

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO-ECONÔMICO
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

JOÃO HENRIQUE GONÇALVES MAZZEU

**ANÁLISE FÍSICO-ESTATÍSTICA DA
ESTABILIDADE DAS DISTRIBUIÇÕES DE SÉRIES
FINANCEIRAS**

DISSERTAÇÃO DE MESTRADO

**FLORIANÓPOLIS
2011**

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO-ECONÔMICO
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

JOÃO HENRIQUE GONÇALVES MAZZEU

**ANÁLISE FÍSICO-ESTATÍSTICA DA
ESTABILIDADE DAS DISTRIBUIÇÕES DE SÉRIES
FINANCEIRAS**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Ciências Econômicas.

Orientador: Prof. Dr. Eraldo Sergio da Silva

**FLORIANÓPOLIS
2011**

Catálogo na fonte pela Biblioteca Universitária
da
Universidade Federal de Santa Catarina

M477a Mazzeu, João Henrique Gonçalves
Análise físico-estatística da estabilidade das
distribuições de séries financeiras [dissertação] / João
Henrique Gonçalves Mazzeu ; orientador, Eraldo Sérgio Barbosa
da Silva. - Florianópolis, SC, 2011.
112 p.: grafs., tabs.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Sócio-Econômico. Programa de Pós-Graduação
em Economia.

Inclui referências

1. Economia - Métodos estatísticos. 2. Econofísica.
3. Econometria. 4. Números primos. I. Silva, Eraldo Sérgio
Barbosa da. II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Economia. III. Título.

CDU 33

JOÃO HENRIQUE GONÇALVES MAZZEU

**ANÁLISE FÍSICO-ESTATÍSTICA DA
ESTABILIDADE DAS DISTRIBUIÇÕES DE SÉRIES
FINANCEIRAS**

Esta dissertação foi julgada adequada para a obtenção do Título de Mestre em Economia, e aprovada em sua forma final pelo Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina.

Prof. Dr. Hoyêdo Nunes Lins
Coordenador

Apresentada à Comissão Examinadora integrada pelos professores:

Prof. Dr. Eraldo Sergio da Silva
Orientador

Prof. Dr. Newton Carneiro Affonso da Costa Jr.
Examinador interno

Prof. Dr. Guilherme Valle Moura
Examinador interno

Prof. Dr. Annibal Dias de Figueiredo Neto
Examinador externo - Universidade de Brasília

DEDICATÓRIA

Dedico este trabalho aos meus pais pelo incentivo, força e compreensão.

AGRADECIMENTOS

Ao professor Dr. Sergio que acreditou na importância e conclusão deste trabalho, e pelas grandes contribuições que me foram prestadas através de seu conhecimento.

Aos professores componentes do corpo docente do programa de Pós-Graduação em Economia, destacando aqueles com quem tive maior oportunidade de contato no decorrer dos trabalhos que desenvolvi neste programa, quais sejam, Prof. Dr. Newton da Costa Jr. e Prof. Dr. André Alves Portela Santos.

Aos comentários e sugestões oferecidos pelos professores componentes da banca examinadora: Prof. Dr. Annibal Dias de Figueiredo Neto, Prof. Dr. Guilherme Valle Moura e Prof. Dr. Newton da Costa Jr.

A todos os colegas sobreviventes e não-sobreviventes da turma de 2009/2010 do mestrado em Economia ao qual apresento este trabalho. Sobretudo ao meu colega da turma 2009, Cleiton Guollo Taufemback, pelas lições iniciais de Matlab.

Reconheço ainda o apoio financeiro concedido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

RESUMO

Este trabalho realiza uma análise estatística de três séries: a primeira é formada pelos retornos da ação PETR4; a segunda, pelos retornos do índice DJIA, compreendendo o período do “flash crash” do dia 6 de maio de 2010; e a terceira, utilizada como variável de controle, é formada pelas primeiras diferenças dos números primos. As duas séries financeiras são coletadas na frequência de 1 minuto. A hipótese de que uma distribuição de Levy-estável não-Gaussiana é adequada para modelar os dados é avaliada e é dada uma atenção particular ao comportamento das caudas das distribuições. Quanto às duas séries financeiras, conclui-se que há um escalonamento não-Gaussiano e que o flash crash não pode ser considerado uma anomalia. Dos estudos das caudas, observa-se que ambas as séries financeiras seguem um padrão de lei de potência fora do regime de Levy, o qual também não é a lei cúbica inversa. Finalmente, mostra-se que a variância dependente do tempo de ambas as séries financeiras, não descrita pela distribuição de Levy-estável, pode ser modelada de uma maneira simples por um processo GARCH(1,1). Por fim, a série dos números primos, utilizada como uma variável de controle não-financeira, não apresentou evidências de escalonamento e de um padrão de lei de potência.

Palavras-chave: flash crash, econofísica, distribuições estáveis, eventos extremos, lei de potência, auto-similaridade, números primos.

Classificação JEL: C15, C16, G01.

ABSTRACT

This work carries out a statistic analysis of three series: the first one is formed by the returns of the PETR4 stocks; the second one by the returns of the DJIA index, comprising the period of the flash crash of May 6, 2010; and the third series, used as a control variable, is formed by the first differences of prime numbers. The two financial series are sampled at a one-minute frequency. The hypothesis of a non-Gaussian Levy-stable distribution to model the data is evaluated and we give particular attention to the distribution tail-bahavior. For the two financial series, we conclude that there is non-Gaussian scaling and that the flash crash cannot be considered an anomaly. From the study of tails, we find that both financial series follow a power-law pattern outside the Levy regime, which is not the inverse cubic law. Besides, we show that the time-dependent variance of both financial series, not tracked by the Levy-stable distribution, can be modeled in a straightforward manner by a GARCH(1,1) process. Finally, the series of prime numbers, used as a non-financial control variable, did not show evidences of scaling and power-law pattern.

Keywords: flash crash, econophysics, stable distributions, extreme events, power-law, auto-similarity, prime numbers.

JEL classification: C15, C16, G01.

SUMÁRIO

1. INTRODUÇÃO	19
2. CARACTERÍSTICAS DAS DISTRIBUIÇÕES ESTÁVEIS	23
2.1 PROPRIEDADE DE ESTABILIDADE.....	27
2.2 TEOREMA DE LIMITE CENTRAL GENERALIZADO (TLCG).....	27
2.3 PARAMETRIZAÇÃO.....	28
3. ECONOFÍSICA DA AÇÃO PETR4	30
3.1 MÉTODO PROPOSTO.....	30
3. 2 ANÁLISE SUMÁRIA DOS RESULTADOS DA AÇÃO PETR4.....	49
4. ECONOFÍSICA DO ÍNDICE DJIA	52
4.1 MÉTODO PROPOSTO.....	53
4.2 ANÁLISE SUMÁRIA DOS RESULTADOS DO ÍNDICE DJIA.....	69
5. ECONOFÍSICA DOS NÚMEROS PRIMOS	71
5.1 MÉTODO PROPOSTO.....	72
5.2 ANÁLISE SUMÁRIA DOS RESULTADOS DOS NÚMEROS PRIMOS.....	88
6. CONCLUSÃO	90
BIBLIOGRAFIA	92

LISTA DE FIGURAS

Figura 1 - Efeito da variação do expoente característico sobre a fdp de Levy-estável com $\beta = 0$, $\gamma = 0.5$ e $\delta = 0$.

Figura 2 - Detalhe das caudas da Figura 1. A distribuição obtém caudas mais pesadas à medida que α diminui.

Figura 3 - Efeito da variação do parâmetro de assimetria sobre a fdp de Levy-estável com $\alpha = 1$, $\gamma = 1$ e $\delta = 0$.

Figura 4 - Efeito da variação do parâmetro de escala sobre a fdp de Levy-estável com $\alpha = 1$, $\beta = 0$ e $\delta = 0$.

Figura 5 - Efeito da variação do parâmetro de posição sobre a fdp de Levy-estável com $\alpha = 1.2$, $\beta = 0$ e $\gamma = 1$.

Figura 6 - Fdp da ação Petr4 para $\Delta t = 1$ minuto, janeiro de 2007 a dezembro de 2008.

Figura 7 - Fdps dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 3, 10, 30, 100, 300, 1000$ minutos.

Figura 8 - Gráfico semilogaritmo das fdps dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 3, 10, 30, 100, 300, 1000$ minutos.

Figura 9 - Gráfico semilogaritmo das fdps dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 1000$ minutos.

Figura 10 - Probabilidade de retorno a origem medida como uma função de Δt . A dependência de lei de potência é apresentada em um gráfico log-log. No eixo y está a $P(Z = 0)$ correspondente a cada intervalo de tempo (de $\Delta t = 10$ min. até $\Delta t = 1000$ min.). A inclinação -0.6143 ± 0.025 sobre três ordens de magnitude é consistente com um escalonamento não-Gaussiano.

Figura 11 - Fdp suavizada dos retornos da ação PETR4 para $\Delta t = 1$ min. (linha contínua preta). A linha vermelha é o ajuste da Levy-estável estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

Figura 12 - P-P plot de variância estabilizada: fda dos retornos da ação Petr4 contra a fda estável estimada pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Figura 13 - Fdp dos retornos normalizados da ação Petr4 para $\Delta t = 1$ min. A fdp de Levy-estável (linha vermelha contínua) é obtida pelos parâmetros $\alpha = 1.424$, $\beta = -0.0112$, $\gamma = 0.01853$ e $\delta = 0.00018$ estimados por máxima verossimilhança. A fdp Gaussiana (linha azul tracejada) também é incluída para comparação.

Figura 14 - As mesmas fdps da Figura 9 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.424$ nas equações 5 e 6.

Figura 15 - Probabilidade de retorno a origem dos retornos da ação PETR4 (linha pontilhada) junto com a probabilidade de retorno a origem Gaussiana (linha tracejada). A distância entre as duas retas é uma medida da natureza não-Gaussiana da fdp.

Figura 16 - Dependência temporal do índice α analisada em uma escala de 24 meses. O parâmetro α ($1.45 < \alpha < 1.8$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

Figura 17 - Dependência temporal do parâmetro γ analisada em uma escala de 24 meses. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Figura 18 - Logaritmo dos retornos absolutos versus $\log(\text{rank} - 1/2)$ da ação PETR4. A inclinação corresponde à estimativa do coeficiente α_T da regressão 8. Um decaimento de lei de potência é observado com um expoente de, aproximadamente, 2.4, fora do regime de Levy e acima de 3 (lei cúbica inversa).

Figura 19 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 20 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 5% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 21 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 10% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 22 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica dos retornos da ação Petr4 para $\Delta t = 1$ min. (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 2.329 \times 10^{-5}$, $a_1 = 0.0915$ e $b_1 = 0.9$.

Figura 23 - Gráfico do índice Dow Jones (DJIA) diário durante o “flash crash” de 6 de maio de 2010.

Figura 24 - Fdp dos retornos do índice DJIA para $\Delta t = 1$ min., 1º de setembro de 2009 a 31 de maio de 2010.

Figura 25 - Fdps dos retornos do índice DJIA calculados em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 10$ minutos.

Figura 26 - Gráfico semilogarítmico das fdps dos retornos do índice DJIA calculados em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 10$ minutos.

Figura 27 - Probabilidade de retorno à origem medida como uma função de Δt . A dependência de lei de potência é apresentada em um gráfico log-log. No eixo y está a $P(Z = 0)$ correspondente a cada intervalo de tempo (de $\Delta t = 10$ min. até $\Delta t = 1000$ min.) A inclinação -0.5947 ± 0.025 sobre três ordens de magnitude é consistente com um escalonamento não-Gaussiano.

Figura 28 - Gráfico da densidade suavizada dos retornos do índice DJIA para $\Delta t = 1$ min (linha preta). A linha vermelha é o ajuste da Levy-estável estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

Figura 29 - P-P plot de variância estabilizada: fda dos retornos do índice DJIA contra a fda estável obtida pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Figura 30 - Fdp dos retornos normalizados do índice DJIA para $\Delta t = 1$ min. A fdp de Levy-estável (linha vermelha contínua) é obtida pelos parâmetros $\alpha = 1.4927$, $\beta = -0.0096$, $\gamma = 2.05441$ e $\delta = 0.04179$ estimados por máxima verossimilhança. A fdp Gaussiana (linha azul tracejada) também é incluída para comparação.

Figura 31 - As mesmas fdps da figura 25 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.4927$.

Figura 32 - As mesmas fdps da figura 26 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.4927$.

Figura 33 - Probabilidade de retorno à origem dos retornos do índice DJIA (linha pontilhada) junto com a probabilidade de retorno à origem Gaussiana (linha tracejada). A distância entre as duas retas é uma medida da natureza não-Gaussiana da fdp.

Figura 34 - Dependência temporal do índice α analisada em uma escala de 9 meses. O parâmetro α ($1.4 < \alpha < 1.8$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

Figura 35 - Dependência temporal do parâmetro γ analisada em uma escala de 9 meses. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Figura 36 - Logaritmo dos retornos absolutos versus $\log(\text{rank} - \frac{1}{2})$ do índice DJIA. A inclinação corresponde à estimativa do coeficiente α_T da regressão 8. Um decaimento de lei de potência é observado com um expoente de aproximadamente, 2.5, fora do regime de Levy e acima de 3 (lei cúbica inversa).

Figura 37 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 38 - Gráfico log-log da fda $P(g)$ para o tamanho da cauda definida por 5% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 39 - Gráfico log-log da fda $P(g)$ para o tamanho da cauda definida por 10% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Figura 40 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica dos retornos do índice DJIA para $\Delta t = 1$ min. (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 0.2348$, $a_1 = 0.0915$ e $b_1 = 0.9$.

Figura 41 - Fdp das diferenças entre dois números primos consecutivos ($\Delta k = 1$).

Figura 42 - Fdps das diferenças dos números primos calculadas em diferentes intervalos, $\Delta k = 1, 2, \dots, 10$.

Figura 43 - Gráfico semilogaritmo das diferenças dos números primos calculadas em diferentes intervalos, $\Delta k = 1, 2, \dots, 10$.

Figura 44 - Gráfico da densidade suavizada das diferenças dos números primos para $\Delta k = 1$. (linha preta). A linha vermelha é o ajuste da Levy estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

Figura 45 - P-P plot de variância estabilizada: fda das diferenças dos números primos (para $\Delta k = 1$) contra a fda estável obtida pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Figura 46 - Fdp das diferenças normalizadas dos números primos para $\Delta k = 1$. A fdp de Levy-estável (linha vermelha) é obtida pelos parâmetros $\alpha = 1.1$, $\beta = 1$, $\gamma = 4.82562$ e $\delta = 8.87558$ estimados por máxima verossimilhança.

Figura 47 - As mesmas fdps da figura 42 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.1$

Figura 48 - As mesmas fdps da figura 43 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.1$.

Figura 49 - Dependência temporal do índice α analisada em uma escala de 20 subconjuntos. O parâmetro α ($1.1 \leq \alpha \leq 1.2$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

Figura 50 - Dependência temporal do parâmetro γ analisada em uma escala de 20 subconjuntos. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Figura 51 - Dependência temporal do índice α calculada pelo método da Função Característica em uma escala de 20 subconjuntos. O parâmetro α ($1.66 < \gamma < 1.74$) permanece dentro do regime de Levy não-Gaussiano ($0 < \gamma < 2$).

Figura 52 - Dependência temporal do parâmetro γ calculada pelo método da Função Característica em uma escala de 20 subconjuntos. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Figura 53 - Logaritmo das diferenças versus $\log(\text{rank} - \frac{1}{2})$ dos números primos. A inclinação corresponde à estimativa do coeficiente α_K da regressão 21. O decaimento possui a forma de uma parábola, característico de uma distribuição gaussiana.

Figura 54 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% das diferenças ordenadas e normalizadas $g(t)$.

Figura 55 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 5% das diferenças ordenadas e normalizadas $g(t)$.

Figura 56 - Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 10% das diferenças ordenadas e normalizadas $g(t)$.

Figura 57 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica das diferenças dos números primos para $\Delta k = 1$ (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 2.0896$, $a_1 = 0.0876$ e $b_1 = 0.9$.

LISTA DE TABELAS

Tabela 1 - Testes de normalidade dos retornos da ação Petr4 para $\Delta t = 1$ minuto.

Tabela 2 - Estimativas dos parâmetros α , β , γ e δ dos retornos da ação PETR4 ($\Delta t = 1$).

Tabela 3 - Estimativas dos expoentes de decaimento dos retornos extremos da ação PETR4: método de Gabaix e Ibragimov (2007).

Tabela 4 - Estimativas dos expoentes de decaimento dos retornos extremos da ação PETR4: método de Gopikrishnan et al. (1998).

Tabela 5 - Testes de normalidade dos retornos do índice DJIA para $\Delta t = 1$ minuto.

Tabela 6 - Estimativas dos parâmetros α , β , γ e δ dos retornos do índice DJIA ($\Delta t = 1$).

Tabela 7 - Estimativas dos expoentes de decaimento dos retornos extremos do índice DJIA: método de Gabaix e Ibragimov (2007).

Tabela 8 - Estimativas dos expoentes de decaimento dos retornos extremos do índice DJIA: método de Gopikrishnan et al. (1998).

Tabela 9 - Testes de normalidade das diferenças dos números primos para $\Delta k = 1$.

Tabela 10 - Estimativas dos parâmetros α , β , γ e δ das diferenças entre os números primos ($\Delta k = 1$).

Tabela 11 - Estimativas dos expoentes de decaimento das diferenças extremas dos números primos para $\Delta k = 1$: método de Gabaix e Ibragimov (2007).

Tabela 12 - Estimativas dos expoentes de decaimento das diferenças extremas dos números primos para $\Delta k = 1$: método de Gopikrishnan et al. (1998).

LISTA DE ABREVIATURAS E SIGLAS

DJIA - Dow Jones Industrial Average

FDA - Função de distribuição acumulada

FDP - Função densidade de probabilidade

I.I.D - Independente e identicamente distribuído

IBOVESPA - Índice da Bolsa de Valores de São Paulo

MQO - Mínimos Quadrados Ordinários

P-P PLOT - Probability-Probability Plot

S&P-500 - Standard & Pools-500

TLC - Teorema de Limite Central

TLCG - Teorema de Limite Central Generalizado

VaR - Value at Risk

1. INTRODUÇÃO

Costuma-se argumentar que os retornos dos ativos financeiros são o resultado cumulativo de um vasto número de peças de informação e de decisões individuais que chegam quase continuamente no tempo (MCCULLOCH, 1996). Como tal, desde o trabalho pioneiro de Louis Bachelier em 1900, os ativos financeiros têm sido modelados pela distribuição Gaussiana. O argumento estatístico mais forte para isto está baseado no Teorema de Limite Central (TLC), o qual afirma que a soma de um número grande de variáveis independentes e identicamente distribuídas (i.i.d) de uma distribuição com variância finita tenderá a ser normalmente distribuída. Entretanto, os retornos dos ativos financeiros geralmente possuem caudas mais pesadas.

Em resposta às evidências empíricas, Mandelbrot (1963) e Fama (1965) propuseram a distribuição estável¹ como um modelo alternativo. Conforme Weron (2004), há no mínimo duas razões boas para modelar variáveis financeiras utilizando distribuições estáveis. Em primeiro lugar, estas últimas são suportadas pelo Teorema de Limite Central Generalizado (TLCG), o qual afirma que leis estáveis são as únicas distribuições limites possíveis para as somas centradas e propriamente normalizadas de variáveis aleatórias i.i.d. Em segundo lugar, distribuições estáveis são leptocúrticas². Como elas podem acomodar as caudas “gordas” e a assimetria, elas fornecem um ajuste muito melhor para as distribuições empíricas.

Em particular, as distribuições estáveis são modelos valiosos para conjuntos de dados cobrindo eventos extremos, como “crashes” de mercado ou catástrofes naturais. Mesmo que elas não sejam universais, elas são uma ferramenta útil nas mãos de um analista que trabalha com finanças ou seguros (BORAK et al. 2005). Ghahfarokhi e Ghahfarokhi (2009) mostraram que medidas de risco podem ser melhoradas pelo uso de uma distribuição estável no lugar de medidas mais convencionais. Especificamente, eles mostraram que estimativas do modelo Value at Risk (VaR) derivadas de uma distribuição estável são factíveis e

¹ Também conhecida como Alfa-estável, Pareto-estável ou Levy-estável.

² Uma distribuição é denominada leptocúrtica quando possui curtose $\kappa > 3$, com isso, ela possui um pico mais agudo do que a Gaussiana ($\kappa = 3$), denominada mesocúrtica, e possui caudas mais “pesadas” ou “gordas”.

melhores do que as estimativas obtidas pela distribuição normal e pela t-student aos níveis de 5% e 10% de significância. Mais aplicações em finanças e economia podem ser encontradas em Borak et al. (2005), Nolan (2003), Scalas e Kim (2006), Mantegna e Stanley (1995) e McCulloch (1996). Tais conjuntos de dados são incompatíveis com um modelo Gaussiano e muitos podem ser bem descritos por uma distribuição estável.

Seguindo as evidências empíricas, este estudo recorre aos conceitos da econofísica para avaliar a hipótese de que uma distribuição estável não-Gaussiana é adequada para modelar os dados de ativos financeiros. A Econofísica é um subcampo interdisciplinar, com um número crescente de praticantes. Muitos processos e conceitos largamente recorrentes na física estatística têm sido utilizados desde muito tempo nos estudos de sistemas financeiros. Auto-similaridade nas distribuições das alterações de preços está implícita no trabalho pioneiro de Bachelier (1900). Escalonamento (no sentido de que caudas de lei de potência estão presentes na distribuição logarítmica das alterações de preços) e distribuições não-Gaussianas auto-similares das alterações logarítmicas de preços foram primeiramente propostos e testados por Mandelbrot (1963) ao modelar e analisar as propriedades estatísticas dos preços do algodão.

Em um estudo mais recente, Mantegna e Stanley (1995) propuseram uma nova abordagem potencialmente poderosa para verificar que a distribuição dos retornos de ações em alta frequência possui um pico muito alto no centro e caudas que são muito “gordas” para serem consistentes com as hipóteses de uma distribuição Gaussiana. Especificamente, Mantegna e Stanley (1995) sugeriram aplicar uma distribuição de Levy-estável no centro da distribuição dos retornos do índice Standard & Poor 500 (S&P 500) e considerar caudas com declínio exponencial. Distribuições estáveis têm a importante propriedade de aditividade: se X e Y são derivados de uma distribuição estável, isto também será assegurado para suas somas. Esta é uma propriedade conveniente para a maioria dos dados econômicos, os quais são o resultado da agregação do tempo (TIMMERMANN, 1995).

Segundo Timmermann (1995), uma modelagem adequada da distribuição dos retornos de ações é importante por causa de seu significado para a gerência de risco dos investidores. Modelos financeiros padrões assumem que os investidores são avessos ao risco. A quantidade de risco que um investidor assume ao reter, por exemplo, o índice de mercado modelado por Mantegna e Stanley (1995), está

refletido pelas distribuições de seus retornos. Portanto, se um investidor, indevidamente, assume que os retornos são derivados de uma distribuição normal, o risco assumido ao reter o índice S&P 500 é, provavelmente, muito maior do que o esperado.

