nas escalas mensais/semanais, b_2^s , também exibe uma ampla variação relativa. Embora individualmente b_2^s não permita uma clara ligação com o estágio econômico do mercado, verificou-se que a razão b_2^s/A fornece uma prova do desvio

dos histogramas empíricos da gaussianidade em escalas temporais diárias (Fig. 4.12).

Os resultados acima são também fundamentados em cálculos analíticos baseados na solução escalada da FDP da EFP assintótica, que relaciona medidas microscópicas a quantidades observáveis macroscópicas, a saber, a variância e o expoente de cauda das FDPs.

Em suma, a presente análise revelou por um lado características universais da dinâmica estocástica dos índices de mercado. Por outro, diferenças entre os mercados através de parâmetros intrínsecos contendo informações importantes sobre a dinâmica do mercado.

7.3

Volumes Negociados no Mercado Brasileiro

Mostramos que o processo estocástico dado pela ocorrência inhomogênea de eventos poissonianos descreve as principais propriedades estatísticas de uma importante observável de mercado, o volume negociado.

A FDP resultante é a q -Gama, estando em muito bom acordo com os histogramas empíricos para os volumes mais significantes. Valores teóricos dos parâmetros foram obtidos das Eqs. (5.3), ao desenharmos o número de estágios efetivos β do ajuste da q -Gama aos histogramas de volume, enquanto α e κ vem do ajuste Gama do histograma das taxas locais. Valores teóricos e empíricos dos parâmetros estão em excelente acordo (ver Fig. 5.2), denotando a consistência da nossa abordagem. Sendo uma medida de granularidade, β afeta a forma das FDPs apenas para valores pequenos. Entretanto, a cauda é determinada apenas por α , cuja finitude reflete o grau de inhomogeneidade do ritmo da atividade do mercado.

Podemos notar que a FDP Gama é observada em muitos fenômenos envolvendo múltiplas tarefas ou estágios [25]. A modificação de seu decaimento exponencial para uma lei de potência pode ser conseguida ao se considerar um processo de Poisson combinado, com taxa variável, que é aleatória por si só (traçada de uma distribuição Gama). Isso constitui uma classe especial de processos de Poisson não homogêneos na qual a taxa não é uma função suave do tempo, mas envolve estocasticidade (como nos processos de Cox), refletindo uma fonte adicional de flutuações aleatórias [25]. Modelos similares em que o parâmetro de taxa da FDP Gama flutua com uma lei de probabilidade Gama tem sido propostos anteriormente em diversos contextos para explicar a presença de caudas pesadas. Exemplos podem ser encontrados em falhas em programas de computador [67], distúrbios ecológicos [68], toques na bola durante uma partida de futebol [69], turbulência [70, 71] e fisiologia [72].

Esta diversidade mostra a universalidade do comportamento da dupla Gama estocástica.

Isso provavelmente ocorre devido ao caráter especial da família Gama de FDPs, que é fechada sob convolução. A aparição frequente de FDPs q -Gama, surgindo como uma FDP misturada, se encaixa dentro do cenário da abordagem introduzida por Wilk e Wlodarczyk [30] e depois generalizada por Beck e Cohen na chamada *superestatística* [31], segundo comentado no Capítulo 2.

O comportamento Gama-Gama tem sido discutido anteriormente também em conexão com a estatística dos volumes [16, 17, 18]. Nele, a estatística Gama condicional deriva de uma equação dinâmica estocástica cuja FDP estacionária associada é a distribuição Gama. Contudo, ela foi proposta *ad hoc*, enquanto outras equações estocásticas igualmente plausíveis compartilham a mesma propriedade [40]. Ainda mais, ficou dúvida se o parâmetro de flutuação intrínseco da Gama foi o volume negociado local médio ou seu inverso. Nosso resultado fornece suporte empírico a última conjectura contra a primeira.

Em resumo, investigamos um banco de dados real dos volumes negociados em uma escala temporal mesoscópica onde a natureza granular das transações e propriedades de agregação estão presentes. Além de fornecer novas interpretações para a estatística observável dos volumes, o presente estudo provê uma demonstração consistente com dados reais sobre como as FDPs com caudas em lei de potência emergem de efeitos de flutuação observáveis, em acordo com a visão superestatística. Nosso resultado, obtido para as escalas temporais da ordem de uma hora, também são consistentes com aquelas observadas para horizontes temporais de minutos [12, 15]. Contudo, uma investigação seguindo as idéias apresentadas aqui, analisando o papel de correlações na escala de ultra alta frequência, torna-se ainda necessária.

7.4

Conclusões Gerais

A dinâmica complexa dos dados financeiros requer o desenvolvimento de métodos de análise avançados e generalizados. Assim, os resultados alcançados são de interesse teórico e prático que podem ir além do problema particular analisado.

No caso de retornos de preços, buscamos obter equações para a evolução temporal das FDPs através da estimativa dos coeficientes de KM. Inicialmente examinamos com mais detalhes as séries temporais diárias e intradiárias do IBOVESPA, um dos grandes mercados emergentes mundiais. Neste cenário,

abordamos uma vasta hierarquia de escalas de tempo, variando de meses a minutos. Obtivemos uma descrição unificada que engloba desde o regime gaussiano até o q -gaussiano. Estendemos essa análise para realizar um estudo comparativo entre diversos mercados mundiais. Identificamos padrões universais assim como também diferenças quantitativas que permitiram classificar os diversos mercados.

No caso dos volumes de negociação, medimos diretamente o nível de atividade do mercado obtendo sua FDP. Mostramos que a forma da distribuição para pequenos volumes é governada pelo grau de granularidade do processo de negociação, enquanto o expoente que controla a cauda da distribuição é uma medida da inhomogeneidade da atividade do mercado. Este resultado provê evidências de como distribuições de probabilidade com cauda em lei de potência, tão comuns nos sistemas complexos podem surgir como consequência das flutuações de um parâmetro intrínseco.

Em suma, este estudo forneceu uma demonstração consistente com dados reais sobre como as FDPs com leis de potência observadas nesse contexto podem emergir através de mecanismos diversos, como processos estocásticos com flutuações aditivo-multiplicativas, ou como resultado de misturas estatísticas.

7.5

Perspectivas futuras

Como possíveis avanços propomos:

- Analisar detalhadamente a dependência em τ dos parâmetros dos coeficientes de KM que aparentemente apresentam oscilações log-periódicas em Δt .
- Obter e interpretar os coeficientes de KM para as séries de volumes negociados.
- Investigar o papel das correlações na escala intradiária sobre os processos estocásticos resultantes de misturas estatísticas.
- Estender os cálculos estocásticos para correção dos coeficientes de KM de ordem superior ao segundo por efeitos de $\Delta\tau$ finito.

Referências Bibliográficas

- [1] L. Borland. Option pricing formulas based on a non-gaussian stock price model. *Phys. Rev. Lett.*, 89:098701, 2002. 1, 3.4
- [2] S. Ghashghaie, W. Breymann, J. Peinke, and P. Talkner. Turbulent cascades in foreign exchange markets. *Nature*, (381):767, 1996. 1, 1.3
- [3] A. Krawiecki, J.A. Holyst, and D. Helbing. Volatility clustering and scaling for financial time series due to attractor bubbling. *Phys. Rev. Lett.*, (89):158701, 2002. 1
- [4] D.L. Cassidy. *Trading on Volume: The Key to Identifying and Profiting from Stock Price Reversals*. McGraw-Hill, New York, 2002. 1.1
- [5] L. Bachelier. Theorie de la especulation. *Ann.Sci. Ecole Norm. Sup.*, (17):21, 1900. 1.2
- [6] E. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383, 1970. 1.2
- [7] B. Mandelbrot. The variation of certain speculative prices. *J. Business*, (36):394–419, 1963. 1.2
- [8] J.L. McCauley and G.H. Gunaratne. *Physica A*, (329):178, 2003. 1.2
- [9] A.C. Silva, R.E. Prange, and V.M. Yakovenko. *Physica A*, (344):227, 2004. 1.2
- [10] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *J. Stat. Phys.*, (52):4791, 1988. 1.2, 2.1.2, 2.3
- [11] C. Tsallis. *Non-Extensive Entropy-Interdisciplinary Applications*. Oxford University Press, Oxford, 2004. 1.2
- [12] C. Tsallis, C. Anteneodo, L. Borland, and R. Osorio. Nonextensive statistical mechanics and economics. *Physica A*, (324):89, 2003. 1.2, 1.2, 3.4, 5.2, 5.3, 7.3

