

Kerolly Kedma Felix do Nascimento

**Simulação de mercados financeiros baseados em agentes
guiados por sistemas inteligentes**

Recife

Março/2017



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**Simulação de mercados financeiros baseados em agentes guiados por sistemas
inteligentes**

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 06/03/2017 pela comissão examinadora

Área de concentração: Biometria e Estatística Aplicada

**Orientador: Dr. Tiago Alessandro Espí-
nola Ferreira**
**Coorientador: Dr. José Ferraz de
Moura Nunes Filho**

Recife

Março/2017

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema Integrado de Bibliotecas da UFRPE
Biblioteca Central, Recife-PE, Brasil

N244s Nascimento, Kerolly Kedma Felix do
Simulação de mercados financeiros baseados em agentes guiados por sistemas inteligentes / Kerolly Kedma Felix do Nascimento. – 2017.
75 f. : il.

Orientador: Tiago Alessandro Espínola Ferreira.
Coorientador: José Ferraz de Moura Nunes Filho.
Dissertação (Mestrado) – Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, BR-PE, 2017.
Inclui referências.

1. Mercados financeiros 2. Particle swarm optimization 3. Simulação 4. Modelagem I. Ferreira, Tiago Alessandro Espínola, orient. II. Nunes Filho, José Ferraz de Moura, coorient. III. Título

CDD 574.018

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

Simulação de mercados financeiros baseados em agentes guiados por sistemas inteligentes

Kerolly Kedma Felix do Nascimento

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 06/03/2017 pela comissão examinadora

Orientador:

**Dr. Tiago Alessandro Espínola
Ferreira**
Orientador

Banca examinadora:

José Ferraz de Moura Nunes Filho
Universidade Federal Rural de
Pernambuco

Pedro Hugo de Figueiredo
Universidade Federal Rural de
Pernambuco

*Dedico esta dissertação a meus pais,
minhas irmãs, parentes e amigos.*

Agradecimentos

Em primeiro lugar, agradeço a Deus pelo dom da minha vida e o de todas as pessoas que tive a oportunidade de conhecer. Agradeço à minha família por serem a base sólida que ao longo de toda minha jornada tem me apoiado, incentivado e acreditado no meu potencial. Por terem compreendido minhas ausências, angústias e preocupações.

Aos meus pais, Tarcísio e Altair, por seu amor incondicional, por todo o empenho e dedicação. Por todos os sacrifícios e renúncias que tiveram que fazer para proporcionarem a mim e às minhas irmãs sempre o melhor que podiam oferecer. A elas, Kellyma e Kenikywaynne, que mesmo distantes fisicamente, sempre se fizeram muito presentes e me apoiaram em todos os momentos.

Agradeço ao professor Dr. Tiago A. E. Ferreira pela oportunidade em participar de seu grupo de pesquisas. Pelo apoio nesse trabalho, orientação, paciência e ensinamentos passados.

Ao Dr. Jader da S. Jale, por toda ajuda na construção dessa dissertação e dos demais trabalhos produzidos nesse percurso, por todas as vezes que pacientemente tirou minhas dúvidas e por todo o conhecimento transmitido.

Ao meu co-orientador, Dr. José Ferraz de Moura Nunes Filho, pelo apoio e disponibilidade.

Às minhas amigas de graduação, Aglaiza e Estela, as quais fazem de tudo para mantermos o vínculo e a amizade construídas durante nosso curso de licenciatura. Em especial, agradeço a Aglaiza por todas as palavras e atitudes tomadas por ela quando decidimos vir estudar em Recife, longe dos nossos familiares e amigos. A Samuel, que me falou sobre o mestrado do PPGBEA, agradeço pela acolhida, por todos os conselhos e sua amizade ao longo desses dois anos. A Leda, pela ajuda, convivência e amizade desde que cheguei em Recife. A Evelyn, por todo o carinho e atenção. A Camiliane por sua valiosa amizade construída no início do nosso curso. A Nathielly e a Herica, que com suas brincadeiras, tornaram meus dias mais alegres e menos difíceis. A Ikaro, por sua amizade e disponibilidade em tirar minhas dúvidas. A Edy, por todos os conselhos e sua disponibilidade em sempre ajudar. A Rodrigo, Albaro, Jonas, Fábio, Isabelly, Filipe, Neidinha, Dalton, Luisa, Glauce, Marcela, Gutemberg, Rivelino, Carlos Renato e aos demais colegas do departamento por toda convivência e ajuda nos momentos difíceis.

Aos professores e funcionários do PPGBEA, em especial, aos professores Tiago, Paulo, Cláudio, Tatjana, Borko, Moacyr, aos secretários Marco e Edivânia e ao pessoal dos serviços gerais.

A CAPES pelo apoio financeiro.

Enfim, tive muita sorte de ser abençoada em conhecer e conviver com cada uma dessas pessoas. Agradeço a todos que direta ou indiretamente contribuíram para a realização deste trabalho.

*“Que os vossos esforços desafiem as impossibilidades,
lembrai-vos de que as grandes coisas do homem
foram conquistadas do que parecia impossível.
(Charles Chaplin)*

Resumo

Muitas são as tentativas realizadas em busca do melhor desenvolvimento de técnicas que auxiliem na tomada de decisão dos participantes dos mercados financeiros. Para isso, pesquisadores têm recorrido ao estudo dos conceitos da Física Estatística e da Econofísica, que utilizam abordagens e teorias da Física no entendimento dos fenômenos econômicos e financeiros. Além disso, têm-se desenvolvido modelos baseados em agentes que buscam reproduzir de forma simplificada a realidade. Nesse trabalho, desenvolvemos um ambiente artificial que simula agentes financeiros interagindo entre si. Nessa abordagem, utilizamos o algoritmo *Particle Swarm Optimization* (PSO) para maximizar a riqueza dos agentes artificiais. O comportamento dos agentes foi analisado em quatro tipos distintos de mercados: seguindo um passeio aleatório, com tendência de alta, com tendência de baixa e com séries financeiras reais. Com as séries de retornos dos preços negociados nesses ambientes, estudamos os chamados fatos estilizados da economia e o ajuste das mesmas pelas funções exponencial e leis de potência. Por fim, com o Coeficiente de Gini avaliamos a distribuição de riqueza entre os agentes financeiros artificiais. Devido aos bons resultados obtidos com o uso do PSO para maximização da riqueza dos agentes financeiros nos mercados estudados, nossa metodologia mostrou-se bastante promissora no entendimento do comportamento dos mercados financeiros.

Palavras-chaves: Mercados Financeiros, *Particle Swarm Optimization*, Simulação, Modelagem.

Abstract

Many attempts have been made to seek the best development of techniques to assist decision-makers in the financial markets. For this, researchers have resorted to the study of the concepts of Statistical Physics and Econophysics, which use approaches and theories of Physics in the understanding of economic and financial phenomena. In addition, we have developed models based on agents that seek to reproduce reality in a simplified way. In this work, we developed an artificial environment that simulates financial agents interacting with each other. In this approach, we use the *Particle Swarm Optimization* (PSO) algorithm to maximize the richness of the artificial agents. The behavior of the agents was analyzed in four different types of markets: following a random trend, with a bullish trend, with a downward trend and with real financial series. With the series of price returns negotiated in these environments, we study the so-called stylized facts of the economy and their adjustment by exponential functions and power laws. Finally, with the Gini coefficient we evaluate the distribution of wealth among the artificial financial agents. Due to the good results obtained with the use of the PSO to maximize the wealth of the financial agents in the markets studied, our methodology was very promising.

Key-words: Financial Markets, Particle Swarm Optimization, Simulation, Modeling.

Lista de Figuras

Figura 1 – Ilustração da movimentação de uma partícula no espaço 2D [1].	8
Figura 2 – Topologias do algoritmo PSO. Topologia Global (a) e uma possível Topologia Local (b) [2]	9
Figura 3 – Índices diários dos preços de fechamento do S&P 500 entre os períodos de 27 de abril de 1993 e 17 de junho de 2016.	13
Figura 4 – Ilustração do PIB do Reino Unido, do Canadá, da África do Sul e do mundo de 1960 a 2015 [3].	13
Figura 5 – Série de preços de fechamento do Ibovespa.	19
Figura 6 – Série de preços de fechamento do CAC 40.	20
Figura 7 – Série temporais de preços de fechamento do IPC.	21
Figura 8 – Série temporais de preços de fechamento da Nasdaq.	22
Figura 9 – Ilustração da Curva de Lorenz [4].	29
Figura 10 – Séries dos preços negociados para os <i>drifts</i> $N(0, 0.01)$, $N(0, 0.05)$, $N(0, 0.10)$ e $N(0, 0.30)$	31
Figura 11 – Autocorrelações dos preços negociados nas simulações com <i>drifts</i> $N(0, 0.01)$, $N(0, 0.05)$, $N(0, 0.10)$ e $N(0, 0.30)$	33
Figura 12 – Séries de preços negociados no mercado simulado para os <i>drifts</i> $N(0.01, 0.01)$, $N(-0.01, 0.05)$, $N(0.05, 0.10)$ e $N(-0.10, 0.10)$	34
Figura 13 – Exemplos de séries temporais de retornos dos preços negociados no mercado simulado.	36
Figura 14 – Exemplos de histogramas de séries temporais de retornos dos preços negociados no mercado simulado: $\sigma = 0.01$ (a), $\sigma = 0.05$ (b), $\sigma = 0.10$ (c), $\sigma = 0.30$ (d), <i>drift</i> (0.05, 0.10) (e) e <i>drift</i> (-0.10, 0.10) (f).	37
Figura 15 – Exemplos de autocorrelações de retornos de preços negociados.	39
Figura 16 – Exemplos de ajustes lineares na escala semi-log.	40
Figura 17 – Exemplos de ajustes lineares na escala log-log.	41
Figura 18 – Exemplos de ajustes pela exponencial sem transformação na escala para as simulações com <i>drifts</i> : $N(0, 0.01)$ (a), $N(0, 0.05)$ (b), $N(0, 0.10)$ (c), $N(0, 0.30)$ (d), $N(0.01, 0.05)$ (e) e $N(-0.10, 0.01)$ (f).	42
Figura 19 – Exemplos de ajustes pela lei de potência sem transformação na escala para as simulações com <i>drifts</i> : $N(0, 0.01)$ (a), $N(0, 0.05)$ (b), $N(0, 0.10)$ (c), $N(0, 0.30)$ (d), $N(0.01, 0.05)$ (e) e $N(-0.10, 0.01)$ (f).	43

Figura 20 – Ilustração dos limiares de venda e de compra para a melhor partícula nos RW's.	44
Figura 21 – Ilustração dos limiares de venda e de compra para a pior partícula nos RW's.	45
Figura 22 – Ilustração de exemplos dos limiares de venda e de compra para a melhor partícula nas simulações com <i>drifts</i> $N(-0.01, 0.05)$ e $N(-0.10, 0.10)$	46
Figura 23 – Ilustração de exemplos dos limiares de venda e de compra para a pior partícula nas simulações com <i>drifts</i> $N(-0.01, 0.05)$ e $N(-0.10, 0.10)$	46
Figura 24 – Ilustração de exemplos dos limiares de venda e de compra para a pior partícula nas simulações com tendência de alta nos mercados.	47
Figura 25 – Ilustração de exemplos dos limiares de venda e de compra para a melhor partícula nas simulações com tendência de alta nos mercados.	47
Figura 26 – Ilustração de exemplos do capital, da quantidade de ações e do <i>fitness</i> para dois mercados com tendência de baixa.	51
Figura 27 – Ilustração de exemplos do capital, quantidade de ações e <i>fitness</i> para dois mercados com tendência de alta.	52
Figura 28 – Ilustração da evolução do capital, da quantidade de ações e do <i>fitness</i> para os mercados com RW.	53
Figura 29 – Ilustração de exemplos da distribuição de riqueza com o Coeficiente de Gini.	55
Figura 30 – Plot 3D com a movimentação da melhor partícula em função dos limiares de venda (δ), de compra (δ) e da quantidade de riqueza (<i>fitness</i>) no caso de alguns mercados simulados.	55
Figura 31 – Séries de retornos dos preços dos mercados reais.	56
Figura 32 – Histograma das séries de retornos dos preços dos mercados reais.	57
Figura 33 – Autocorrelações das séries de retornos dos preços dos mercados reais.	58
Figura 34 – Ajustes lineares nas escalas semi-log (CAC 40 (a), Ibovespa (c), IPC (e) e Nasdaq (g)) e log-log (CAC 40 (b), Ibovespa (d), IPC (f) e Nasdaq (h)).	59
Figura 35 – Ajustes pela função exponencial (CAC 40 (a), Ibovespa (c), IPC (e) e Nasdaq (g)) e lei de potência (CAC 40 (b), Ibovespa (d), IPC (f) e Nasdaq (h)).	60
Figura 36 – Limiar de venda (δ) e de compra (λ) das melhores partículas nas iterações.	61
Figura 37 – Ilustração da quantidade de capital das séries de retornos dos preços dos índices Ibovespa e CAC 40, respectivamente.	62
Figura 38 – Ilustração da quantidade de capital das séries de retornos dos preços dos índices IPC e Nasdaq, respectivamente.	63
Figura 39 – Coeficiente de Gini da quantidade de <i>fitness</i> (riqueza) nos mercados reais.	64

Figura 40 – Plot 3D com a movimentação da melhor partícula em função dos limiares de venda (δ), de compra (δ) e da quantidade de riqueza (<i>fitness</i>) no caso dos mercados reais.	65
---	----