No Brasil, Gleria et al. (2002) empregaram as ferremantas desenvolvidas na física estatística e apresentadas na análise de Mantegna e Stanley (1995) e mostraram que o escalonamento da função densidade de probabilidade (fdp) dos retornos logarítmicos diários do índice Bovespa (IBOVESPA), para o período 1968-1998, pode ser modelado por um processo de Levy-estável na região modal da distribuição empírica para um período não excedendo a vinte dias de negociações. Conforme Gleria et al. (2002), seus resultados sugerem que há a presença de um processo similar e universal nos mercados de ações e as propriedades estatísticas dos mercados de ações avançados parecem se generalizar para os mercados emergentes, um resultado longe de ser óbvio.

Seguindo os resultados encontrados por Mantegna e Stanley (1995, 2000) e Gleria et al. (2002), o presente estudo tem como objetivo estudar as propriedades estatísticas de três bases de dados diferentes: a primeira composta pelos preços de fechamento da ação PETR4, pois como foi mencionado anteriormente, os retornos dos ativos financeiros são o resultado cumulativo de um vasto número de peças de informação e de decisões individuais que chegam quase continuamente no tempo. Portanto, na presença de caudas “pesadas” é natural assumir que os retornos de ações são, aproximadamente, governados por uma distribuição estável não-Gaussiana. A segunda base de dados é composta pelos fechamentos do índice Dow Jones (DJIA) compreendendo o período do crash do mercado de ações que ocorreu em 6 de maio de 2010, envolvendo ações corporativas americanas. Pois, como também já foi enfatizado, distribuições estáveis são modelos valiosos para conjuntos de dados cobrindo eventos extremos, como crashes de mercado ou catástrofes naturais, pelo fato dessas distribuições serem capazes de acomodar caudas pesadas e assimetria. Além disso, se os mercados de ações são vistos como um sistema complexo, não há necessidade de se encontrar uma causa para explicar o crash (BAK e PACZUSKI, 1995). Eventos extremos não precisam ter uma causa específica quando uma confluência de fatores está envolvida. A terceira base de dados é formada pelos primeiros 1.000.000 de números primos, tal variável será utilizada para controle, a fim de verificarmos se a análise estatística proposta pode ser aplicada em

qualquer série. Além disso, segundo Wells (2005), os números primos são tão irregulares que é atentador pensar neles como um tipo de sequência aleatória, cujo caso seria possível aplicar a teoria de probabilidade e estatística para estudá-los. Veja-se por exemplo o trabalho de Kumar et al. (2003), os quais estudaram as propriedades estatísticas das segundas diferenças na sequência dos números primos para verificar algum tipo de aleatoriedade em suas distribuições.

Neste sentido será estudada a hipótese de que uma distribuição de Levy-estável não-Gaussiana é adequada para modelar as três bases de dados e será feito um estudo minucioso das caudas de cada uma das três distribuições.

Para alcançar tal objetivo será empregado o método de análise estatística proposto por Mantegna e Stanley (1995, 2000), complementado pelas técnicas de estimação e diagnósticos sugeridos por Nolan (1999). Em seu artigo, Nolan propõe um método numérico consistente para estimar os parâmetros da distribuição estável e, além disso, sugere fazer testes de diagnósticos baseados na inspeção visual para verificar a não-estabilidade dos dados.

Assim, este estudo está organizado da seguinte forma. O Capítulo 2 faz uma introdução sobre as distribuições estáveis e descreve suas características básicas. O Capítulo 3 apresenta o método de análise estatística proposto e o aplica sobre a série dos retornos da ação PETR4, e de forma paralela são explicados alguns conceitos adicionais. O Capítulo 4 aplica a análise proposta sobre a série dos retornos do índice DJIA no período incluindo o crash de 6 de maio de 2010. O Capítulo 5 repete a análise sobre a série das diferenças dos números primos. Por fim, o Capítulo 5 apresenta as considerações finais.

2. CARACTERÍSTICAS DAS DISTRIBUIÇÕES ESTÁVEIS

A descoberta da distribuição estável é geralmente atribuída a Levy (1925), e, portanto, o nome Levy-estável tem sido designado a esta distribuição por Mandelbrot (1963). O estudo de Levy investigou a distribuição de rendas e encontrou distribuições muito mais gordas do que seria sugerido por uma distribuição Gaussiana. Entretanto, o trabalho de Levy não recebeu muita atenção na literatura financeira até Mandelbrot publicar seu estudo.

Uma distribuição estável, também conhecida como Alfa-estável, Pareto-estável ou Levy-estável, pode ser vista como uma generalização da distribuição Gaussiana que permite maior concentração próxima da média, mais valores extremos e assimetria. Com exceção da distribuição Gaussiana ($\alpha = 2$), as distribuições estáveis são leptocúrticas, com isso suas caudas são mais “pesadas”. A distribuição estável depende de quatro parâmetros, α , β , γ e δ , cujas interpretações seguem abaixo.

α , ($0 < \alpha \leq 2$), é o parâmetro de estabilidade básico, também conhecido como índice da cauda, expoente da cauda ou expoente característico. Ele determina o peso das caudas, ou seja, quanto menor o seu valor, maior a frequência e o tamanho de eventos extremos. Observa-se na Figura 1 que quanto menor o valor de α , maior a leptocurtose da fdp. A fdp com $\alpha = 2$ corresponde à distribuição Gaussiana. Quando $\alpha < 2$, a variância é infinita, como se pode observar na Figura 2, quanto menor o valor de α , mais pesada é a cauda da distribuição. Dessa forma, quando $\alpha < 2$, as caudas da distribuição estável exibem um comportamento assintótico de lei de potência³, aproximadamente, do tipo $|x|^{-(1+\alpha)}$. Especificamente, $E\{|x|^n\}$ diverge quando $n \geq 2$. Como já foi mencionado, todos os processos de Levy-estável com $\alpha < 2$ têm variância infinita. Isto significa que os processos estocásticos estáveis não-Gaussianos não possuem uma escala característica (MANTEGNA e STANLEY, 2000).

³ Uma lei de potência é uma forma assumida por um número notável de regularidades, ou “leis”, encontradas na economia e em finanças. É uma relação do tipo $Y = kX^\alpha$, onde Y e X são variáveis de interesse, α é o expoente de lei de potência e k é uma constante (GABAIX, 2008).

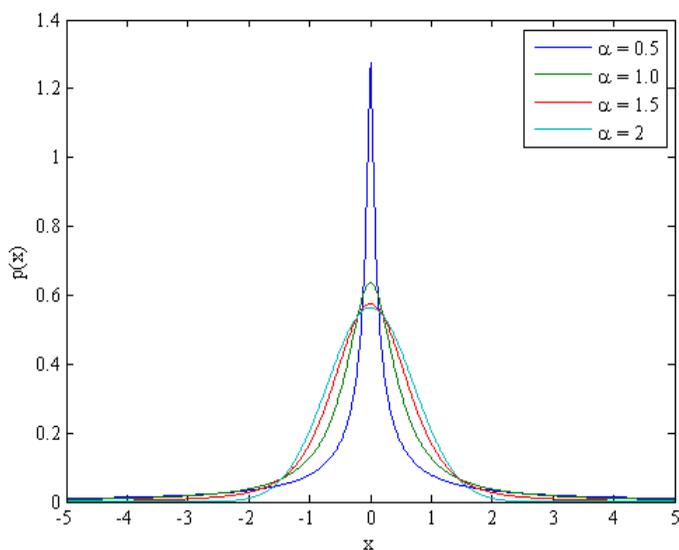


Figura 1 - Efeito da variação do expoente característico sobre a fdp de Levy-estável com $\beta = 0$, $\gamma = 0.5$ e $\delta = 0$.

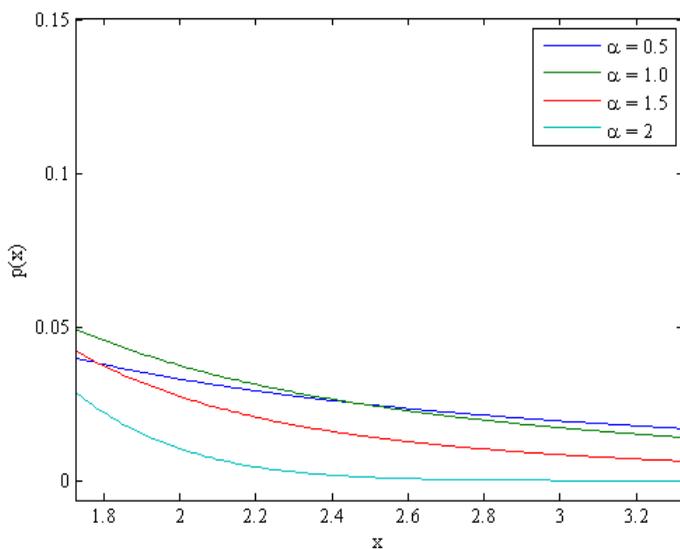


Figura 2 - Detalhe das caudas da Figura 1. A distribuição obtém caudas mais pesadas à medida que α diminui.

β , ($-1 \leq \beta \leq 1$), é o parâmetro de assimetria. Portanto, um beta igual a zero implica que a distribuição é simétrica. Um beta negativo ou positivo implica que a distribuição é assimétrica à esquerda ou à direita, respectivamente (Figura 3).

À medida que α se aproxima de 2, β perde seu efeito e a distribuição estável se aproxima da distribuição Gaussiana, independente de β .

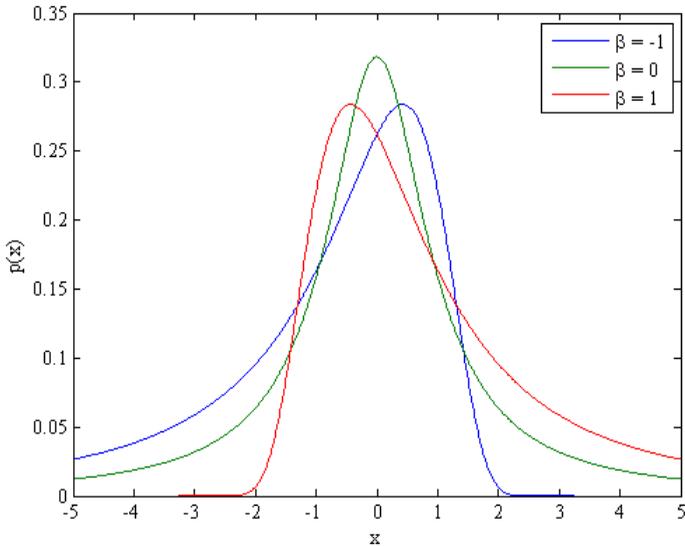


Figura 3 - Efeito da variação do parâmetro de assimetria sobre a fdp de Levy-estável com $\alpha = 1$, $\gamma = 1$ e $\delta = 0$.

O parâmetro γ é positivo e mede a dispersão ou a escala da distribuição. Ele é análogo à variância de uma distribuição normal. Como pode ser visto na Figura 4, à medida que γ aumenta, a fdp se expande em torno de δ .

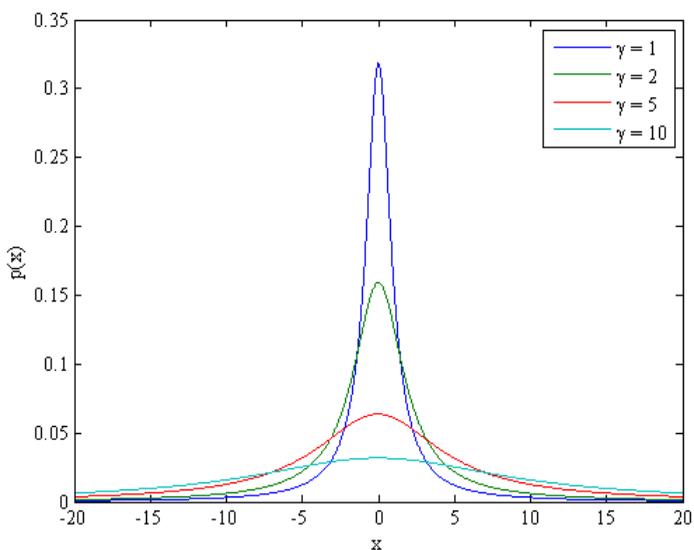


Figura 4 - Efeito da variação do parâmetro de escala sobre a fdp de Levy-estável com $\alpha = 1$, $\beta = 0$ e $\delta = 0$.

O parâmetro δ é um número real qualquer e pode ser visto como uma medida de posição. Ele é análogo à média de uma distribuição normal. Pela Figura 5 observa-se que à medida que δ aumenta, a distribuição se desloca para direita. Quando $\alpha > 1$, a média da distribuição existe e é igual a δ , mas se $\alpha \leq 1$, a média é indefinida.

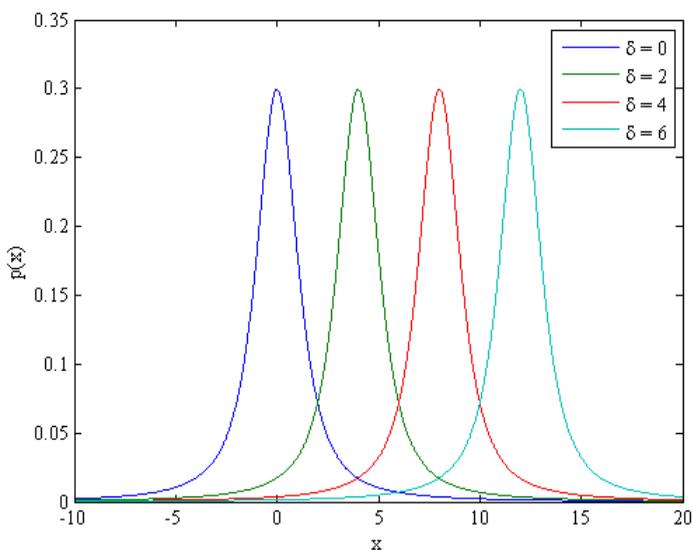


Figura 5 - Efeito da variação do parâmetro de posição sobre a fdp de Levy-estável com $\alpha = 1.2$, $\beta = 0$ e $\gamma = 1$.

2.1. PROPRIEDADE DE ESTABILIDADE

Se um número de variáveis aleatórias i.i.d possui uma distribuição estável, então uma combinação linear dessas variáveis terá a mesma distribuição, com exceção dos parâmetros de dispersão e posição que poderão apresentar valores diferentes.

Para ser mais preciso, as variáveis aleatórias X_1, \dots, X_n são i.i.d e simetricamente estáveis com o mesmo expoente característico α , se e somente se, para quaisquer constantes $\alpha_1, \dots, \alpha_n$, a combinação linear $\sum_{i=1}^n \alpha_i X_i$ tem a mesma distribuição estável. Deste modo, a distribuição estável tem a propriedade de manter sua forma funcional inalterada à medida que se altera o valor de n (GHAHFAROKHI e GHAHFAROKHI, 2009).

2.2. TEOREMA DE LIMITE CENTRAL GENERALIZADO (TLCG)

É uma generalização do Teorema de Limite Central (TLC), o qual afirma que a família das distribuições estáveis contém todas as distribuições limites das somas de variáveis aleatórias i.i.d. O TLC

afirma que a soma de n variáveis aleatórias com variância finita tenderá para uma distribuição normal à medida que n aumenta. Por sua vez, Gnedenko e Kolmogorov (1954) propõem um teorema de limite que afirma que a soma de n variáveis aleatórias com distribuições de caudas de Lei de Potência com decaimento $1/|X|^{\alpha+1}$ (e, portanto, tendo variância infinita) tenderá para uma distribuição de Levy-estável $f(x; \mu, 0, c, 0)$ à medida que n aumenta. (VOIT, 2003, apud GHAHFAROKHI e GHAHFAROKHI, 2009).

2.3. PARAMETRIZAÇÃO

Devido à falta de formas analíticas para o cálculo de suas densidades, a distribuição de Levy-estável é, mais convenientemente, descrita por sua função característica $\phi(t)$ – a transformada inversa de Fourier da função densidade de probabilidade. Entretanto, segundo Nolan (2001) há no mínimo dez diferentes parametrizações para as distribuições estáveis. Todas envolvem diferentes aplicações da função característica e são úteis por várias razões técnicas. Descreveremos, aqui, duas parametrizações diferentes, as quais denotamos por $S(\alpha, \beta, \gamma, \delta_1; 1)$ e $S(\alpha, \beta, \gamma, \delta_0; 0)$.

A primeira é a mais conhecida na literatura, logo é importante entendê-la. Assim, a função característica de $X \sim S(\alpha, \beta, \gamma, \delta_1; 1)$, isto é, uma variável aleatória estável com parâmetros α , β , γ e δ_1 , é dada por Samorodnitsky e Taqqu (1994 apud Borak et al., 2005):

$$\ln\phi(t) = \begin{cases} -\gamma^\alpha |t|^\alpha \left\{ 1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} \right\} + i\delta_1 t, & \alpha \neq 1 \\ -\gamma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \ln|t| \right\} + i\delta_1 t, & \alpha = 1 \end{cases} \quad (1)$$

No entanto, para objetivos numéricos é frequentemente aconselhado usar a parametrização de Nolan (2001) ($S(\alpha, \beta, \gamma, \delta_0; 0)$):

$$\ln\phi_0(t) = \begin{cases} -\gamma^\alpha |t|^\alpha \left\{ 1 + i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} [(\gamma|t|)^{1-\alpha} - 1] \right\} + i\delta_0 t, & \alpha \neq 1 \\ -\gamma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \ln(\gamma|t|) \right\} + i\delta_0 t, & \alpha = 1 \end{cases} \quad (2)$$

Esta parametrização é uma variação da parametrização(M) de Zolotarev (1986). Suas funções características são juntamente contínuas em todos os quatro parâmetros.

Os parâmetros α , β e γ possuem as mesmas definições nas duas parametrizações, enquanto os parâmetros de posição estão relacionados por:

$$\delta_1 = \begin{cases} \delta_0 - \beta \left(\tan \frac{\pi\alpha}{2} \right) \gamma, & \alpha \neq 1 \\ \delta_0 - \beta \frac{2}{\pi} \gamma \ln \gamma, & \alpha = 1 \end{cases} \quad (3)$$

Portanto, no caso simétrico ($\beta = 0$), as duas parametrizações coincidem, mas diferem no parâmetro de posição quando $\beta \neq 0$.

Cabe lembrar que a forma analítica da distribuição de Levy-estável é conhecida somente para três casos especiais:

- Levy-Smirnov ($\alpha = 1/2$, $\beta = 1$);
- Lorentziana ou Cauchy ($\alpha = 1$, $\beta = 0$);
- Gaussiana ($\alpha = 2$).

3. ECONOFÍSICA DA AÇÃO PETR4

Segundo Borak et al. (2005), muitas técnicas das finanças modernas dependem fortemente da hipótese de que as variáveis sob investigação seguem uma distribuição Gaussiana. Entretanto, séries temporais observadas em finanças – e também em outras aplicações – se desviam, frequentemente, do modelo Gaussiano, na medida em que suas distribuições marginais têm caudas pesadas e, possivelmente, assimétricas. Em tais situações, a adequação da hipótese Gaussiana, comumente adotada, é altamente questionada. Portanto, na presença de caudas pesadas é natural assumir que os retornos de ações são aproximadamente governados por uma distribuição estável não-Gaussiana.

Com isso, a análise empírica deste estudo se inicia com uma das ações mais líquidas do mercado brasileiro e de grande peso na composição do índice IBOVESA, a PETR4. Os dados da ação provêm da BOVESPA e foram coletados na frequência de 1 minuto, compreendendo o período de janeiro de 2007 a dezembro de 2008, totalizando 207,239 observações. Os dados ainda foram ajustados por splits. Assim, com essa amostra, realizou-se uma análise estatística seguindo o método proposto de Mantegna e Stanley (1995, 2000) com o objetivo de avaliar a hipótese de que uma distribuição de Levy-estável é adequada para modelar os dados. Além disso, foi dada uma atenção particular às caudas da distribuição e à dependência temporal da variância dos dados.

A próxima seção apresenta o método proposto e as conclusões estão resumidas na seção 3.2.

3.1. MÉTODO PROPOSTO

A análise foi realizada através de dez etapas. Na primeira etapa foi construído um gráfico da fdp dos retornos, Z , definidos como

$$Z_{\Delta t}(t) \equiv Y(t + \Delta t) - Y(t) \quad (4)$$

Onde Y é uma observação do preço de fechamento da ação PETR4 no período t . Assim, primeiramente, fez-se um gráfico da fdp $P(Z)$ dos retornos $Z_{\Delta t}(t)$ usando o intervalo de $\Delta t = 1$ minuto (Figura 6). Para uma melhor visualização, o gráfico foi construído com o logaritmo de $P(Z)$ no eixo y e os retornos Z foram divididos pelo seu desvio-padrão

($\sigma = 0.0523$), e foram representados no eixo x. Uma distribuição Gaussiana (linha azul tracejada) também foi construída para comparação usando o mesmo valor da variância. Nota-se que a fdp dos dados (em círculos) é quase simétrica, altamente leptocúrtica e não-Gaussiana para as pequenas alterações da ação. Além disso, a curtose apresentou um valor bem acima de 3 ($\kappa = 445.77$), e a Gaussianidade também não foi detectada em nenhum dos testes padrões apresentados na Tabela 1.

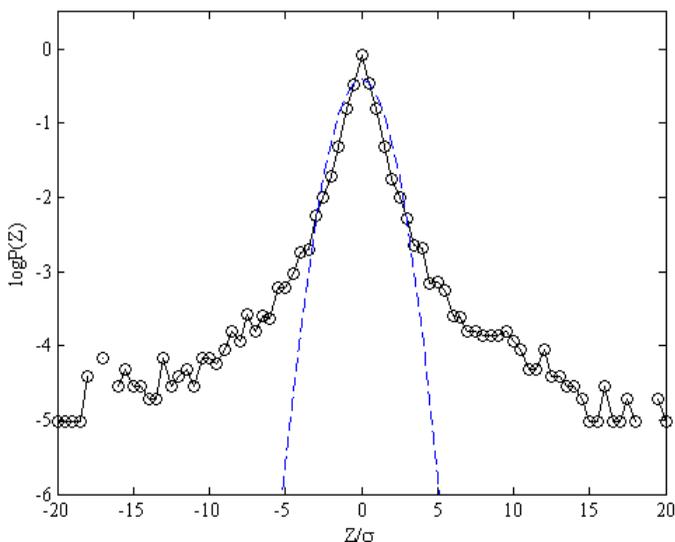


Figura 6 - Fdp da ação Petr4 para $\Delta t = 1$ minuto, janeiro de 2007 a dezembro de 2008.

Tabela 1 - Testes de normalidade dos retornos da ação Petr4 para $\Delta t = 1$ minuto.

Lilliefors	Cramer-von Mises	Anderson-Darling	Curtose
0.149623(0.0000)	1900.093 (0.0000)	10342.97 (0.0000)	445.77

Nota: p-valores em parênteses. Todos os p-valores são, aproximadamente, zeros, portanto, a hipótese nula de normalidade é rejeitada para qualquer nível de significância padrão (0.01, 0.05 ou 0.1).

Como o formato de $P(Z)$ muda com o tempo? Isto foi avaliado (na segunda etapa) considerando os valores crescentes de Δt na definição dos retornos (equação 4). Geralmente, as fdps mudam sua forma funcional e sua escala à medida que o tempo se altera. Entretanto, a fdp Gaussiana é estável, ou seja, ela não altera sua forma funcional, mas somente sua escala (ela se torna mais achatada à medida que o intervalo de tempo aumenta). Como os dados da Figura 6 são compatíveis com uma fdp estável não-Gaussiana, o comportamento temporal de $P(Z)$ foi avaliado considerando-se vários subconjuntos de $Z_{\Delta t}(t)$ extraídos em função dos valores crescentes de Δt . O número de observações em cada subconjunto reduziu de 207.239 ($\Delta t = 1$ min.) para 206.240 ($\Delta t = 1000$ min.). Figura 7 mostra as fdps para $\Delta t = 1, 3, 10, 30, 100, 300, 1000$. Figuras 8 e 9 mostram as fdps em um gráfico semilogarítmico. A partir daí, dois padrões emergem: (1) as fdps se espalham à medida que Δt aumenta, como em um processo aleatório (Figuras 8 e 9); (2) os picos das fdps ($Z = 0$) diminuem quando Δt aumenta (Figura 7), os quais são sugestivos de um escalonamento não-Gaussiano.