- [13] P. Gopikrishnan, V. Plerou, X. Gabaix, and H.E. Stanley. Statistical properties of share volume traded in financial markets. *Phys. Rev. E*, (62):R4493, 2000. 1.2
- [14] X. Gabaix, P. Gopikrishnan, V. Plerou, and H.E. Stanley. Institutional investors and stock market volatility. *Quarterly Journal of Economics*, (121):461, 2006. 1.2
- [15] R. Osorio, L. Borland, and C. Tsalli. *Nonextensive Entropy - Interdisciplinary Applications*. Oxford University Press, New York, 2004. 1.2, 5.2, 5.3, 7.3
- [16] S.M.D. Queirós. On the emergence of a generalised gamma distribution. application to traded volume in financial markets. *Europhysics Letters*, (71):339, 2005. 1.2, 5.2, 7.3
- [17] J. de Souza, L.G. Moyano, and S.M.D. Queirós. On statistical properties of traded volume in financial markets. *Eur. Phys. J. B.*, (50):165, 2006. 1.2, 7.3
- [18] S.M.D. Queirós and L.G. Moyano. Yet on statistical properties of traded volume: Correlation and mutual information at different value magnitudes. *Physica A*, (383):10, 2007. 1.2, 7.3
- [19] S.M.D. Queirós, L.G. Moyano, J. de Souza, and C. Tsallis. A nonextensive approach to the dynamics of financial observables. *Eur. Phys. J. B.*, (55):161, 2007. 1.2
- [20] Friedrich R., Peinke J., and Renner C. How to quantify deterministic and random influences on the statistics of the foreign exchange market. *Phys. Rev. Lett.*, 84:5224, 2000. 1.3, 2.2.4, 3.2, 3.2, 3.2, 3.3, 3.5, 4.4
- [21] W. Feller. *An introduction to probability theory and its applications*. Wiley, New York, 1968. 2, 2.1.1, 2.2.6, 5.3
- [22] C. Anteneodo. *Processos Estocásticos*. V Escola do CBPF, 2004. 2
- [23] N.G.V. Kampen. *Stochastic processes in physics and chemistry*. Amsterdam, New York: North-Holland, 1981. 2, 2.2.1, 1, 2.2.1
- [24] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer-Verlag, Berlin, 1984. 2, 2.2, 2.2.1, 2.2.4, 2.2.4, 2.2.5, 3.2, 3.4, 6.1

- [25] R.B. Cooper. *Introduction to queuing theory*. Elsevier North Holland, Inc., New York, 1981. 2.1.1, 7.3
- [26] C. Renner, Peinke J., and R. Friedrich. Evidence of markov properties of high frequency exchange rate data. *Physica A*, 298:499, 2001. 2.2.4, 3.5, 4.4
- [27] F. Farahpour, Z. Eskandari, A. Bahraminasab, G.R. Jafari, F. Ghasemi, M. Sahimi, and M.R.R. Tabar. A langevin equation for the rates of currency exchange based on the markov analysis. *Physica A*, (385):601, 2007. 2.2.4, 3.5
- [28] F. Ghasemi, M. Sahimi, J. Peinke, R. Friedrich, G.R. Jafari, and M.R.R. Tabar. Markov analysis and kramers-moyal expansion of nonstationary stochastic processes with application to the fluctuations in the oil price. *Phys. Rev. E*, (75):060102, 2007. 2.2.4, 3.5, 4.3
- [29] A.A.G. Cortines, R. Riera, and C. Anteneodo. From short to fat tails in financial markets: a unified description. *Eur. Phys. J. B*, (60):385, 2007. 2.2.4, 3.5, 4.4
- [30] G. Wilk and Z. Wlodarczyk. Interpretation of the nonextensivity parameter q in some applications of tsallis statistics and levy distributions. *Phys. Rev. Lett.*, (84):2770, 2000. 2.3, 2.3, 5.3, 7.3
- [31] C. Beck and E.G.D. Cohen. *Physica A*, (322):267, 2003. 2.3, 7.3
- [32] C. Tsallis and A.M.C. Souza. Constructing a statistical mechanics for beck-cohen superstatistics. *Phys. Rev. E*, (67):026106, 2003. 2.3
- [33] J.B. McDonald and R.J. Butler. Some generalized mixture distributions with an application to unemployment duration. *Review of Economics and Statistics*, (69):232, 1987. 2.3
- [34] B.A. Maguire, E.S. Pearson, and A.H.A. Wynn. Evidence of increment of efficiency of the mexican stock market through the analysis of its variations. *Biometrika*, (39):168, 1952. 2.3
- [35] K. Ivanova, M. Ausloos, and H. Takayasu. Deterministic and stochastic influences on japan and us stock and foreign exchange markets: A fokker-planck approach. *cond-mat*, (0301268), 2003. 3.2, 3.5, 4.4
- [36] M. Karth and J. Peinke. *Complexity*. John Wiley & Sons Inc., New York, 2002. 3.2

- [37] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numeric Recipes. The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1986. 3.4, 5.3
- [38] A. Schenzle and H. Brand. Multiplicative stochastic-processes in statistical physics. *Phys. Rev. A*, (20):1628, 1979. 3.4
- [39] C. Anteneodo and C. Tsallis. Multiplicative noise: a mechanism leading to nonextensive statistical mechanics. *J. Math. Phys.*, (44):5194, 2003. 3.4, 4.5
- [40] C. Anteneodo and R. Riera. Additive-multiplicative stochastic models of financial mean-reverting processes. *Phys. Rev. E*, (72):026106, 2005. 3.4, 4.5, 7.3
- [41] V. Plerou, P. Gopikrishnan, L.A.N. Amaral, M. Meyer, and H.E. Stanley. Scaling of the distribution of price fluctuations of individual companies. *Phys. Rev. E*, (60):6519, 1999. 3.4
- [42] R. Huisman, K.G. Koedijk, C.J.M. Kool, and F. Palm. Tail-index estimates in small samples. *Journal of Business and Economic Statistics*, (19):208, 2001. 3.4
- [43] N. Kozuki and N. Fuchikami. Dynamical model of financial markets: fluctuating 'temperature' causes intermittent behavior of price changes. *Physica A*, (329):222, 2003. 3.4
- [44] M. Ausloos and K. Ivanova. Dynamical model and nonextensive statistical mechanics of a market index on large time windows. *Phys. Rev. E*, (68):046122, 2003. 3.4, 3.5, 4.4
- [45] A.P. Nawroth and J. Peinke. Small scale behavior of financial data. *Eur. Phys. J. B*, (50):147, 2006. 3.4
- [46] H.F. Coronel-Brizio, A.R. Hernandez-Montoya, and R. Huerta-Quintanilla. Evidence of increment of efficiency of the mexican stock market through the analysis of its variations. *Physica A*, (380):391, 2007. 4.1
- [47] N.B. Ferreira, R. Menezes, and D.A. Mendes. Asymmetric conditional volatility in international stock markets. *Physica A*, (382):73, 2007. 4.1
- [48] J.H. Park, S.K. Nam, and K.S. Eom. Market efficiency in kosdaq: A volatility comparison between main boards and new markets using a

- permanent and transitory component model. *Asia-Pacific Journal of Financial Studies*, (36):533, 2007. 4.1
- [49] S. Edwards and R. Susmel. Interest-rate volatility in emerging markets. *Review of Economics and Statistics*, (85):328, 2003. 4.1
- [50] T. Di Matteo, T. Astel, and M.M. Dacorogna. Scaling behaviors in differently developed markets. *Physica A*, (324):183, 2003. 4.1, 4.5
- [51] T. Di Matteo, T. Astel, and M.M. Dacorogna. Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *J. Banking and Finance*, (29):827, 2005. 4.1, 4.5
- [52] D.O. Cajueiro and B.M. Tabak. The hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. *Physica A*, (336):521, 2004. 4.1, 4.5
- [53] D.O. Cajueiro and B.M. Tabak. Ranking efficiency for emerging markets. *Chaos Solutions and Fractals*, (22):349, 2004. 4.1, 4.5
- [54] L. Zunino, B.M. Tabak, D.G. Pérez, M. Garavaglia, and O.A. Rosso. Inefficiency in latin-american market indices. *Eur. Phys. J. B*, (60):111, 2007. 4.1, 4.5
- [55] B.L. Lan and Y.O. Tan. Statistical properties of stock market indices of different economies. *Physica A*, (375):605, 2007. 4.1
- [56] T. Qiu, B. Zheng, and F. Ren et al. Statistical properties of german dax and chinese indices. *Physica A*, (378):387, 2007. 4.1
- [57] R. Davidson and J.G. MacKinnon. *Estimation and Inference in Econometrics*. Oxford University Press, New York, 1993. 4.3, A.1.4
- [58] R.L. Costa and G.L. Vasconcelos. Long-range correlations and nonstationarity in the brazilian stock market. *Physica A*, (329):231, 2003. 4.5
- [59] H. Takayasu, A.H. Sato, and M. Takayasu. Stable infinite variance fluctuations in randomly amplified langevin systems. *Phys. Rev. Lett.*, (79):966, 1997. 4.5
- [60] D. Sornette. Multiplicative processes and power laws. *Phys. Rev. E*, (57):4811, 1998. 4.5

- [61] A.A.G. Cortines, C. Anteneodo, and R. Riera. Stock index dynamics worldwide: a comparative analysis. *Eur. Phys. J. B*, (65):289, 2008. 4.7
- [62] A.A.G. Cortines, R. Riera, and C. Anteneodo. Measurable inhomogeneities in stock trading volume flow. *Europhysics Letters*, (83):30003, 2008. 5.4
- [63] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin, 1992. 6.2, 6.3, 6.3, A.1.3, A.1.3
- [64] P. Sura and J. Barsugli. A note on estimating drift and diffusion parameters from timeseries. *Physics Letters A*, (305):304, 2002. 6.3, 6.3
- [65] C. Anteneodo and R. Riera. Arbitrary-order corrections for finite-time drift and diffusion coefficients. *Physical Review E*, (80):031103, 2009. 6.3, 6.4, A.1.3
- [66] C. Anteneodo, R. Riera, and A.A.G. Cortines. *em preparação*, 2009. 6.4
- [67] B. Littlewood and J.L. Verall. Bayesian reliability growth model for computer software. *Applied Statistics*, (22):332, 1973. 7.3
- [68] R.W. Katz, G.S. Brush, and M.B. Parlange. Statistics of extremes: Modeling ecological disturbances. *Ecology*, (86):1124, 2005. 7.3
- [69] R.S. Mendes, L.C. Malacarne, and C. Anteneodo. Statistics of football dynamics. *Eur. Phys. J. B*, (57):357, 2007. 7.3
- [70] C. Beck. Statistics of three-dimensional lagrangian turbulence. *Phys. Rev. Lett.*, (98):064502, 2007. 7.3
- [71] A.M. Reynolds, N. Mordant, A.M. Crawford, and E. Bodenschatz. On the distribution of lagrangian accelerations in turbulent flows. *New J. Phys.*, (7):58, 2005. 7.3
- [72] L.L. Chen and C. Beck. A superstatistical model of metastasis and cancer survival. *Physica A*, (387):3162, 2008. 7.3
- [73] W. Enders. *Applied econometric time series*. Wiley, New York, 1995. A.1.4