Lista de tabelas

Tabela 1 – Estatísticas da série Ibovespa.	19
Tabela 2 – Estatísticas da série CAC 40.	20
Tabela 3 – Estatísticas da série IPC.	21
Tabela 4 – Estatísticas da série Nasdaq.	22
Tabela 5 – Teste de tendência de Mann-Kendall (<i>Rank Test</i>) para os preços médios em cada RW.	32
Tabela 6 – Teste de Dickey-Fuller Aumentado (ADF) para os RW's.	32
Tabela 7 – Teste de tendência de Mann-Kendall (<i>Rank Test</i>) para os preços médios em cada <i>drift</i>	35
Tabela 8 – Teste de Normalidade de Jarque-Bera para os preços das ações em cada <i>drift</i>	38
Tabela 9 – Teste de homogeneidade das variâncias (Levene) entre os preços médios dos RW's.	39
Tabela 10 – Teste de Normalidade de Jarque-Bera para os limiares δ e λ em cada RW.	48
Tabela 11 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.05$ nos RWs.	48
Tabela 12 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.10$ nos RWs.	48
Tabela 13 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.30$ nos RWs.	49
Tabela 14 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.05$ vs $\sigma = 0.10$ nos RWs.	49
Tabela 15 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.05$ vs $\sigma = 0.30$ nos RWs.	49
Tabela 16 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.10$ vs $\sigma = 0.30$ nos RWs.	50
Tabela 17 – Teste de tendência e homogeneidade para as médias dos limiares δ e λ de acordo com os RWs.	50
Tabela 18 – Valores médios para δ , λ , <i>fitness</i> e índice de Gini (\bar{g}) em relação aos melhores e piores agentes em 30 simulações de cada <i>drift</i> negativo.	54
Tabela 19 – Valores médios para δ , λ , <i>fitness</i> e índice de Gini (\bar{g}) em relação aos melhores e piores agentes em 30 simulações de cada <i>drift</i> positivo.	54

Tabela 20 – Coeficiente de Gini para os mercados reais 63

Sumário

1	Introdução	1
1.1	Introdução	1
1.2	Objetivo	2
1.3	Estrutura da Dissertação	3
2	Revisão de Literatura	5
2.1	Mercados de Ações	5
2.2	Algoritmo Particle Swarm Optimization (PSO)	7
2.3	Fatos Estilizados	10
2.3.1	Assimetria e Curtose	11
2.3.2	Estacionariedade	12
2.3.3	Autocorrelação Linear	14
2.3.4	Heterocedasticidade	15
2.4	Medidas de Desigualdade	15
2.5	Distribuição Exponencial e de Lei de Potência	16
3	Metodologia	18
3.1	Base de Dados	18
3.1.1	Séries Artificiais	18
3.1.2	Série do Índice Ibovespa - Brasil	18
3.1.3	Série do Índice CAC 40 - França	19
3.1.4	Série do Índice IPC - México	20
3.1.5	Série do Índice Nasdaq - Estados Unidos da América	21
3.2	Ambiente Artificial Proposto para Simulação do Mercado Acionário	22
3.3	Testes Estatísticos	26
3.3.1	Teste de Mann-Kendall	26
3.3.2	Teste de Levene	27
3.3.3	Teste de Jarque-Bera	28
3.3.4	Teste de Dickey-Fuller Aumentado	28
3.4	Coeficiente de Gini	29
3.5	Ajuste Exponencial e por Lei de Potência	30
4	Resultados	31
4.1	Mercado Simulado com PSO	31
4.1.1	Validação do modelo	35

4.1.2	Análises dos Parâmetros dos Agentes Financeiros no Ambiente Artificial	44
4.2	Aplicação em Mercados Reais	56
5	Conclusões	66
5.1	Trabalhos Futuros	68
	Referências Bibliográficas	69

1 Introdução

1.1 Introdução

Os mercados de ações têm atraído o interesse tanto de investidores quanto de pesquisadores, na busca pelos melhores métodos para compreensão desse tema e para garantir um bom êxito nas tomadas de decisões. Essas são baseadas em estratégias técnicas e/ou fundamentalistas. As estratégias técnicas têm a finalidade de definir a tendência dos mercados a partir da observação das informações passadas existentes. Por outro lado, as estratégias fundamentalistas buscam reduzir as informações financeiras de uma empresa a um valor relacionado ao seu desempenho e, com isso, determinar se haverá valorização nas ações da empresa [5].

Em ambas as estratégias, são empregadas as séries temporais financeiras. Dizemos que uma sequência qualquer de observações ordenadas no tempo, capaz de transmitir informações de um certo fenômeno, é uma série temporal [6, 7]. Um exemplo disso são as séries econômicas e financeiras. Elas são consideradas sistemas complexos por serem formadas por elementos interdependentes que podem sofrer influência de fatores internos e externos ao ambiente e por apresentarem um comportamento não linear, advindo das interações entre os elementos do sistema e do ambiente, dificultando a determinação da composição da série [8].

O problema para conseguir uma boa previsibilidade do mercado de ações ocorre, principalmente, devido a não-estacionariedade das séries dos dados e a sua volatilidade, sendo necessário o emprego de grandes esforços na busca por boas previsões das séries financeiras [9]. Através das estimativas, busca-se entender o comportamento dos mercados mensurando os riscos, as perdas e os ganhos. Nesse contexto, desde seu surgimento nos anos de 1990, a Econofísica tem sido bastante relevante na literatura científica. Ela vem empregando os conceitos físicos na procura pelas respostas às questões econômicas [10].

Além disso, uma forma que pode ser usada para compreender os mercados financeiros é por meio da análise dos Mercados Financeiros Artificiais, utilizando modelagem baseada em agentes. Nesta abordagem, observa-se como ocorrem as interações individuais dos agentes e, a partir do perfil e da dinâmica das microestruturas, traçam-se os comportamentos macros do sistema, isto é, dos mercados. Esses modelos de mercados buscam reproduzir de forma simplificada as características dos mercados reais, através de

experimentos computacionais.

Na construção do mercado simulado, nesse trabalho, a ideia consiste em ter um sistema formado por agentes, que evolua reproduzindo o seu comportamento. Essa construção é baseada em um modelo de otimização inspirado no comportamento de um bando de pássaros ou enxames, onde há interação entre as partículas e troca de informação. Com isso, espera-se observar as informações do mercado financeiro artificial, que incluem dados dos agentes que o compõe e informações de compra e venda realizadas.

São usados algoritmos inteligentes para a adaptação dos agentes. Em particular, nessa pesquisa, usa-se um tipo de algoritmo de inteligência de enxame, o *Particle Swarm Optimization* (PSO) [11, 12]. Esse algoritmo possui fácil implementação e tem sido aplicado com bastante êxito a processos de otimização em geral [13], além da modelagem de mercados [14].

Com a aplicação do PSO procura-se a observação de alguns fatos estilizados. Os chamados fatos estilizados são propriedades estatísticas observadas empiricamente no mercado real, sendo esperado reproduzi-los no mercado simulado. Estes incluem características como assimetria, excesso de curtose e ausência de autocorrelação na série de retornos, por exemplo [15]. Para validação do modelo de mercado gerado no ambiente artificial, investiga-se se nele são cumpridos os fatos estilizados da economia.

Posteriormente, este trabalho também incluiu análises quanto à distribuição de riqueza entre os agentes financeiros. Essa, pode ser medida através do Coeficiente de Gini [16], uma medida estatística de dispersão. Para efeitos de comparação, as investigações e análises foram realizadas tanto no mercado real como no mercado artificial.

Por fim, foram utilizadas a função exponencial e a lei de potência para modelar as séries de retornos dos preços das ações e verificar que para eventos extremos é possível realizar ajustes aos dados. As leis de potência vêm sendo bastante estudadas e utilizadas na modelagem de séries de retornos e também no ajuste da distribuição de riqueza. Foram observadas, por exemplo, leis de potência na função densidade de probabilidade (fdp) das ações dos Estados Unidos [17], da bolsa de valores de São Paulo [18] e de Milão [19]. Bem como, na distribuição de riqueza entre as pessoas mais ricas dos Estados Unidos no ano de 1996 [20].

1.2 Objetivo

O objetivo geral dessa dissertação é utilizar simulação computacional com a utilização de agentes, cujos comportamentos são regidos por uma dinâmica que busca o maior acúmulo de riquezas, utilizando o algoritmo de PSO com a intenção de modelar mercados

financeiros e, através das séries temporais dos índices financeiros, avaliar a distribuição de riqueza entre os agentes econômicos.

Dessa maneira, a seguir estão listados os objetivos específicos almejados com esse trabalho:

- Desenvolver através do algoritmo PSO um mercado simulado formado por agentes;
- Observar o comportamento dos agentes financeiros nos variados mercados financeiros artificiais gerados;
- A partir do estudo dos conceitos da Econofísica e da Física Estatística, analisar a volatilidade entre os mercados;
- Estudar empiricamente a função densidade de probabilidade (fdp) das séries de retorno geradas no ambiente artificial e nos mercados reais;
- Medir a distribuição de riqueza entre os agentes;
- Observar se a fdp das séries de retorno é melhor descrita através da distribuição de lei de potência ou da distribuição exponencial.

1.3 Estrutura da Dissertação

A estrutura desta dissertação está dividida em 5 capítulos, conforme as descrições a seguir:

Capítulo 1 - Introdução: como já exposto, neste capítulo uma breve introdução sobre o problema abordado é apresentada, bem como os objetivos buscados com os estudos e pesquisas realizadas.

Capítulo 2 - Revisão de Literatura: neste capítulo é abordada a relação entre o mercado de ações e seus índices, a técnica inteligente do algoritmo empregado para geração do ambiente artificial, definições para análises iniciais que possam validar o modelo e abordagens da função exponencial e da lei de potência quanto à modelagem da função densidade de probabilidade das séries de retorno.

Capítulo 3 - Metodologia: neste capítulo é apresentado o conjunto de dados artificiais e reais com suas descrições, o algoritmo proposto para a simulação, os fundamentos e equações para análises estatísticas e a modelagem adotada para o ajuste das séries de retornos.

Capítulo 4 - Resultados: neste capítulo são mostrados os resultados e as discussões referentes à pesquisa aqui apresentada. Na primeira parte, encontram-se todos

os resultados para o mercado simulado, enquanto que na segunda parte podem ser vistos aqueles encontrados com os dados dos mercados reais.

Capítulo 5 - Conclusões: neste capítulo são apresentadas as conclusões dos experimentos e técnicas executadas, além da proposta dos trabalhos futuros.

2 Revisão de Literatura

2.1 Mercados de Ações

No mercado financeiro ocorrem as negociações de compra e venda de ativos, os quais são obtidos pelos agentes econômicos a partir da troca de unidades monetárias. Esses agentes podem ser pessoas físicas, empresas privadas ou públicas e instituições financeiras. Caracterizamos os ativos como sendo itens que possuem valor econômico e que permitem a realização de pagamentos com transações. Observada a liquidez dos ativos, que é a simplicidade de convertê-los em dinheiro, eles podem ser classificados em [21]:

- Líquidos;
- Plenamente líquidos;
- Ilíquidos.

De forma simplificada, vemos que um ativo é mais fácil de ser vendido à medida que o valor da sua liquidez aumenta. Os preços dos ativos são definidos a partir de estratégias de negociação daqueles que participam do mercado (agentes), visando a maximização dos lucros e/ou a minimização das perdas financeiras individuais, gerando um cenário competitivo entre eles.

Dentre os ativos que podem ser negociados, existem as ações de uma dada empresa. Essas, correspondem às menores parcelas do capital social que a empresa possui e não asseguram remuneração fixa aos seus investidores [22]. Ao adquirir uma ação e investir nela, os acionistas de uma empresa não têm como saber exatamente qual será o retorno adquirido e assumem o risco de obter sucesso ou fracasso nesse investimento.

Em oposição ao que acontecia antes do século XV [23], quando através da expansão comercial surgiram as primeiras bolsas de valores, com características de "feira" (pouco conforto e bastante gritaria), atualmente as negociações das ações das empresas podem acontecer na bolsa de valores no segmento do mercado de ações, que é um local físico ou eletrônico. Através das corretoras de valores, que são agentes credenciados na bolsa de valores, as transações de compra e venda entre os investidores acontecem.

No tocante ao Sistema Financeiro Internacional, pode-se dizer que o mercado de capitais tem se desenvolvido ao receber grande influência das bolsas de valores, propiciando

uma forte integração dos mercados financeiros como um todo. Quando falamos em mercado de capitais, estamos nos referindo ao ambiente que possibilita o processo de capitalização, interferindo direta e positivamente no desenvolvimento financeiro do país. Através dos índices das ações, é possível examinar por meio de um único valor, o processo evolutivo dos seus preços. Entre as principais formas de metodologias usadas para obtenção dos índices acionários, destacam-se as seguintes formas de ponderação [24]:

- Índice ponderado por liquidez - leva em conta a negociabilidade das ações no mercado. O peso da ação nesse índice diminui, à medida que sua liquidez torna-se menor;
- Índice ponderado por pesos iguais - é calculado por uma média aritmética onde a quantidade de ações é o divisor e, como o nome sugere, os preços delas possuem o mesmo peso;
- Índice de retorno ponderado por pesos iguais ou índice de retorno simples - é obtido pelo ajuste de seu valor anterior à média dos retornos de cada ação no grupo;
- Índice ponderado por capitalização de mercado - proporcionalmente à capitalização de mercado, esse índice é calculado através da soma dos preços de mercado das ações.

Nesse ambiente, os investidores visam ter o mínimo de risco envolvido nas negociações e para tanto, definem estratégias que permitam evitar efeitos desastrosos nos mercados, como altas e quedas abruptas de preços, antes de tomarem a decisão de investir ou não em algo [25]. Nesse sentido, recorrem aos métodos de previsão e de predição, explorando também os fatores que interferem nas ações dos agentes financeiros, mensurando como se dá a distribuição de riqueza entre eles.

Através da realização da boa previsibilidade da volatilidade de um ativo, torna-se possível identificar os momentos mais incertos no mercado. Esses momentos ocorrem quando há uma maior variação nos preços e, conseqüentemente, quando seus retornos possuem uma maior variância. Uma maneira de estimar a volatilidade, é através da obtenção do desvio padrão ou da média de diferenças na série, por exemplo. Observações empíricas mostram que à medida que a volatilidade de uma dada ação aumenta, maior o risco em investir nela [26, 27]

Assim, as séries temporais econômicas e financeiras compõem um dos mais populares sistemas complexos conhecidos. Isso se deve ao fato, de que as razões que influenciam os mercados não podem ser determinadas com exatidão, bem como a amplitude dessas influências e as respostas dadas pelo mercado a elas. Além disso, o mercado acionário sofre a ação direta dos agentes financeiros, que tentam entender o sistema e aplicar as melhores estratégias na busca por alcançar seus objetivos [8, 28].

2.2 Algoritmo Particle Swarm Optimization (PSO)

Os algoritmos de inteligência de enxame (*Swarm Intelligence*) são baseados em uma população de indivíduos com capacidade de interação com o meio ambiente e entre si [29]. Individualmente, eles utilizam regras simples de comportamento e como resultado dessas interações, sob a ótica global, surgem os comportamentos inteligentes e complexos de todo o enxame [29, 30]. Os exemplos mais conhecidos de inteligência de enxames são as colônias de cupins, de formigas e os bandos de aves.