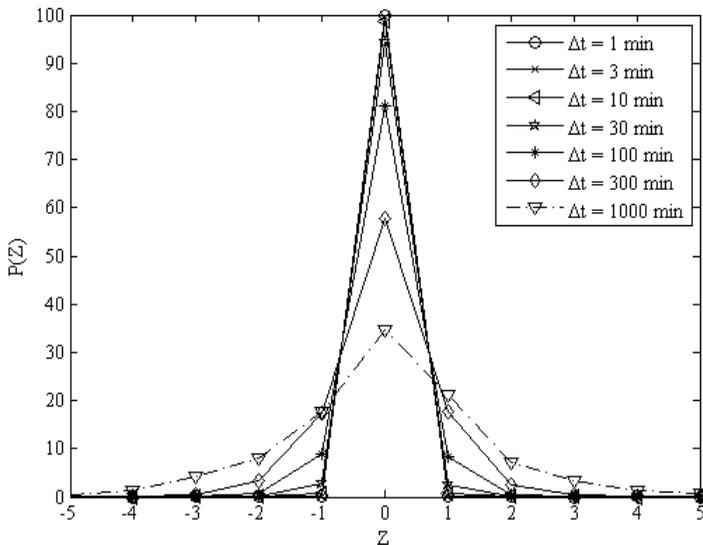


Figura 7 - Fdps dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 3, 10, 30, 100, 300, 1000$ minutos.

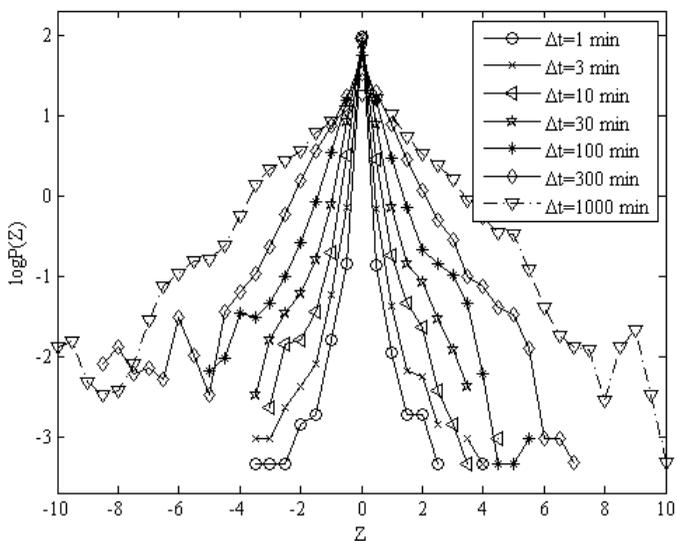


Figura 8 - Gráfico semilogarítmico das fdp dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 3, 10, 30, 100, 300, 1000$ minutos.

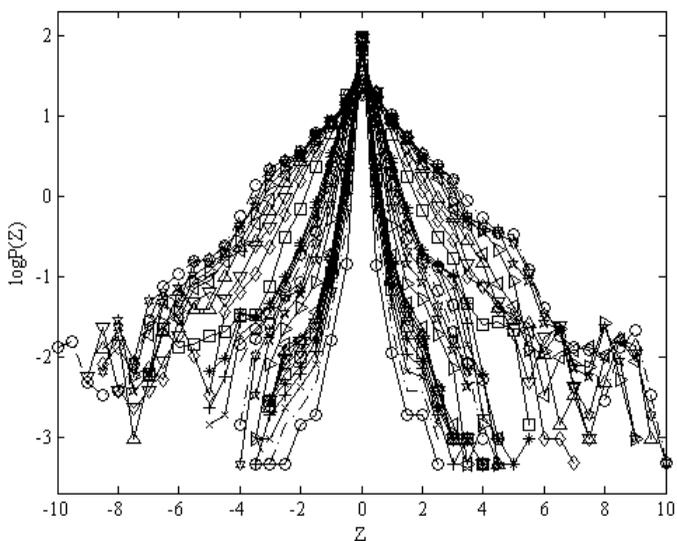


Figura 9 - Gráfico semilogarítmico das fdp dos retornos da ação Petr4 calculadas em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 1000$ minutos.

Escalonamento significa que as distribuições estáveis são auto-similares⁴ (MANTEGNA e STANLEY, 2000). As variáveis escalonadas para um processo de Levy-estável de índice α são

$$\tilde{Z} \equiv \frac{Z}{(\Delta t)^{1/\alpha}} \quad (5)$$

e

$$\tilde{P}(\tilde{Z}) \equiv P(Z) \cdot (\Delta t)^{1/\alpha} \quad (6)$$

Dessa forma, a auto-similaridade de uma distribuição estável é revelada pelas equações (5) e (6).

Assim, na terceira etapa verificou-se a existência e as características do escalonamento da fdp dos retornos aplicando o método da probabilidade de retorno a origem ($P_{\Delta t}(Z = 0)$) em função de Δt . Apresentando os resultados em um gráfico log-log, observa-se que o comportamento dos dados segue uma lei de potência (Figura 10). Este resultado é compatível com uma fdp de Levy-estável. Por fim, a existência de um escalonamento na Figura 10 justifica o gráfico escalonado das fdps apresentado nas Figuras 7, 8 e 9.

⁴ Auto-similaridade implica que a forma funcional permanece a mesma, qualquer que seja a escala utilizada. Isto significa que, seja para o curto ou longo prazo, a fdp será a mesma. Isto é uma característica dos fractais: a auto-semelhança.

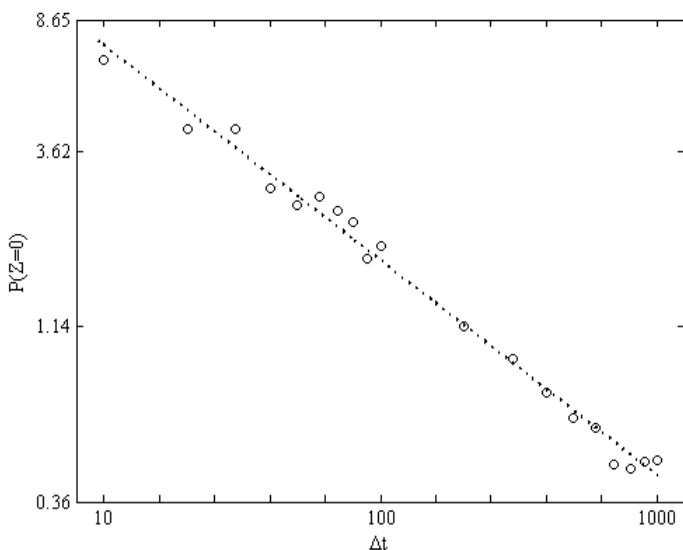


Figura 10 - Probabilidade de retorno a origem medida como uma função de Δt . A dependência de lei de potência é apresentada em um gráfico log-log. No eixo y está a $P(Z=0)$ correspondente a cada intervalo de tempo (de $\Delta t=10$ min. até $\Delta t=1000$ min.). A inclinação -0.6143 ± 0.025 sobre três ordens de magnitude é consistente com um escalonamento não-Gaussiano.

Na quarta etapa os parâmetros da distribuição de Levy-estável foram estimados. Entretanto, nesta etapa nos desviamos da análise apresentada por Mantegna e Stanley (1995, 2000) e calculamos os parâmetros α , β , γ e δ por três métodos de estimação: Método da Função Característica de Koutrovelis (1980, 1981) e de Kogon e Williams (1998), Método de Quantis de McCulloch (1986) e o Método de Máxima Verossimilhança de Nolan (1997)⁵. Os três métodos estão disponíveis no programa Stable.exe desenvolvido por Nolan (1997) (disponível em <http://academic2.american.edu/~jpnolan/>). Cabe lembrar que as estimativas dos três métodos são feitas com base na parametrização S0, pois de acordo com Nolan (2001), esta é mais adequada para cálculos numéricos e para modelar os dados do que as

⁵ Para maiores detalhes sobre os três métodos de estimação, ver Nolan (1997, 2001).

representações padrões. Ao menos que se tenha uma razão específica para usar outra parametrização, aconselha-se usar S0.

A Tabela 2 apresenta os parâmetros estimados pelos três métodos propostos.

Tabela 2 - Estimativas dos parâmetros α , β , γ e δ dos retornos da ação PETR4 ($\Delta t = 1$).

Método	α	β	γ	δ
Máxima	1.424	- 0.0112	0.01853	0.00018
Verossimilhança	(± 0.0068)	(± 0.0119)	(± 0.00009)	(± 0.00014)
Quantis	1.5635	0	0.02068	0.00000
Função	1.5968	- 0.0286	0.02021	0.00026

Nota: os intervalos com 95% de confiança estão em parênteses

Antes de prosseguir com a análise de Mantegna e Stanley (1995, 2000), foram realizados alguns diagnósticos para tentar detectar a não-estabilidade dos dados, pois como tão importante são os métodos de estimação, são os diagnósticos para avaliar a qualidade do ajuste nos dados. Segundo Nolan (1999), assim como qualquer outra família de distribuições, não é possível provar que um conjunto de dados é estável ou não. O melhor que se pode fazer é determinar se os dados são consistentes com a hipótese de estabilidade. Em primeiro lugar, Nolan (1999) sugere comparar os valores dos parâmetros entre os três métodos de estimação. Se os valores estiverem próximos, isto suporta a ideia de que os dados seguem uma distribuição estável. Se os parâmetros diferirem significativamente e a amostra for grande, então isto demonstra que os dados não seguem uma distribuição estável, porque os três métodos de estimação são consistentes em amostras grandes. Como se pode observar na Tabela 2, os três métodos apresentam valores próximos, suportando a hipótese de estabilidade dos dados.

Em segundo lugar, Nolan (1999) sugere fazer um gráfico da densidade suavizada dos dados. Se houver muitas oscilações no gráfico, então os dados não vêm de uma distribuição estável. Deste modo, para gráficos de densidade, os dados foram suavizados pela função kernel Gaussiana, obtendo-se uma densidade suavizada plausivelmente estável (Figura 11). No mesmo gráfico, as distribuições Gaussiana (linha azul tracejada) e a Levy-estável (linha vermelha contínua) foram construídas utilizando-se os parâmetros estimados por máxima verossimilhança, pois a comparação entre a fdp suavizada dos dados e o ajuste proposto

oferece um bom sentido de quão bem a distribuição teórica se ajusta à região central da fdp empírica.

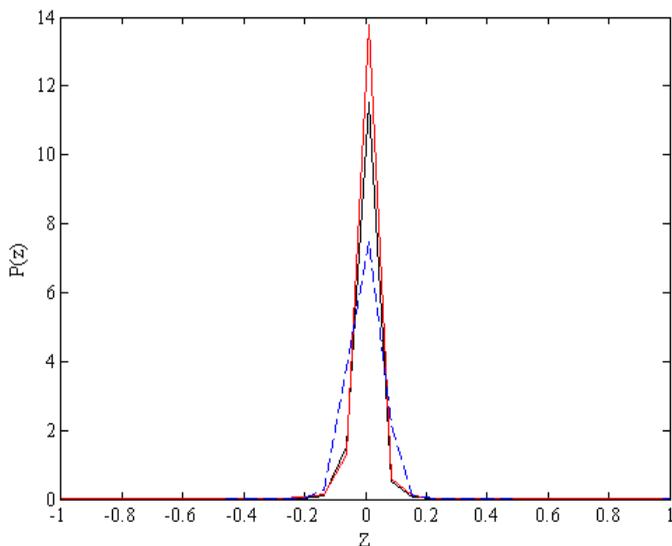


Figura 11 - Fdp suavizada dos retornos da ação PETR4 para $\Delta t = 1$ min. (linha preta contínua). A linha vermelha é o ajuste da Levy-estável estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

E por último, Nolan (1999) recomenda comparar os dados empíricos com a distribuição estável ajustada usando um p-p plot⁶ de variância estabilizada de Michael (1983). Tal p-p plot tem a propriedade de detectar um ajuste próximo aos extremos de forma mais precisa do que p-p plots padrões. Desta forma, quanto mais próximo o ajuste dos pontos na linha diagonal, melhor o ajuste da distribuição estável. Como se observa na Figura 12, o p-p plot mostrou um ajuste razoável da distribuição estável, o segmento horizontal no centro do p-p plot pode ser devido aos períodos em que a ação PETR4 não variou por sucessivos minutos. Logo, o resultado do conjunto dos diagnósticos oferece suporte para a hipótese de estabilidade dos dados, pois os parâmetros apresentaram valores próximos entre os três métodos de estimação, a

⁶ O p-p plot (probability-probability plot) é um gráfico que apresenta os valores da função distribuição acumulada (fda) empírica contra os valores da fda teórica. É utilizado para determinar o quão bem uma distribuição específica se ajusta aos dados observados.

fdp suavizada dos dados se apresentou plausivelmente estável e o p-p plot mostrou um ajuste razoável entre a distribuição estável e os dados empíricos.

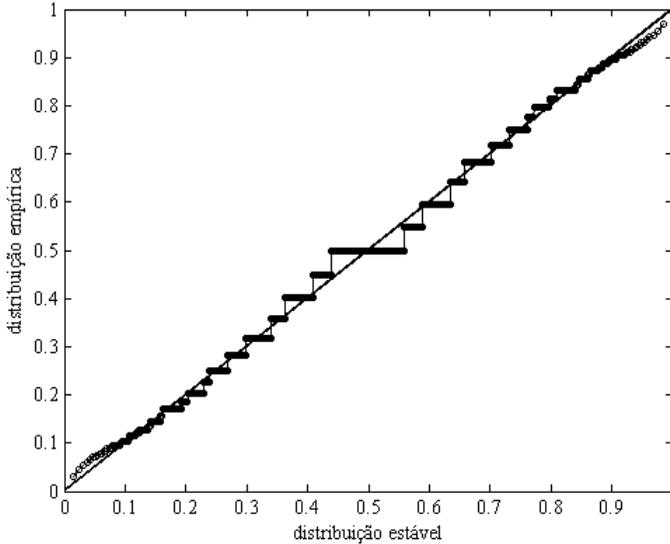


Figura 12 - P-P plot de variância estabilizada: fda dos retornos da ação Petr4 contra a fda estável estimada pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Desta maneira, realizados os diagnósticos para examinar a qualidade do ajuste da distribuição estável, procedemos com a análise de Mantegna e Stanley (1995, 2000). Continuamos a análise com os parâmetros estimados por máxima verossimilhança pelo fato deste método ter a vantagem de produzir intervalos de confiança para amostras grandes aos parâmetros. Assim, utilizando esses parâmetros, na quinta etapa a fdp de Levy-estável teórica foi construída em um gráfico semilogarítmico (Figura 13). Um ótimo ajuste foi observado na região central da fdp, mas os dados extremos se localizaram abaixo da Levy teórica, indicando quebra de escala. Isto é de se esperar, porque a Levy teórica apresenta variância infinita e o segundo momento dos dados é finito. O ajuste com a fdp empírica se apresenta melhor para o conjunto de observações

$$|Z| < 2.72\sigma \quad (7)$$

Desta forma, os dados da fdp empírica ficam abaixo da Levy teórica quando $|Z| \geq 2.72\sigma$. E as caudas decaem precisamente porque a variância da fdp empírica é finita. Logo, pode-se concluir, parcialmente, que (1) os dados dos retornos podem ser modelados por uma distribuição de Levy-estável na região central da fdp e (2) as caudas da fdp empírica merecem um exame minucioso.

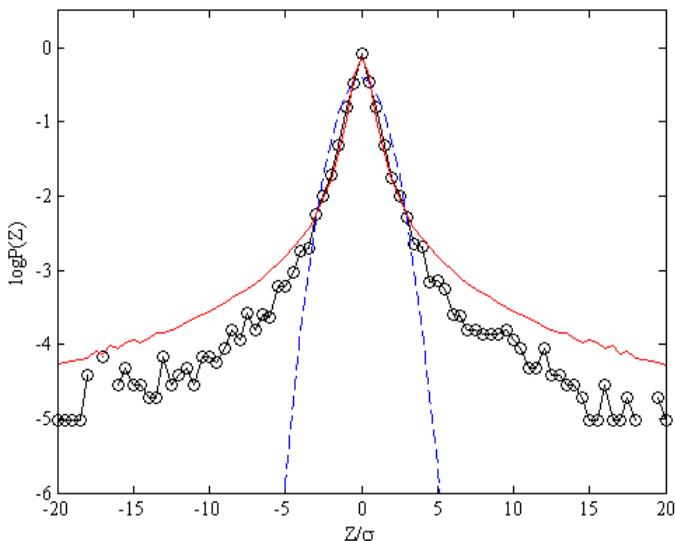


Figura 13 - Fdp dos retornos normalizados da ação Petróleo 4 para $\Delta t = 1$ min. A fdp de Levy-estável (linha vermelha contínua) é obtida pelos parâmetros $\alpha = 1.424$, $\beta = -0.0112$, $\gamma = 0.01853$ e $\delta = 0.00018$ estimados por máxima verossimilhança. A fdp Gaussiana (linha azul tracejada) também é incluída para comparação.

Vimos na Figura 10 que os máximos das fdps apresentam um escalonamento para intervalos $\Delta t \leq 1000$ minutos. Logo, na sexta etapa, a existência de um escalonamento foi estudada para as outras regiões da fdp empírica e não somente para a região onde $Z = 0$. Para isso, as equações (5) e (6) foram consideradas juntamente com $\alpha = 1.424$. Figura 14 mostra que todas as fdps empíricas (para todos os valores de Δt) colapsam na fdp de $\Delta t = 1$ minuto, o que revela também a auto-similaridade presente nas fdps. Novamente, o ajuste se apresenta melhor na região central das fdps, sugerindo, então, que o escalonamento quebre para longos intervalos de tempo. Logo, pode-se concluir que (1) há um escalonamento não-Gaussiano para os retornos da ação PETR4; e

(2) que este escalonamento, que por sua vez implica em auto-similaridade dos dados, é válido para um intervalo de tempo finito porque a variância dos dados é finita.

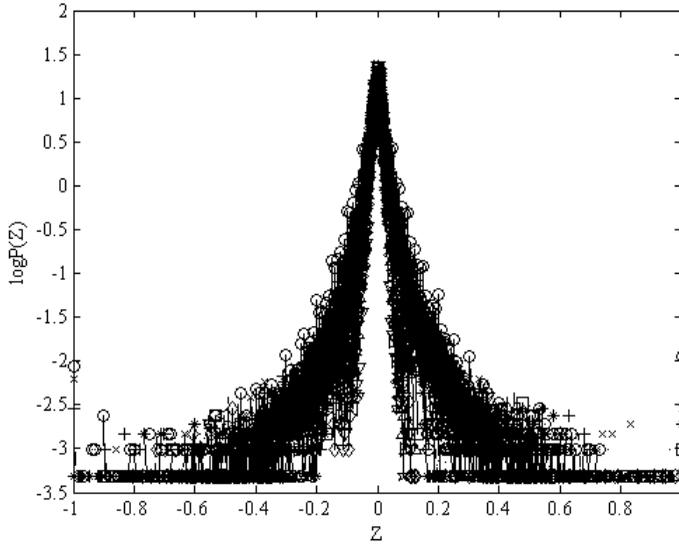


Figura 14 - As mesmas fdps da Figura 9 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.424$ nas equações (5) e (6).

Na sétima etapa o escalonamento não-Gaussiano foi comparado com o escalonamento Gaussiano para encontrar o intervalo onde o primeiro escalonamento realmente ocorre. Tomamos a probabilidade de retorno a origem empírica $P_{\Delta t}(Z=0)$ e a comparamos com a probabilidade de retorno a origem gaussiana $P_G(Z=0)$. Esta última foi calculada utilizando-se os valores da média e variância dos retornos para cada Δt . A Figura 15 mostra o gráfico log-log do escalonamento não-gaussiano em $P_{\Delta t}(Z=0)$ (linha pontilhada) e do escalonamento Gaussiano em $P_G(Z=0)$ (linha tracejada). Observa-se que, para dado Δt , a diferença entre $P_{\Delta t}(Z=0)$ e $P_G(Z=0)$ diminui quando $30 < \Delta t < 1000$ minutos. A reta dos dados encontra a reta Gaussiana em, aproximadamente, 10^4 minutos. Portanto, o escalonamento não-Gaussiano ocorre, aproximadamente, entre 1 e 10.000 minutos.

Concluindo, então, que este escalonamento é válido para um intervalo de tempo que é grande, mas finito.

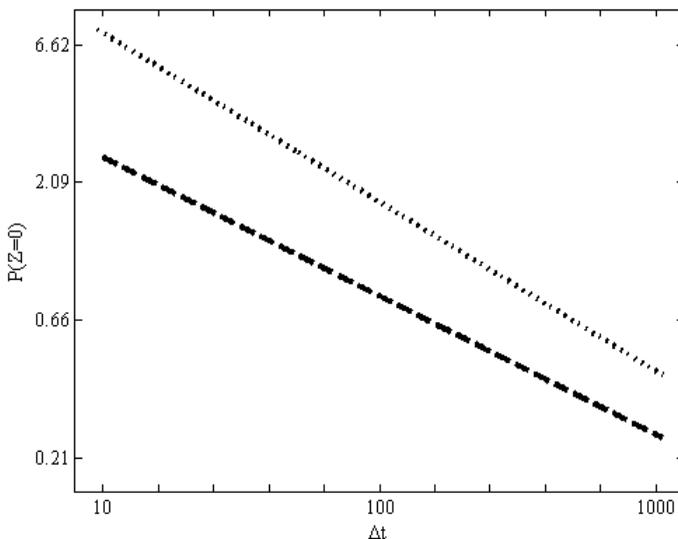


Figura 15 - Probabilidade de retorno à origem dos retornos da ação PETR4 (linha pontilhada) junto com a probabilidade de retorno a origem Gaussiana (linha tracejada). A distância entre as duas retas é uma medida da natureza não-Gaussiana da fdp.

Na oitava etapa, a robustez dos parâmetros estimados α e γ foi checada. Desta forma, a base de dados original ($\Delta t = 1$ min.) foi dividida em 24 subconjuntos (meses) e os parâmetros α e γ de cada subconjunto foram estimados por máxima verossimilhança. Figura 16 mostra que o α permanece dentro do regime de Levy não-Gaussiano ($1.45 < \alpha < 1.8$), confirmando, então, que a hipótese de uma distribuição de Levy-estável não pode ser rejeitada dos dados. Figura 17 mostra que o γ apresenta fortes flutuações ($0.01 < \gamma < 0.045$). Cabe lembrar, que o γ fornece uma medida de volatilidade do processo e que γ aumenta à medida que σ aumenta⁷. Logo, como o parâmetro γ é instável, a fdp de Levy-estável falha em descrever a volatilidade presente nos dados.

⁷ A relação entre os dois parâmetros é dada por $\gamma = \frac{\sigma^2}{2}$ (MANTEGNA E STANLEY, 2000).