A

Apêndices

A.1

Considerações Teóricas Complementares

A.1.1

Sobre a solução invariante da EFP

Vamos buscar soluções invariantes para a EFP (3.11), de acordo com o obtido empiricamente para escalas τ longas e curtas, assumindo comportamento de $b_0(\tau)$ dado pela Eq. (3.10). Para tal, consideramos soluções da forma:

$$P(r, \tau) \equiv \frac{1}{\sigma(\tau)} p\left(\frac{r}{\sigma(\tau)}\right). \quad (\text{A.1})$$

Fazendo $x = r/\sigma(\tau)$, podemos definir a função τ -invariante:

$$p(x) \equiv \sigma(\tau) P(\sigma(\tau)x, \tau). \quad (\text{A.2})$$

Substituindo a Eq. (A.1) na Eq. (3.11), obtemos:

$$\frac{\partial}{\partial \tau} P(r, \tau)|_{r=cte} = \frac{\sigma'(\tau)}{\sigma^2(\tau)} [1 + x d_x] p(x). \quad (\text{A.3})$$

As derivadas em r da EFP (3.11), podem ser reescritas como

$$-\partial_r (D^{(1)}(r, \tau) P(r, \tau)) = \frac{a_1}{\sigma(\tau)} [1 + x d_x] p(x), \quad (\text{A.4})$$

e

$$\partial_r^2 (D^{(2)}(r, \tau) P(r, \tau)) = \frac{2b_2}{\sigma(\tau)} [1 + 2x d_x] p(x) + \frac{1}{\sigma(\tau)} \left[\frac{b_0(\tau)}{\sigma^2(\tau)} + 2b_2 x^2 \right] d_x^2 p(x). \quad (\text{A.5})$$

Das Eqs. (A.3), (A.4) e (A.5), obtêm-se que:

$$\left[\frac{b_0(\tau)}{\sigma^2(\tau)} + b_2 x^2 \right] d_x^2 p(x) + \left[a_1 + 4b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)} \right] x d_x p(x) + \left[a_1 + 2b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)} \right] p(x) = 0. \quad (\text{A.6})$$

De acordo com a Eq. (A.2), considerando o *ansatz* q -gaussiano (3.12), obtêm-se a forma τ -invariante:

$$p(x) \equiv C_q \frac{1}{[\mathcal{A} + b_2 x^2]^\theta}, \quad (\text{A.7})$$

onde $\theta = \frac{\mu+1}{2}$, C_q é uma constante de normalização e $\mathcal{A} = b_0(\tau)/\sigma^2(\tau)$ é independente de τ .

Substituindo a Eq. (A.7) na Eq. (A.6), e agrupando os termos equivalentes:

$$\frac{4\theta(\theta+1)b_2x^2 - [a_1 + 4b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)}]2b_2\theta x^2}{[\mathcal{A} + b_2x^2]^{1+\theta}} + \frac{a_1 + 2b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)} - 2b_2\theta}{[\mathcal{A} + b_2x^2]^\theta} = 0. \quad (\text{A.8})$$

Igualando a zero os coeficientes do polinômio, temos para os dois termos da Eq. (A.8):

$$a_1 + 2b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)} = 2b_2\theta. \quad (\text{A.9})$$

Logo,

$$\theta = \frac{a_1 + 2b_2 + \frac{\sigma'(\tau)}{\sigma(\tau)}}{2b_2}. \quad (\text{A.10})$$

A partir da Eq. (3.12), encontramos:

$$\mu = 1 + \frac{a_1 + \frac{\sigma'(\tau)}{\sigma(\tau)}}{b_2}. \quad (\text{A.11})$$

Considerando a Eq. (3.10) para $b_0(\tau)$:

$$\frac{b'_0(\tau)}{b_0(\tau)} = -\gamma \ln 2 \equiv -B. \quad (\text{A.12})$$

Do fato que $b_0(\tau) \propto \sigma^2(\tau)$:

$$\frac{\sigma'(\tau)}{\sigma(\tau)} = \frac{1}{2} \frac{b'_0(\tau)}{b_0(\tau)} = -\frac{B}{2}. \quad (\text{A.13})$$

Logo, usando a Eq. (A.11)

$$\mu = 1 + \frac{a_1 - B/2}{b_2}, \quad (\text{A.14})$$

que equivale à Eq. (3.13).

Vamos agora proceder à dedução da Eq. (3.14). A Eq. (3.12) pode ser reescrita como:

$$P(r, \tau) \sim \left(1 + \frac{b_2}{b_0(\tau)} r^2\right)^{-(\mu+1)/2}. \quad (\text{A.15})$$

Relacionando os expoentes das Eqs. (A.15) e (2.10), temos:

$$\begin{aligned} \frac{1 + \mu}{2} &= \frac{1}{q - 1}, \\ q &= \frac{3 + \mu}{1 + \mu}. \end{aligned} \quad (\text{A.16})$$

Assim, podemos estabelecer a relação:

$$\frac{q-1}{5-3q} = \frac{1}{\mu-2}. \quad (\text{A.17})$$

Comparando-se a Eq. (A.15) com a expressão da q -gaussiana dada pela Eq. (2.14), e utilizando a Eq. (A.17), obtém-se:

$$\frac{1}{(\mu-2)\sigma^2(\tau)} = \frac{b_2}{b_0(\tau)}. \quad (\text{A.18})$$

Usando a Eq. (A.14):

$$\sigma^2(\tau) = \frac{b_0(\tau)}{a_1 - B/2 - b_2}, \quad (\text{A.19})$$

que equivale à Eq. (3.14).

Assim, obtivemos a Eq. (3.12) como a solução invariante da EFP (3.11) com os vínculos dados pelas Eqs. (3.13) e (3.14).

A.1.2

Influência do reescalonamento no cálculo dos coeficientes de KM

Considere um retorno r , que desejamos reescalonar por um fator σ . Para o valor reescalonado $r' = r/\sigma$, temos, da definição do coeficiente de KM de primeira ordem (2.34) e (2.35):

$$D^{(1)'}(r') = \frac{1}{\sigma} D^{(1)}(r). \quad (\text{A.20})$$

Considerando a dependência linear em r , $D^{(1)}(r) = a_1 r + a_0$, tem-se que:

$$D^{(1)'}(r') = \frac{a_1 r + a_0}{\sigma} = a_1 \frac{r}{\sigma} + \frac{a_0}{\sigma} = a_1' r' + a_0'.$$

Assim, o coeficiente reescalonado $D^{(1)'}(r')$ será dado pelos parâmetros $a_1' = a_1$ e $a_0' = a_0/\sigma$.

Para o coeficiente de difusão $D^{(2)}$,

$$D^{(2)'}(r') = \frac{1}{\sigma^2} D^{(2)}(r). \quad (\text{A.21})$$

Considerando comportamento parabólico em r , $D^{(2)}(r) = b_2 r^2 + b_1 r + b_0$, tem-se que:

$$\begin{aligned} D^{(2)'}(r') &= \frac{b_2 r^2 + b_1 r + b_0}{\sigma^2} \\ &= b_2 \frac{r^2}{\sigma^2} + \frac{b_1 r}{\sigma} + \frac{b_0}{\sigma^2} = b_2' r'^2 + b_1' r' + b_0' \end{aligned} \quad (\text{A.22})$$

Logo, o coeficiente reescalado $D^{(2)'}(r')$ terá seus parâmetros modificados por $b'_2 = b_2$, $b'_1 = b_1/\sigma$ e $b'_0 = b_0/\sigma^2$.

A.1.3

Expansão de Itô-Taylor para coeficientes de KM[61]

Apresentaremos nesta seção, algumas etapas do cálculo da expansão de Itô-Taylor até a segunda ordem, para as correções em τ -finito dos coeficientes de KM dadas pelas Eqs. (6.3) e (6.4), onde $\tau \equiv \Delta t$.

Consideramos nesta seção, por simplicidade de notação, $D^{(1)}(x(t), t) \equiv A(t)$ e $\sqrt{2D^{(2)}(x(t), t)} \equiv B(t)$.

A partir da Eq. (2.36), temos:

$$dx(t) = A(t)dt + B(t)dW, \quad (\text{A.23})$$

ou ainda, na forma integral

$$\Delta x(t) = x(t + \tau) - x(t) = \int_t^{t+\tau} A(t')dt' + \int_t^{t+\tau} B(t')dW(t'). \quad (\text{A.24})$$

Chamaremos as integrais da Eq. (A.24), respectivamente de I e II:

$$I = \int_t^{t+\tau} A(t')dt',$$

$$II = \int_t^{t+\tau} B(t')dW(t').$$

Na Eq. (A.24) acima, vemos que é necessário conhecermos valores das funções A e B para valores intermediários de tempo t' .