Sendo baseado em população [31], foi desenvolvido o algoritmo *Particle Swarm Optimization* (PSO). Trata-se de um método de otimização proposto inicialmente por Kennedy e Eberhart em 1995 e inspirado no comportamento social de bandos de pássaros à procura de alimento ou de abrigo [32, 33]. Essa procura e a interação existente entre os pássaros enquanto voam é modelada como uma técnica de otimização [13].

Este algoritmo tem sido aplicado com bastante êxito a diversos problemas de busca e otimização, tanto em engenharia como em computação, e em processos de otimização em geral. É bastante simples, rápido e utiliza pouca memória. Na sua versão tradicional, observa-se que se trata de um procedimento iterativo, anárquico e coletivo, com foco na cooperação entre os indivíduos [34]. Entre as principais definições dadas pelos criadores do PSO em seu desenvolvimento, destacam-se:

- Partícula - possível solução para o problema de otimização. Movimenta-se pelo espaço de busca, geralmente multidimensional, levando em consideração a sua própria experiência e a experiência da sua vizinhança e tendem a imitar apenas as partículas que são melhores do que elas (estão melhor posicionadas);
- População - conjunto de partículas que possuem a mesma função objetivo;
- *Fitness* da partícula - indica o quanto uma partícula está bem posicionada em relação à qualidade da solução do problema no espaço de busca;
- Otimizar - encontrar a melhor solução possível para o problema, que pode ser de minimização ou de maximização de uma função *fitness*;
- *Pbest* - termo que indica o conhecimento da melhor posição já encontrada por cada partícula;
- *Gbest* - termo que indica a melhor posição já encontrada pelas partículas até o momento (o posicionamento global de todo o enxame).

Conforme ilustrado na Figura 1, na definição da nova posição de uma partícula, leva-se em consideração sua posição atual. A nova posição é obtida a partir de um vetor velocidade, resultante da combinação dos vetores que indicam o termo social, o termo cognitivo e o termo de inércia da partícula. O termo social indica a melhor posição da vizinhança, enquanto o termo cognitivo aponta a melhor posição da partícula.

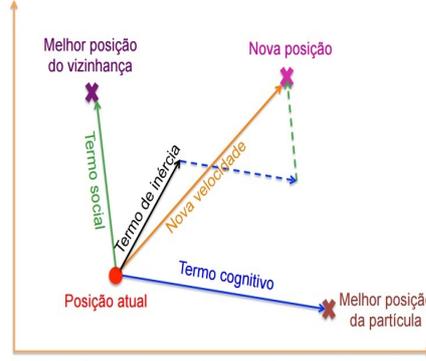


Figura 1 – Ilustração da movimentação de uma partícula no espaço 2D [1].

Assim, cada partícula tem um vetor de velocidade e um vetor de posição associado a ela. De acordo com a Equação 1, a velocidade da partícula é ajustada a cada iteração.

$$v_{i,k}(t+1) = w * v_{ik}(t) + c_1 * r_{1,k}(t)(p_{i,k}(t) - x_{i,k}(t)) + c_2 * r_{2,k}(t)(p_g(t) - x_{i,k}(t)) \quad (1)$$

Em que $v_{i,k}$ representa a associação entre a k -ésima dimensão do vetor velocidade e sua i -ésima partícula. Em cada dimensão $k \in N$, atualiza-se a velocidade separadamente. $x_{i,k}$ representa a posição na dimensão k da i -ésima partícula. w é conhecido como peso inercial e tem a capacidade de controlar a convergência das partículas, equilibrando a capacidade de procura do local e do global. Ele pode ser um valor fixo ou variar de acordo com a Equação 2, por exemplo. c_1 e c_2 são coeficientes de aceleração, cognitivo e social respectivamente, que influenciam a velocidade com que são movidas as partículas no espaço de busca. Enquanto c_1 controla a direção da melhor posição individual de cada partícula, c_2 o faz na direção da vizinhança ou da melhor posição global. r_1 e r_2 são números aleatórios gerados de uma distribuição uniforme em $[0,1]$ e produzem uma influência estocástica sobre o algoritmo. Diz-se que o termo $c_1 * r_{1,k}(t)(p_{i,k}(t) - x_{i,k}(t))$ está associado a cognição e considera as experiências passadas das partículas. Por outro lado, o termo $c_2 * r_{2,k}(t)(p_g(t) - x_{i,k}(t))$ está associado ao social e a partícula leva em conta a melhor solução que há ao seu redor [35, 36].

$$w = w_{min} + t \frac{w_{máx} - w_{min}}{t_{máx}} \quad (2)$$

onde t representa o número da iteração, $t_{m\acute{a}x}$ é a quantidade máxima de iterações que o algoritmo irá realizar, $w_{m\acute{a}x}$ o peso máximo e w_{min} o peso mínimo.

A partir do novo vetor velocidade obtido, a posição da partícula é atualizada através da Equação 3:

$$x_i(t + 1) = x_i(t) + v_i(t) \quad (3)$$

Assim, a heurística que o algoritmo segue é a de que a melhor solução para o problema estará próxima à melhor posição global que uma partícula encontrou e com isso, as demais tenderão a buscar a melhor solução em torno do *gbest*. Há basicamente duas formas de organização da população no PSO tradicional, dadas pela topologia global e pela topologia local, conforme exemplificado na Figura 2.

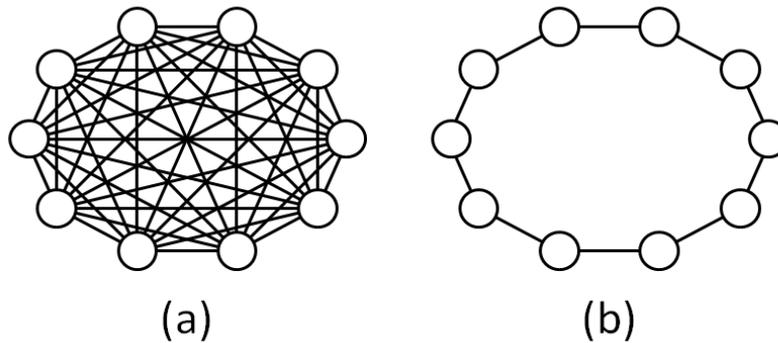


Figura 2 – Topologias do algoritmo PSO. Topologia Global (a) e uma possível Topologia Local (b) [2]

Os círculos representam as partículas que compõem o enxame. Enquanto na topologia global todas elas detêm informação das demais, na topologia local, as partículas só conhecem as informações de suas vizinhas mais próximas (a da direita e a da esquerda no exemplo da Figura 1(b)) [2]. Escolher a topologia adequada ajuda a evitar que o algoritmo tenha uma convergência prematura em mínimos/máximos locais [37].

Abaixo, o Algoritmo 1 descreve o pseudo-código da implementação do algoritmo PSO tradicional [38].

Encontram-se na literatura trabalhos em que o algoritmo PSO é usado na modelagem e previsão dos índices acionários [14, 39, 40]. Em [41], para a previsão dos preços das ações no mercado S&P500, é feito um estudo comparativo entre cinco diferentes algoritmos. Observam-se os resultados encontrados com a implementação da estratégia *Weight Reward Strategy* (WRS) para definição de uma técnica de negociação de ações, em [42]. A estratégia WRS é resultante da combinação entre diversas regras de negociação de *Moving Average* (MA) e *Trading Range Break-out* (TRB). Na análise MA, são consideradas duas janelas

Algorithm 1 Pseudo-código do Algoritmo PSO

Entrada: Parâmetros c_1 , c_2 , q_{iter} , i (quantidade de partículas) e $f(x)$

2: **Saída:** g_{best}

/* Inicializar as posições e velocidades iniciais das i partículas aleatoriamente */

4: **while** $t < t_{máx}$ (critério: pseudotempo ou contador de iteração)

for (laço para todas as i partículas e todas as k dimensões)

6: Gerar as novas velocidades $v_{i,k}(t + 1)$ usando a Equação 1

 Gerar as novas posições $x_i(t + 1)$ usando a Equação 3

8: Avaliar a função *fitness* de acordo com as novas posições $x_i(t + 1)$

 Encontrar a melhor posição atual para cada partícula x_i

10: **end for**

 Encontrar o melhor global atual (g_{best})

12: **end while**

 Exibir os resultados finais: x_i e g_{best}

móveis de dias e as médias dos preços das ações em cada dia de negociação são calculadas, enquanto são emitidos sinais de compra e venda das ações. Já em TRB, as normas que regem as negociações são mais simples. Os sinais de compra ou venda são emitidos após a observação do preço da ação no dia atual e a comparação com o preço acionário dos últimos dias. O PSO foi usado para otimizar a estratégia WRS, maximizando o lucro líquido gerado por essa estratégia.

Em [43] e [44], através da combinação do algoritmo baseado em enxames de partículas e das redes neurais artificiais, é descrita uma estratégia para tomada de decisão de investimento. Nessa estratégia, a otimização por enxames de partículas é empregada no treinamento das redes neurais que, por sua vez, são usadas na análise dos retornos históricos das ações e na estimativa de possíveis lucros futuros.

2.3 Fatos Estilizados

Os fatos estilizados econométricos são propriedades estatísticas universais observadas nas séries temporais financeiras reais [15] e podem ser usadas para ratificar os mercados financeiros artificiais. Através do estudo dessas séries financeiras é possível que sejam obtidas informações relevantes que contribuam para uma boa tomada de decisão daqueles que participam dos mercados financeiros.

Nesse sentido, normalmente as séries de retornos dos ativos negociados são empregadas pela comunidade científica da Econofísica [45, 46, 47, 48, 27, 49], pois as mesmas não sofrem influência dos efeitos de escala e isso simplifica a busca pelas características universais. Uma forma bastante comum utilizada para sua obtenção é apresentada na

Equação 4:

$$r(t) = \log p_t - \log p_{t-\Delta t} \quad (4)$$

em que $r(t)$ representa o retorno do preço do ativo, p_t é o valor de fechamento do índice no instante t (medido em dias de pregão) e Δt é o intervalo de amostragem, considerado nesse trabalho igual a um dia.

Nessa pesquisa, definimos o retorno diário normalizado como:

$$R_n(t) = \frac{|\log p_t - \log p_{t-\Delta t}|}{\sigma} \quad (5)$$

Em que σ é o desvio padrão da série de retornos.

Entre os principais fatos estilizados observados nos diferentes tipos de mercados financeiros e mencionados por Cont em seu trabalho publicado em 2001 [15], destacam-se a assimetria, o excesso de curtose, a ausência de autocorrelação, a não estacionariedade e a heterocedasticidade.

2.3.1 Assimetria e Curtose

Considerando uma série temporal $\{Z_t | t = 1, 2, \dots, N\}$ e assumindo que ela possui média μ e desvio padrão σ , então dizemos que o r -ésimo momento central de $\{Z_t | t = 1, 2, \dots, N\}$ com $\mu_2 = \sigma^2$, é $\mu_r = \langle (z - \mu)^r \rangle$, em que z corresponde a cada observação da amostra considerada. Com isso, podemos definir os coeficientes de assimetria e curtose, respectivamente, como [50]:

$$\alpha_3 = \frac{\mu_3}{\sigma^3} = \frac{\langle (z - \mu)^3 \rangle}{\langle (z - \mu)^2 \rangle^{3/2}} \quad (6)$$

e

$$\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{\langle (z - \mu)^4 \rangle}{\langle (z - \mu)^2 \rangle^2} \quad (7)$$

A assimetria indica as distribuições assimétricas, ou seja, aquelas que apresentam uma cauda mais longa para a esquerda (assimetria negativa) ou para a direita (assimetria positiva). Por outro lado, a curtose informa sobre o grau de achatamento das distribuições de frequências. Através dessa medida de dispersão, busca-se caracterizar o nível de concentração dos valores ao redor de sua média. À medida que esses valores crescem, maior se torna o valor da curtose e a curva associada a essa distribuição tem sua parte central mais pontiaguda, sendo essa curva chamada de leptocúrtica.

Em geral, a assimetria e a curtose são medidas em relação a uma distribuição Gaussiana (Normal), que apresenta $\alpha_3 = 0$ e $\alpha_4 = 3$. Caso a função de distribuição tenha o

mesmo achatamento da distribuição normal, ela é chamada de função mesocúrtica. Porém, se ela for mais achatada, será platicúrtica.

Observações empíricas das séries de retornos de variados mercados de ativos apontaram para funções de distribuição leptocúrticas [15, 27]. Entretanto, para valores progressivamente maiores de Δt , pelo Teorema Central do Limite, a distribuição dos retornos converge para uma Gaussiana. Esse fato é conhecido como Gaussianidade Agregativa [15].

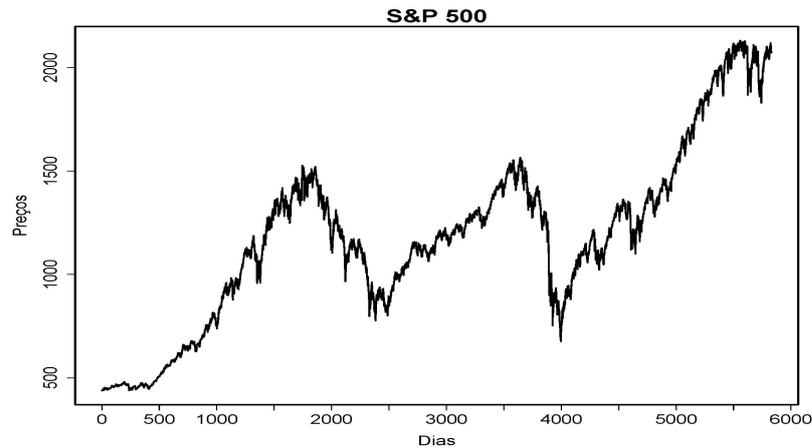
As caudas "pesadas" avaliadas nas séries dos mercados financeiros têm sido motivo de intensas investigações e tentativas de modelagem das distribuições, onde distribuições como Lévy-Estável [51], Lévy truncada [51, 27] e t-Student [52], têm sido aplicadas na tentativa de se ajustarem aos dados econômicos. Particularmente, as distribuições de Lévy têm sido aplicadas à modelagem de mercados por sua boa aderência, quanto à assimetria e às caudas longas observadas na função de distribuição desses dados [53, 54].