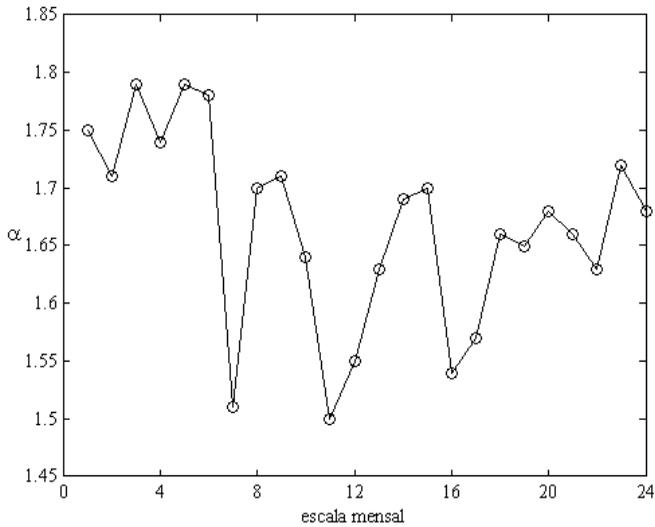


Figura 16 – Dependência temporal do índice α analisada em uma escala de 24 meses. O parâmetro α ($1.45 < \alpha < 1.8$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

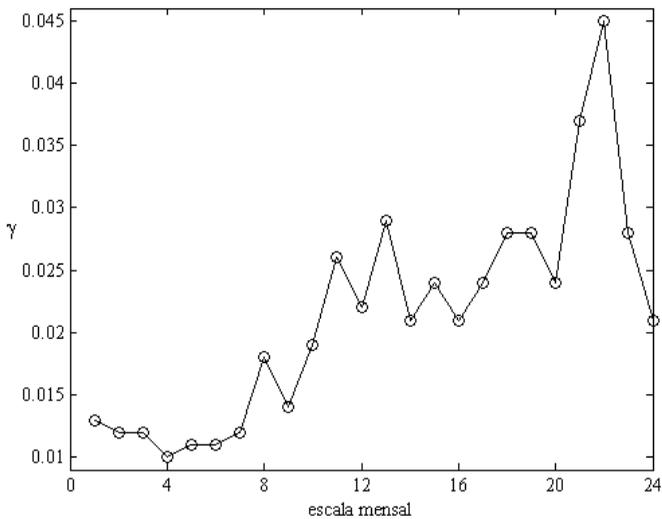


Figura 17 - Dependência temporal do parâmetro γ analisada em uma escala de 24 meses. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Para comparar esses resultados com os trabalhos anteriores, nota-se que os valores de $\alpha = 1.424$ e $\gamma = 0.01853$ para a ação PETR4 correspondem aos valores encontrados por Mantegna e Stanley (1995) e por Gleria et al. (2002), os quais reportaram valores de $\alpha = 1.40$ e $\gamma = 0.00375$ para o índice S&P-500 e valores de $\alpha = 1.66$ e $\gamma = 0.000093$ para o índice Bovespa, respectivamente.

Devido ao fato de que a fdp da Levy-estável de índice 1.424 não é capaz de explicar as raras ocorrências de grandes retornos positivos ou negativos da ação PETR4, os quais ocorrem quando $|Z| \geq 2.72\sigma$, na nona etapa, as propriedades das caudas da fdp empírica para $\Delta t = 1$ minuto foram estudadas. Para isso, o decaimento das caudas foi estimado por dois principais métodos existentes.

A primeira técnica de estimação dos expoentes das caudas é mais simples (embora robusta) e consiste em executar uma regressão por mínimos quadrados ordinários (MQO) para os tamanhos dos valores extremos absolutos ordenados de forma decrescente ($Z_{(1)} \geq \dots \geq Z_{(n)}$), ou seja,

$$\log\left(\text{rank} - \frac{1}{2}\right) = c - \alpha_T \log Z(t) \quad (8)$$

Onde α_T é uma estimativa do expoente de decaimento, cujo erro padrão assintótico é dado por $\sqrt{\frac{2}{t}}\alpha_T$.

A Tabela 3 apresenta os resultados para α_T utilizando-se a equação (8) para os tamanhos das caudas correspondentes a 1%, 5% e 10% da série dos retornos ordenados. A Figura 18 mostra o complemento da fda dos retornos absolutos. A inclinação corresponde à estimativa do coeficiente α_T da regressão representada pela equação (8). A linha reta apresenta evidências de um decaimento de lei de potência fora do regime de Levy de $0 < \alpha \leq 2$. O expoente de decaimento ($\alpha_T \approx 2.4$) obtido para 10% da série inteira dos retornos ordenados, é menor do que 3, e portanto, uma lei de potência cúbica inversa⁸ para os valores extremos está ausente nos dados.

Resumindo, (1) o decaimento da cauda segue uma lei de potência de expoente, aproximadamente, 2.4, sugerindo um padrão para a ação PETR4; (2) o expoente da cauda se encontra fora do regime de Levy, confirmando, então, que a fdp de Levy-estável não é capaz de explicar

⁸ É uma lei de potência com expoente igual a 3.

os eventos extremos; e (3) a ação PETR4 não pode ser acomodada por uma lei cúbica inversa.

Tabela 3 - Estimativas dos expoentes de decaimento dos retornos extremos da ação PETR4: método de Gabaix e Ibragimov (2007).

Tamanho da cauda	α_T (cauda positiva)	α_T (cauda negativa)
1%	2.17 (0.0874)	2.12 (0.1035)
5%	2.39 (0.043)	2.39 (0.0522)
10%	2.34 (0.0298)	2.44 (0.0377)

Nota: erro padrão em parênteses

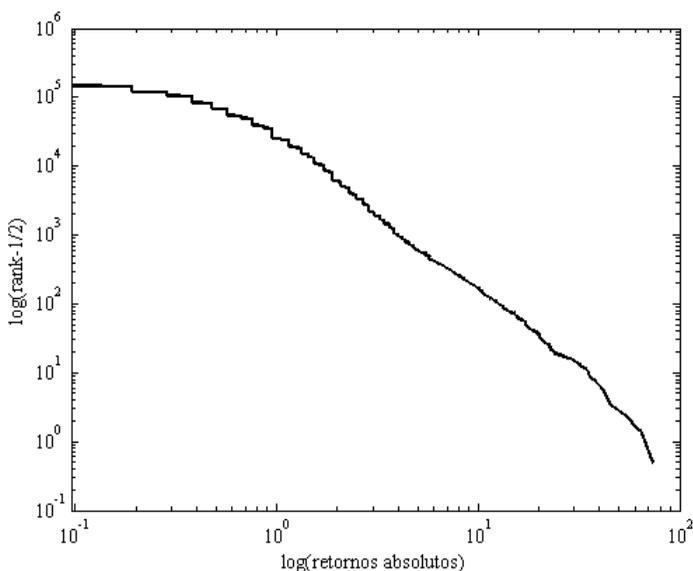


Figura 18 - Logaritmo dos retornos absolutos versus $\log\left(\text{rank} - \frac{1}{2}\right)$ da ação PETR4. A inclinação corresponde à estimativa do coeficiente α_T da regressão 8. Um decaimento de lei de potência é observado com um expoente de, aproximadamente, 2,4, fora do regime de Levy e acima de 3 (lei cúbica inversa).

Para confirmar a análise das caudas consideramos o segundo método, de Gopikrishnan et al. (1998). Neste método os retornos Z para $\Delta t = 1\text{min}$ foram divididos pelo desvio padrão calculado dos dados ($\sigma = 0.0523$), como apresentado na equação (9):

$$g(t) = \frac{Z(t)}{\sigma} \quad (9)$$

Em seguida, os retornos normalizados absolutos foram ordenados de forma decrescente e calculou-se a distribuição de probabilidade cumulativa de se observar uma alteração $g(t)$ maior ou igual a g para os retornos extremos, isto é, $P(g) \equiv P\{g(t) \geq g\}$. Dessa forma o expoente de decaimento foi obtido pela inclinação de $\log P(g)$ em função de $\log(g)$. De fato, os dados são bem ajustados pela seguinte lei de potência:

$$P(g) \approx g^{-\alpha_T} \quad (10)$$

O qual, em logaritmo, é representada como

$$\log P(g) = -\alpha_T \log g \quad (11)$$

A Tabela 4 mostra que os expoentes de decaimento apresentaram, em média, valores um pouco menores do que aqueles calculados pelo método anterior de Gabaix e Ibragimov (2007). Figuras 19, 20 e 21 apresentam as leis de potência para os tamanhos das caudas definidas por 1%, 5% e 10% da série dos retornos ordenados.

Tabela 4 - Estimativas dos expoentes de decaimento dos retornos extremos da ação PETR4: método de Gopikrishnan et al. (1998).

Tamanho da cauda	α_T (cauda positiva)	α_T (cauda negativa)
1%	2.2826 (0.037)	2.2184 (0.039)
5%	2.3612 (0.04)	2.3017 (0.034)
10%	2.3138 (0.038)	2.2628 (0.045)

Nota: erro padrão em parênteses

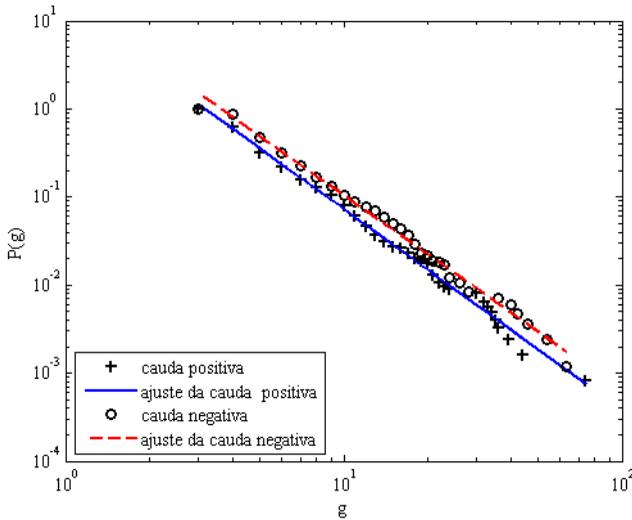


Figura 19 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

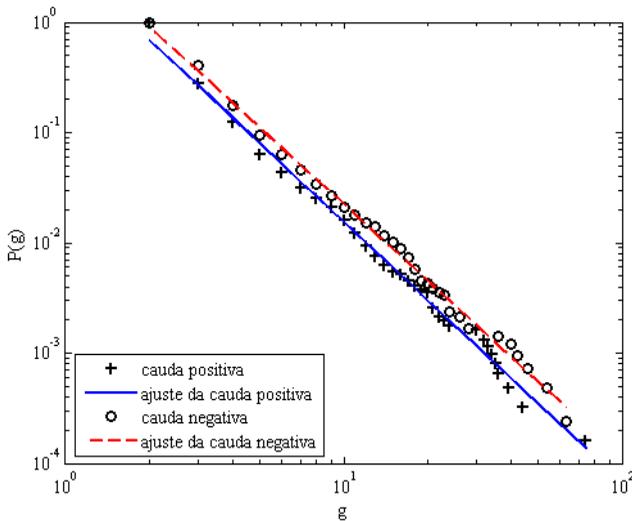


Figura 20 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 5% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

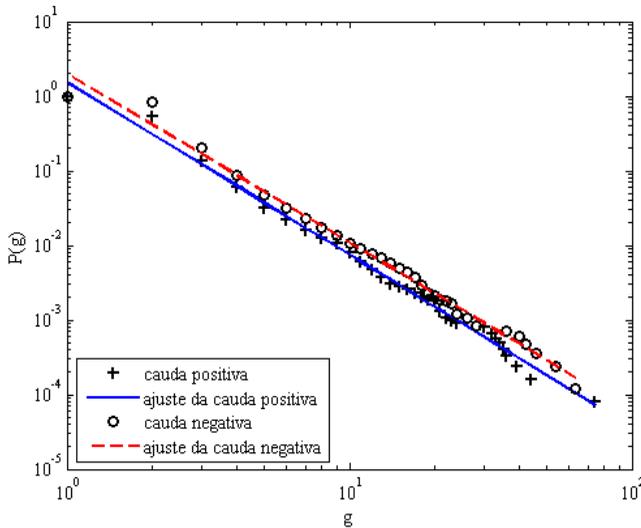


Figura 21 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 10% dos retornos ordenados e normalizados $g(t)$ da ação PETR4: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Na última etapa desta análise, a dependência temporal de σ foi estudada. Como descrito anteriormente, como o parâmetro γ é instável, a fdp da Levy falha em descrever a volatilidade dos dados. Logo, um candidato natural para capturar esta volatilidade dependente do tempo é um modelo GARCH(1,1)

$$\sigma_t^2 = a_0 + a_1 x_{t-1}^2 + b_1 \sigma_{t-1}^2 \quad (12)$$

Onde x_t é uma variável aleatória com média zero e variância σ_t^2 e a_0 , a_1 e b_1 são os parâmetros de controle. No modelo GARCH(1,1), embora σ seja localmente não-estacionário, ele é assintoticamente estacionário. A variância assintótica incondicional é dada por

$$\sigma^2 = \frac{a_0}{1 - a_1 - b_1} \quad (13)$$

E a curtose é expressa como

$$\kappa = 3 + \frac{6a_1^2}{1 - 3a_1^2 - 2a_1b_1 - b_1^2} \quad (14)$$

Utilizando-se as equações (13) e (14), considerando os valores de $\sigma = 0.0523$ e $\kappa = 445.77$, calculados dos dados empíricos para $\Delta t = 1$ min, e o valor de $b_1 = 0.9$, obtemos $a_0 = 2.329 \times 10^{-5}$ e $a_1 = 0.0915$. O valor de $b_1 = 0.9$ foi escolhido por este ser comumente usado na literatura financeira (MANTEGNA e STANLEY, 2000). Logo, a variância condicional é descrita por:

$$\sigma_t^2 = 2.329 \times 10^{-5} + 0.0915x_{t-1}^2 + 0.9\sigma_{t-1}^2 \quad (15)$$

Cabe notar que os parâmetros α_0 e α_1 apresentaram valores muito próximos de Mantegna e Stanley (2000), os quais reportaram valores de 2.30×10^{-5} e 0.09105 , respectivamente.

A Figura 22 compara a fdp empírica dos retornos da ação PETR4 para $\Delta t = 1$ min com o processo GARCH(1,1) caracterizado por $a_0 = 2.329 \times 10^{-5}$, $a_1 = 0.0915$ e $b_1 = 0.9$. Tal processo foi obtido por inúmeras simulações realizadas com os parâmetros calculados acima. Assim, observa-se que o ajuste entre as duas distribuições é muito próximo, tanto nas caudas quanto na região central.

Como o processo GARCH(1,1) tem variância finita, o TLC se aplica para a soma de suas variáveis aleatórias. Isto significa que o processo GARCH(1,1) converge, assintoticamente, para a bacia de atração Gaussiana. Contudo, como este processo é instável, ele não apresenta escalonamento. Aqui, tentou-se estudar as propriedades de escalonamento do processo GARCH(1,1) pela probabilidade de retorno a origem, mas segundo Mantegna e Stanley (2000), ainda não existe uma resposta teórica no momento. Várias simulações numéricas deste processo foram realizadas utilizando-se os mesmos parâmetros de controle usados na Figura 22, contudo, os valores de $P(0)$ não se comportaram de forma escalonada.

Logo, o processo GARCH(1,1) é capaz de descrever a fdp dos retornos da ação PETR4 para $\Delta t = 1$ min, mas falha em descrever, apropriadamente, as propriedades de escalonamento das fdps em diferentes horizontes de tempo, presentes nos dados (MANTEGNA e STANLEY, 2000).

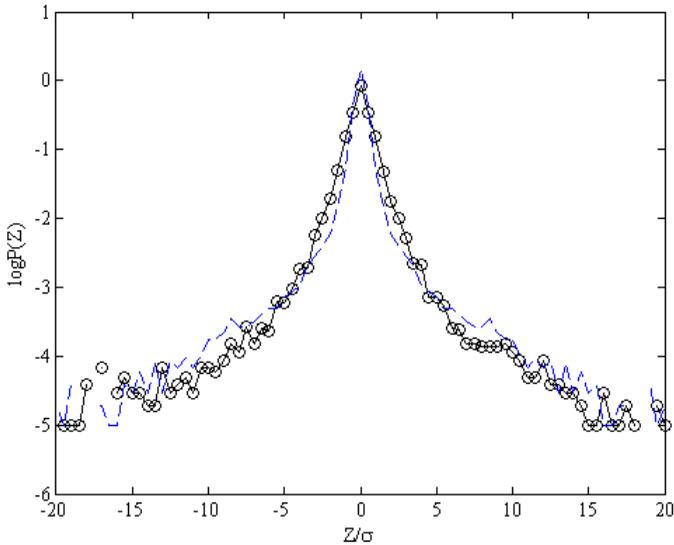


Figura 22 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica dos retornos da ação Petr4 para $\Delta t = 1$ min. (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 2.329 \times 10^{-5}$, $a_1 = 0.0915$ e $b_1 = 0.9$.

3. 2. ANÁLISE SUMÁRIA DOS RESULTADOS DA AÇÃO PETR4

O estudo da ação PETR4, cujos dados foram coletados na frequência de 1 minuto e compreendem o período de janeiro de 2007 a dezembro de 2008, nos permitem concluir que:

1. A fdp da ação PETR4 é quase simétrica, altamente leptocúrtica e não-Gaussiana para pequenas alterações da ação (há um excesso de curtose, $\kappa = 445.77$);
2. As fdps se espalham à medida que o intervalo de tempo na definição dos retornos aumenta, como em um processo aleatório. Além disso, os picos das fdps diminuem com o aumento do intervalo de tempo, seguindo um padrão que é indicativo de um escalonamento não-Gaussiano;
3. De fato, um escalonamento não-Gaussiano foi observado para os máximos das fdps ao aplicar o método de probabilidade de retorno a origem;

4. Um parâmetro de índice $\alpha = 1.424$ pôde ser calculado dos dados. Como $0 < \alpha < 2$, este resultado é sugestivo de uma fdp de Levy-estável gerando os retornos da ação PETR4. O ajuste é bom para a região central da fdp empírica, mas não para as caudas, demandando, portanto, uma atenção maior ao comportamento das últimas;
5. Isto também significa que o escalonamento deveria quebrar para longos intervalos de tempo na definição dos retornos. De fato, ao comparar o escalonamento não-Gaussiano com o escalonamento Gaussiano pela probabilidade de retorno a origem, observa-se que o primeiro ocorre, aproximadamente, entre 1 e 10.000 minutos;
6. Os parâmetros α e γ apresentaram valores próximos de trabalhos anteriores;
7. O índice α permaneceu dentro do regime de Levy-estável ao se considerar subamostras derivadas da base de dados original. Entretanto, o parâmetro γ sofreu grandes flutuações entre as subamostras. Como o parâmetro γ é instável, a fdp da Levy falha em descrever a volatilidade presente nos dados, demandando, então, uma análise mais apurada de σ ;
8. Como a fdp de Levy de índice $\alpha = 1.424$ não é capaz de explicar os grandes retornos positivos e negativos da ação PETR4, o decaimento das caudas da fdp empírica foi estimado por dois métodos principais. Ambos constataram que o decaimento segue uma lei de potência de expoente maior que dois, mas menor que três, sugerindo, então, um padrão de lei de potência para a ação PETR4, mas que não é a lei cúbica inversa;
9. Como a fdp de Levy-estável não é capaz de explicar o fato de que σ é dependente do tempo, o candidato natural para complementar a análise é o processo GARCH(1,1). Como resultado, os parâmetros do processo GARCH(1,1) foram calculados para modelar a evolução temporal da volatilidade do retornos da ação PETR4. Mas, como distribuições não-estáveis não apresentam escalonamento, o modelo GARCH(1,1) não é capaz de detectar o escalonamento não-Gaussiano presente nos dados;
10. Resumindo, a hipótese de que os retornos da ação PETR4 são gerados por uma distribuição de Levy-estável não pode ser rejeitada. Um escalonamento está presente e é claramente não-Gaussiano. Um estudo mais apurado dos eventos extremos da distribuição empírica mostra a presença de um decaimento de lei de potência fora do regime de Levy, mas que não é a lei cúbica inversa. Distribuições de Levy são caracterizadas por possuir variância infinita, e, portanto, pela sua definição, elas não podem modelar dados de variância finita. A variância dependente do tempo da ação PETR4 foi, portanto, modelada por um

processo GARCH(1,1). Não obstante, como este é um processo instável ele não apresenta escalonamento. O escalonamento não-Gaussiano dos dados sugere, então, que um tipo específico de distribuição de Levy-estável (em vez de qualquer distribuição não-estável, como o processo GARCH(1,1)) é a principal fonte de geração dos retornos da ação PETR4.

4. ECONOFÍSICA DO ÍNDICE DJIA

O “flash crash” é o termo usado para descrever a queda do mercado acionário no dia 6 de maio de 2010 envolvendo ações americanas. O índice da média industrial Dow Jones (DJIA) sofreu seu maior declínio intradiário, 998.5 pontos. A maioria das perdas ocorreu entre 14:40h e 15:00h, com um pico em 14:45h (Figura 23). As ações da Accenture, por exemplo, foram rapidamente negociadas por um centavo. A queda foi seguida por uma recuperação quase imediata.

A origem da queda permanece desconhecida, mas alguns observadores apontam para algumas possíveis causas, incluindo operações automatizadas por computador e erro de operadores humanos. Um boato inicial de que um operador tinha digitado uma ordem de venda de 16 bilhões de ações da Procter & Gamble, em vez de 16 milhões, foi posteriormente descartado pelas autoridades reguladoras. Em 1º de outubro de 2010, a Comissão de Valores Mobiliários dos Estados Unidos emitiu um relatório acusando uma ordem de venda executada por um grupo de fundos mútuos (Waddel & Reed), que começou a vender US\$ 4.1 bilhões de contratos futuros “E-mini” através de uma operação robotizada, tendo em conta apenas o volume, e não o tempo ou o preço. Alguns analistas culpam uma ordem de sweep intermercado, a ansiedade sobre o pacote de resgate da Grécia, o resultado das eleições britânicas, e simplesmente, dois dias anteriores de declínio. Mesmo que não seja a causa principal, a negociação robotizada através de plataformas de comércio eletrônico (tais como o Direct Edge e BATS), que executam operações em milissegundos, certamente desempenhou um papel no aumento do crash. A causa ainda é incerta, mas uma confluência de fatores (que não precisa de uma causa particular) é uma explicação mais provável para o crash. O que é certo, entretanto, é que os reguladores não agiram preventivamente; eles reagiram somente na sequência do crash. A Comissão de Valores Mobiliários sugeriu posteriormente um sistema de mercado a nível de “disjuntores”, o qual exigiria que todas as negociações parem ou abranham por alguns minutos se o mercado experimentar uma certa taxa de declínio. Mas se esta estratégia funciona, ainda há de ser testada no futuro.

Se os mercados de ações são vistos como sistemas complexos, não há necessidade de uma causa para explicar o crash (BAK e PACZUSKI, 1995). Eventos extremos não precisam de uma causa quando uma confluência de fatores está envolvida. Aqui, compartilha-se

esse ponto de vista. Como consequência, estamos interessados em avaliar a hipótese de uma forma não-Gaussiana com caudas pesadas e leptocúrtica para a distribuição dos retornos do índice. Para isso, os dados de fechamento do índice DJIA foram coletados na frequência de 1 minuto entre o período de 1º setembro de 2009 a 31 maio de 2010, totalizando 65,534 observações, tais dados provêm da base ProfitChart da empresa Nelógica e foram ajustados por proventos. Assim, com essa amostra, realizou-se uma análise estatística, seguindo o método de Mantegna e Stanley (1995, 2000), com o objetivo de avaliar a hipótese de que uma distribuição de Levy-estável é adequada para modelar os dados. Além disso, foi dada uma atenção particular às caudas da distribuição e à dependência temporal da variância dos dados.