Para tal, vamos derivar Fórmula de Itô [63] que fornece uma expressão para variações infinitesimais de funções arbitrárias F da variável estocástica X_t , $F(X_t, t)$, considerando até a 1ª ordem em dt:

$$dF = \partial_t F(X_t, t)dt + \partial_{X_t} F(X_t, t)dX_t + \frac{1}{2}\partial_{X_t X_t}^2 F(X_t, t)(dX_t)^2. \quad (\text{A.25})$$

Substituindo (A.23) em (A.25):

$$dF = \partial_t F dt + \partial_{X_t} F [A(t)dt + B(t)dW] + \frac{1}{2}\partial_{X_t X_t}^2 F [A^2(t)(dt)^2 + A(t)B(t)dt dW + B^2(t)(dW)^2]. \quad (\text{A.26})$$

Usando que $\langle (dW)^2 \rangle = dt$ e reterdo apenas termos até a ordem dt:

$$dF = \left[\partial_t F + A(t) \partial_{X_t} F + \frac{1}{2} B^2(t) \partial_{X_t X_t}^2 F \right] dt + B(t) \partial_{X_t} F dW, \quad (\text{A.27})$$

que reproduz a Eq. (6.6).

Reescrevendo a Eq. (A.27):

$$dF = L_0 F dt + L_1 F dW, \quad (\text{A.28})$$

com os operadores L_0 e L_1 definidos por:

$$\begin{aligned} L_0 &\equiv \partial_t F dt + A(t) \partial_{X_t} + \frac{1}{2} B^2(t) \partial_{X_t X_t}^2, \\ L_1 &\equiv B(t) \partial_{X_t}. \end{aligned} \quad (\text{A.29})$$

A fórmula de Itô apresentada na Eq. (A.27), na forma integral, se escreve como:

$$F(X_{t'}, t') = F(X_t, t) + \int_t^{t'} dF(t'') dt''. \quad (\text{A.30})$$

Considerando a Eq. (A.30) para $F(X_{t'}, t') \equiv A(t')$ na expressão I da Eq. (A.24):

$$\begin{aligned} I &= \int_t^{t+\tau} A(t') dt' = \int_t^{t+\tau} dt' A(t) + \int_t^{t+\tau} dt' \int_t^{t'} L_0 A(t'') dt'' + \\ &+ \int_t^{t+\tau} dt' \int_t^{t'} L_1 A(t'') dW(t''). \end{aligned} \quad (\text{A.31})$$

Utilizando as definições dos operadores L_0 e L_1 dados pela Eq. (A.24)(com $\partial A / \partial t = 0$),

$$\begin{aligned} I &= \int_t^{t+\tau} dt' A(t) + \int_t^{t+\tau} dt' \int_t^{t'} A(t'') \partial_{X_{t''}} A(t'') dt'' \\ &+ \int_t^{t+\tau} dt' \int_t^{t'} \frac{1}{2} B^2(t'') \partial_{X_{t''} X_{t''}}^2 A(t'') dt'' \\ &+ \int_t^{t+\tau} dt' \int_t^{t'} B(t'') \partial_{X_{t''}} A(t'') dW(t''). \end{aligned} \quad (\text{A.32})$$

Analogamente, utilizando a Eq. (A.30) para $F(X_{t'}, t') \equiv B(t')$ na expressão II da Eq. (A.24):

$$\begin{aligned}
 II = & \int_t^{t+\tau} B(t') dW(t'') = \int_t^{t+\tau} dW(t') B(t) + \int_t^{t+\tau} dW(t') \int_t^{t'} L_0 B(t'') dt'' + \\
 & + \int_t^{t+\tau} dW(t') \int_t^{t'} L_1 B(t'') dW(t''). \quad (A.33)
 \end{aligned}$$

Utilizando as definições dos operadores L_0 e L_1 dados pela Eq. (A.24) (com $\partial B/\partial t = 0$),

$$\begin{aligned}
 II = & \int_t^{t+\tau} dW(t') B(t) + \int_t^{t+\tau} dW(t') \int_t^{t'} A(t'') \partial_{X_{t''}} B(t'') dt'' \\
 & + \int_t^{t+\tau} dW(t') \int_t^{t'} \frac{1}{2} B^2(t'') \partial_{X_{t''} X_{t''}}^2 B(t'') dt'' \\
 & + \int_t^{t+\tau} dW(t') \int_t^{t'} B(t'') \partial_{X_{t''}} B(t'') dW(t''). \quad (A.34)
 \end{aligned}$$

Novamente, os valores de A e B em instantes intermediários t'' podem ser expandidos usando iterativamente a Fórmula de Itô (A.30), obtendo-se correções de ordem superior.

Truncando a expansão em (A.32) e (A.34), as operações no instante t'' são tomadas no instante inicial t. Assim,

$$\begin{aligned}
 I = & A(t) \int_t^{t+\tau} dt' + A(t) \partial_{X_t} A(t) \int_t^{t+\tau} dt' \int_t^{t'} dt'' + \frac{1}{2} B^2(t) \partial_{X_t X_t}^2 A(t) \times \\
 & \times \int_t^{t+\tau} dt' \int_t^{t'} dt'' + B(t) \partial_{X_t} A(t) \int_t^{t+\tau} dt' \int_t^{t'} dW(t''), \quad (A.35)
 \end{aligned}$$

e

$$\begin{aligned}
 II = & B(t) \int_t^{t+\tau} dW(t') + A(t) \partial_{X_t} B(t) \int_t^{t+\tau} dW(t') \int_t^{t'} dt'' + \frac{1}{2} B^2(t) \partial_{X_t X_t}^2 B(t) \times \\
 & \times \int_t^{t+\tau} dW(t') \int_t^{t'} dt'' + B(t) \partial_{X_t} B(t) \int_t^{t+\tau} dW(t') \int_t^{t'} dW(t''). \quad (A.36)
 \end{aligned}$$

Para o cálculo do valor médio $\langle x(t+\tau) - x(t) \rangle$ em (6.3) e (6.4), utilizamos em (A.35) os valores médios das seguintes integrais estocásticas [63]:

$$I_0 = \int_t^{t+\tau} dt' = \tau \quad (\text{A.37})$$

$$I_{0,0} = \int_t^{t+\tau} dt' \int_t^{t'} dt'' = \frac{\tau^2}{2} \quad (\text{A.38})$$

$$I_{0,1} = \left\langle \int_t^{t+\tau} dt' \int_t^{t'} dW(t'') \right\rangle = 0. \quad (\text{A.39})$$

Analogamente, em (A.36):

$$I_1 = \left\langle \int_t^{t+\tau} dW(t') \right\rangle = 0 \quad (\text{A.40})$$

$$I_{1,0} = \left\langle \int_t^{t+\tau} dW(t') \int_t^{t'} dt'' \right\rangle = 0 \quad (\text{A.41})$$

$$I_{1,1} = \left\langle \int_t^{t+\tau} dW(t') \int_t^{t'} dW(t'') \right\rangle = \tau. \quad (\text{A.42})$$

Finalmente, agregando as integrais I e II, obtemos o valor médio de (A.24):

$$\begin{aligned} \langle x(t+\tau) - x(t) \rangle &= AI_0 + AA'I_{0,0} + BA'I_{1,0} + \frac{1}{2}B^2A''I_{0,0} + BI_1 + AB'I_{0,1} \\ &\quad + BB'I_{1,1} + \frac{1}{2}B^2B''I_{0,1} + \text{termos de ordem mais alta} \\ &= AI_0 + BI_1 + (AA' + \frac{1}{2}B^2A'')I_{0,0} + (AB' + \frac{1}{2}B^2B'')I_{0,1} \\ &\quad + BA'I_{1,0} + BB'I_{1,1} + \text{termos de ordem mais alta} \quad (\text{A.43}) \end{aligned}$$

Utilizando os valores médios das integrais estocásticas (A.37) a (A.42), obtém-se que:

$$\langle x(t+\tau) - x(t) \rangle = A\tau + (AA' + \frac{1}{2}BA'')\frac{\tau^2}{2}. \quad (\text{A.44})$$

Logo:

$$D^{(1)}(x, t) = \frac{1}{\tau} \langle x(t+\tau) - x(t) \rangle = A + (AA' + \frac{1}{2}BA'')\frac{\tau}{2}. \quad (\text{A.45})$$

A substituição das definições de A e B reproduz a Eq. (6.8). Demonstração análoga nos leva à Eq. (6.9).

Para considerar aproximações de ordens maiores, devemos seguir os mesmos procedimentos apresentados nesta seção, contudo, aparecem nos cálculos, integrais de mais altas ordens. Embora seja possível resolvê-las

manualmente, o grande número de termos nas equações de ordens maiores do que três, sugere a utilização de programas específicos para o tratamento algébrico de equações, como o Maple©, distribuído pela Maplesoft. Ele pode ser utilizado, por exemplo, para a generalização das expressões para coeficientes de KM de terceira ou quarta ordem. Mais informações podem ser obtidas em [65].

A.1.4

Testes estatísticos de estacionariedade

Realizamos dois testes *estatísticos* para verificar a estacionariedade das séries financeiras analisadas. Utilizamos para esse fim, o programa Eviews©. Os testes adotados foram:

- (i) Augmented Dickey-Fuller (ADF);
- (ii) Philips-Perron (PP).