2.3.2 Estacionariedade

Define-se como processo estocástico todo fenômeno ou sistema que evolui no decorrer do tempo de forma imprevisível e pode ser descrito por uma estrutura de probabilidade [6]. Desta maneira, são exemplos de processos estocásticos:

- Eletrocardiograma;
- Variação dia a dia do índice Ibovespa;
- Produto Interno Bruto (PIB) de um país;
- Variação mensal no tamanho do estoque de uma certa empresa;
- Quantidade de motos que trafegam em um determinado cruzamento;
- Comportamento das partículas de um gás suscetíveis a choques aleatórios; etc.

Abaixo, encontram-se algumas ilustrações desses exemplos. Na Figura 3 estão ilustrados os 5746 valores diários dos preços de fechamento do S&P 500 no período de 27 de abril de 1993 a 17 de junho de 2016.



Fonte: próprio autor

Figura 3 – Índices diários dos preços de fechamento do S&P 500 entre os períodos de 27 de abril de 1993 e 17 de junho de 2016.

Na Figura 4, está representado o Produto Interno Bruto (PIB) do Reino Unido, do Canadá, da África do Sul e do mundo no período de 1960 a 2015.

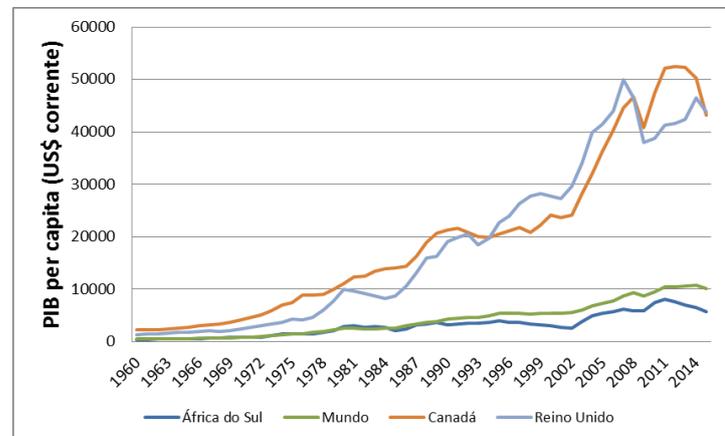


Figura 4 – Ilustração do PIB do Reino Unido, do Canadá, da África do Sul e do mundo de 1960 a 2015 [3].

Um processo estocástico é dito ser estacionário quando ele apresenta equilíbrio estatístico, ou seja, quando não possui tendência e seus dados oscilam em torno de uma média e uma variância constantes. Em contrapartida, caso suas propriedades se alterem ao longo do tempo, como as séries temporais com tendência exponencial ou com tendência linear e as séries sazonais, ele será não-estacionário [7, 6].

Pode-se definir duas formas de estacionariedade:

- Estacionariedade forte: todos os momentos estatísticos da fdp do processo são constantes;
- Estacionariedade fraca: os dois primeiros momentos da fdp do processo são constantes.

De forma prática, considera-se a estacionariedade fraca como critério para definição de uma série estacionária, mesmo porque o único processo conhecido que é estacionário forte é o ruído branco (uma sequência de erros ou choques aleatórios, cujas variáveis aleatórias seguem uma distribuição normal de média zero, variância constante e covariâncias nulas) [55]. Dessa maneira, a estacionariedade de uma série Z_t estará assegurada ao serem satisfeitas as seguintes propriedades [56]:

$$\text{Média : } E(Z_t) = \mu \quad (8)$$

$$\text{Variância : } \text{var}(Z_t) = E(Z_t - \mu)^2 = \sigma^2 \quad (9)$$

$$\text{Covariância : } \gamma_k = E[(Z_t - \mu)(Z_{t+k} - \mu)] \quad (10)$$

Em que k corresponde ao atraso (defasagem ou *lag*). De fato, há um grande empenho por parte dos pesquisadores em utilizar séries temporais estacionárias, mas boa parte das séries econômicas já observadas são não estacionárias [56].

2.3.3 Autocorrelação Linear

A função de autocorrelação é um parâmetro numérico que mensura o grau de dependência na relação entre as observações atuais e as anteriores de uma série temporal em um dado instante de tempo. Através dela pode ser analisado o nível de irregularidade proveniente de um sinal. A função de autocorrelação é dada pela razão entre a autocovariância e a variância de um conjunto de dados [6, 56], conforme apresentado abaixo:

$$r_k = \hat{\rho}_k = \frac{c_k}{c_0} \quad (11)$$

Na qual c_k corresponde à função de autocovariância e é dada por:

$$c_k = \hat{\gamma}_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \quad k = 0, 1, \dots, m \quad (12)$$

onde N é a quantidade de observações, z_t indica uma observação da série e \bar{z} é a média do processo. A variância c_0 é definida como:

$$c_0 = \frac{\sum (z_t - \bar{z})^2}{N} \quad (13)$$

Nas séries financeiras é interessante estudar e avaliar se há autocorrelações, pois em caso de existência seria possível desenvolver, a partir dos dados passados, as estratégias mais adequadas para obtenção de melhores resultados nos mercados. No entanto, esse fato iria se contrapor à "hipótese de mercado eficiente", onde é esperado das funções de autocorrelação das séries de retorno valores muito próximos a zero, caracterizando ausência de autocorrelação linear [15].

A ausência de autocorrelação nas séries de retornos vem sendo investigada [54] desde que um de seus precursores, Fama, a percebeu ao estudar as séries de preços de mercados [57, 58]. Sob a ótica da estatística, se uma série for independente ou seja, se não existir nenhuma correlação entre as observações em diferentes tempos da série, sua correlação será nula [15].

2.3.4 Heterocedasticidade

No campo das finanças, a variância é bastante empregada para a obtenção da volatilidade das séries de retornos. A volatilidade caracteriza a dispersão e a amplitude com que ocorrem as alterações de preços nas séries. Sendo assim, através dela pode-se observar como variam os retornos em relação à média do processo.

Além disso, a volatilidade também tem sido amplamente utilizada como medida de risco. À medida que a volatilidade aumenta, cresce também o risco da exposição do agente financeiro que optou pelo investimento nessa série financeira.

Portanto, se a variância é uma função que depende do tempo, a volatilidade também dependerá e o mercado é dito ser heterocedástico. A heterocedasticidade caracteriza a grande dispersão dos dados, fenômeno estatístico que apresenta, para todas as observações, variâncias distintas em torno de uma média. Nas séries financeiras de retornos, observam-se agrupamentos de volatilidades.

2.4 Medidas de Desigualdade

Existem várias formas que podem ser usadas para medir a desigualdade. Entre elas, destacam-se o coeficiente de Gini e as medidas que foram propostas por Theil.

O Theil-T (primeiro índice de Theil) e o Theil-L (segundo índice de Theil) fazem parte de uma família de índices que obedecem à fórmula geral abaixo [4]:

$$D(c) = \frac{1}{nc(c-1)} \sum_{i=1}^n \left(\left(\frac{r_i}{\mu} \right)^c - 1 \right), \quad \forall c \neq 0, 1 \quad (14)$$

em que n é a quantidade de indivíduos da população, r_i é a renda do indivíduo i e μ é a média de rendimentos da população.

Se $c \rightarrow 0$, então $D(0)$ dado pela Equação 15 representa o Theil-L.

$$D(0) = L = Ln(\mu) - Ln(g) \quad (15)$$

Onde g corresponde a média geométrica, dada por:

$$g = \sqrt[n]{\prod_{i=1}^n r_i} \quad (16)$$

Se $c \rightarrow 1$, então $D(1)$ dado pela Equação 17 representa o Theil-T.

$$D(1) = T = \frac{1}{n} \sum_{i=1}^n \frac{r_i}{\mu} Ln \left(\frac{r_i}{\mu} \right) \quad (17)$$

Mesmo havendo vários índices de desigualdade, como já mencionado, o que vem sendo amplamente utilizado é o coeficiente de Gini. Essa medida foi desenvolvida pelo primeiro presidente do *Istituto Centrale di Statistica*, Corrado Gini [59].

O coeficiente de Gini não é restrito apenas a economia, mas também possui aplicação em outras áreas, como a saúde [60], educação [61], agricultura [62] e ecologia [63], por exemplo.

Essa métrica varia de 0 a 1, de forma que 0 representa a igualdade perfeita e 1 a desigualdade total. Pode-se tomar como exemplo, a distribuição de rendimentos em uma população qualquer. Se a medida do coeficiente fosse igual a 1, implicaria que somente uma pessoa está concentrando todo o rendimento, mas se fosse igual a 0, então toda a população estaria recebendo o mesmo salário.

2.5 Distribuição Exponencial e de Lei de Potência

Encontram-se na literatura diversas abordagens adotadas na investigação das séries de retorno, incluindo diferentes formas de ajustar uma função a esses dados. Fama [57], Mandelbrot [64] e Mantegna e Stanley [65] utilizaram uma distribuição de Lévy. Cont *et al.* [66] adotaram uma distribuição Exponencial truncada. Com uma distribuição q-Gaussiana, Queiros *et al.* [67] fazem um bom ajuste das séries.

No entanto, muitos econofísicos têm utilizado as leis de potência para realizar o ajuste das séries de retornos dos dados [49, 67, 68]. Através dessa função podem ser obtidas informações referentes ao comportamento das bolsas de valores, as oscilações dos seus

índices e a volatilidade dos mercados, por exemplo. Fisicamente, conforme a Equação 18, pode-se descrever uma lei de potência como a relação entre duas variáveis que representam a série de retornos (r) e o valor do ajuste ($f(r)$).

$$f(r) = ar^k \quad (18)$$

onde a e k são a constante de proporcionalidade e o expoente da lei, respectivamente.

Com relação à distribuição exponencial, cuja Equação 19 descreve seu comportamento, muitos ajustes e aproximações também têm sido obtidos para a série de retornos [69, 70, 71], visando caracterizar a volatilidade dos dados.

$$f(r) = a.e^{br} \quad (19)$$

Kelly e Jiang utilizaram as propriedades de lei de potência na investigação do comportamento da cauda do retorno dos dados da S&P500 [72]. Com a série de retornos normalizada pelo desvio padrão, Lima [73] realizou uma comparação entre os ajustes da lei de potência e da função exponencial em alguns mercados. Neto *et al.* [26] realiza o ajuste usando quatro tipos diferentes de algoritmo (*Least Squares*, *Trust Region*, *Levenberg-Marquardt* e *Maximum Likelihood Estimation*) tanto para com a função exponencial, como para a lei de potência. Gresnigt *et al* [74] utilizam os retornos no desenvolvimento de uma modelagem capaz de prever a probabilidade de ocorrer eventos extremos na série de dados. Eles levam em consideração a capacidade da função exponencial em extrair informações, quanto à influência dos eventos anteriores sobre os novos.

Quanto à análise de distribuição de riqueza, Levy e Solomon [20] utilizaram leis de potência para estudar a distribuição de riqueza entre algumas pessoas ricas dos Estados Unidos. Recentemente, Brzezinski [75] investigou essa distribuição não somente entre os americanos, mas entre os russos e os chineses também. Dragulescu *et al* [76] analisa o ajuste realizado pela função exponencial a toda a série de dados da distribuição de renda individual dos estados dos EUA e a utilização da lei de potência somente na cauda da distribuição desses dados.

Com o auxílio das leis de potência, é possível observar de que maneira os eventos extremos podem afetar na dinâmica de sistemas complexos, tais como o mercado de ações, as redes sociais e os terremotos [77, 49]. Nesse sentido, é necessário levar em conta a probabilidade desses eventos ocorrerem [47]. Com isso, podemos relacionar o evento à sua intensidade: quanto menor for a probabilidade de um evento desse tipo ocorrer, maior é a intensidade do mesmo [73].

3 Metodologia

3.1 Base de Dados

Os dados dessa pesquisa são compostos pelas séries temporais financeiras do mercado simulado e do mercado real. As séries temporais do mercado simulado são referentes aos preços das ações negociadas nesse ambiente, que é inspirado no modelo tipo *Random Walk* e visa reproduzir o comportamento de mercados. Já as séries reais, correspondentes aos preços de fechamento de diferentes índices financeiros, foram coletadas do site <<http://finance.yahoo.com/>> . Foram escolhidos índices de mercados de países desenvolvidos e em desenvolvimento, são eles: Nasdaq (Estados Unidos da América), CAC 40 (França), Ibovespa (Brasil) e IPC (México). Os dois últimos correspondem a índices de países em desenvolvimento, cujas economias são mais instáveis em relação aos países dos outros dois índices mencionados. Toda a estatística e análise gráfica foi realizada com o auxílio dos softwares R *Core Team* 2015 e Matlab (R2016a). A implementação do mercado simulado foi realizada no Microsoft Visual 2012[®] e em Linux.

3.1.1 Séries Artificiais

As séries financeiras artificiais referem-se aos preços das ações negociadas no mercado simulado, correspondendo a 20.000 dias de negociação, e foi observada a quantidade de riqueza dos agentes financeiros na última rodada de negociações. Além disso, temos também informações referentes aos limiares de venda e de compra do melhor e do pior agente financeiro em cada rodada de negociação. Quando nos referirmos ao pior agente, estaremos falando daquele que detêm menos riqueza. De forma análoga, quando nos referirmos ao melhor agente, estaremos mencionando aquele que possui mais riqueza.

3.1.2 Série do Índice Ibovespa - Brasil

O Índice Bovespa (Ibovespa) é o principal parâmetro para avaliação do desempenho da bolsa de valores do mercado acionário brasileiro, a Bolsa de Valores de São Paulo (BOVESPA). A qual em maio de 2008, passou a se chamar BM&FBovespa por conta de sua integração à bolsa de derivativos BM&F. Desde sua implantação, o Ibovespa não passou por muitas mudanças em sua metodologia, conservando a base de sua série histórica [78, 79].

Visando fazer uma boa avaliação da evolução das ações na Bovespa, o Ibovespa é ponderado por liquidez. Nesse sentido, à medida que aumenta a liquidez da ação no mercado, aumenta também o seu respectivo peso relacionado ao índice. Isso significa, que conforme as negociações variam no mercado, as ações que compõem a carteira teórica podem ganhar relevância [24].

Obtivemos os valores de fechamento diário do Ibovespa para o período de 27 de abril de 1993 a 01 de agosto de 2016, totalizando 5777 observações. Na Figura 5 observamos o comportamento do Ibovespa durante o período mencionado. As estatísticas descritivas a respeito da série podem ser visualizadas na Tabela 1.