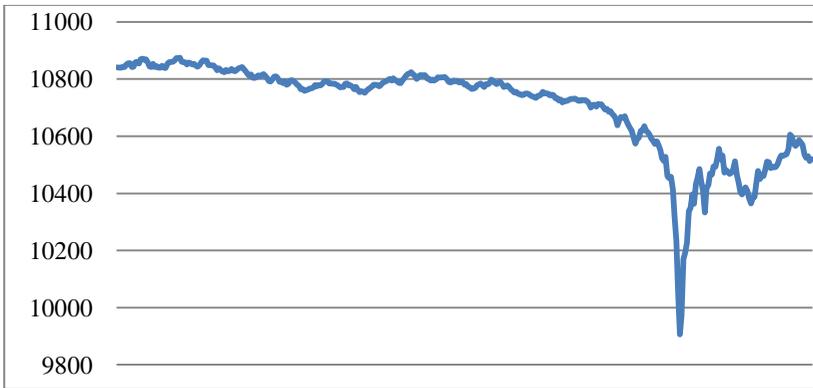


Figura 23 - Gráfico do índice Dow Jones (DJIA) diário durante o “flash crash” de 6 de maio de 2010.

4.1. MÉTODO PROPOSTO

A mesma análise empregada no capítulo 3 foi aplicada no índice DJIA. Desta maneira, alguns conceitos teóricos já apresentados não foram novamente descritos.

Assim, na primeira etapa fez-se o gráfico da fdp dos retornos, Z , para $\Delta t = 1$ min., definidos na equação (4) (Figura 24). Observa-se na Figura 24 que a fdp dos dados (em círculos) é quase simétrica, altamente leptocúrtica e não-Gaussiana para pequenas alterações do índice, cuja distribuição se apresentou de modo similar à fdp da ação PETR4 (Figura 6). Além disso, a curtose apresentou um valor bem acima de 3 ($\kappa =$

208.5), e a Gaussianidade também não foi detectada em nenhum dos testes padrões apresentados na Tabela 5.

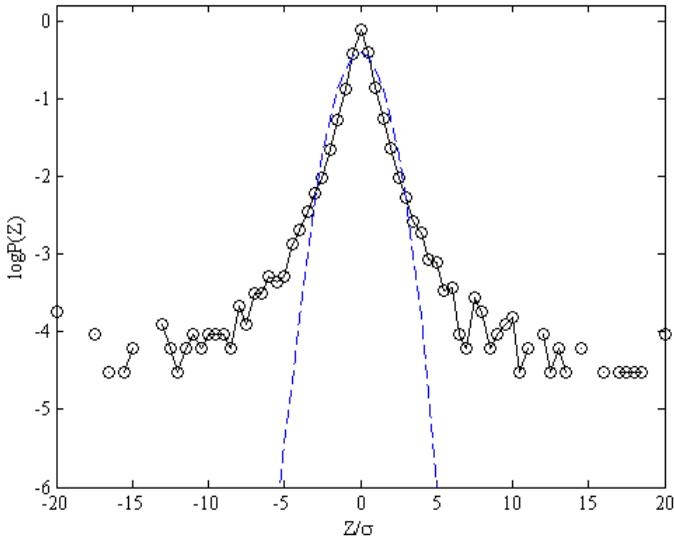


Figura 24 - Fdp dos retornos do índice DJIA para $\Delta t = 1$ min., 1º de setembro de 2009 a 31 de maio de 2010.

Tabela 5 - Testes de normalidade dos retornos do índice DJIA para $\Delta t = 1$ minuto.

Lilliefors	Cramer-von Mises	Anderson-Darling	Curtose
0.1215 (0.0000)	454.8362 (0.0000)	2553.2260 (0.0000)	208.50

Nota: p-valores em parênteses. Todos os p-valores são próximos de zero, portanto, a hipótese nula de normalidade é rejeitada para qualquer nível de significância padrão (0.01, 0.05 ou 0.1).

Na segunda etapa, a fim de se verificar como a forma funcional de $P(Z)$ se altera com o tempo, o comportamento temporal de $P(Z)$ foi avaliado considerando-se vários subconjuntos de $Z_{\Delta t}(t)$ extraídos em função dos valores crescentes de Δt . O número de observações em cada subconjunto reduziu de 65.534 ($\Delta t = 1$ min.) para 65.525 ($\Delta t = 10$ min.) e para 64.535 ($\Delta t = 1000$ min.). Figura 25 mostra as fdps para $\Delta t = 1, 2, \dots, 10$ (Figura 26 mostra as fdps correspondentes em um gráfico semilogarítmico). A partir daí, dois padrões emergem: (1) as fdps se espalham à medida que Δt aumenta, como em um processo aleatório

(Figura 26); (2) os picos das fdps ($Z=0$) diminuem quando Δt aumenta (Figura 25), seguindo, aproximadamente, intervalos visíveis, os quais são sugestivos de um escalonamento não-Gaussiano.

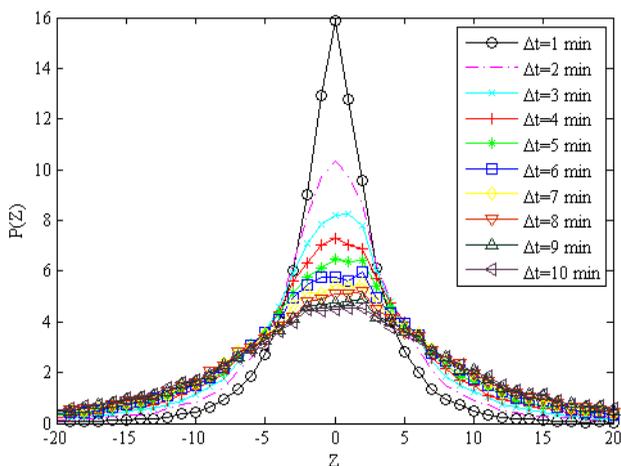


Figura 25 - Fdps dos retornos do índice DJIA calculados em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 10$ minutos.

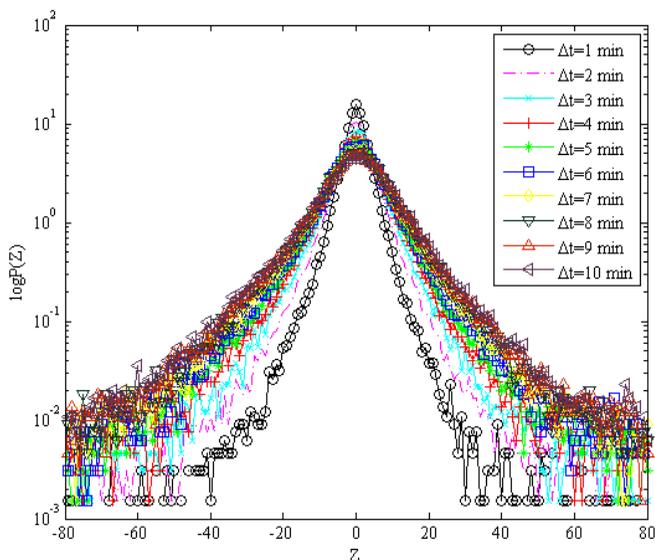


Figura 26 - Gráfico semilogarítmico das fdps dos retornos do índice DJIA calculados em diferentes horizontes de tempo, $\Delta t = 1, 2, \dots, 10$ minutos.

Em seguida, na terceira etapa verificou-se a existência e as características do escalonamento da fdp dos retornos aplicando o método de probabilidade de retorno a origem $P_{\Delta t}(Z = 0)$ em função de Δt . Colocando os resultados em um gráfico log-log, vemos que o comportamento dos dados segue uma lei de potência (Figura 27). A existência de um escalonamento na Figura 27 justifica o gráfico escalonado das fdp apresentado nas Figuras 25 e 26.

Este resultado é compatível com uma fdp de Levy-estável, cujo comportamento também foi verificado para a ação PETR4 (Figura 10).

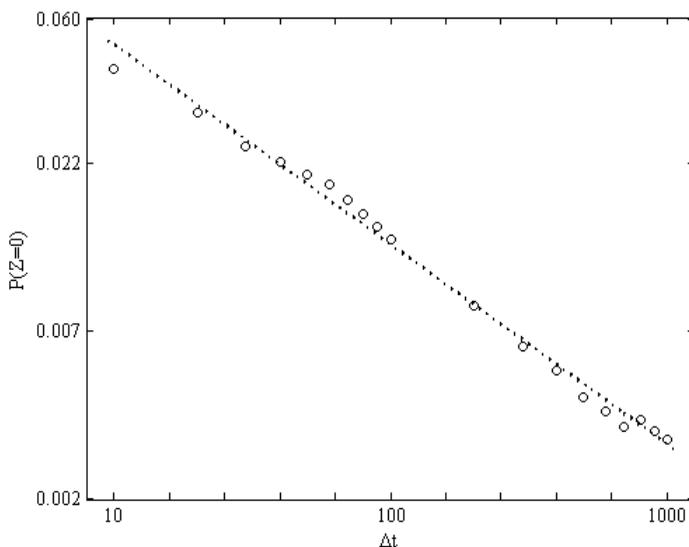


Figura 27 - Probabilidade de retorno à origem medida como uma função de Δt . A dependência de lei de potência é apresentada em um gráfico log-log. No eixo y está a $P(Z = 0)$ correspondente a cada intervalo de tempo (de $\Delta t = 10$ min. até $\Delta t = 1000$ min.) A inclinação -0.5947 ± 0.025 sobre três ordens de magnitude é consistente com um escalonamento não-Gaussiano.

Na quarta etapa os parâmetros α , β , γ e δ foram estimados pelos três métodos de estimação apresentados no capítulo 3, e os diagnósticos de Nolan (1999) foram aplicados para analisar a qualidade do ajuste. Na Tabela 6 verifica-se que os métodos de Máxima Verossimilhança e de Quantis apresentam valores próximos, e os valores estimados pelo método da Função Característica não diferem muito, suportando, então,

a hipótese de estabilidade dos dados. Em seguida, fez-se um gráfico da densidade suavizada dos dados pela função kernel Gaussiana. Observa-se na figura 28 que não há oscilações na densidade do índice, o qual se apresentou plausivelmente estável. Além disso, a distribuição de Levy-estável (linha vermelha contínua), obtida pelos parâmetros estimados por máxima verossimilhança, oferece um ajuste melhor do que a distribuição Gaussiana (linha azul tracejada) na região central da fdp dos dados.

Tabela 6 - Estimativas dos parâmetros α , β , γ e δ dos retornos do índice DJIA ($\Delta t = 1$).

Método	α	β	γ	δ
Máxima verossimilhança	1.4927 (± 0.0122)	-0.0096 (± 0.0231)	2.05441 (± 0.01708)	0.04179 (± 0.02749)
Quantis	1.4158	0.0013	1.98077	-0.00077
Função	1.6017	0.0171	2.14271	0.03195

Nota: intervalos com 95% de confiança em parênteses

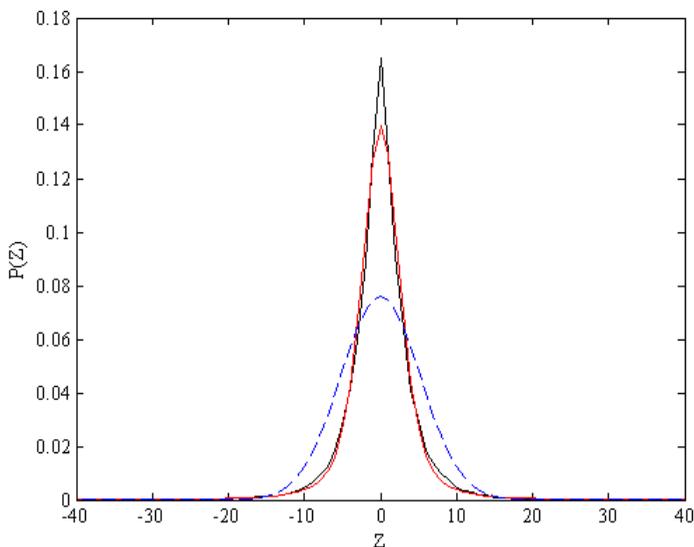


Figura 28 - Gráfico da densidade suavizada dos retornos do índice DJIA para $\Delta t = 1$ min (linha preta). A linha vermelha é o ajuste da Levy-estável estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

Na última etapa dos diagnósticos foi construído o p-p plot de variância estabilizada de Michael (1983) a fim de avaliar o ajuste da distribuição estável em todas as regiões da fdp empírica. O p-p plot mostrou um ótimo ajuste na região central e um ajuste razoável nas regiões extremas dos dados empíricos (Figura 29). Desta forma, o resultado do conjunto dos diagnósticos oferece suporte para a hipótese de estabilidade dos dados, pois os parâmetros da fdp de Levy-estável apresentaram valores próximos entre os três métodos de estimação, a fdp suavizada dos dados se apresentou plausivelmente estável e o p-p plot mostrou um bom ajuste entre a distribuição estável e os dados empíricos.

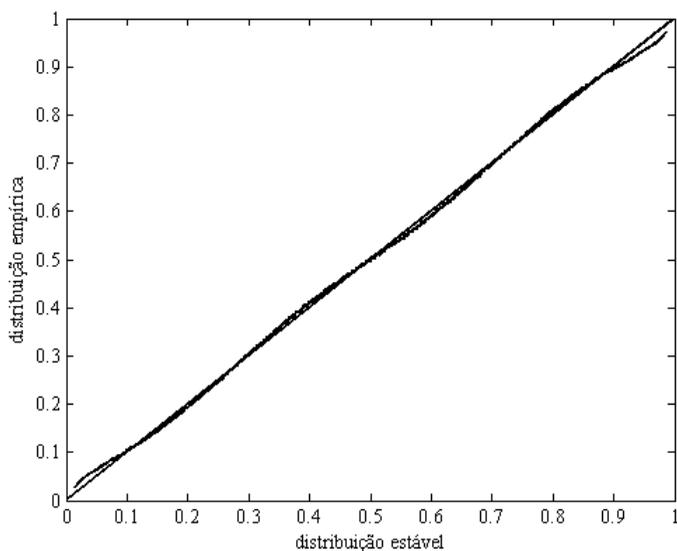


Figura 29 - P-P plot de variância estabilizada: fda dos retornos do índice DJIA contra a fda estável obtida pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Finalizados os diagnósticos, prosseguimos com a análise de Mantegna e Stanley (1995, 2000) considerando os parâmetros estimados por máxima verossimilhança, pois como já foi mencionado, este método tem a vantagem de produzir intervalos de confiança para amostras grandes aos parâmetros. Deste modo, na quinta etapa, a fdp da Levy-estável teórica foi construída em um gráfico semilogarítmico (Figura 30). Um ótimo ajuste foi observado na região central da fdp, mas os dados

extremos se localizaram abaixo da Levy teórica, indicando quebra de escala. O ajuste com a fdp empírica se apresenta melhor para o conjunto de observações

$$|Z| < 2.56\sigma \quad (16)$$

Desta forma, os dados da fdp empírica ficam abaixo da Levy teórica quando $|Z| \geq 2.56\sigma$. E as caudas decaem precisamente porque a variância da fdp empírica é finita. Logo, pode-se concluir, parcialmente, que (1) os dados dos retornos podem ser modelados por uma distribuição de Levy-estável na região central da fdp e (2) as caudas merecem um exame minucioso.

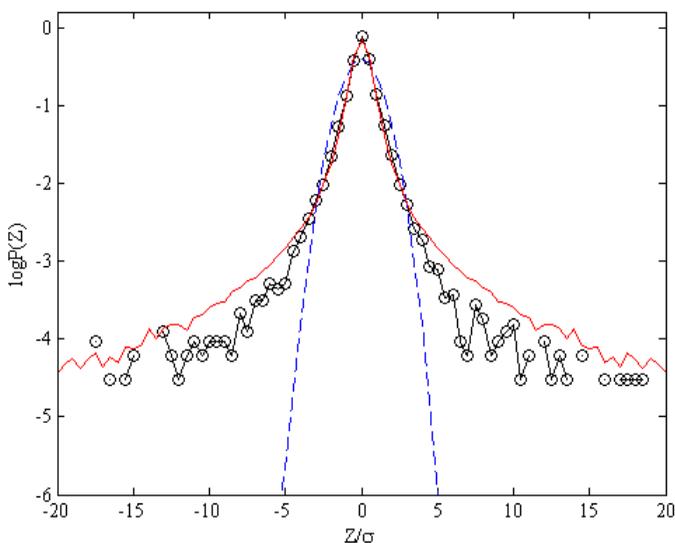


Figura 30 - Fdp dos retornos normalizados do índice DJIA para $\Delta t = 1$ min. A fdp de Levy-estável (linha vermelha contínua) é obtida pelos parâmetros $\alpha = 1.4927$, $\beta = -0.0096$, $\gamma = 2.05441$ e $\delta = 0.04179$ estimados por máxima verossimilhança. A fdp Gaussiana (linha azul tracejada) também é incluída para comparação.

Na sexta etapa, a existência do escalonamento foi estudada para as outras regiões da fdp empírica e não somente para a região onde $Z = 0$. Para isso, as equações (5) e (6) foram consideradas juntamente com $\alpha = 1.493$. Figuras 31 e 32 mostram que todas as fdps empíricas (para

todos os valores de Δt) colapsam na fdp de $\Delta t = 1$ minuto, revelando a auto-similaridade dos dados. Novamente, o ajuste se apresenta melhor na região central das fdps (Figura 32), sugerindo, então, que o escalonamento quebre para longos intervalos de tempo. Logo, da mesma forma que os resultados encontrados para a ação PETR4 (Figura 14), concluímos que (1) há um escalonamento não-Gaussiano para os retornos do índice DJIA; e (2) que este escalonamento é válido para um intervalo de tempo finito porque a variância dos dados é finita.

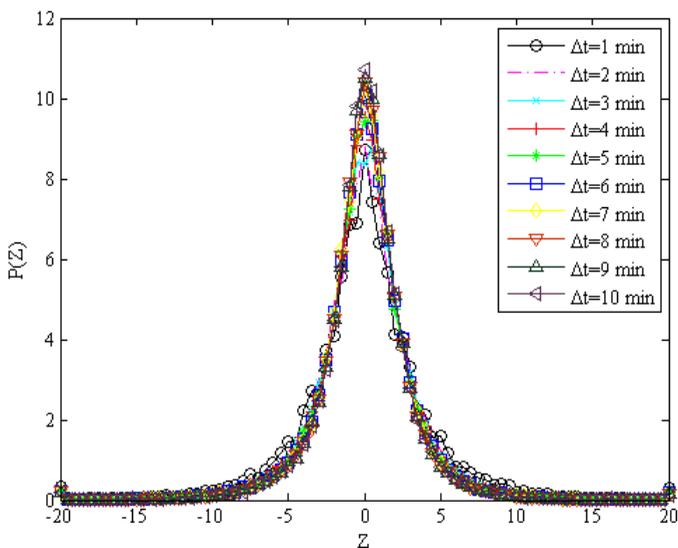


Figura 31 - As mesmas fdps da figura 25 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.4927$.

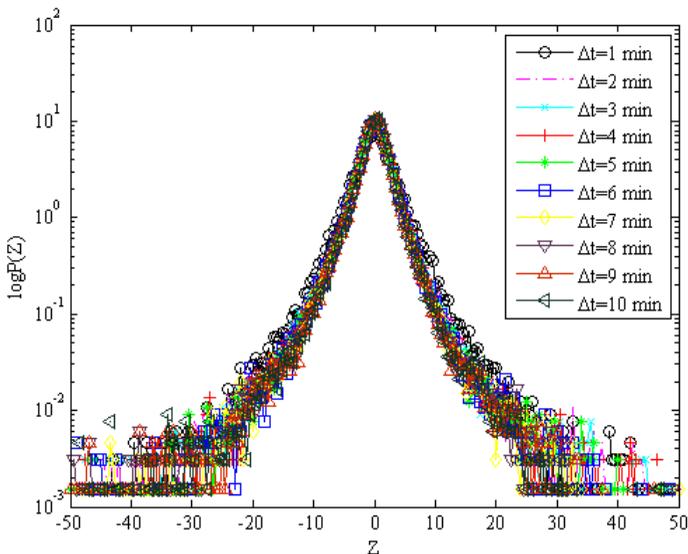


Figura 32 - As mesmas fdps da figura 26 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.4927$.

Na sétima etapa, a fim de encontrar o intervalo onde os escalonamentos das Figuras 31 e 32 realmente ocorrem, o escalonamento não-Gaussiano foi comparado com o escalonamento Gaussiano. Tomamos a probabilidade de retorno empírica $P_{\Delta t}(Z=0)$ e a comparamos com a probabilidade de retorno Gaussiana $P_G(Z=0)$. Esta última foi calculada utilizando-se os valores empíricos da média e variância dos retornos para cada Δt .

A Figura 33 mostra o gráfico log-log do escalonamento não-Gaussiano em $P_{\Delta t}(Z=0)$ (linha pontilhada) e do escalonamento Gaussiano em $P_G(Z=0)$ (linha tracejada), vemos que, para dado Δt , a diferença entre $P_{\Delta t}(Z=0)$ e $P_G(Z=0)$ diminui quando $30 < \Delta t < 1000$ minutos. A reta dos dados encontra a reta gaussiana, aproximadamente, em 10^4 minutos.

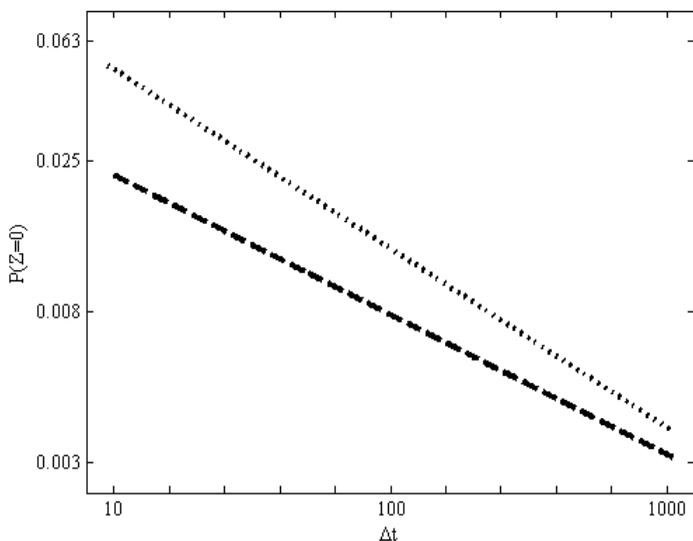


Figura 33 - Probabilidade de retorno à origem dos retornos do índice DJIA (linha pontilhada) junto com a probabilidade de retorno à origem Gaussiana (linha tracejada). A distância entre as duas retas é uma medida da natureza não-Gaussiana da fdp.

Portanto, o escalonamento não-Gaussiano ocorre, aproximadamente, entre 1 e 10.000 minutos. Concluindo, então, que este escalonamento é válido para um intervalo de tempo que é grande, mas finito. Cabe lembrar que tal resultado da Figura 33 também foi encontrado na análise da ação PETR4 (Figura 15).