Os testes são baseados em uma regressão estatística da série analisada. Para ilustrá-lo, consideremos a equação estocástica:

$$Y_t = \alpha + \rho Y_{t-1} + u_t, \quad (\text{A.46})$$

onde u_t é um ruído branco. Se $|\rho| > 1$, a série é explosiva. Subtraindo Y_{t-1} nos dois membros da Eq. (A.46), temos

$$\Delta Y_t = \alpha + \delta Y_{t-1} + u_t, \quad (\text{A.47})$$

com $\delta = \rho - 1$. O teste considera $\delta = 0$ a hipótese nula. A hipótese alternativa é $\delta < 0$. É realizado um teste do tipo t , generalizado por Dickey-Fuller, no parâmetro δ estimado. Se $t >$ valor crítico do teste, então não há rejeição da hipótese nula, ou seja, existe uma raiz unitária (série não-estacionária). Se $t <$ valor crítico do teste, há rejeição da hipótese nula, ou seja, não existe raiz unitária (série estacionária). Assim, o teste gera valores críticos para p-valor de 1%, 5% e 10%, nas quais a hipótese nula é rejeitada.

A diferença entre os testes está no tratamento de séries com correlações de mais altas ordens. O teste ADF adiciona termos γ na equação de diferenças:

$$\Delta Y_t = \alpha + \beta T + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_{p-1} \Delta Y_{t-p-1} + u_t. \quad (\text{A.48})$$

Por outro lado, o teste PP faz uma correção não-paramétrica ao teste- t no coeficiente δ para levar em conta a correlação em u_t . O programa Eviews usa o estimador de correlação de Newey-West ω , que é aplicado ao teste t do PP.

Em ambos os testes, as séries de retorno para os mercados analisados e a série de volumes para o IBOVESPA foram classificadas como bastante estacionárias, ou seja o valor de t foi bem menor do que os valores de referência. Para verificar a consistência desses testes, os aplicamos também nas séries de preços, e os resultados, como esperado, classificaram estas séries como não-estacionárias.

Apresentamos alguns resultados dos testes ADF na tabela A.1.4. $CV(*)$ representa os valores críticos para a rejeição da hipótese de raiz nula. Valores menores que $CV(*)$ classificam a série como estacionária. O prefixo D significa série de incrementos de preços para 1 dia e Dlog significa retorno logarítmico para 1 dia.

Tabela A.1: Teste ADF para séries de preços, incrementos e retornos logarítmicos para os índices Merval, IBOVESPA e AOX e série de volumes para o BOVESPA, respectivamente.

Variável	ADF Test	1%CV(*)	5%CV(*)	10%CV(*)	Estacionária
Merval	-1,06	-3,97	-3,41	-3,13	NÃO
DMerval	-23,2	-3,97	-3,41	-3,13	SIM
DlogMerval	-22,0	-3,97	-3,41	-3,13	SIM
BOVESPA	0,58	-3,97	-3,41	-3,13	NÃO
DBOVESPA	-27,8	-3,97	-3,41	-3,13	SIM
DlogBOVESPA	-28,3	-3,97	-3,41	-3,13	SIM
AOX	-1,95	-3,97	-3,41	-3,13	NÃO
DAOX	-16,5	-3,97	-3,41	-3,13	SIM
DlogAOX	-16,5	-3,97	-3,41	-3,13	SIM
Volume	-39,64	-3,43	-2,86	-2,56	SIM

Os testes PP apresentados na tabela A.1.4 apontaram resultados ainda melhores em relação à presença de estacionaridade nas séries de retorno do que os testes ADF:

Em suma, os testes estatísticos ADF e PP classificam, como deveriam, todas as séries de preço como séries não-estacionárias. As séries de retornos para 1 dia ($\tau = 5$) por outro lado, são todas classificadas como bastante estacionárias, fato que valida nossas aproximações para o cálculo dos coeficientes de KM.

Uma análise mais extensa sobre os testes pode ser encontrada em [57, 73].

Tabela A.2: Teste PP para séries de preços, incrementos e retornos logarítmicos para os índices Merval, BOVESPA e AOX e série de volumes para o BOVESPA, respectivamente.

Variável	PP Test	1%CV(*)	5%CV(*)	10%CV(*)	Estacionária
Merval	-1,01	-3,97	-3,41	-3,13	NÃO
DMerval	-49,0	-3,97	-3,41	-3,13	SIM
DlogMerval	-47,8	-3,97	-3,41	-3,13	SIM
BOVESPA	0,64	-3,97	-3,41	-3,13	NÃO
DBOVESPA	-55,5	-3,97	-3,41	-3,13	SIM
DlogBOVESPA	-53,8	-3,97	-3,41	-3,13	SIM
AOX	-2,30	-3,97	-3,41	-3,13	NÃO
DAOX	-40,6	-3,97	-3,41	-3,13	SIM
DlogAOX	-40,8	-3,97	-3,41	-3,13	SIM
Volume	-81,1	-3,43	-2,86	-2,56	SIM

A.2 Análises Complementares sobre o IBOVESPA

A.2.1 Limite para escalas curtas

Apresentamos nessa seção, o comportamento dos coeficientes $\{a_i, b_j\}$ ao realizarmos o limite $\Delta\tau \rightarrow 0$. As dependências em τ dos parâmetros \tilde{a}_1 e \tilde{b}_0 são utilizadas como exemplos e exibidas nas Figs. A.1 e A.2, para diferentes valores de $\Delta\tau$.

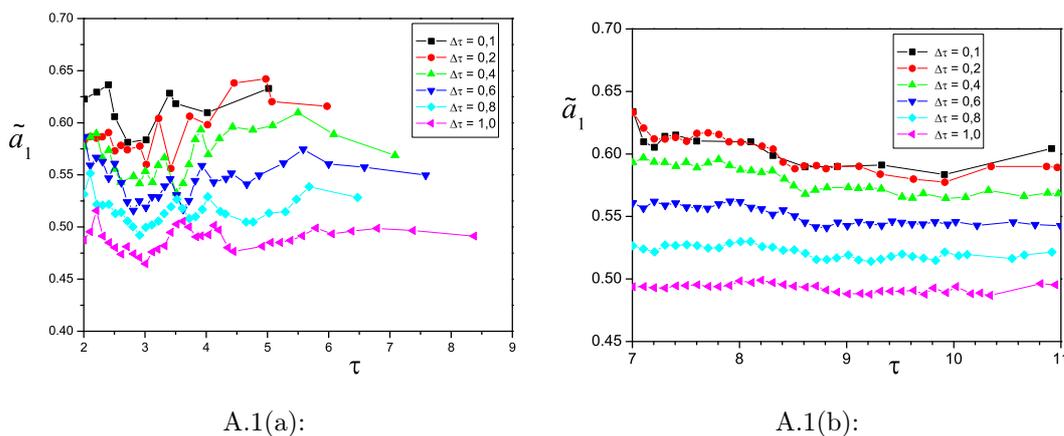


Figura A.1: Parâmetro \tilde{a}_1 calculado para diversos $\Delta\tau$ a partir das séries de 15 minutos (a) e de 30 segundos (b) apresentadas nesta tese para o IBOVESPA.

Levando-se em conta as restrições de acessibilidade empírica do limite de $\Delta\tau \rightarrow 0$, apresentados na seção (3.2), nota-se a partir das Figs. A.1 e A.2, que dentro das flutuações estatísticas, o comportamento limite dos parâmetros podem ser estimados a partir dos resultados para $\Delta\tau = 0,2$. Assim, nesta

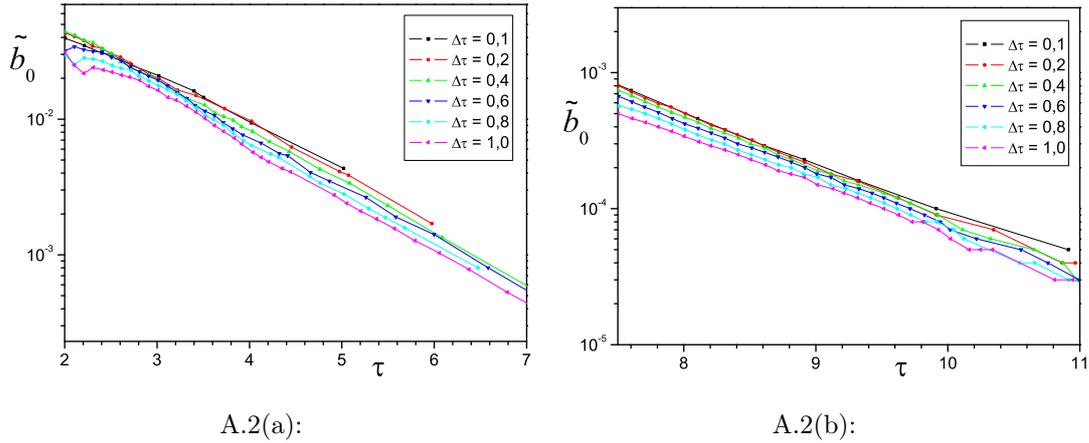


Figura A.2: Parâmetro \tilde{b}_0 calculado para diversos $\Delta\tau$ a partir das séries de 15 minutos (a) e de 30 segundos (b) apresentadas nesta tese para o IBOVESPA.

tese, nossa análise foi feita considerando os resultados obtidos para $\Delta\tau = 0,2$ como aqueles que definem $\{a_i, b_j\}$ em todos os mercados e escalas temporais utilizadas.

A.2.2

Informações sobre as séries de dados utilizadas

Ao analisar o IBOVESPA ao longo de diversas escalas de tempo, utilizamos três séries de dados, diária, e intradiárias, com ticks de 15 minutos e 30 segundos. Elas foram apresentadas no Capítulo 3. Como escala temporal definimos uma escala logarítmica reversa apresentada na Eq. (3.1).

Para realizar a evolução temporal das FDPs de retorno em direção a menores escalas de tempo, da ordem de minutos, precisamos definir a quantidade média de minutos por 32 dias presentes nas séries intradiárias. Assim, precisamos fazer uma relação entre os retornos construídos a partir das três séries, utilizando a escala τ dada pela Eq. (3.1).