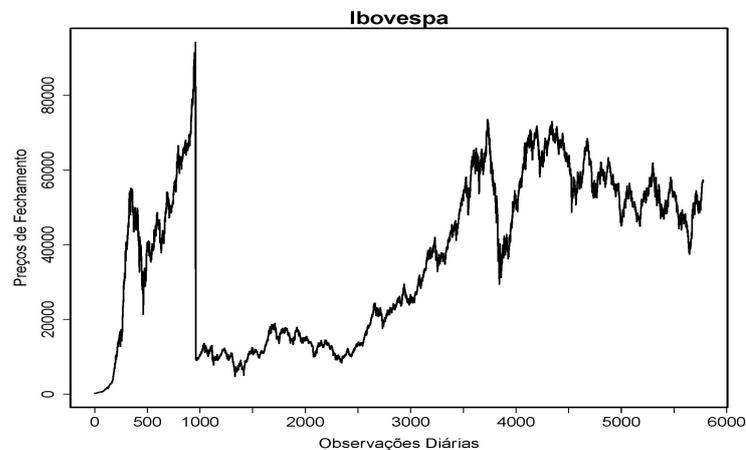


Figura 5 – Série de preços de fechamento do Ibovespa.

Tabela 1 – Estatísticas da série Ibovespa.

Estatística	Valor
Mínimo	237
Máximo	94200
Média	36700
Mediana	39500
Desvio-Padrão	21369
Amplitude Total	93953

3.1.3 Série do Índice CAC 40 - França

A *Cotation Assistée en Continu* 40 (CAC 40) corresponde ao índice da bolsa composta pelas 40 empresas mais importantes da França. Esse índice é obtido através da média ponderada de capitalização dos preços que integram as ações dessas empresas. A divulgação dos valores do índice é realizada pela *Euronext Paris* e supervisionado pelo

Conseil Scientifique des Indices, o qual atua de forma independente, tendo como objetivo adaptar o índice de modo que capte as mudanças no mercado [80, 81].

Usamos os registros diários dos valores de fechamento das ações negociadas durante o período de 04 de janeiro de 2000 a 30 de dezembro de 2016, totalizando 4276 pontos. A Figura 6 mostra o gráfico da série das ações do CAC 40 no período adotado.

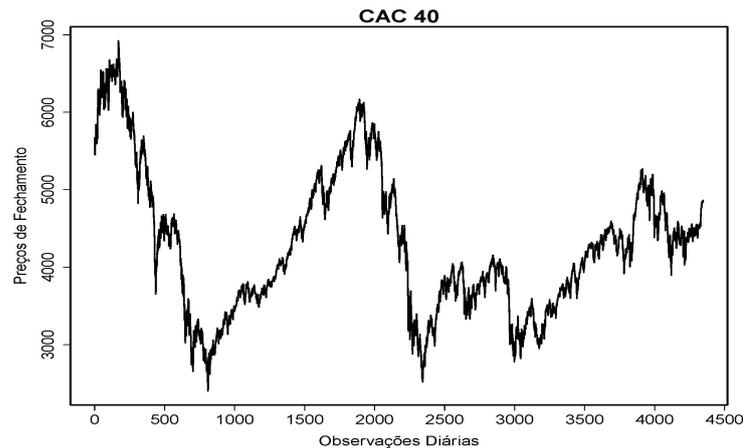


Figura 6 – Série de preços de fechamento do CAC 40.

A Tabela 2 exhibe os valores referentes a estatísticas básicas da série dos preços de fechamento das ações do CAC 40.

Tabela 2 – Estatísticas da série CAC 40.

Estatística	Valor
Mínimo	2400
Máximo	6920
Média	4280
Mediana	4190
Desvio-Padrão	911.9
Amplitude Total	4519

3.1.4 Série do Índice IPC - México

O *Índice de Precios y Cotizaciones* (IPC) é um índice ponderado, o qual monitora o comportamento das empresas mais influentes da *Mexican Stock Exchange* (Bolsa Mexicana de Valores - BMV), que na América Latina é a segunda maior bolsa de valores por capitalização. As negociações ocorrem eletronicamente, através do sistema BMV-SENTRA, e entre os principais derivativos negociados estão os títulos corporativos e governamentais, as debêntures (títulos de crédito) e as ações [82].

Os dados da série utilizada nesse trabalho para o período compreendido entre 03 de janeiro de 2000 e 30 de dezembro de 2016 foram calculados com base nos fechamentos diários do índice IPC, totalizando 4277 pontos, em que a Figura 7 exhibe o gráfico da série diária, enquanto a Tabela 3 mostra suas estatísticas básicas.

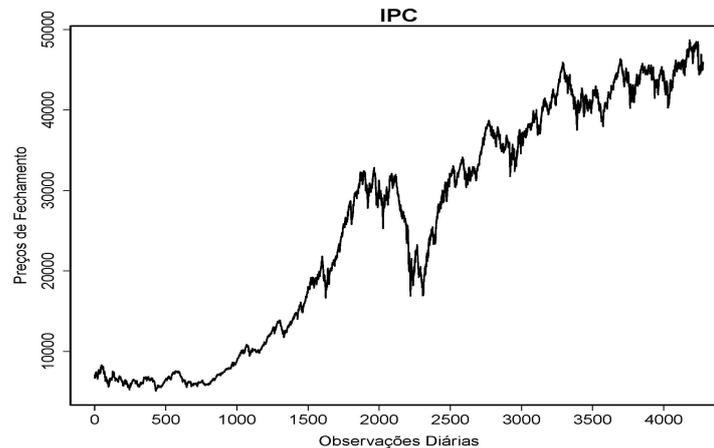


Figura 7 – Série temporais de preços de fechamento do IPC.

Tabela 3 – Estatísticas da série IPC.

Estatística	Valor
Mínimo	5080
Máximo	48700
Média	25700
Mediana	28500
Desvio-Padrão	14488
Amplitude Total	43613

3.1.5 Série do Índice Nasdaq - Estados Unidos da América

Na *National Association of Securities Dealers Automatic Quotation System* (Nasdaq) ocorrem as cotações eletrônicas do mercado norte-americano, o qual pertencem cerca de 4000 ações ordinárias [83]. Essas ações possibilitam a participação dos investidores nas tomadas de decisão das empresas, além da concessão dos lucros referentes a elas [84]. O índice Nasdaq é um dos maiores do mundo em volume de transações, seu cálculo é obtido pela soma dos preços de mercado das ações, consistindo num índice de ponderação por capitalização.

Nossos dados referentes ao índice Nasdaq, consistem em observações diárias abrangendo o período de 08 de julho de 2003 até 17 de junho de 2016 (um total de 3261

observações diárias), onde a Figura 8 e a Tabela 4 apresentam, respectivamente, o gráfico da série e a estatística descritiva dos dados.

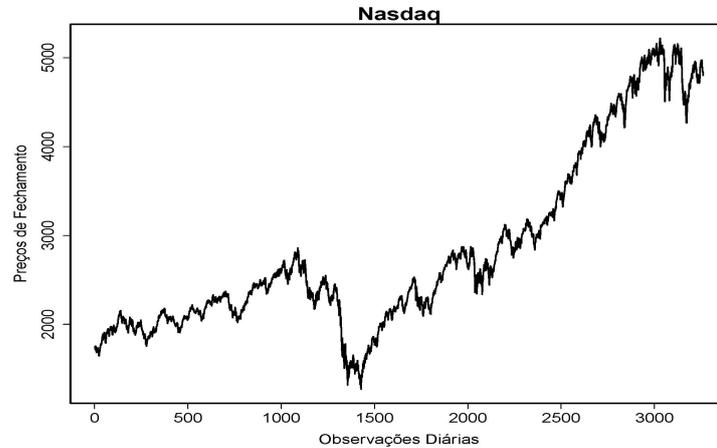


Figura 8 – Série temporais de preços de fechamento da Nasdaq.

Tabela 4 – Estatísticas da série Nasdaq.

Estatística	Valor
Mínimo	1270
Máximo	5270
Média	2850
Mediana	2480
Desvio-Padrão	1013
Amplitude Total	3950

3.2 Ambiente Artificial Proposto para Simulação do Mercado Acionário

Para analisar as particularidades do mercado de ações, foi criado um ambiente artificial visando a simulação da dinâmica dos mercados. Nesse ambiente, as partículas são compreendidas como os agentes financeiros que compõem os mercados. Sendo assim, ao adotarmos o algoritmo PSO para fazer a maximização da riqueza no nosso sistema financeiro, assumimos que as informações de posição das partículas do algoritmo podem representar os limiares de venda e de compra dos agentes financeiros e, conseqüentemente, ao interagirem entre si em busca da melhor solução para a otimização, realizam trocas de informações. As informações compreendem aos preços das ações, as negociações realizadas e a riqueza de cada agente econômico no mercado. Vale ressaltar, que os preços históricos da

ações influenciam na definição de cada novo preço. O Algoritmo 2 descreve o pseudo-código do modelo proposto.

O algoritmo foi implementado em linguagem C e na primeira rodada de simulações, foi compilado no Microsoft Visual Studio 2012[®]. Devido às posteriores alterações impostas ao código, o tempo de execução aumentou e a segunda rodada de simulações ocorreu com o desenvolvimento e a compilação das simulações em um servidor Linux com grande desempenho computacional.

Algorithm 2 Pseudo-código do Algoritmo PSO para o Mercado Simulado

```

/*Inicialização do espaço de busca e das partículas*/
2: for  $i = 1$  até  $N_{particulas}$  do
     $\delta[i] \leftarrow \mathcal{U}(0, 1)$ 
4:    $\lambda[i] \leftarrow \mathcal{U}(0, 1)$ 
    end for
6: /*Simulações do PSO (s)*/
    for  $s = 1$  até  $N_{pso}$  do
8:     if  $s = 1$  then
         $preço[s] \leftarrow 1$ 
10:    else
         $preço[s] \leftarrow preço[s - 1] + \mathcal{N}(\mu, \sigma)$ 
12:    end if
        /*Para as negociações (l)*/
14:    for  $l = 1$  até  $N_{neg}$  do
        Sorteio dos agentes vendedores e dos agentes compradores
16:        Atualização dos preços de venda e de compra de cada agente ( $i$ )
             $Preço\_Venda[i] = preço[s] + \text{limiar de venda } (\delta[i])$ 
18:             $Preço\_Compra[i] = preço[s] + \text{limiar de compra } (\lambda[i])$ 
        end for
20:    if  $Preço\_Venda[i] \leq Preço\_Compra[i]$  then
        Agente comprador (ac) compra a ação negociada e transfere o capital para o
        agente vendedor (av)
22:        Agente vendedor recebe capital do comprador e transfere a ação para ele
    end if
24:     $preço\_mercado[s] \leftarrow Preço\_Médio\_Negociado()$ 
        /* Usando o PSO para maximizar a riqueza dos agentes*/
26:    for  $i = 1$  até  $N_{particulas}$  do
         $fitness[i] \leftarrow capital[i] + preço\_mercado[s] * ações[i]$ 
28:    end for
        for  $i = 1$  até  $N_{particulas}$  do
30:            A partir de novos números aleatórios para os limiares, atualiza as coordenadas
            das partículas  $x_i = (\delta_i, \lambda_i)$ 
        end for
32: end for
return Informações da melhor e da pior partícula para cada  $s$ -iteração

```

A quantidade de partículas que compõe o enxame foi denotada por $N_{particulas}$. N_{ps0} representa a quantidade de iterações do algoritmo PSO, onde em cada uma delas ocorrem n rodadas de negociação (N_{neg}). O preço da ação negociada em cada simulação é representado por $preço[s]$.

A simulação inicia com a definição aleatória de uma população composta por 1000 partículas e seu lançamento no espaço de busca. Conforme já mencionado, os agentes financeiros do mercado simulado são representados por essas partículas. Dessa população, dois grupos para os agentes são gerados ao acaso e com igual probabilidade em sua definição, ou seja, são sorteados 50% de agentes com intenção de venda e 50% de agentes com intenção de compra. A princípio, cada um deles possui 500 ações e capital inicial igual a 1.

Para cada partícula i , são sorteados aleatoriamente de uma distribuição uniforme entre zero e um, dois conjuntos de parâmetros que representam os limiares de venda e de compra dos agentes financeiros, $\delta[i]$ e $\lambda[i]$, respectivamente. Esses limiares podem variar de 0.001 a 0.999 e são eles que definem a dimensão do espaço de busca, nesse caso, trata-se de um problema bidimensional. Desse modo, cada partícula é representada pelo par $x_i = (\delta_i, \lambda_i)$.

Escolhe-se de forma aleatória um agente do grupo dos vendedores e um agente do grupo dos compradores. Eles irão interagir entre si e decidir, através da análise dos seus preços de venda e de compra, se haverá negociação entre eles ou não. Os limiares de venda e de compra são observados e há negociação se o limiar de venda for menor ou igual ao limiar de compra. Assim, a iniciativa para a negociação é dada pelo agente comprador e é ele quem decide se haverá compra da ação. Esta ocorrerá, se o preço máximo ofertado por quem quer vender é igual ao preço predefinido por quem quer comprar.

Observados os limiares, se houver negociação, é transferido o capital de um agente para o outro e, também, a ação negociada. Em seguida, os dois agentes voltam aos seus respectivos grupos e essa interação entre as partículas está concluída.

Em cada simulação do PSO, o preço de mercado da ação é um número positivo gerado e atualizado por uma distribuição gaussiana $N(\mu, \sigma)$ e pelo preço da ação na iteração anterior do PSO. Os preços de venda e de compra são dados, respectivamente, pelo preço de mercado mais o limiar de venda e o preço de mercado mais o limiar de compra.

Nas primeiras simulações, o preço foi atualizado por uma distribuição gaussiana $N(\mu, \sigma)$, com média (μ) igual a zero e desvio-padrão (σ) assumindo os valores 0.01, 0.05, 0.10 ou 0.30, e produzia uma perturbação estocástica nos preços das ações do sistema artificial.