Na oitava etapa, a robustez dos parâmetros estimados, α e γ , foi checada. Desta forma, a base de dados original foi dividida em 9 subconjuntos (meses) e os parâmetros α e γ de cada subconjunto foram estimados pelo método de máxima verossimilhança.

Figura 34 mostra que o α permanece dentro do regime de Levy não-Gaussiano ($1.4 < \alpha < 1.8$), confirmando, então, que a hipótese de uma distribuição de Levy-estável não pode ser rejeitada dos dados. Figura 35 mostra que o γ apresenta grandes flutuações ($1.43 < \gamma < 4.31$). Logo, como o parâmetro γ é instável, a fdp da Levy falha em descrever a volatilidade presente nos dados.

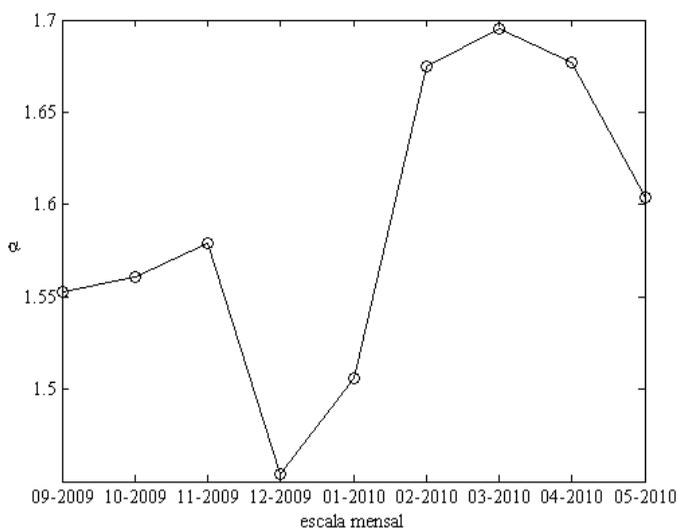


Figura 34 - Dependência temporal do índice α analisada em uma escala de 9 meses. O parâmetro α ($1.4 < \alpha < 1.8$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

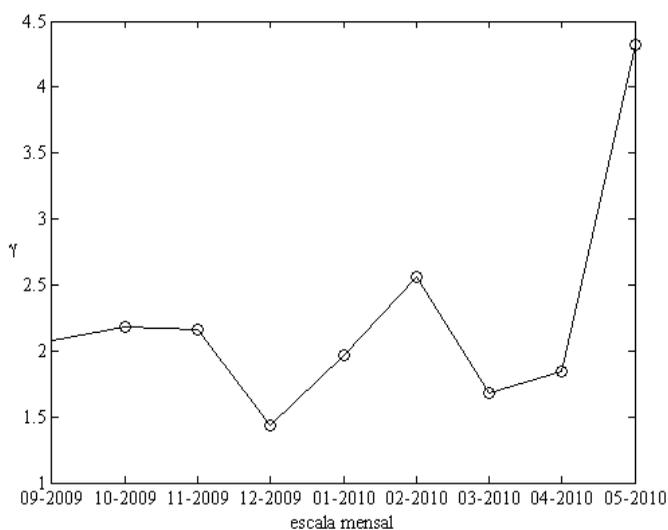


Figura 35 - Dependência temporal do parâmetro γ analisada em uma escala de 9 meses. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Para comparar esses resultados com os trabalhos anteriores, nota-se que o valor do parâmetro $\alpha = 1.493$ para o índice DJIA se aproxima do valor encontrado no trabalho pioneiro do índice S&P-500 (MANTEGNA e STANLEY, 1995), o qual reportou um valor de $\alpha = 1.40$. Este resultado também é similar ao do trabalho do índice Bovespa (GLERIA et al., 2002), com $\alpha = 1.66$. Entretanto, o valor de $\gamma = 2.0544$ para o índice DJIA é maior do que os dos índices S&P-500 ($\gamma = 0.00375$) e Bovespa ($\gamma = 0.00093$). Especula-se que este valor alto para γ pode estar relacionado à nossa amostra selecionada.

Prosseguindo com a análise, vimos anteriormente na Figura 30 que a fdp de Levy-estável de índice 1.493 não é capaz de explicar as raras ocorrências de grandes retornos positivos ou negativos do índice DJIA, os quais ocorrem quando $|Z| \geq 2.56\sigma$. Logo, na nona etapa, o expoente de decaimento das caudas da distribuição dos retornos do índice DJIA para $\Delta t = 1\text{min}$ foi estimado pelos dois métodos descritos no capítulo 3.

Desta forma, seguindo os mesmos procedimentos do método de Gabaix e Ibragimov (2007) descritos no capítulo 3, os tamanhos dos retornos extremos absolutos ordenados de forma decrescente foram regredidos de acordo com a equação (8). A Tabela 7 apresenta os resultados de α_T utilizando-se a equação (8) para os tamanhos das caudas correspondentes a 1%, 5% e 10% da série dos retornos ordenados. A Figura 36 mostra o complemento da fda dos retornos absolutos. A inclinação corresponde à estimativa do coeficiente α_T da regressão representada pela equação (8). A linha reta apresenta evidências de um decaimento de lei de potência fora do regime de Levy de $0 < \alpha \leq 2$. O expoente de decaimento ($\alpha_T \approx 2.5$) obtido para 10% da série inteira dos retornos ordenados é menor do que 3, e portanto, uma lei de potência cúbica inversa para os valores extremos está ausente nos dados.

Resumindo, (1) o decaimento da cauda segue uma lei de potência de expoente, aproximadamente, 2.5, sugerindo um padrão para o “flash crash”; (2) o expoente da cauda se encontra fora do regime de Levy, confirmando, então, o fato de que a fdp de Levy-estável não é capaz de explicar os eventos extremos; e (3) o flash crash não pode ser acomodado por uma lei cúbica inversa.

Tabela 7 - Estimativas dos expoentes de decaimento dos retornos extremos do índice DJIA: método de Gabaix e Ibragimov (2007).

Tamanho da cauda	α_T (cauda positiva)	α_T (cauda negativa)
1%	2.30 (0.1778)	2.20 (0.1738)
5%	2.53 (0.0875)	2.41 (0.0851)
10%	2.54 (0.0620)	2.38 (0.0595)

Nota: erro padrão em parênteses

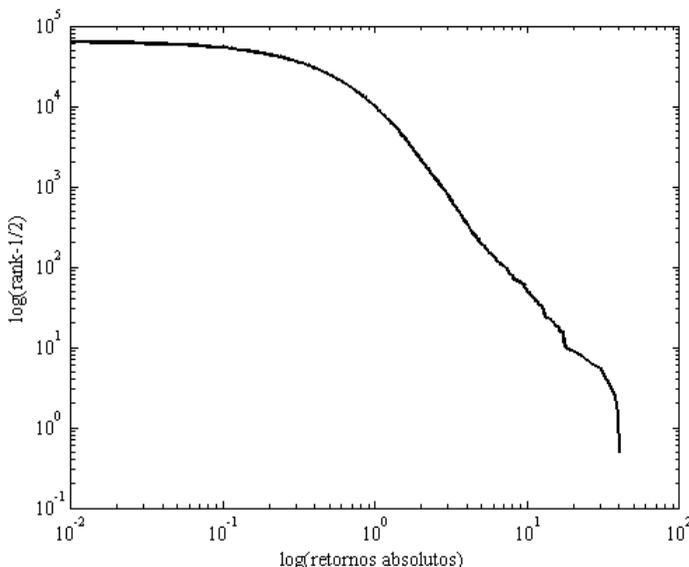


Figura 36 - Logaritmo dos retornos absolutos versus $\log\left(\text{rank} - \frac{1}{2}\right)$ do índice DJIA. A inclinação corresponde à estimativa do coeficiente α_T da regressão 8. Um decaimento de lei de potência é observado com um expoente de aproximadamente, 2.5, fora do regime de Levy e acima de 3 (lei cúbica inversa).

Para confirmar a análise das caudas, o método de Gopikrishnan et al. (1998) foi aplicado, no qual o expoente de decaimento, α_T , é obtido, calculando-se a inclinação de $\log P(g)$ em função de $\log(g)$. Onde $P(g)$ é a distribuição de probabilidade cumulativa de se observar uma alteração $g(t)$ maior ou igual a g para os retornos extremos, isto é, $P(g) \equiv P\{g(t) \geq g\}$, e g é dado pela equação (9).

A Tabela 8 mostra que os expoentes de decaimento apresentam valores um pouco menores do que aqueles calculados pelo método de

Gabaix e Ibragimov (2007). Figuras 37, 38, e 39 apresentam as leis de potência para os tamanhos das caudas definidas por 1%, 5% e 10% da série dos retornos ordenados.

Tabela 8 - Estimativas dos expoentes de decaimento dos retornos extremos do índice DJIA: método de Gopikrishnan et al. (1998).

Tamanho da cauda	α_T (cauda positiva)	α_T (cauda negativa)
1%	2.23 (0.0622)	2.12 (0.0553)
5%	2.39 (0.0652)	2.26 (0.0534)
10%	2.33 (0.0681)	2.21 (0.0593)

Nota: erro padrão em parênteses

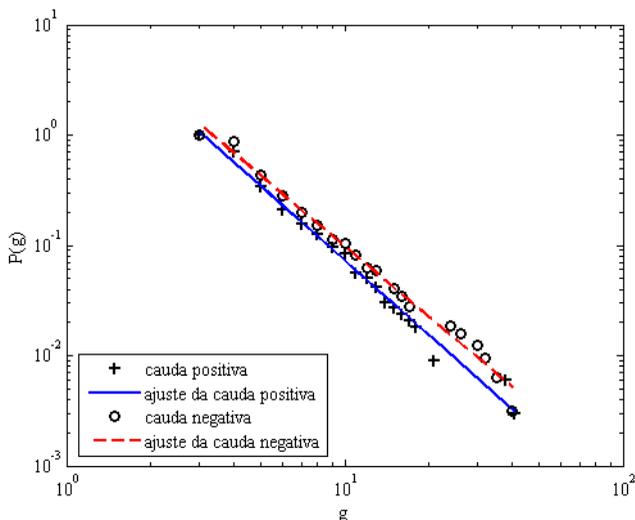


Figura 37 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

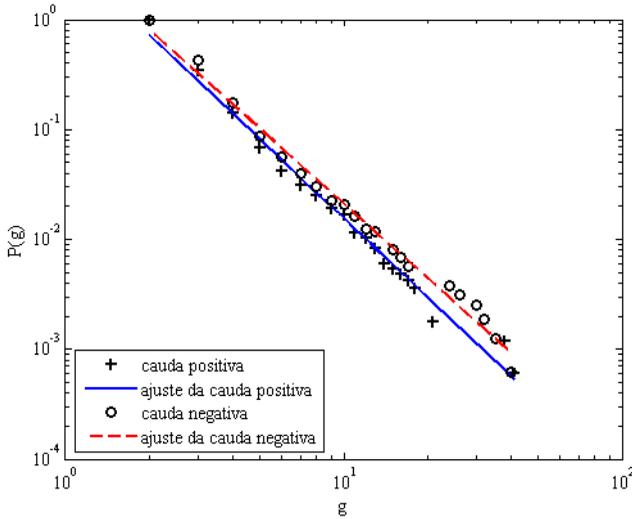


Figura 38 – Gráfico log-log da fda $P(g)$ para o tamanho da cauda definida por 5% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

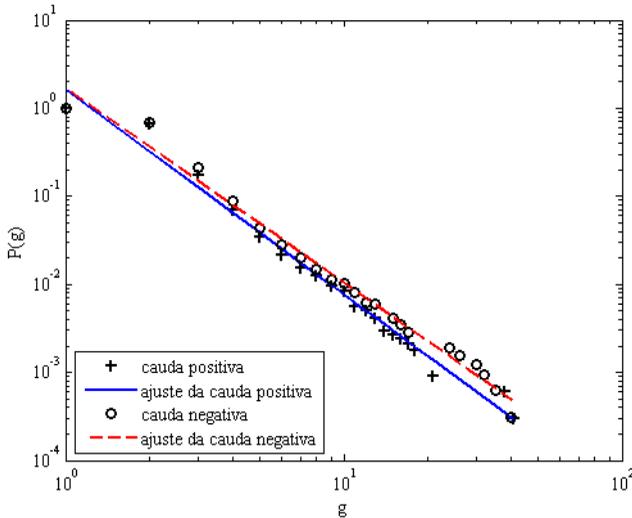


Figura 39 – Gráfico log-log da fda $P(g)$ para o tamanho da cauda definida por 10% dos retornos ordenados e normalizados $g(t)$ do índice DJIA: uma lei de potência emerge dos dados com um expoente de decaimento maior que 2, mas menor que 3.

Por fim, na última etapa, a dependência temporal de σ foi analisada pelo modelo GARCH(1,1) (equação 12). Pois, como foi descrito anteriormente, o parâmetro γ é instável (Figura 35), e, portanto, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Utilizando-se as equações (13) e (14), considerando os valores de $\sigma = 5.243$ e $\kappa = 208.5$, calculados dos dados empíricos para $\Delta t = 1$ min., e o valor de $b_I = 0.9$, obtemos $a_0 = 0.2348$ e $a_1 = 0.0915$. Novamente, o valor de $b_I = 0.9$ foi escolhido por este ser comumente usado na literatura financeira (MANTEGNA e STANLEY, 2000). Logo, a variância condicional é descrita por:

$$\sigma_t^2 = 0.2348 + 0.0915x_{t-1}^2 + 0.9\sigma_{t-1}^2 \quad (17)$$

Figura 40 compara a fdp empírica do índice DJIA para $\Delta t = 1$ min. com o processo GARCH(1,1) caracterizado por $a_0 = 0.2348$, $a_1 = 0.0915$ e $b_I = 0.9$. Tal processo foi obtido por inúmeras simulações realizadas com os parâmetros calculados acima. Assim, observa-se que o processo GARCH(1,1) fornece um ajuste bom na região central e razoável nas caudas.

Em suma, o processo GARCH(1,1) é capaz de descrever a fdp dos retornos do índice DJIA para $\Delta t = 1$ min., mas falha em descrever, apropriadamente, as propriedades de escalonamento das fdps, em diferentes horizontes de tempo, presentes nos dados. Pois, como tal processo é instável, ele não apresenta escalonamento.

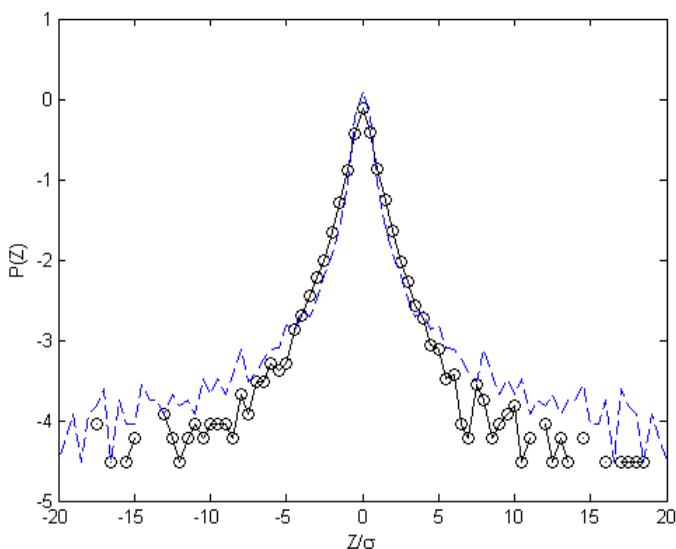


Figura 40 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica dos retornos do índice DJIA para $\Delta t = 1$ min. (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 0.2348$, $a_1 = 0.0915$ e $b_1 = 0.9$.

4.2. ANÁLISE SUMÁRIA DOS RESULTADOS DO ÍNDICE DJIA

A análise do flash crash apresentou resultados similares aos da ação PETR4, os quais seguem:

1. A fdp do índice DJIA produziu um formato similar à fdp da ação PETR4, sendo quase simétrica, altamente leptocúrtica e não-Gaussiana para pequenas alterações do índice ($\kappa = 208.5$);
2. As fdps se espalham à medida que o intervalo de tempo na definição dos retornos aumenta, como em um processo aleatório. Além disso, os picos das fdps diminuem com o aumento do intervalo de tempo, seguindo um padrão que é indicativo de um escalonamento não-Gaussiano;
3. De fato, um escalonamento não-Gaussiano foi observado para os máximos das fdps ao aplicar o método de probabilidade de retorno à origem;
4. A fdp de Levy-estável, com $\alpha = 1.493$, proporcionou um ótimo ajuste na região central da fdp empírica, mas não tão bom para as caudas,

demandando, portanto, uma atenção maior para o comportamento das últimas;

5. Isto também significa que o escalonamento deveria quebrar para longos intervalos de tempo na definição dos retornos. De fato, ao comparar o escalonamento não-Gaussiano com o escalonamento Gaussiano pela probabilidade de retorno a origem, observa-se que o primeiro ocorre, aproximadamente, entre 1 e 10.000 minutos;
6. Um parâmetro $\gamma = 2.0544$ foi estabelecido, o qual é maior do que aqueles normalmente encontrados para outros índices de mercado de ações. Todavia, como o γ está relacionado à volatilidade, a seleção da amostra compreendendo o episódio do crash pode explicar o alto valor deste parâmetro;
7. O parâmetro α permaneceu dentro do regime de Levy entre as subamostras. Por outro lado, o parâmetro γ sofreu fortes flutuações. Logo, como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade presente nos dados, demandando, então, uma análise minuciosa de σ ;
8. Tal como na ação PETR4, o decaimento das caudas da distribuição do índice DJIA segue uma lei de potência de expoente maior que dois, mas menor que três, sugerindo, então, um padrão de lei de potência para o flash crash, mas que não é a lei cúbica inversa;
9. Para modelar a volatilidade dependente do tempo, utilizou-se o modelo GARCH(1,1), tal modelo proporcionou um bom ajuste para a fdp empírica para $\Delta t = 1$ min. No entanto, como o processo GARCH(1,1) é instável, ele não é capaz de detectar o escalonamento não-Gaussiano presente nos dados;
10. Resumindo, a hipótese de que os retornos do índice DJIA são gerados por uma distribuição de Levy-estável não pode ser rejeitada. Um escalonamento está presente e é claramente não-Gaussiano. Um estudo mais aprofundado dos eventos extremos da distribuição empírica mostra a presença de um decaimento de lei de potência fora do regime de Levy, mas que não é a lei cúbica inversa. A volatilidade presente nos retornos do índice pôde ser modelada por um processo GARCH(1,1). Não obstante, como este é um processo instável, ele não apresenta escalonamento. O escalonamento não-Gaussiano dos dados sugere, então, que um tipo específico de distribuição de Levy-estável (em vez de qualquer distribuição não-estável, como o processo GARCH(1,1)) é a principal fonte de geração dos retornos do índice DJIA.

5. ECONOFÍSICA DOS NÚMEROS PRIMOS

Um time de físicos pôde ter se deparado com uma descoberta surpreendente sobre uma das questões mais profundas e mais estudadas na matemática pura: se ou não os números primos aparecem aleatoriamente na sequência dos números inteiros.

Kumar et al. (2003) avaliam que eles podem ter encontrado um tipo de ordem entre a distribuição dos números primos, os números que não podem ser divididos por qualquer outro número menor, com exceção de 1.

Os primeiros poucos primos são 2, 3, 5, 7, 11 e 13; o maior conhecido atualmente tem mais de quatro milhões de dígitos. Ninguém ainda provou que suas ocorrências seguem algum padrão, ou se definitivamente não há padrão.

Kumar et al. (2003) estudaram os incrementos dos intervalos entre os primos consecutivos. Por exemplo, os intervalos entre os primeiros poucos primos são 1, 2, 2, 4 e 2. Os incrementos são as diferenças entre esses intervalos consecutivos: +1, 0, +2 e -2.

Se os primos realmente se distribuem aleatoriamente, então os incrementos também deveriam se distribuir aleatoriamente. Mas o time de Kumar concluiu que não. Os incrementos têm uma certa previsibilidade. Valores positivos são geralmente seguidos por valores negativos correspondentes, por exemplo, os primeiros +2, -2 incrementos. Além disso, eles encontraram uma oscilação de período 3. Ou seja, os incrementos que são múltiplos de 6 são menos frequentes do que outros incrementos. Isto se encaixa na análise estatística anterior de Wolf (1999), que encontrou oscilações de período 6 nas diferenças entre dois primos consecutivos (WELLS, 2005).

No entanto, os resultados de Kumar et al. (2003) não são suportados por qualquer tipo de prova matemática rigorosa. E, infelizmente, eles não podem lançar alguma luz sobre um dos maiores problemas da matemática: a hipótese de Riemann. Esta conjectura da teoria dos números está intimamente relacionada com a distribuição dos números primos.

Mas, esses resultados empíricos podem ser de grande importância para avançar na compreensão da natureza dos números primos, bem como nos processos físicos e biológicos onde os números primos desempenham um papel importante. (KUMAR et al. 2003)

Desta maneira, com base no trabalho de Kumar et al. (2003), o qual motiva os estudos dos números primos utilizando-se de métodos da

física estatística, este capítulo se concentra nas propriedades estatísticas desses números, com o objetivo de verificar a hipótese de uma forma não-gaussiana com cauda pesada e leptocúrtica para a distribuição das diferenças entre os primos consecutivos. A amostra compreende os primeiros 1,000,000 de primos, cujos dados foram coletados em <http://primes.utm.edu/lists/small/millions/>. Assim, com essa amostra, realizou-se a análise proposta de Mantegna e Stanley (1995, 2000) com o objetivo de avaliar a hipótese de que uma distribuição de Levy-estável é adequada para modelar os dados. Desta forma, este capítulo também contribui com a literatura ao tentar verificar a consistência dos números primos com alguma distribuição teórica.

Por fim, a série dos números primos servirá como uma variável de controle não-financeira, pois conforme já foi mencionado, podemos verificar se a análise estatística proposta pode ser aplicada em qualquer série.

5.1. MÉTODO PROPOSTO

Desta vez, a análise de Mantegna e Stanley (1995, 2000) é realizada através de oito etapas. Na primeira etapa, em vez de calcularmos os retornos, Z , calculamos as diferenças entre os números primos, mas cuja fórmula de cálculo é análoga a dos retornos da equação 4.

$$D_{\Delta k}(k) = N_p(k + \Delta k) - N_p(k) \quad (18)$$

Onde k denota o k -ésimo elemento da sequência, e de forma análoga ao intervalo de tempo (Δt), Δk denota o intervalo da distância entre os números primos.

Dessa forma, foi construída a fdp $P(D)$ para as diferenças $D_{\Delta k}(k)$ utilizando-se $\Delta k = 1$, ou seja, a fdp das diferenças entre dois números primos consecutivos (Figura 41). Como propriedade dos números primos, a frequência das diferenças entre dois números primos consecutivos decresce exponencialmente em um gráfico semilogarítimo. Assim, observa-se que a fdp dos números primos é assimétrica à direita e não-Gaussiana para $\Delta k = 1$. A curtose está um pouco acima de 3 ($\kappa = 7.9106$) e a Gaussianidade também não é detectada em nenhum dos testes padrões apresentados na Tabela 9.

Tabela 9 - Testes de normalidade das diferenças dos números primos para $\Delta k = 1$.

Lilliefors	Cramer-von Mises	Anderson-Darling	Curtose
0.165205 (0.000)	6909.409 (0.000)	40700.31 (0.000)	7.9106

Nota: p-valores em parênteses. Todos os p-valores são, aproximadamente, zeros, portanto, a hipótese nula de normalidade é rejeitada para qualquer nível de significância padrão (0.01, 0.05 ou 0.1).