Primeiramente, os dados intradiários foram verificados quanto à origem das fontes, i.e., se os valores apresentavam correspondência ao final do dia com os dos dados diários. As séries intradiárias também foram todas deflacionadas pelo IGP-DI diário, ou seja, aplicamos para todos os dados de alta frequência de cada dia, o mesmo índice deflator relativo aquela data.

A série de preços de fechamento intradiário não possui o mesmo número de ticks em cada dia e logo, em cada ano. Assim, calculamos a média de ticks intradiários para cada período estudado. Dessa forma, foi possível construir com uma melhor aproximação as FDPs de retorno para as escalas diárias a partir das escalas intradiárias.

Apresentamos na tabela A.3, o número médio de horas por dia de pregão na BOVESPA utilizando a série de dados de 15 minutos para o período entre 01/1998 e 06/2003.

Ano	dias	marcadores	média de ticks	média de horas
1998	247	7057	28,57	7h 9min
1999	246	7521	30,57	7h 39min
2000	248	7189	28,99	7h 15min
2001	244	7315	29,98	7h 30min
2002	247	7151	28,95	7h 14min
2003	61	1751	28,7	7h 11min
Total	1293	37984	29,38	7h 21min

Tabela A.3: Número médio de horas/dia de pregão da série de preços catalogadas a cada 15 minutos para o período entre 01/1998 e 06/2003 no IBOVESPA.

Verificamos na tabela A.3 que em média, não há grande diferença entre as horas médias de pregão por dia entre os anos estudados. Além disso, a série de dados de 15 minutos total possui 1293 dias e 37984 ticks. Dessa forma temos 7,34h /dia ou 29,38 ticks /dia. Para calcular os retornos reescalados, consideramos 32 dias de negócio possuindo em média 14101 minutos.

Analogamente, para a série de dados de 30 segundos, temos na tabela A.4, o número médio de minutos por dia de pregão para o período entre 01 de novembro de 2002 e 19 de julho de 2006. De acordo com esses valores, podemos perceber que a média de minutos por dia de negociação também não varia muito significativamente no período analisado. Considerando o total de ticks por dia, calculamos os retornos reescalados a partir da série de preços de 30 segundos, consideramos 32 dias de negócio possuindo em média 13799 minutos para toda a série.

Ano	dias	marcadores	média minutos/dia
2002	39	33865	434,17
2003	248	215224	433,92
2004	249	215340	432,41
2005	249	212539	426,78
2006	136	117342	431,40
Total	921	794310	431,22

Tabela A.4: Número médio de min/dia de pregão da série de preços catalogados a cada 30 segundos para o período entre 01/11/02 e 19/07/06 do IBOVESPA.

A partir da definição do número médio de minutos para 32 dias de pregão, de cada amostra de dados intradiários, construímos a tabela A.5 utilizando a Eq. (3.1). Assim, as escalas em vermelho na tabela A.5 foram construídas com $\Delta t_0 = 14101$ minutos. Analogamente, os dados em azul, com $\Delta t_0 = 13799$

minutos. Esses valores de τ são os utilizados nas FDPs teóricas e empíricas da Fig. 3.9.

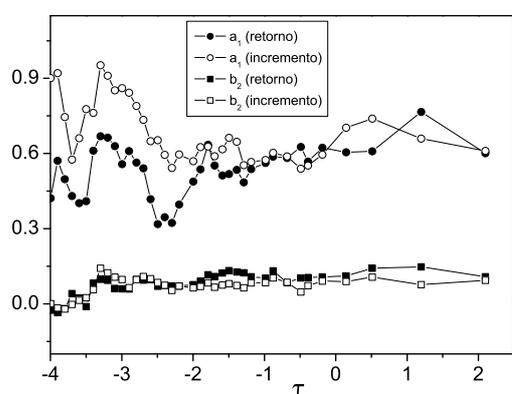
Δt	τ
128 dias	-2
64 dias	-1
32 dias	0
16 dias	1
8 dias	2
4 dias	3
2 dias	4
1 dia	5
225 minutos	5,97
105 minutos	7,07
60 minutos	7,88
30 minutos	8,88
13 minutos	10,05
8 minutos	10,75
4 minutos	11,75
2 minutos	12,75
1 minutos	13,75
30 segundos	14,75

Tabela A.5: Valores de τ segundo a Eq. (3.1) utilizados nas FDPs teóricas e empíricas apresentadas na Fig. 3.9 (de baixo para cima). As cores representam as séries de dados utilizadas para a construção do histograma empírico: preto - diária, vermelho - 15 minutos, azul - 30 segundos.

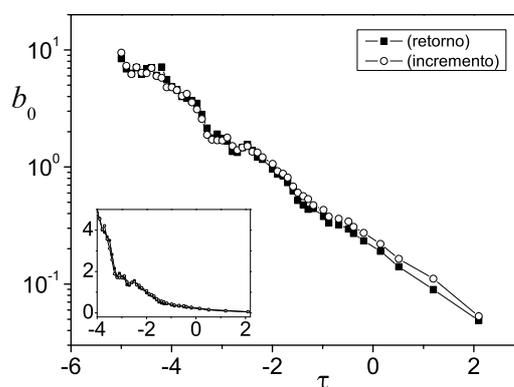
A.2.3

Comparação dos resultados para retornos logarítmicos e incrementos

Investigamos a robustez dos parâmetros de KM segundo a utilização de incrementos de preços ou retornos logarítmicos. Para tal, além dos resultados para retornos logarítmicos apresentados no Capítulo 3 para o IBOVESPA, também calculamos os incrementos, definidos pela Eq. (1.1). Nesta seção, a época utilizada foi de um período pós-Real, de 4 de julho de 1994 até 28 de junho de 2007, totalizando 3214 dados em 13 anos. Os coeficientes de KM foram calculados em uma região de $\sigma = 1,75$ e com $\Delta\tau = 0,2$. Apresentamos os resultados na Fig.A.3.



A.3(a):



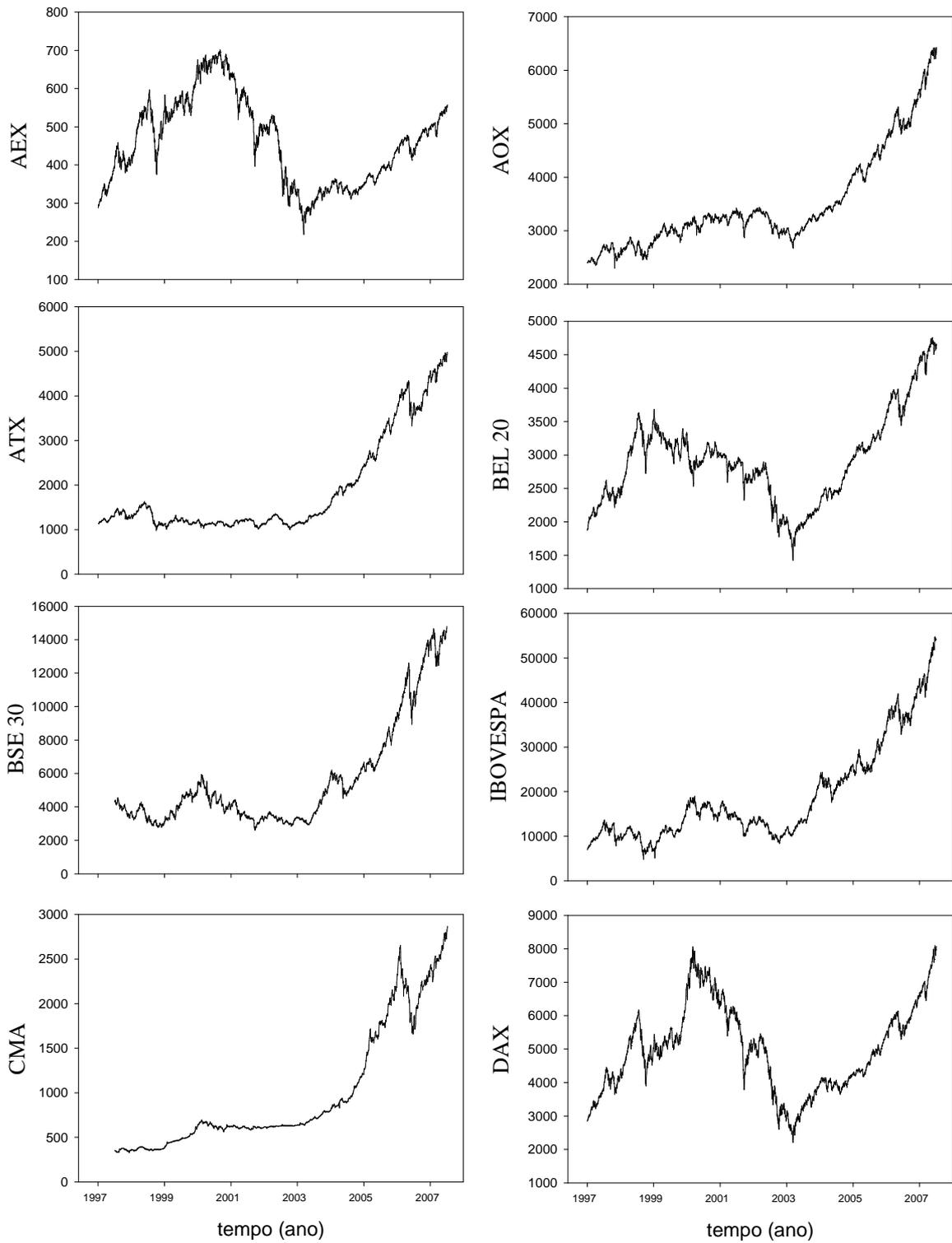
A.3(b):