Posteriormente, na segunda rodada impomos que o capital inicial de cada agente fosse de 1000 e passou-se a atualizar o preço das novas simulações por uma distribuição gaussiana $N(\mu, \sigma)$, em que foi introduzida uma constante que gera uma tendência determinística, também chamada de *drift*. Com essa constante pode-se impor tendência no mercado simulado. De forma geral, as equações para os preços das ações, quanto à determinação de tendências, foram utilizadas da seguinte maneira:

- Sem tendência:

$$preço[s + 1] = preço[s] + N(\mu, \sigma) \quad (20)$$

- Com tendência:

$$preço[s + 1] = preço[s] + \theta + N(\mu, \sigma) \quad (21)$$

em que θ é uma constante. Se $\theta > 0$, haverá tendência de alta. Se $\theta < 0$, haverá tendência de baixa. Vale ressaltar, que adicionar uma constante a uma distribuição resultará em uma outra distribuição, isto é,

$$N(\mu, \sigma) + constante = N(\mu + constante, \sigma) \quad (22)$$

Assim, θ assumiu os valores ± 0.01 , ± 0.05 ou ± 0.10 . As variações entre θ e σ servem para avaliarmos o comportamento do mercado simulado em diferentes situações.

Em cada simulação do algoritmo, são realizadas 20000 iterações do PSO e em cada uma delas 1000 rodadas de negociação. Cada iteração do PSO corresponde a um dia de negociação, o que equivale a dizer que as 1000 rodadas de compra e venda dos títulos entre os agentes financeiros acontecem de forma simultânea.

No caso dos mercados reais, os preços são pré-definidos e atualizados de acordo com as séries históricas diárias de preços para cada índice utilizado. Dessa maneira, a quantidade de simulações do PSO equivale a quantidade de observações diárias de cada índice do mercado real. Foram utilizadas 1000 partículas representando os agentes financeiros participantes desses mercados e houve 1000 rodadas de negociação em cada iteração do PSO.

Ao final de cada iteração, são atualizados os limiares de venda e de compra e analisa-se a quantidade de riqueza que cada agente possui. Ela é dada pela função *fitness* e expressa pelo somatório da quantidade de capital que o agente possui e a quantidade de ações vezes o preço de mercado da ação, conforme apresentado na Equação 23.

$$fitness[i] \leftarrow capital[i] + preço_mercado[s] * ações[i] \quad (23)$$

Cada configuração foi simulada 30 vezes e extraímos o comportamento médio de cada uma delas. Totalizando, obtivemos quatro tipos de mercado satisfazendo um modelo de *Random Walk* (RW) para a primeira versão das simulações. Para a segunda versão, foram nove tipos de mercado com tendência de alta e nove tipos de mercado com tendência de baixa. Além de quatro tipos de ambientes gerados com os preços históricos reais.

Dessa forma, o objetivo do nosso algoritmo é maximizar a quantidade de riqueza que os agentes possuem e através do PSO, gerar interação entre as partículas (os agentes financeiros). O algoritmo é desenvolvido para procurar a configuração mais eficiente para as posições financeiras dos agentes (posição de venda e de compra). Após cada iteração do PSO, todas as informações sobre os agentes, as negociações, os preços das ações e as posições das coordenadas $(\delta[p], \lambda[p])$ de cada partícula são analisadas.

3.3 Testes Estatísticos

Realizaram-se os testes de Mann-Kendall, Levene, Jarque-Bera e Dickey-Fuller Aumentado para verificação das hipóteses levantadas e validação dos resultados.

3.3.1 Teste de Mann-Kendall

O teste não paramétrico de Mann-Kendall é proposto para avaliar a existência de tendência estatisticamente significativa em uma série temporal. Ao realizar o teste proposto por Mann (1945), estamos testando as hipóteses:

$$\begin{cases} H_0 : \text{Não há tendência} \\ H_1 : \text{Há tendência} \end{cases}$$

Dessa forma, considerando x_1, x_2, \dots, x_n observações de uma série temporal, o teste estatístico de Mann-Kendall à luz de H_0 , é dado por [85, 86]:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \quad (24)$$

Em que as amostras x_j e x_i possuem o mesmo tamanho, com $j > i$ e

$$\text{sgn}(x_j - x_i) = \begin{cases} -1, & \text{se } (x_j - x_i) < 0 \\ 0, & \text{se } (x_j - x_i) = 0 \\ 1, & \text{se } (x_j - x_i) > 0 \end{cases} \quad (25)$$

Além disso, $S \sim N(\mu, \sigma^2)$, onde $\mu = 0$ e

$$\sigma^2 = \frac{n(n-1)(2n-5) - \sum_{i=1}^m t_i(i)(i-1)(2i+5)}{18} \quad (26)$$

Em que n é a quantidade de observações, m a quantidade de grupos que têm observações iguais e t_i a quantidade de observações iguais no grupo i .

No limite de grandes amostras, a estatística do teste pode ser calculada como:

$$Z = \begin{cases} \frac{S+1}{\sqrt{\text{Var}(S)}}, & \text{se } S < 0 \\ 0, & \text{se } S = 0 \\ \frac{S-1}{\sqrt{\text{Var}(S)}}, & \text{se } S > 0 \end{cases} \quad (27)$$

A partir da Equação 25, pode-se concluir que os dados crescem ao longo do tempo, apontando uma tendência positiva, caso o valor de S seja positivo. Entretanto, haverá tendência negativa, caso o valor de S seja negativo. Além disso, se o valor de S for significativamente diferente de zero, pode-se rejeitar a hipótese nula para um certo nível de significância, levando à conclusão de que há tendência na série de dados.

3.3.2 Teste de Levene

Utiliza-se o teste paramétrico de Levene para averiguar a homogeneidade das variâncias de duas ou mais amostras. Com ele, pode-se observar se há variabilidade das amostras em relação à resultante do procedimento a ser testado. Nos testes paramétricos, as variâncias amostrais não precisam ser idênticas, contudo, devem ser homogêneas [87].

Dessa maneira, ao realizarmos o teste de Levene estamos testando as seguintes hipóteses:

$$\begin{cases} H_0 : \sigma_E^2 = \sigma_C^2 \\ H_1 : \sigma_E^2 \neq \sigma_C^2 \end{cases}, \text{ onde E indica o grupo experimental e C indica o grupo controle.}$$

Define-se a partir da Equação 28, a estatística do teste de Levene (W) [88].

$$W = \frac{(N - k) \sum_{i=1}^k N_i (Z_i - Z_{..})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - Z_i)^2} \quad (28)$$

Em que W indica o resultado do teste, N a quantidade total de casos em todos os grupos, N_i a quantidade de casos no i -ésimo grupo, k a quantidade de grupos diferentes em que estão inclusos os casos amostrais e Y_{ij} o valor da variável medida para o j -ésimo caso do i -ésimo grupo. Z_{ij} é dado por:

$$Z_{ij} = \begin{cases} |Y_{ij} - \bar{Y}_i|, \text{ onde } \bar{Y}_i \text{ é a média do } i\text{-ésimo grupo} \\ |Y_{ij} - \tilde{Y}_i|, \text{ onde } \tilde{Y}_i \text{ é a mediana do } i\text{-ésimo grupo} \end{cases} \quad (29)$$

3.3.3 Teste de Jarque-Bera

O teste de Jarque-Bera é utilizado para verificar se os dados de uma determinada série temporal são distribuídos normalmente ou não. Para testar a hipótese de normalidade, ele se baseia na curtose e na assimetria da amostra. Assim, as hipóteses do teste são as seguintes:

$$\left\{ \begin{array}{l} H_0 : \text{Os dados são normalmente distribuídos, com assimetria igual a zero e} \\ \quad \text{curtose igual a três} \\ H_1 : \text{Os dados não são normalmente distribuídos} \end{array} \right.$$

A estatística do teste de Jarque-Bera é definida conforme a Equação 30 [89]:

$$JB = \frac{n}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right) \quad (30)$$

Em que n é a quantidade de observações e S e K são, respectivamente, a assimetria e a curtose da amostra, dadas pelas Equações 31 e 32 [90]:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^3}{\left(\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 \right)^{3/2}} \quad (31)$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^4}{\left(\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 \right)^2} \quad (32)$$

Para uma distribuição gaussiana (normal), a assimetria é 0 e a curtose é 3. O excesso de curtose, que é calculado através da Equação 33, na distribuição gaussiana é 0.

$$E_K = K - 3 \quad (33)$$

Em que K é a curtose.

Então a estatística da Equação 30, no limite de grandes amostras, será comparada a uma distribuição de qui-quadrado de grau 2.

3.3.4 Teste de Dickey-Fuller Aumentado

O teste de Dickey-Fuller aumentado, também conhecido como teste ADF, trata-se de uma versão aumentada do teste para raízes unitárias de Dickey-Fuller. É utilizado para verificar a existência de estacionariedade em séries temporais e busca-se realizar com ele, o estudo sobre a regressão abaixo [56]:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \epsilon_t$$

Em que β_1 representa o intercepto (*drift* da série), β_2 o coeficiente de tendência, δ o coeficiente de presença de raiz unitária, m o número de defasagens consideradas na série, ϵ_t um ruído branco e $\Delta Y_{t-1} = Y_{t-1} - Y_{t-2}$.

A hipótese nula a ser avaliada nesse teste é $H_0 : \delta = 0$ e a estatística do teste (T) é definida como:

$T = \frac{\hat{\delta}}{se(\hat{\delta})}$, em que $\hat{\delta}$ indica um estimador para δ e $se(\hat{\delta})$ um estimador para o desvio padrão do erro de σ .

3.4 Coeficiente de Gini

Na abordagem dada ao coeficiente de Gini é utilizada a Curva de Lorenz, que trata-se de um gráfico de dispersão, onde está representada a distribuição acumulada de uma variável (ϕ) em relação a um conjunto determinado (p), conforme Figura 9.

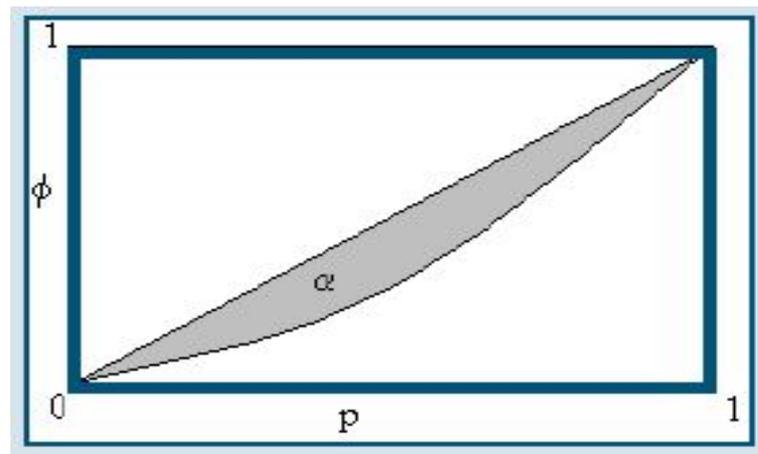


Figura 9 – Ilustração da Curva de Lorenz [4].

A Curva de Lorenz é crescente e vai do ponto $(0, 0)$ ao ponto $(1, 1)$. A reta que liga esses pontos representa a distribuição onde todos têm o mesmo rendimento e é chamada Reta da Igualdade Perfeita. Dessa forma, quanto mais afastada estiver a curva de Lorenz de uma distribuição da Reta da Igualdade Perfeita, maior será a área entre elas e também o Coeficiente de Gini, significando uma maior desigualdade.

Uma boa aproximação para o Coeficiente de Gini é dada pela Equação 34:

$$G = 1 - \sum_{k=1}^{n-1} (X_{k+1} - X_k)(Y_{k+1} + Y_k) \quad (34)$$

Em que X corresponde a distribuição acumulada da variável e Y a distribuição do conjunto determinado (no nosso caso, equivalente a riqueza dos agentes econômicos).

O Coeficiente de Gini dá um enfoque maior às mudanças que acontecem no centro da distribuição [91].

Para análise da distribuição de riqueza nos nossos mercados, utilizamos as informações da última iteração do algoritmo PSO, o qual é visto como sendo o último dia em que houve negociação. Observamos os dados de cada agente financeiro e a riqueza total de cada um deles, que leva em consideração a quantidade de ações e a quantidade de capital que os mesmos possuem.

3.5 Ajuste Exponencial e por Lei de Potência

Com as séries de retornos dos preços dos ativos negociados, realizamos a aproximação da distribuição dos retornos pela lei de potência e pela função exponencial. Para isso, normalizamos as séries conforme a Equação 5 e, em seguida, calculamos a probabilidade de sobrevivência dos retornos. Para o ajuste, utilizamos os algoritmos *Linear Least Squares* (LLS), *Levenberg-Marquardt* (LM) e *Trust Region* (TR) na *toolbox curve fitting* do *software* Matlab. Consideramos o *Mean Square Error* (MSE) na avaliação dos erros dos ajustes e observamos os valores da fdp das séries sem transformação de escala e nas escalas semi-log e log-log.

O algoritmo *Linear Least Squares* busca encontrar um bom ajuste linear aos dados [92], através de um valor que minimize a soma dos quadrados dos resíduos. Ao realizarmos esse ajuste na escala semi-log, estamos fazendo um ajuste pela função exponencial, pois nessa escala funções exponenciais comportam-se como retas. De forma análoga, ao aplicar a escala log-log estamos adotando uma aproximação pelas leis de potência.

O método de *Levenberg-Marquardt* é adotado em problemas de mínimos quadrados não-lineares. Trata-se de uma combinação do método de gradiente descendente e do método de Gauss-Newton. Se os parâmetros se distanciam da solução ótima, o primeiro método força a descida de gradiente. Em contrapartida, o segundo método atua quando o valor ótimo encontra-se próximo aos parâmetros [93].

Já o algoritmo *Trust Region*, que é um método não-linear robusto, baseia-se na ideia de aproximação da solução para o problema, minimizando a região de busca. Com isso, o ajuste é realizado a cada iteração e ao encontrar uma boa região, ela é expandida. Caso contrário, ela será reprimida [94].

4 Resultados

4.1 Mercado Simulado com PSO

Pode-se observar na Figura 10 o preço médio obtido a partir de 30 simulações para cada *drift* dos RW's no ambiente artificial. Nessa primeira rodada, para a geração dos preços que seguem uma distribuição gaussiana $N(\mu, \sigma)$, foi fixada a média em 0 e o desvio-padrão variou entre os valores 0.01, 0.05, 0.10 e 0.30.