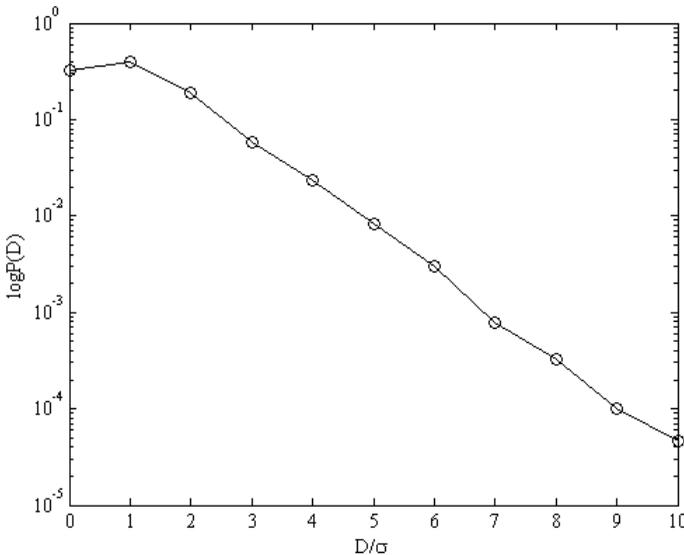


Figura 41 - Fdp das diferenças entre dois números primos consecutivos ($\Delta k = 1$).

Na segunda etapa, a fim de verificar como a forma funcional de $P(D)$ se altera com Δk , o comportamento de $P(D)$ foi avaliado considerando-se vários subconjuntos de $D_{\Delta k}(k)$ extraídos em função dos valores crescentes de Δk . O número de dados em cada subconjunto reduziu de 999,999 ($\Delta k = 1$) para 999,000 ($\Delta k = 1000$). Figura 42 mostra as fdps para $\Delta k = 1, 2, \dots, 10$ (a Figura 43 mostra as fdps correspondentes em um gráfico semilogarítimo). A partir daí, duas considerações são feitas: (1) as fdps se espalham à medida que Δk aumenta (Figura 43), contudo a forma funcional de $P(D)$ se altera com o aumento de Δk , indicando evidências de não-estabilidade da fdp dos números primos, passando de assimétrica à direita para simétrica; (2) à medida que Δk

aumenta, a distância entre as fdps e a fdp de $\Delta k = 1$ tende a crescer (Figura 42). Esses resultados parecem indicar que os números primos não são gerados por uma distribuição estável, visto a alteração da forma funcional das fdps para diferentes Δks , e que não há algum escalonamento presente em sua sequência. Contudo, antes de qualquer conclusão, prosseguimos com a análise a fim de verificarmos as propriedades estatísticas de tais números.

Neste capítulo, pulamos para etapa de estimação dos parâmetros, α , β , γ e δ , pois não foi possível calcular a probabilidade de retorno à origem, pelo simples fato de não haver $D = 0$.

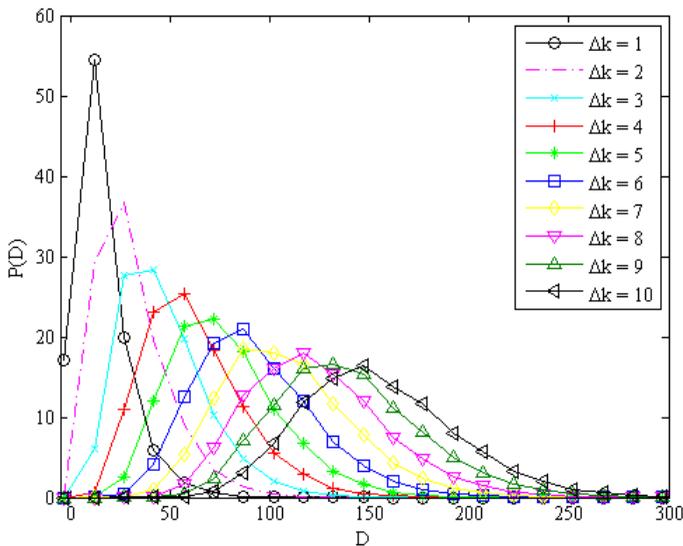


Figura 42 - Fdps das diferenças dos números primos calculadas em diferentes intervalos, $\Delta k = 1, 2, \dots, 10$.

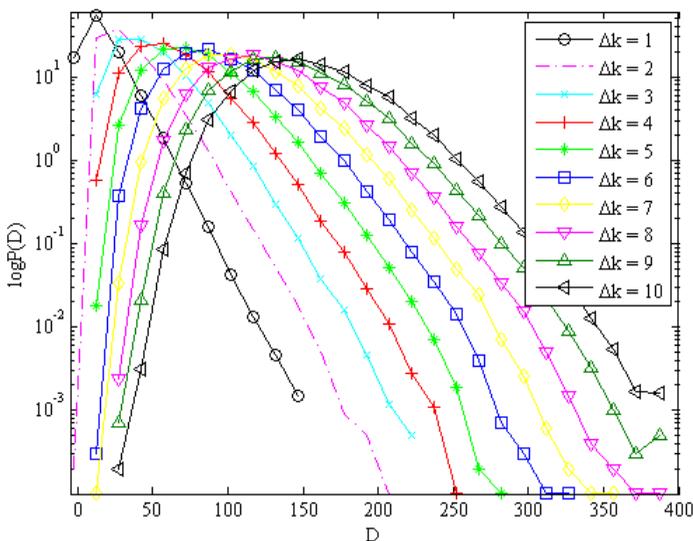


Figura 43 - Gráfico semilogarítmico das diferenças dos números primos calculadas em diferentes intervalos, $\Delta k = 1, 2, \dots, 10$.

Na Tabela 10, verifica-se que a estimação de α por máxima verossimilhança apresenta um valor bem diferente das estimativas por quantis e pela função característica. Isto é uma indicação de que os dados não são estáveis, pois se sua distribuição fosse estável, então todos os estimadores consistentes deveriam produzir valores próximos quando a amostra for grande. Tal resultado foi confirmado pelo gráfico da densidade suavizada dos dados pela função kernel Gaussiana. Observa-se na Figura 44 que há muitas oscilações na densidade dos dados e de que a fdp de Levy-estável, obtida pelos parâmetros estimados por máxima verossimilhança, ofereceu um péssimo ajuste na região modal da fdp dos dados. Por fim, o p-p plot mostrou um péssimo ajuste da distribuição estável, tanto na região modal quanto nas regiões extremas (Figura 45). Desta forma, os resultados do conjunto dos diagnósticos não oferecem suporte para a hipótese de estabilidade dos dados, pois os parâmetros apresentaram valores, significativamente, diferentes entre os três métodos de estimação, a fdp suavizada dos dados não se apresentou plausivelmente estável e o p-p plot mostrou um péssimo ajuste entre a distribuição estável e a distribuição empírica.

Entretanto, a análise deve ser completada, pois de acordo com Nolan (1999), os diagnósticos aqui discutidos são uma tentativa de detectar a não-estabilidade. Como já foi citado: assim como qualquer outra família de distribuições, não é possível provar que um determinado conjunto de dados é ou não estável.

Tabela 10 - Estimativas dos parâmetros α , β , γ e δ das diferenças entre os números primos ($\Delta k = 1$).

Método	α	β	γ	δ
Máxima verossimilhança	1.1 (± 0.0022)	1 (± 0.000)	4.82562 (± 0.01019)	8.87558 (± 0.01624)
Quantis	1.6426	1	7	12
Função	1.6782	1	6.76339	9.76081

Nota: intervalos com 95% de confiança em parênteses.

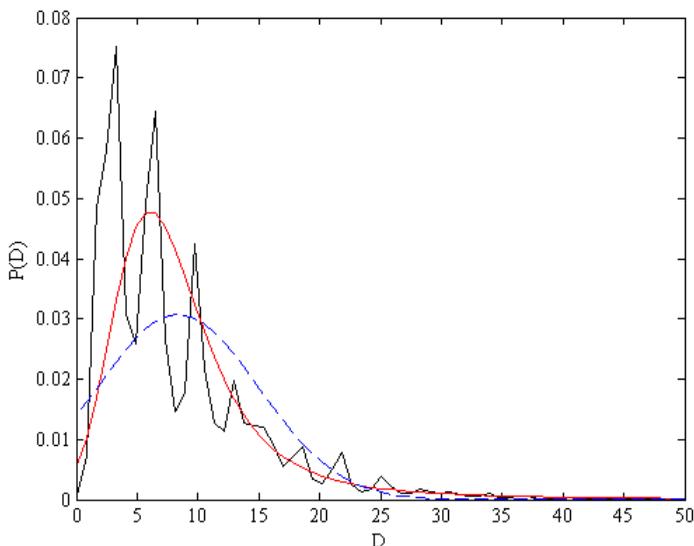


Figura 44 - Gráfico da densidade suavizada das diferenças dos números primos para $\Delta k = 1$. (linha preta). A linha vermelha é o ajuste da Levy estimado pelo método de máxima verossimilhança. A linha azul tracejada é o ajuste da normal obtido pela média e variância da série empírica.

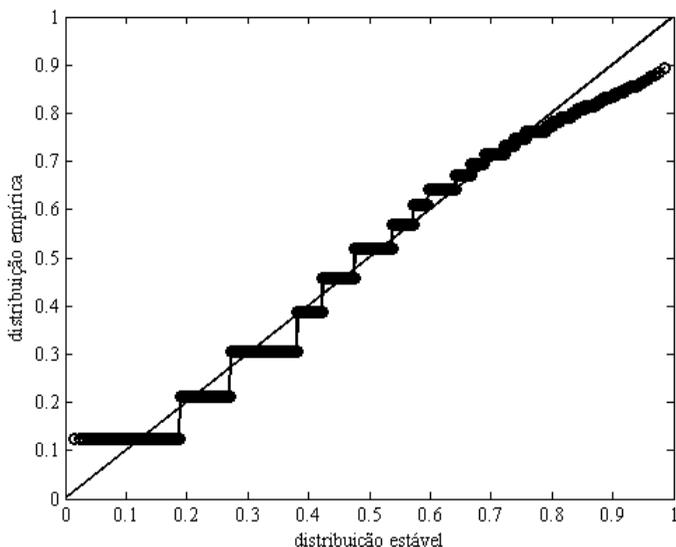


Figura 45 - P-P plot de variância estabilizada: fda das diferenças dos números primos (para $\Delta k = 1$) contra a fda estável obtida pelo método de máxima verossimilhança. A linha diagonal é a linha de referência.

Dessa maneira, na quarta etapa, a fdp de Levy-estável teórica foi construída em um gráfico semilogaritmo (Figura 46). Um ajuste razoável foi observado na região modal da fdp, e, além disso, os dados extremos se localizaram abaixo da Levy teórica, sugerindo que a variância dos dados é finita. Portanto, apesar de não se poder concluir, parcialmente, que as diferenças podem ser modeladas por uma distribuição de Levy-estável na região modal de sua fdp, a cauda da distribuição empírica, ainda sim, merece um exame mais apurado, o qual será discutido na sétima etapa.

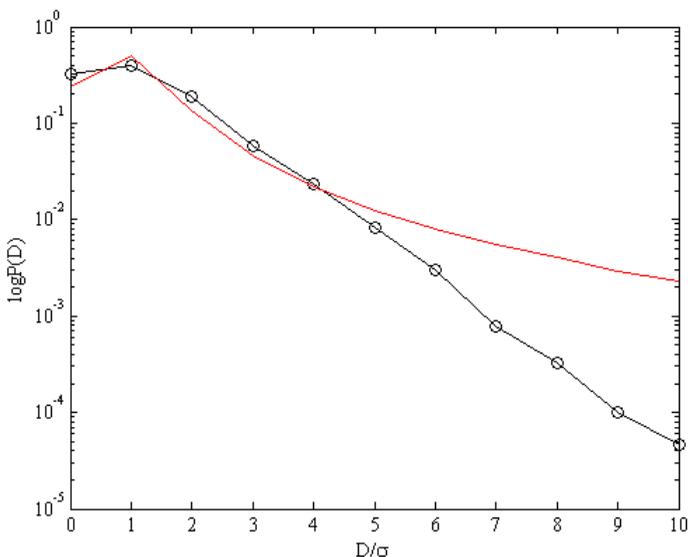


Figura 46 - Fdp das diferenças normalizadas dos números primos para $\Delta k = 1$. A fdp de Levy-estável (linha vermelha) é obtida pelos parâmetros $\alpha = 1.1$, $\beta = 1$, $\gamma = 4.82562$ e $\delta = 8.87558$ estimados por máxima verossimilhança.

Na quinta etapa tentou-se verificar algum padrão de escalonamento com as diferenças escalonadas pelas equações (19) e (20) juntamente com $\alpha = 1.1$. Figuras 47 e 48 não mostram evidências de algum escalonamento, pois as fdps empíricas (para os valores crescentes de Δk) não colapsam na fdp de $\Delta k = 1$.

$$\tilde{D} \equiv \frac{D}{(\Delta k)^{1/\alpha}} \quad (19)$$

e

$$\tilde{P}(\tilde{D}) \equiv P(D) \cdot (\Delta k)^{1/\alpha} \quad (20)$$

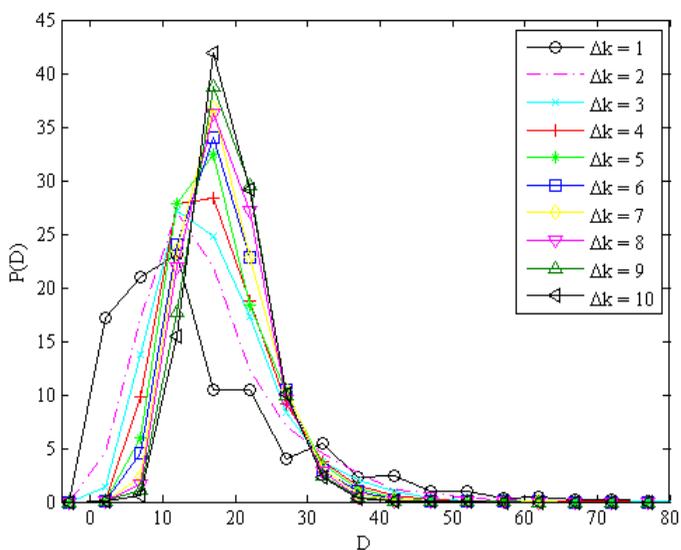


Figura 47 - As mesmas fdps da figura 42 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.1$

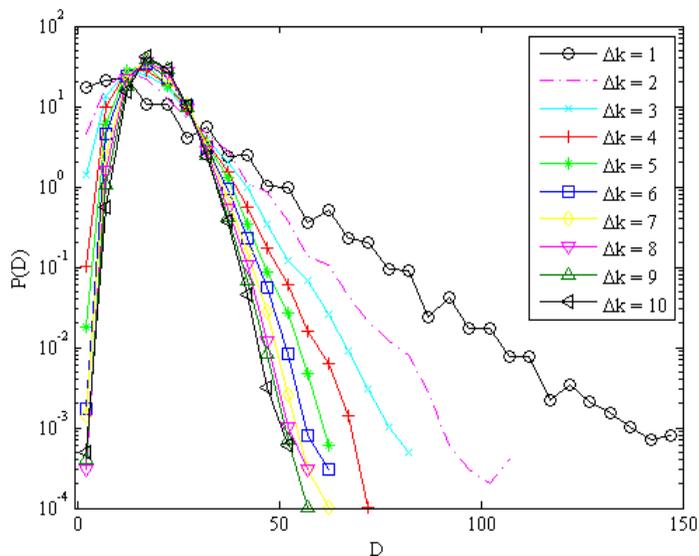


Figura 48 - As mesmas fdps da figura 43 construídas em unidades escalonadas. O escalonamento foi feito utilizando o valor de $\alpha = 1.1$.

Como não foi detectado algum padrão de escalonamento entre as diferenças dos dados considerados, partimos diretamente para a checagem da robustez dos parâmetros estimados, α e γ . Na sexta etapa, a base de dados das diferenças para $\Delta k = 1$ foi dividida em 20 subconjuntos (cada um representando 5% do total da série), e os parâmetros α e γ de cada subconjunto foram estimados pelo método de Máxima Verossimilhança.

Figura 49 mostra que o α permanece dentro do regime de Levy não-Gaussiano, alternando basicamente entre dois valores: 1.1 e 1.2. Por sua vez, o γ apresenta uma tendência ascendente, apesar de apresentar algumas oscilações (Figura 50). Logo, duas considerações podem ser feitas: (1) o parâmetro α é aproximadamente constante; e (2) o parâmetro γ é instável e não parece alcançar um limite. Portanto, a fdp de Levy-estável falha em descrever tal instabilidade presente nos dados.

Para fins de dúvidas, reportamos a evolução dos parâmetros α e γ calculados pelo método da Função Característica, cujos resultados foram similares às estimativas por máxima verossimilhança. O α , apesar de apresentar uma evolução diferente da Figura 49, variou menos, no intervalo ($1.66 < \alpha < 1.74$), e o γ apresentou a mesma tendência ascendente, embora com menos oscilações (Figuras 51 e 52).

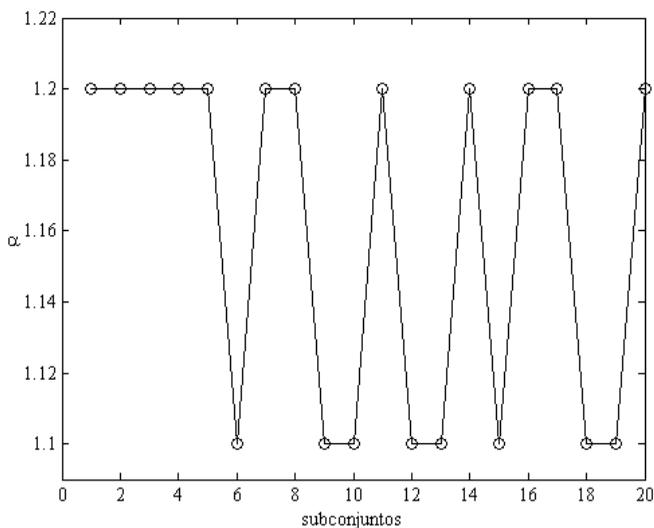


Figura 49 – Dependência temporal do índice α analisada em uma escala de 20 subconjuntos. O parâmetro α ($1.1 \leq \alpha \leq 1.2$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

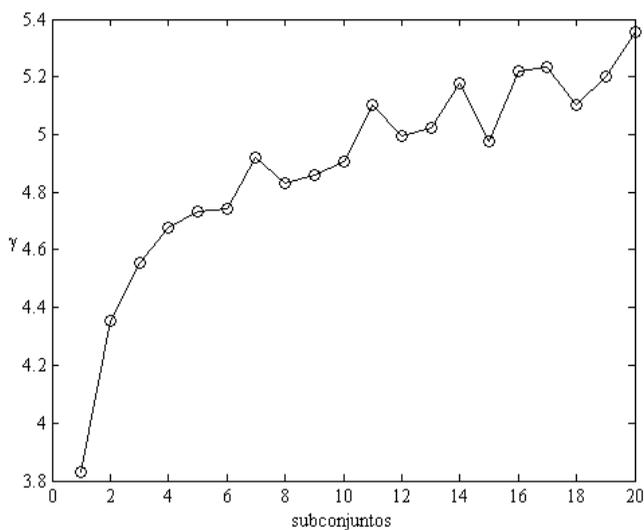


Figura 50 - Dependência temporal do parâmetro γ analisada em uma escala de 20 subconjuntos. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

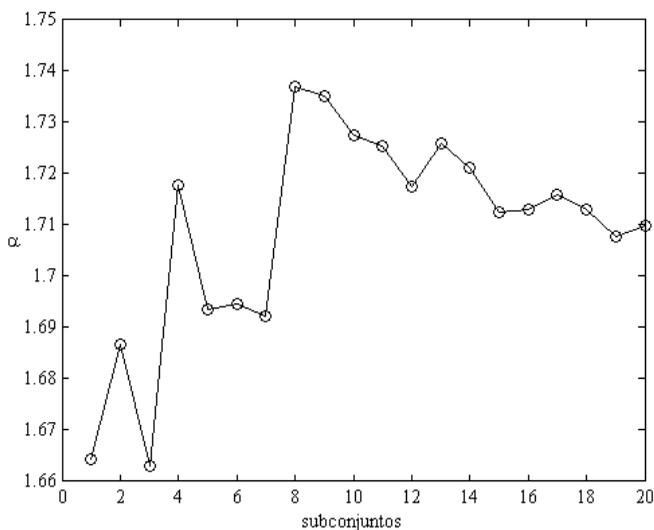


Figura 51 - Dependência temporal do índice α calculado pelo método da Função Característica em uma escala de 20 subconjuntos. O parâmetro α ($1.66 < \alpha < 1.74$) permanece dentro do regime de Levy não-Gaussiano ($0 < \alpha < 2$).

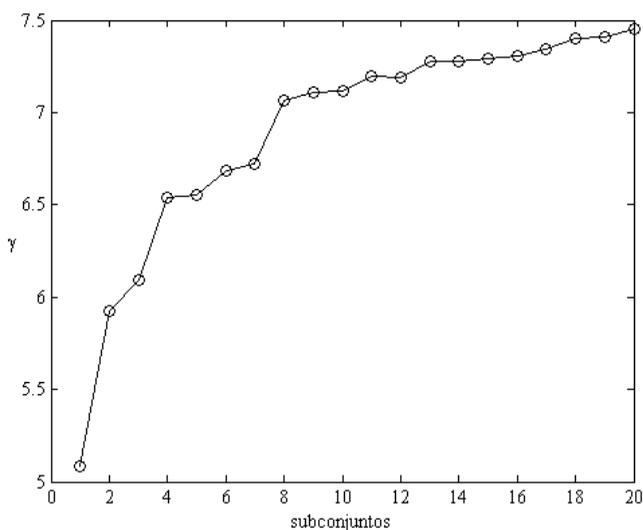


Figura 52 - Dependência temporal do parâmetro γ calculado pelo método da Função Característica em uma escala de 20 subconjuntos. Como γ é instável, a fdp de Levy-estável falha em descrever a volatilidade dos dados.

Na sétima etapa, os expoentes de decaimento da cauda da fdp das diferenças dos números primos para $\Delta k = 1$ foram estimados pelos dois métodos descritos no capítulo 3. Como estamos trabalhando com o conceito de diferenças em vez de retornos, alteramos somente a notação das equações, sendo que os procedimentos para as estimativas são os mesmos.

Dessa forma, pelo método de Gabaix e Ibragimov (2007), os tamanhos das diferenças extremas ordenadas de forma decrescente foram regredidos por MQO, de acordo com a equação (21).

$$\log\left(\text{rank} - \frac{1}{2}\right) = c - \alpha_K \log D(k) \quad (21)$$

Onde o erro padrão assintótico de α_K é dado por $\sqrt{\frac{2}{n}}\alpha_K$, e n é o tamanho da série das diferenças extremas.

Tabela 11 apresenta os resultados de α_K utilizando-se a equação (21) para os tamanhos da cauda correspondentes a 1%, 5% e 10% da série das diferenças ordenadas. A Figura 53 mostra o complemento da fda das diferenças. A inclinação corresponde à estimativa do coeficiente α_K da regressão representada pela equação (21). No entanto, observa-se que os extremos não decaem linearmente seguindo uma lei de potência, cujo decaimento apresenta mais a forma de uma parábola, característico de uma distribuição Gaussiana. Além disso, à medida que o tamanho da série dos extremos aumenta, o coeficiente α_K decresce, atingindo o valor da lei cúbica inversa somente para 27% dos valores ordenados (Tabela 11).