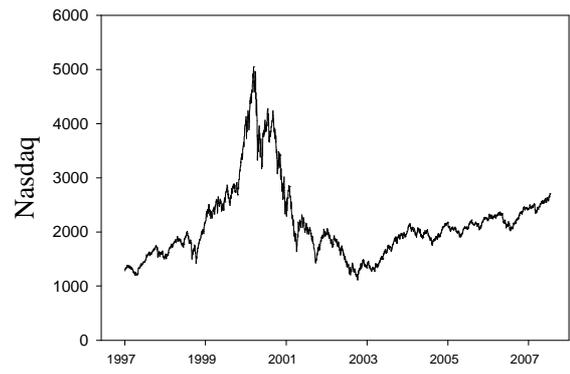
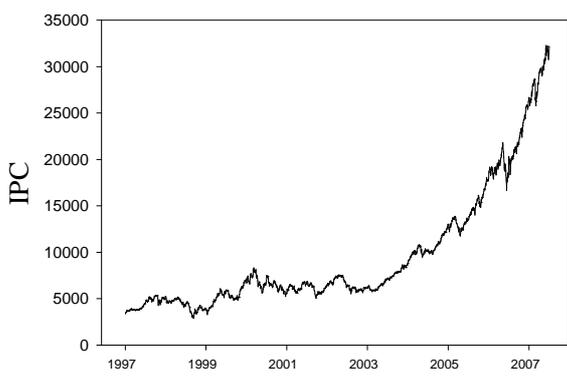
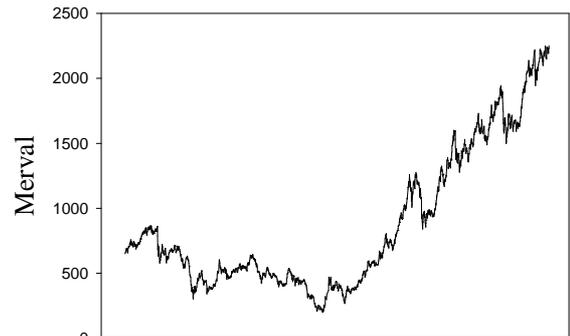
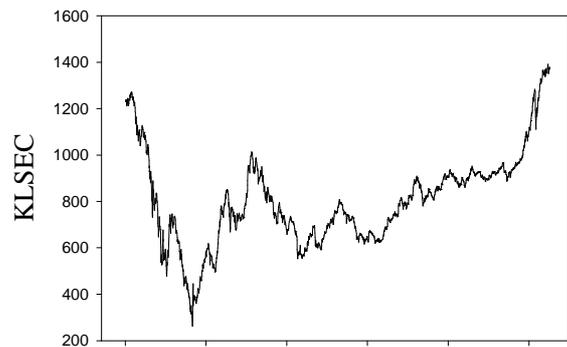
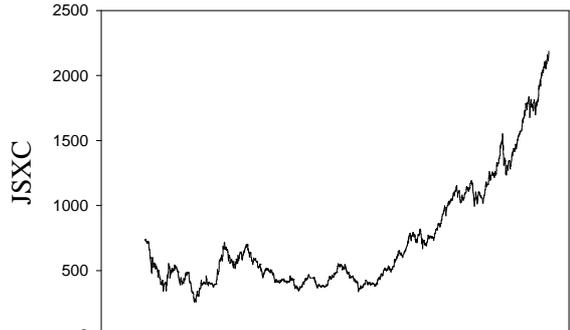
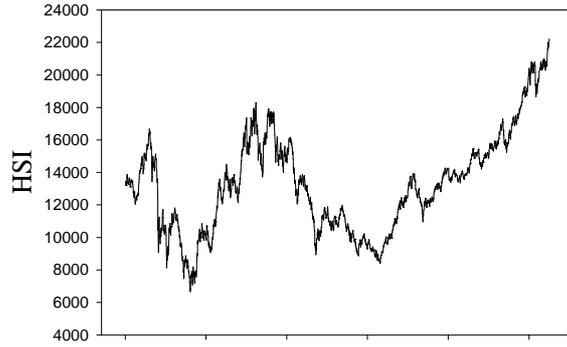
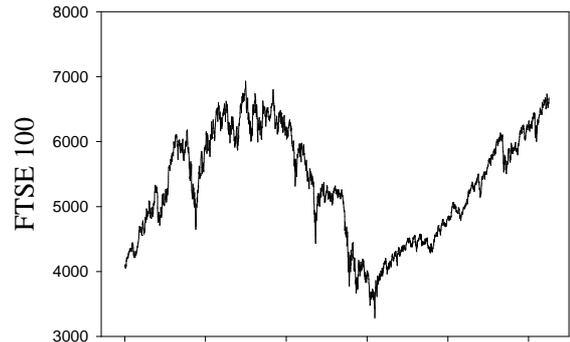
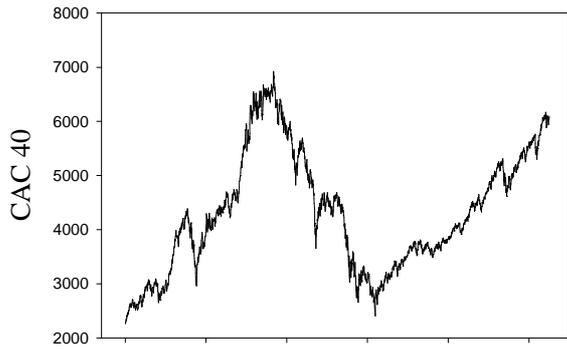
Figura A.3: Coeficientes a_1 e b_2 (a) e b_0 (b) calculados a partir de séries de retornos e incrementos diários. Eles foram calculados em uma região de $\sigma = 1,75$ e com $\Delta\tau = 0,2$.

Podemos verificar que os coeficientes calculados não mudam significativamente entre as séries de retorno e incrementos. Comparando os coeficientes calculados a partir das séries de incremento e de retorno, verificamos que o coeficiente a_1 possui nível maior para escalas menores em τ , onde a maior magnitude dos incrementos influencia mais fortemente os parâmetros. Essa diferença diminui com o aumento das escalas τ . O termo b_2 possui nível apenas ligeiramente diferente entre retornos e incrementos. Podemos perceber também que o coeficiente b_0 possui mesma inclinação para escalas $\tau < 2$.

A.3

Séries Temporais dos Índices dos Mercados Mundiais





tempo (ano)

tempo (ano)

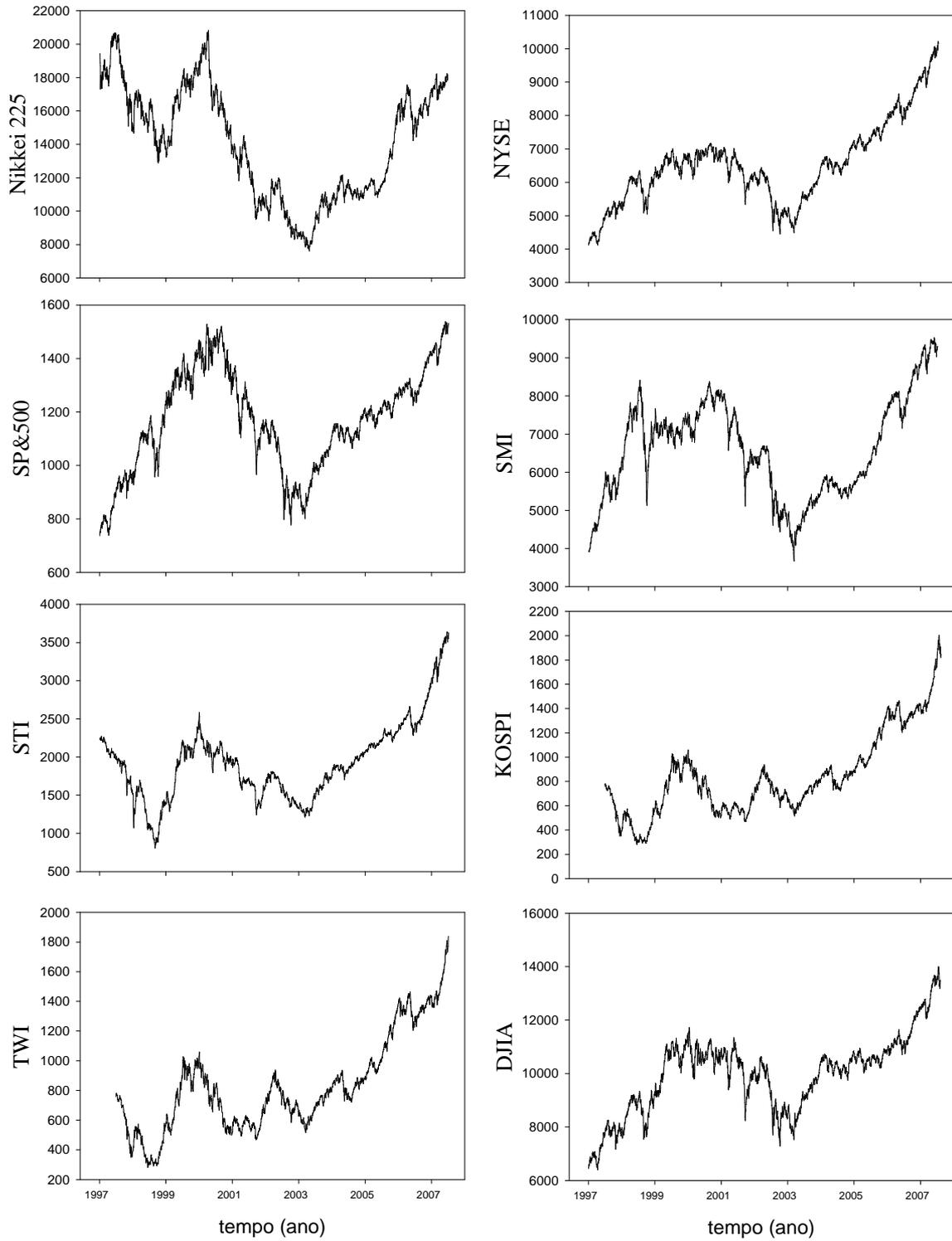


Figura A.4: Séries de preço dos índices mundiais apresentados.