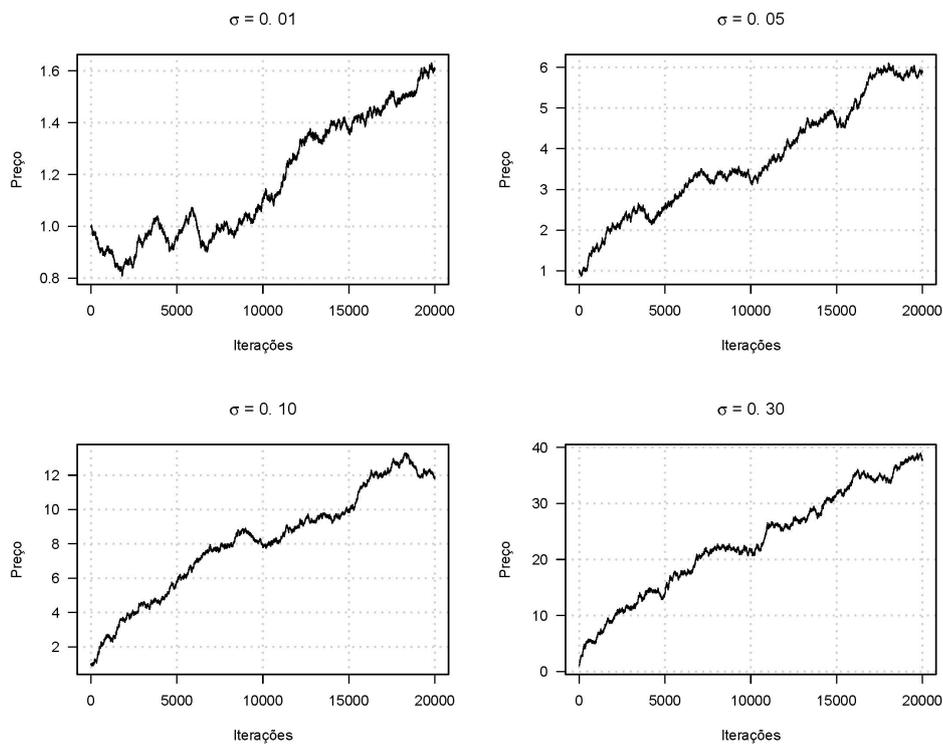


Figura 10 – Séries dos preços negociados para os *drifts* $N(0, 0.01)$, $N(0, 0.05)$, $N(0, 0.10)$ e $N(0, 0.30)$.

Observações empíricas sobre essas séries sugerem que elas sejam de processos não estacionários, caracterizando um passeio aleatório, também conhecido como *Random Walk* (RW). Inicialmente, realizamos o teste de Mann-Kendall para verificar se as séries de preços obtidas possuem tendência. A Tabela 5 exhibe os resultados encontrados.

Tabela 5 – Teste de tendência de Mann-Kendall (*Rank Test*) para os preços médios em cada RW.

σ	Estatística	p -valor	Hipótese Aceita
0.01	180.2	< 0.0001*	H_1 (Há tendência)
0.05	190.78	< 0.0001*	H_1 (Há tendência)
0.10	193.86	< 0.0001*	H_1 (Há tendência)
0.30	196.08	< 0.0001*	H_1 (Há tendência)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

De acordo com a Tabela 5, a hipótese nula de que não há tendência nas séries não foi aceita em todos os casos. Isso indica, que todas as séries temporais analisadas apresentam tendência. Isso se deve ao fato, de que apesar dos preços serem gerados aleatoriamente, como impomos apenas valores positivos para a série, os preços sempre tenderão a crescer.

Aprofundando um pouco mais as investigações sobre a existência de aleatoriedade nessas séries, aplicou-se o teste de raiz unitária Dickey-Fuller, o qual consiste em verificar se o conjunto de dados é estacionário ou não. Caso seja, a série não segue um *Random Walk*. Os resultados obtidos são exibidos na Tabela 6.

Tabela 6 – Teste de Dickey-Fuller Aumentado (ADF) para os RW's.

σ	Estatística	Valor Crítico	Hipótese Aceita
0.01	-3.227249	-3.41	H_0 (Tem raiz unitária)
0.05	-2.752709	-3.41	H_0 (Tem raiz unitária)
0.10	-2.184279	-3.41	H_0 (Tem raiz unitária)
0.30	-4.104844	-3.41	H_1 (Estacionária)

Devido ao baixo poder do teste, quando $\sigma = 0.30$ o teste de Dickey-Fuller a 5%, não detectou raiz unitária na série de preços. Logo, a mesma não segue um RW. Para os demais casos, a hipótese nula foi aceita e podemos assumir com mais segurança que as séries seguem um RW.

Além disso, a estacionariedade também pode ser verificada com a função de autocorrelação. Conforme a Figura 11, foi observada uma elevada resistência no decaimento da função de autocorrelação das séries, o que pode indicar presença de raiz unitária, caracterizando falta de estacionariedade.

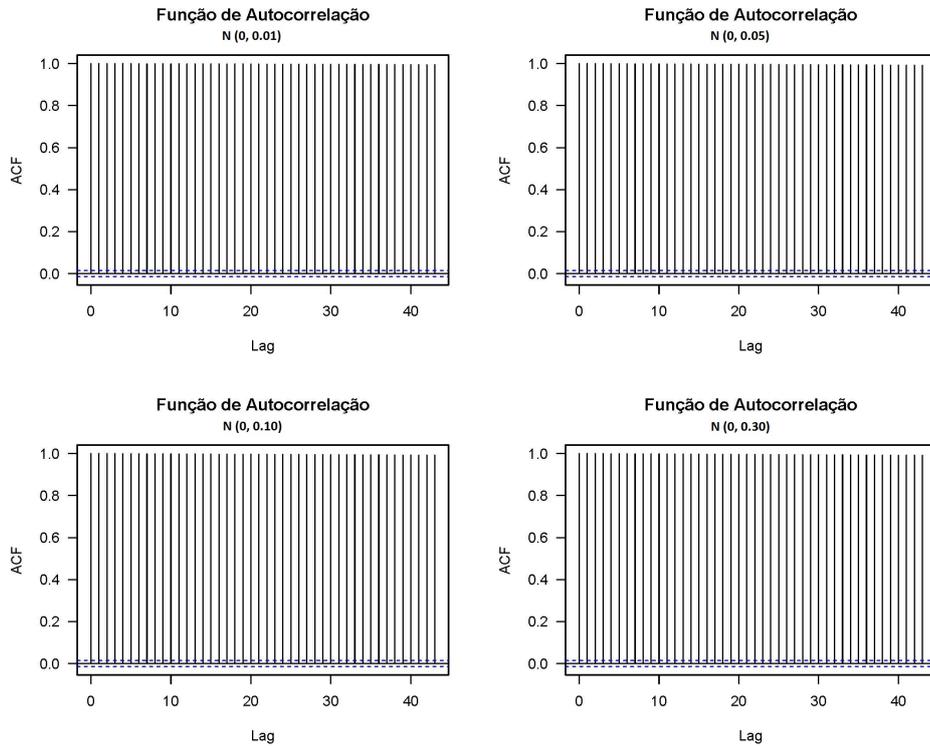


Figura 11 – Autocorrelações dos preços negociados nas simulações com *drifts* $N(0, 0.01)$, $N(0, 0.05)$, $N(0, 0.10)$ e $N(0, 0.30)$.

Dessa maneira, com os resultados obtidos nos testes de estacionariedade e tendência podemos concluir que as simulações realizadas na primeira rodada são de um processo não estacionário.

Na segunda rodada de simulações, os preços das ações que anteriormente seguiam uma distribuição gaussiana com média fixa, começaram a ser gerados por essa mesma distribuição, mas com o nível variando entre os valores ± 0.01 , ± 0.05 e ± 0.10 . Algumas das séries produzidas são descritas na Figura 12.

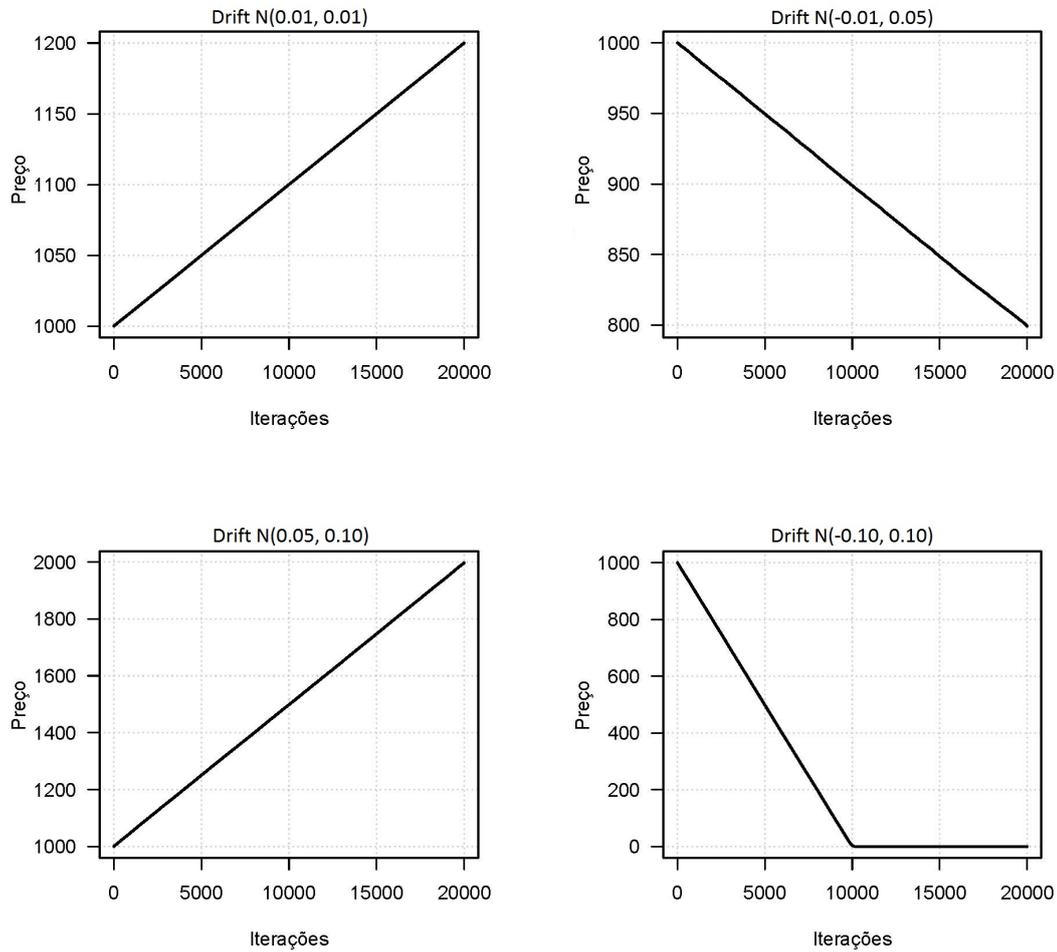


Figura 12 – Séries de preços negociados no mercado simulado para os *drifts* $N(0.01, 0.01)$, $N(-0.01, 0.05)$, $N(0.05, 0.10)$ e $N(-0.10, 0.10)$

Para confirmar a existência de tendência crescente ou decrescente nas séries de preços obtidas nessa segunda rodada de simulações, o teste de Mann-Kendall foi realizado e indica que para os *drifts* positivos as séries apresentam tendência de alta. Em contrapartida, para os *drifts* negativos há ocorrência de tendência de baixa. Ver Tabela 7.

Tabela 7 – Teste de tendência de Mann-Kendall (*Rank Test*) para os preços médios em cada *drift*.

$N(\mu, \sigma)$	Estatística	p -valor	Hipótese Aceita
$N(0.01, 0.01)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.01, 0.05)$	212.1	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.01, 0.10)$	212.05	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.05, 0.01)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.05, 0.05)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.05, 0.10)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.10, 0.01)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.10, 0.05)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(0.10, 0.10)$	212.11	$< 0.0001^*$	H_1 (Tendência Positiva)
$N(-0.01, 0.01)$	-212.11	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.01, 0.05)$	-212.1	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.01, 0.10)$	-212.05	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.05, 0.01)$	-212.11	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.05, 0.05)$	-212.11	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.05, 0.10)$	-212.11	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.10, 0.01)$	-159.87	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.10, 0.05)$	-160.25	$< 0.0001^*$	H_1 (Tendência Negativa)
$N(-0.10, 0.10)$	-161.34	$< 0.0001^*$	H_1 (Tendência Negativa)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

4.1.1 Validação do modelo

Com os resultados obtidos nas simulações, buscou-se inicialmente validar o modelo do mercado simulado a partir da observação dos fatos estilizados. Como já ressaltado anteriormente, estes, são propriedades estatísticas observadas em séries financeiras reais que desejamos reproduzir nas séries financeiras artificiais. A seguir, são apresentados os fatos estilizados identificados nas séries financeiras simuladas, médias do comportamento de cada situação.

Após a obtenção das séries de retornos, de acordo com a Equação 4, foram feitos os seus respectivos gráficos e gerados os histogramas das séries, com o objetivo de observar através de suas funções de distribuição, as probabilidades de perdas e ganhos nos mercados. Exemplos desses gráficos e histogramas podem ser visualizados nas Figuras 13 e 14, respectivamente.