Resumindo, as diferenças extremas entre dois números primos consecutivos não seguem uma lei de potência, pois apresentaram valores do expoente α_K muito acima do regime de Levy para 1%, 5% e 10% das diferenças ordenadas, e por isso não podem ser acomodadas por uma distribuição estável.

Tabela 11 - Estimativas dos expoentes de decaimento das diferenças extremas dos números primos para $\Delta k = 1$: método de Gabaix e Ibragimov (2007).

Tamanho da cauda	α_K
1%	-6.5876 (0.0932)
5%	-4.8782 (0.0309)
10%	-4.1384 (0.0185)
27%	-3.0061 (0.0082)
30%	-2.8423 (0.0073)

Nota: erro padrão em parênteses

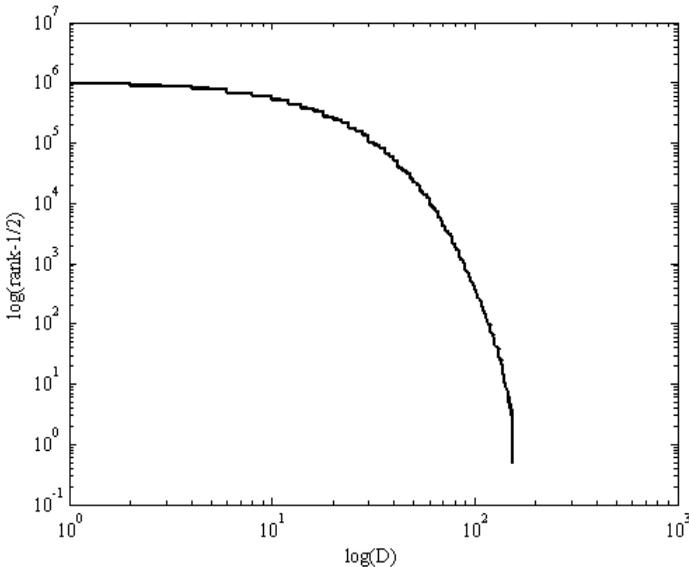


Figura 53 - Logaritmo das diferenças versus $\log\left(\text{rank} - \frac{1}{2}\right)$ dos números primos. A inclinação corresponde à estimativa do coeficiente α_K da regressão 21. O decaimento possui a forma de uma parábola, característico de uma distribuição Gaussiana.

Para confirmar a análise da cauda consideramos o método de Gopikrishnan et al. (1998). Estudamos novamente as diferenças D para $\Delta k = 1$ e dividimos seus valores pelo desvio padrão calculado dos dados ($\sigma = 12.9817$), como apresentado na equação (22):

$$g(k) = \frac{D(k)}{\sigma} \quad (22)$$

Em seguida, as diferenças normalizadas foram ordenadas de forma decrescente e calculou-se a distribuição de probabilidade cumulativa de se observar uma alteração $g(k)$ maior ou igual a g para as diferenças extremas, isto é, $P(g) \equiv P\{g(k) \geq g\}$. Dessa forma, o expoente de decaimento foi obtido pela inclinação de $\log P(g)$ em função de $\log(g)$.

$$\log P(g) = -\alpha_K \log g \quad (23)$$

A Tabela 12 mostra que os expoentes de decaimento apresentam valores maiores que aqueles calculados pelo método de Gabaix e Ibragimov (2007) e observa-se que o expoente α_K decresce a uma taxa menor à medida que aumentamos o tamanho da série dos extremos, se aproximando de 3 somente para 100% da série das diferenças. Além disso, Figuras 54, 55 e 56 não apresentam leis de potência para os tamanhos da cauda definidos por 1%, 5% e 10% da série das diferenças ordenadas, pois os dados não decaem linearmente. Tais resultados confirmam que a frequência dos extremos não segue uma lei de potência.

Portanto, somando os resultados encontrados até aqui, podemos concluir que a distribuição dos números primos não é consistente com a hipótese de estabilidade, pois a fdp de Levy-estável não proporcionou um bom ajuste na região modal dos dados (Figura 46) e não foi capaz de explicar seus valores extremos (Figuras 54, 55 e 56).

Tabela 12 - Estimativas dos expoentes de decaimento das diferenças extremas dos números primos para $\Delta k = 1$: método de Gopikrishnan et al. (1998).

Tamanho da cauda	α_K
1%	-8.4935 (0.5862)
5%	-6.7133 (0.6209)
10%	-5.67 (0.6374)

Nota: erro padrão em parênteses

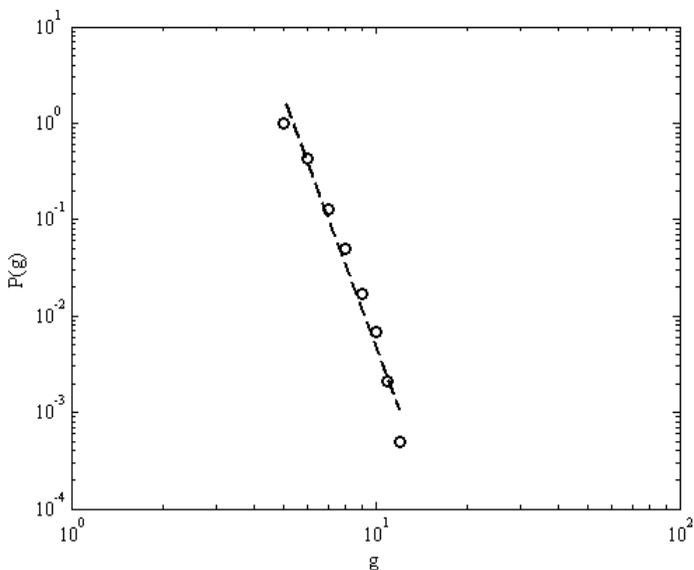


Figura 54 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 1% das diferenças ordenadas e normalizadas $g(t)$.

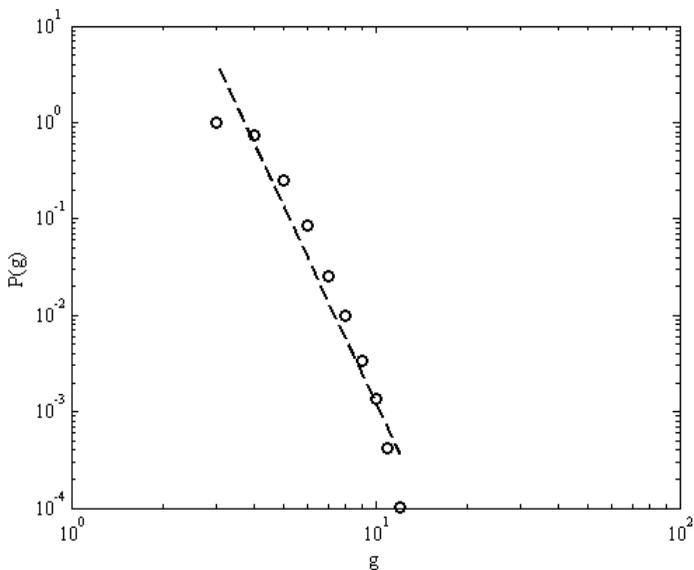


Figura 55 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 5% das diferenças ordenadas e normalizadas $g(t)$.

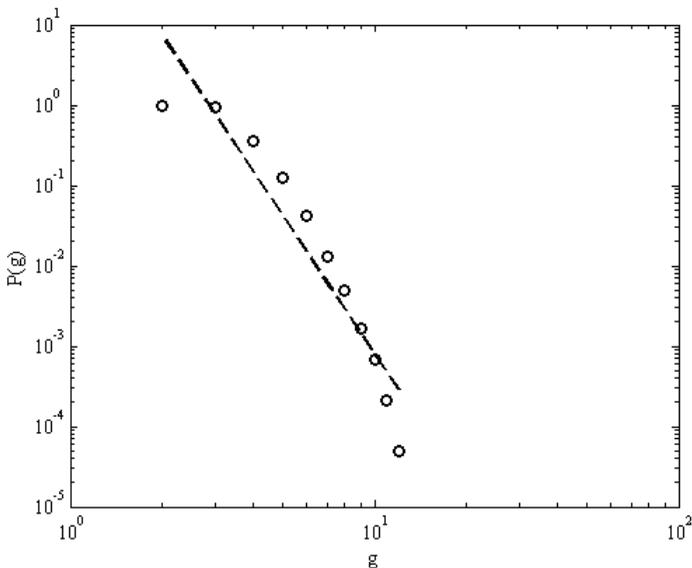


Figura 56 – Gráfico log-log da fda de $P(g)$ para o tamanho da cauda definida por 10% das diferenças ordenadas e normalizadas $g(t)$.

Finalizando a análise, na última etapa, a volatilidade dos números primos foi modelada por um processo GARCH(1,1), definido como:

$$\sigma_k^2 = a_0 + a_1 x_{k-1}^2 + b_1 \sigma_{k-1}^2 \quad (24)$$

Utilizando-se as equações (13) e (14), considerando os valores de $\sigma = 12.9817$ e $\kappa = 7.9106$, calculados dos dados empíricos para $\Delta k = 1$, e o valor de $b_1 = 0.9$, obtemos $a_0 = 2.0897$ e $a_1 = 0.0876$. Logo, a variância condicional é descrita por:

$$\sigma_k^2 = 2.0897 + 0.0876 x_{k-1}^2 + 0.9 \sigma_{k-1}^2 \quad (25)$$

Figura 57 compara a fdp empírica das diferenças dos primos para $\Delta k = 1$ com o processo GARCH(1,1) caracterizado por $a_0 = 2.0897$, $a_1 = 0.0876$ e $b_1 = 0.9$. Tal processo foi obtido por inúmeras simulações realizadas com os parâmetros calculados acima. Assim, observa-se que o processo GARCH(1,1) fornece um ajuste inferior ao da fdp de Levy-estável (Figura 46), tanto na região modal quanto na cauda da fdp empírica.

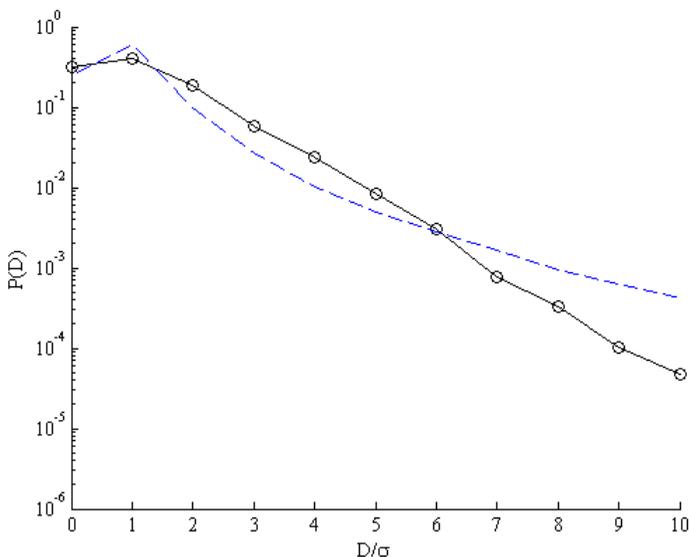


Figura 57 - Ajuste do processo GARCH(1,1) (linha azul tracejada) à fdp empírica das diferenças dos números primos para $\Delta k = 1$ (círculos). A fdp incondicional do processo GARCH(1,1) é caracterizada por $a_0 = 2.0896$, $a_1 = 0.0876$ e $b_1 = 0.9$.

5.2. ANÁLISE SUMÁRIA DOS RESULTADOS DOS NÚMEROS PRIMOS

O estudo da distribuição dos números primos, cuja amostra compreende os primeiros 1,000,000 de primos, nos permitem concluir que:

1. A fdp das diferenças entre dois números primos consecutivos é assimétrica à direita e não-Gaussiana ($\kappa = 7.9106$);
2. A forma funcional da fdp das diferenças se altera com o aumento de Δk , indicando não-estabilidade dos números primos. Além disso, não há evidências de algum tipo de escalonamento presente nos dados;
3. A fdp de Levy-estável de $\alpha = 1.1$ proporcionou um ajuste razoável na região modal da fdp empírica, mas a cauda da Levy se localizou abaixo da cauda dos dados, indicando que a variância da fdp empírica é finita;
4. O parâmetro α permaneceu, aproximadamente, constante entre os subconjuntos. O parâmetro γ , por sua vez, apresentou uma tendência

ascendente, apesar de apresentar algumas oscilações. Logo, como γ é instável, a fdp de Levy falha em descrever a volatilidade presente nos dados;

5. Os métodos de estimação do decaimento da cauda da distribuição empírica constataram que as diferenças extremas não seguem uma lei de potência, pois os dados não decaem linearmente. Sugerindo, então, que não há um padrão de lei de potência para os números primos;
6. Utilizou-se o modelo GARCH(1,1) para tentar descrever a evolução temporal de σ . Entretanto, o processo GARCH(1,1) forneceu um ajuste inferior ao da distribuição de Levy-estável, tanto na região modal quanto na região extrema da fdp empírica;
7. Resumindo, a hipótese de que as diferenças dos números primos são geradas por uma distribuição de Levy-estável não pode ser suportada, pois não há evidências de algum tipo de escalonamento presente nos dados. Além disso, o decaimento da cauda não segue uma lei de potência, cujos expoentes apresentaram valores muito acima do regime de Levy. Portanto, pode-se concluir que a distribuição dos números primos não é consistente com a hipótese de estabilidade.

6. CONCLUSÃO

Este trabalho estudou as propriedades estatísticas dos retornos de alta frequência da ação PETR4 e do índice DJIA e das diferenças dos números primos. E foi avaliado se as três séries são consistentes com a hipótese de estabilidade. Como resultado concluiu-se que os retornos de alta frequência das duas séries financeiras podem ser modelados por uma distribuição de Levy-estável na região modal de suas fdps. Um escalonamento não-Gaussiano foi verificado em ambas as séries financeiras. Os expoentes característicos dos retornos da ação PETR4 e do índice DJIA, de $\alpha = 1.42$ e $\alpha = 1.49$, respectivamente, foram compatíveis com os estudos anteriores de Mantegna e Stanley (1995) e Gleria et al. (2002), os quais reportaram valores de $\alpha = 1.40$ e $\alpha = 1.66$ para os índices S&P-500 e IBOVESPA, respectivamente. Além disso, o flash crash não pode ser considerado uma anomalia. Um estudo mais apurado das caudas das duas séries financeiras indicou a presença de um decaimento de lei de potência acima do regime de Levy, mas não sendo a lei cúbica inversa. Isto significa que os eventos extremos das duas séries financeiras não podem ser modelados por uma distribuição de Levy-estável. Como a volatilidade dependente do tempo das duas séries financeiras não pôde ser descrita pela fdp de Levy-estável, utilizou-se o modelo GARCH(1,1) para modelar a volatilidade, o qual, então, forneceu um bom ajuste tanto para a região central quanto para os extremos. Mas, como o GARCH(1,1) é um processo instável, ele não apresenta escalonamento, sendo válido somente para um intervalo de tempo pequeno.

Em suma, um escalonamento não-Gaussiano nos dados financeiros sugere que um tipo específico de distribuição de Levy-estável (em vez de qualquer distribuição não-estável, como o processo GARCH(1,1)) é a principal fonte de geração dos retornos da ação PETR4 e do índice DJIA.

Dado a não rejeição da hipótese de estabilidade para os dados financeiros, como tema futuro sugere-se modelar os resíduos de alguns modelos econométricos importantes na literatura financeira, tais como o CAPM e o GARCH, por uma distribuição estável e verificar sua eficiência em prever os retornos das ações e diminuir os erros de apreçamento.

Com relação aos números primos, concluiu-se que suas primeiras diferenças não são consistentes com a hipótese de estabilidade, pois a distribuição de Levy-estável não proporcionou um bom ajuste na região

modal da fdp empírica e não houve evidências de algum padrão de escalonamento presente nos dados. Além disso, o decaimento da cauda empírica não segue uma lei de potência, cujos expoentes apresentaram valores muito acima do regime de Levy. Somando-se a isso, os resultados do conjunto dos diagnósticos também não ofereceram suporte para a hipótese de estabilidade dos dados. Apesar de tal série apresentar uma variância não constante, o processo GARCH(1,1) forneceu um ajuste inferior ao da distribuição de Levy-estável, tanto na região modal quanto na região extrema da fdp empírica.

Por fim, acredita-se que a presente pesquisa tenha contribuído com a literatura financeira ao estudar as propriedades estatísticas dos dados de uma ação individual do mercado brasileiro (PETR4), do índice DJIA compreendendo o crash mais recente e dos números primos utilizados como variável de controle não-financeira com o objetivo de avaliar a hipótese de estabilidade dos dados. Pois, de acordo com Mantegna e Stanley (2000), compreender as propriedades estatísticas dos retornos de ativos é essencial por razões fundamentais, referentes ao entendimento das dinâmicas de mercado, e por razões aplicadas, relacionadas aos problemas práticos de apreçamento de opções e para a gestão de carteiras. Tal conhecimento é crucial para a construção e para o teste de um modelo estatístico de um mercado financeiro. Apesar de mais de 50 anos de esforços, este objetivo ainda não foi alcançado.

BIBLIOGRAFIA

AKGIRAY, V. Conditional Heteroscedasticity in time series of stock returns: evidence and forecasts. **Journal of Business**, v. 62, p. 55-80, 1989.

BACHELIER, L. J. B. **Théorie de la Speculation**. Paris: Gauthier-Villars, 1900.

BAK, P.; PACZUSKI, M. Complexity, contingency and criticality. **Proceedings of the National Academy of Sciences of the USA**, v. 92, p. 6689-6696, 1995.

BORAK, S.; HÄRDLE, W.; WERON, R. Stable distributions. In: CIZEK, P.; HÄRDLE, W.; WERON, R. **Statistical Tools for Finance and Insurance**. Berlin: Springer-Verlag, p. 21-44, 2005.

BORAK, S.; MISIOREK, A.; WERON, R. Models for Heavy-tailed Asset Returns. **Munich Personal Repec Archive**. Germany, 35 p., 2010. Trabalho não publicado.

FAMA, E. F. The behavior of stock market prices. **Journal of Business**, v. 38, p. 34–105, 1965.

GABAIX, X. Power Laws in Economics and Finance. **NBER Working Papers**, v. 14299, p. 01-48, 2008.

GABAIX, X; IBRAGIMOV, R. Rank $- 1/2$: a simple way to improve the OLS estimation of tail exponents. **NBER Working Paper**, v. 342, p. 01-32, 2007.

GHAHFAROKHI, M. B.; GHAHFAROKHI, P. B. Applications of Stable Distributions in Time Series Analysis, Computer Sciences and Financial Markets. **World Academy of Science, Engineering and Technology**, v. 49, p. 1027-1031, 2009.

GLERIA, I.; MATSUSHITA, R.; SILVA, Da S. Scaling power laws in the Sao Paulo Stock Exchange. **Economics Bulletin**, vol. 7, n. 3 p. 01–12, 2002.

GNEDENKO, B. V.; KOLMOGOROV, A. N. **Limit Distributions for Sums of Independent Random Variables**. Cambridge MA: Addison-Wesley, 1954.

GOPIKRISHNAN, P. et al. Inverse cubic law for the distribution of stock price variations. **European Journal of Physics B**, v. 3, p. 139-140, 1998.

KOGON, S. M.; WILLIAMS, D. B. Characteristic function based estimation of stable parameters. In: ADLER, R.; FELDMAN, R.; TAQQU, M. **A Practical Guide to Heavy Tailed Data**. Boston: Birkhäuser, p. 311-335, 1998.

KOUTROUVELIS, I. A. Regression type estimation of the parameters of stable laws, **JASA**, v. 75, p. 918-928, 1980.

KOUTROUVELIS, I. A. An iterative procedure for the estimation of the parameters of stable laws. **Communications in Statistics – Simulation and Computation**, v.10, p.17-28, 1981.

KUMAR, P.; IVANOV, P.; C. STANLEY, H. E. Information entropy and correlations in prime numbers. Disponível em: <http://xxx.lanl.gov/abs/cond-mat/0303110> (2003).

LEVY, P. **Calcul des probabilités**. Paris: Gauthier-Villars, 1925.

MANDELBROT, B. B. The variation of certain speculative prices. **Journal of Business**, v. 36, p. 394–419, 1963.

MANTEGNA, R.; STANLEY, H. E. Scaling behaviour in the dynamics of an economic index. **Nature**, v. 376, p. 46-49, 1995.

MANTEGNA, R.; STANLEY, H. E. Modeling of financial data: Comparison of the truncated Lévy flight and the ARCH(1) and GARCH(1,1) processes. **Physica A**, v. 254, p. 77-84, 1998.

MANTEGNA, R.; STANLEY, H. E. **An Introduction to Econophysics: Correlations and Complexity in Finance**. Cambridge: Cambridge University Press, 2000.

MCCULLOCH, J. H. Simple consistent estimators of stable distribution parameters. **Communications in Statistics – Simulation and Computation**, v. 15, p. 1109-1136, 1986.

MCCULLOCH, J. H. Financial applications of stable distributions. In: MADDALA, G. S.; RAO, C. R. **Handbook of Statistics 14, Statistical Methods in Finance**. Elsevier, p. 393–425, 1996.

MICHAEL, J. R. The stabilized probability plot. **Biometrika**, v. 70, p.11-17, 1983.

NOLAN, J. P. Numerical Calculation of Stable Densities and Distribution Functions. **Communications in Statistics – Stochastic Models**, v. 13, p. 759–774, 1997.

NOLAN, J. P. Fitting data and assessing goodness of fit with stable distributions. In: **Proceedings of the Conference on Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics**. American University, Washington, DC, 1999.

NOLAN, J. P. Maximum Likelihood Estimation and Diagnostics for Stable Distributions. In: BARNDORFF-NIELSEN, O. E.; MIKOSCH, T.; RESNICK, S. **Lévy Processes**. Boston: Birkhäuser, 2001.

NOLAN, J. P. Modeling Financial Data with Stable Distributions. In: RACHEV, S. T.; ZIEMBA, W. T. **Handbooks of Finance**. North Holland, 2003.

PODOBNIK, B. et al. ARCH-GARCH approaches to modeling high-frequency financial data. **Physica A**, v. 344, p. 216-220, 2004.

SAMORODNITSKY, G.; TAQQU, M. S. **Stable Non-Gaussian Random Processes**. New York: Chapman and Hall, 1994.

SCALAS, E.; KIM, K. The art of fitting financial time series with Levy stable distributions. **Munich Personal Repec Archive**. Germany, 17 p., 2006. Trabalho não publicado.

TIMMERMANN, A. Scales and Stock Markets. **Nature**, v. 376, p. 18-19, 1995.

VOIT J. **The Statistical Mechanics of Financial Markets** (Texts and Monographs in Physics). Springer-Verlag, 2003.

WELL, D. **Prime numbers: the most mysterious figures in math**. USA: John Wiley & Sons, 2005.

WERON, R. Computationally intensive Value at Risk calculations. In: GENTLE, J. E.; HARDLE, W.; MORI, Y. **Handbook of Computational Statistics**. Berlin: Springer, p. 911–950, 2004.

WOLF, M. Applications of statistical mechanics in number theory. **PHYSICA A**, v. 274, n. 1, p. 149-157, 1999.

ZOLOTAREV, V. M. One-dimensional stable distributions. Translations of Mathematical Monographs, v. 65. **American Mathematical Society**, 1986.