A.4 Análise Estatística da Quantidade de Transações

Também investigamos a série dos lotes de ações ou número de transações, realizadas por período. Chamaremos esta quantidade simplesmente de *quantidade de transações* (Q_T). A série analisada foi do período de 2006, sendo composto de 3600 períodos de 30 minutos. Estes dados foram obtidos diretamente da BOVESPA. Apresentamos esta série na Fig. A.5, comparando-a com a série de volumes negociados para o mesmo período. Também obtivemos a partir da série de 30 minutos, novas séries em escalas maiores, agregando os dados.

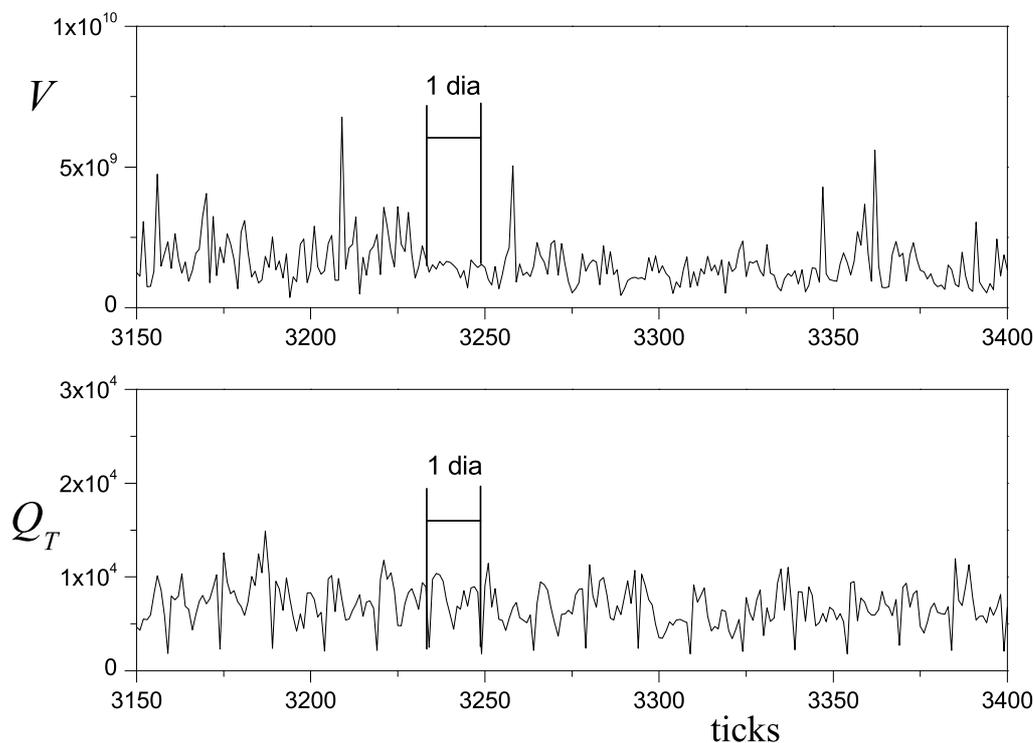


Figura A.5: Series temporais de volume e de quantidade de transações em um trecho (para melhor visualização do intervalo relativo a 1 dia) no ano de 2006 para a escala de 30 minutos.

Podemos constatar que a série de volumes é a mais volátil. A série de transações possui um ciclo diário (que compreende 15 períodos de 30 minutos) mais notável.

Ao contrário da FDP dos volumes vista no Capítulo 5, observamos que nas diferentes escalas de agregação adotadas, a FDP da quantidade de transações segue o comportamento de uma distribuição Gama (ou uma distribuição q -Gama com q próximo de 1). Podemos verificar este ajuste na Fig. A.6. Se acumularmos esta grandeza em função do número de períodos

de transação (tempo), obtemos um gráfico com uma inclinação praticamente única, como podemos comprovar na Fig. A.7. O valor dessa inclinação é de $1,7 \cdot 10^{-4}$, que multiplicada pelo valor acumulado máximo $2,14 \cdot 10^7$ nos fornece o número total de dados de forma coerente.

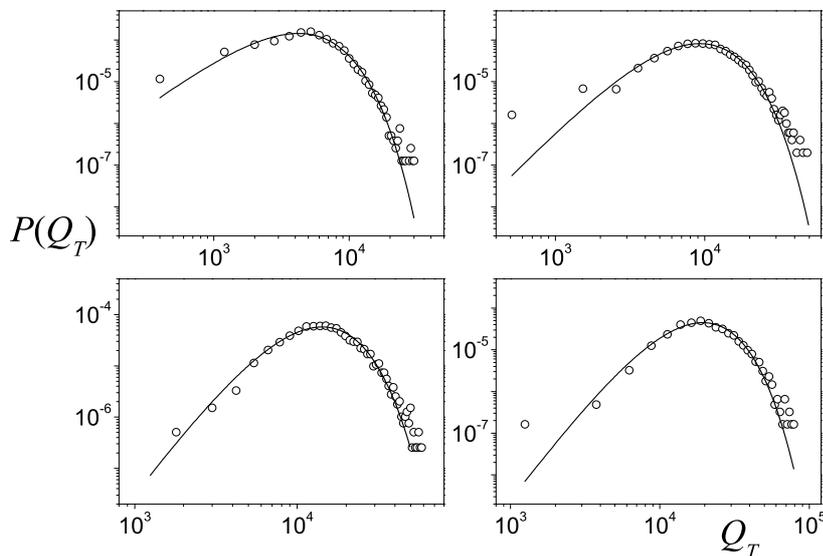


Figura A.6: Ajustes da função Gama obtidos para as FDPs da quantidade de transações para diferentes escalas para a BOVESPA no ano de 2006: 30min(a), 60min(b), 90min(c), 120min(d). A FDP Gama ajustada foi calculada com peso estatístico e para todo o intervalo de dados.

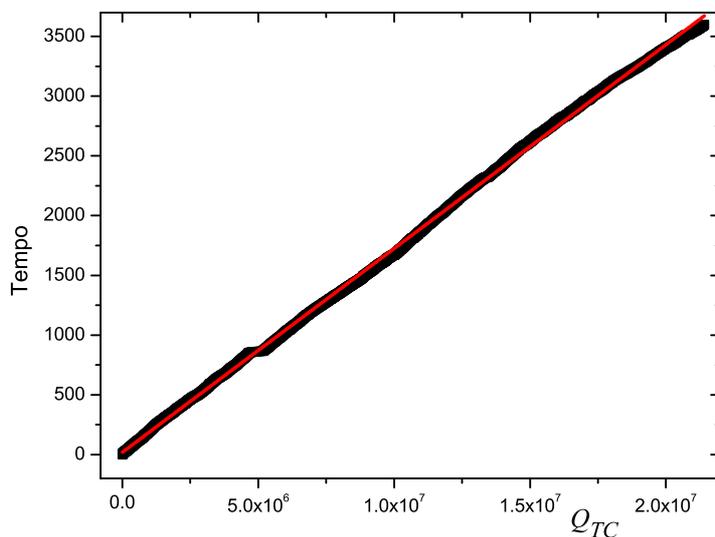


Figura A.7: Número de ticks de 30 minutos em função da quantidade de transações acumulada Q_{TC} na BOVESPA durante o ano de 2006.

Os parâmetros α e κ obtidos das distribuições Gama ajustadas para a distribuição da quantidade de transações para diversas escalas são apresentados na tabela A.6.

Escala	α	κ	10^{-4} média	$10^{-4}\sigma$
30min	7,39	802	0,59	0,32
60min	9,08	1269	1,18	0,58
90min	11,71	1467	1,77	0,80
120min	13,57	1694	2,37	1,01
150min	17,68	1622	2,96	1,21
180min	19,78	1752	3,55	1,39
210min	21,70	1858	4,14	1,60
240min	22,90	2000	4,73	1,75

Tabela A.6: Parâmetros das FDPs Gamas ajustadas para as distribuições de Q_T em diversas escalas.

Aparentemente há dois regimes diferentes, para escalas maiores e menores de 120 minutos. Contudo, essa diferença de regimes pode estar acontecendo devido a dificuldades no ajuste das distribuições ou pobreza de dados para escalas maiores.

A.5

Lista de abreviaturas

- ADF - Augmented Dickey-Fuller
- BOVESPA - Bolsa de Valores de São Paulo
- CK - Chapman-Kolmogorov
- EFP - Equação de Fokker Planck
- EL - Equação de Langevin
- FDP - Função de densidade de probabilidade
- IBOVESPA - Índice de Bolsa de Valores de São Paulo
- IGP-DI - Índice geral de preços - disponibilidade interna
- KM - Kramers-Moyal
- PE - Processo estocástico
- PP - Philips-Perron

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)