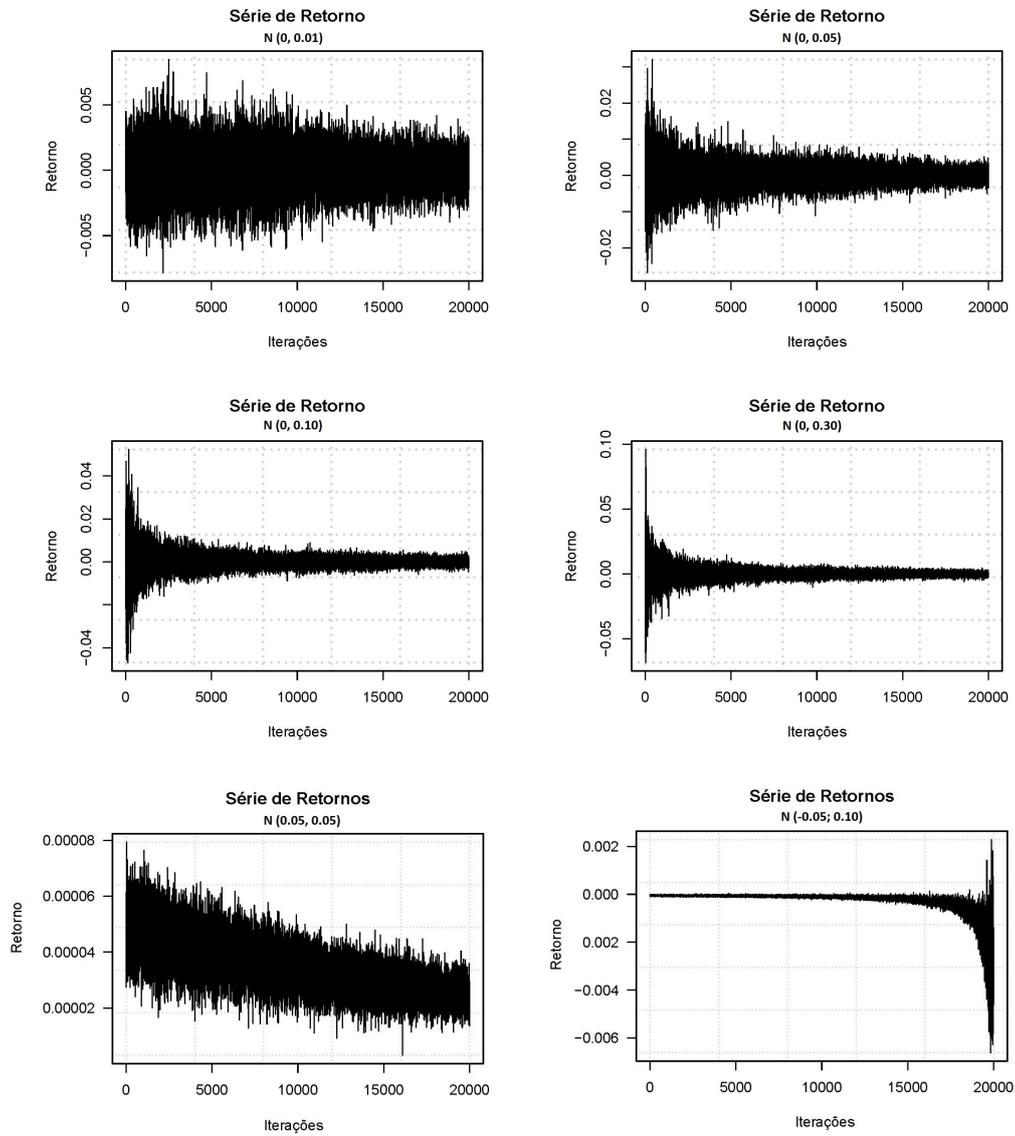


Figura 13 – Exemplos de séries temporais de retornos dos preços negociados no mercado simulado.

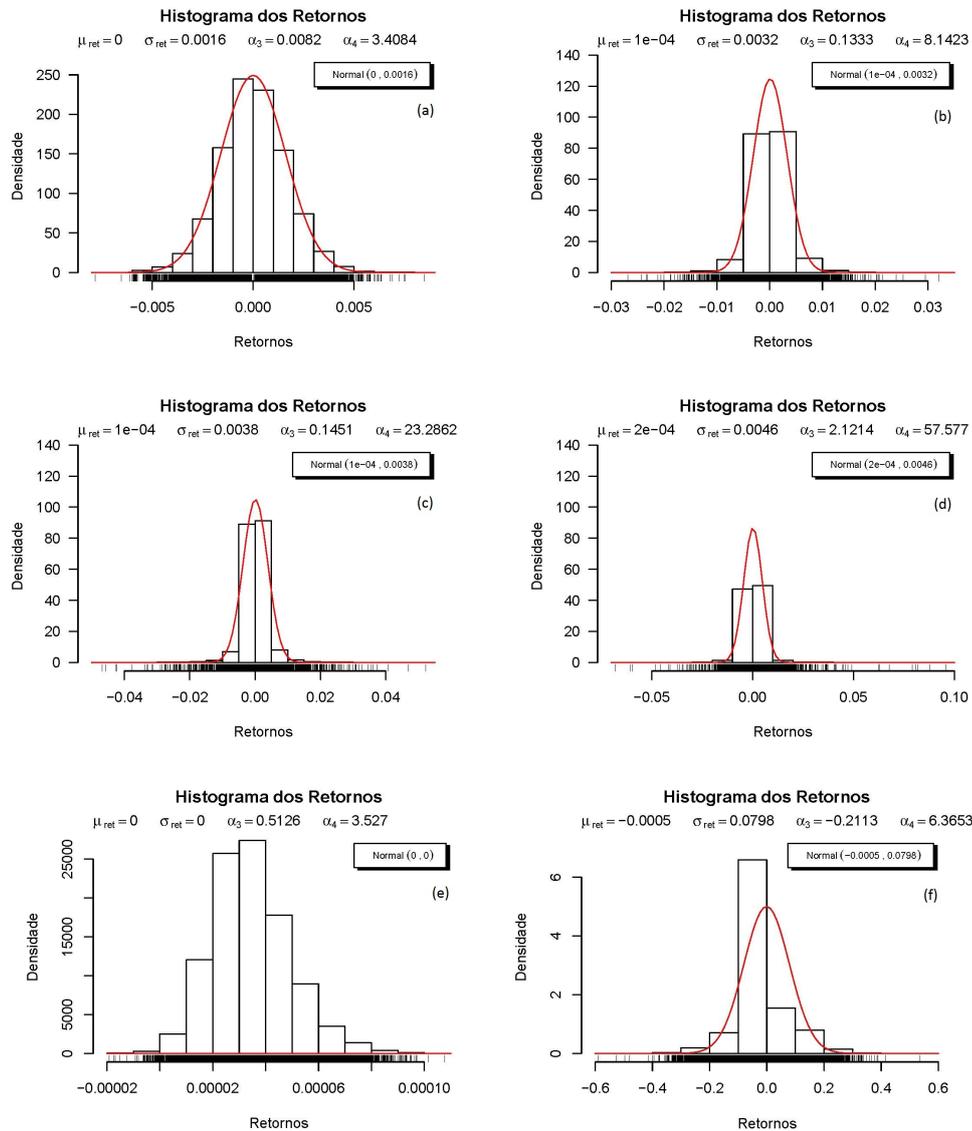


Figura 14 – Exemplos de histogramas de séries temporais de retornos dos preços negociados no mercado simulado: $\sigma = 0.01$ (a), $\sigma = 0.05$ (b), $\sigma = 0.10$ (c), $\sigma = 0.30$ (d), $drift(0.05, 0.10)$ (e) e $drift(-0.10, 0.10)$ (f).

Na Figura 14, a curva em vermelho representa o ajuste realizado por uma densidade normal aos dados dos retornos. As distribuições normais usadas para o ajuste estão indicadas. Observa-se que para o $drift(0.05, 0.10)$ na Figura 13 (e), devido a baixíssima variabilidade encontrada nos dados (infinitesimais), a estimação dos parâmetros da normal resultaram em média e variância nulas.

Durante muitos anos, acreditou-se que as distribuições das séries de retornos de preços teriam o formato de uma distribuição gaussiana. Contudo, Mandelbrot observou em seu trabalho publicado em 1963 [64], que em janelas de observações curtas, há excesso

de curtose e que, portanto, as distribuições não seguiriam uma distribuição gaussiana.

Nas nossas séries financeiras se verificou excesso de curtose e assimetria, negativa para os casos de mercado com tendência de baixa, e positiva para os casos de mercados com tendência de alta e para os mercados gerados pelo modelo RW. O excesso de curtose sugere que a distribuição dos retornos apresenta cauda mais pesada que a distribuição normal, enquanto a assimetria indica se os valores predominam acima ou abaixo da média. Quando há assimetria negativa, há maior probabilidade de ocorrerem os valores na cauda esquerda da distribuição. Por outro lado, quando há assimetria positiva, a probabilidade de ocorrerem os valores na cauda direita da distribuição é maior. As caudas pesadas na distribuição e a assimetria dos dados são fenômenos observados em dados financeiros [64] e satisfazem aos fatos estilizados da economia [15].

Dada a assimetria e o excesso de curtose observados nas séries, o teste de normalidade de Jarque-Bera foi realizado e comprovou a não normalidade dos dados ao rejeitar a hipótese nula. A Tabela 8 expõe os resultados para o teste de Jarque-Bera.

Tabela 8 – Teste de Normalidade de Jarque-Bera para os preços das ações em cada *drift*.

<i>drift</i>	Estatística	<i>p</i> -valor	Hipótese Aceita
$N(0, 0.01)$	1974.5	$< 0.0001^*$	H_1 (Não Normal)
$N(0, 0.05)$	770.97	$< 0.0001^*$	H_1 (Não Normal)
$N(0, 0.10)$	789.93	$< 0.0001^*$	H_1 (Não Normal)
$N(0, 0.30)$	878.09	$< 0.0001^*$	H_1 (Não Normal)
$N(0.01, 0.01)$	1201	$< 0.0001^*$	H_1 (Não Normal)
$N(0.01, 0.05)$	1192.2	$< 0.0001^*$	H_1 (Não Normal)
$N(0.01, 0.10)$	1205.3	$< 0.0001^*$	H_1 (Não Normal)
$N(0.05, 0.01)$	1199.6	$< 0.0001^*$	H_1 (Não Normal)
$N(0.05, 0.05)$	1205	$< 0.0001^*$	H_1 (Não Normal)
$N(0.05, 0.10)$	1193.6	$< 0.0001^*$	H_1 (Não Normal)
$N(0.10, 0.01)$	1200.1	$< 0.0001^*$	H_1 (Não Normal)
$N(0.10, 0.05)$	1198.5	$< 0.0001^*$	H_1 (Não Normal)
$N(0.10, 0.10)$	1197.1	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.01, 0.01)$	1203	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.01, 0.05)$	1220.1	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.01, 0.10)$	1207.4	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.05, 0.01)$	1201.3	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.05, 0.05)$	1204.1	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.05, 0.10)$	1190.9	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.10, 0.01)$	3203.2	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.10, 0.05)$	3198.2	$< 0.0001^*$	H_1 (Não Normal)
$N(-0.10, 0.10)$	3219.3	$< 0.0001^*$	H_1 (Não Normal)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Indicado na Tabela 9, realizamos o teste de Levene para verificar a homocedasticidade entre as séries. Segundo os resultados adquiridos, não se pode assumir a igualdade de variâncias entre os preços gerados pelos respectivos RW's utilizados nas simulações.

Tabela 9 – Teste de homogeneidade das variâncias (Levene) entre os preços médios dos RW's.

σ	Estatística	p -valor	Hipótese Aceita
0.01 – 0.05	24037	$< 0.0001^*$	H_1 (Não homogêneas)
0.01 – 0.10	29166	$< 0.0001^*$	H_1 (Não homogêneas)
0.01 – 0.30	42082	$< 0.0001^*$	H_1 (Não homogêneas)
0.05 – 0.10	8807.3	$< 0.0001^*$	H_1 (Não homogêneas)
0.05 – 0.30	32029	$< 0.0001^*$	H_1 (Não homogêneas)
0.10 – 0.30	18745	$< 0.0001^*$	H_1 (Não homogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

A Figura 15 mostra exemplos de autocorrelações dos retornos das ações, onde evidencia-se ausência de autocorrelação linear. Para lag 0 temos $fac=1$, mas para lag $k>1$ então $fac = 0$, como esperado teoricamente para modelos tipo RW. O *drift* $N(0.10, 010)$, por exemplo, indica alguma correlação entre a primeira lag .

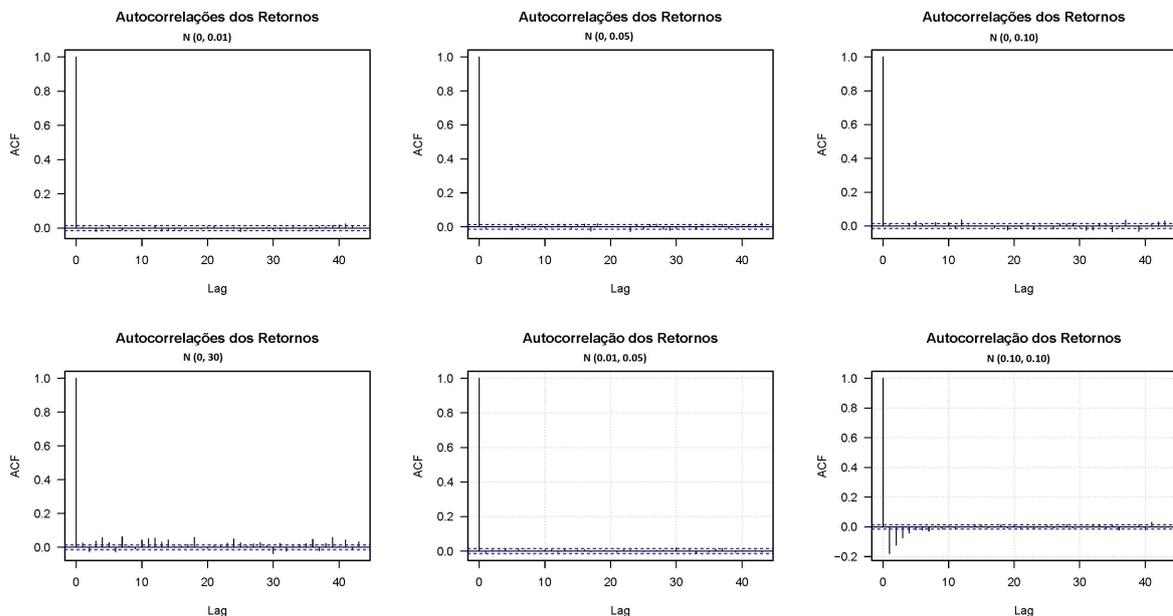


Figura 15 – Exemplos de autocorrelações de retornos de preços negociados.

A independência nas séries implica que elas apresentam ausência de autocorrelação. Se os retornos fossem autocorrelacionados, seria possível que os agentes do mercado

financeiro elaborassem estratégias preditivas para ganhar dinheiro [15]. Assim, haveria garantia de que os retornos passados poderiam ser usados com êxito na predição dos retornos futuros.

Tendo em vista o comportamento das nossas séries de dados, buscamos realizar a modelagem desses valores através da aproximação pelas funções exponencial e lei de potência. Exemplos dos comportamentos obtidos para o ajuste linear nas escalas semi-log e log-log, exponencial sem transformação de escala e lei de potência sem transformação de escala encontram-se a seguir. A linha azul reflete o ajuste das funções, enquanto os pontos pretos indicam os valores das séries de dados.

Como na escala semi-log o ajuste linear equivale ao ajuste da exponencial e na escala log-log equivale ao realizado pelas leis de potência, pois as funções exponenciais em escala semi-log comportam-se como retas e em escala log-log as leis de potência são retas [26], realizamos essas aproximações.

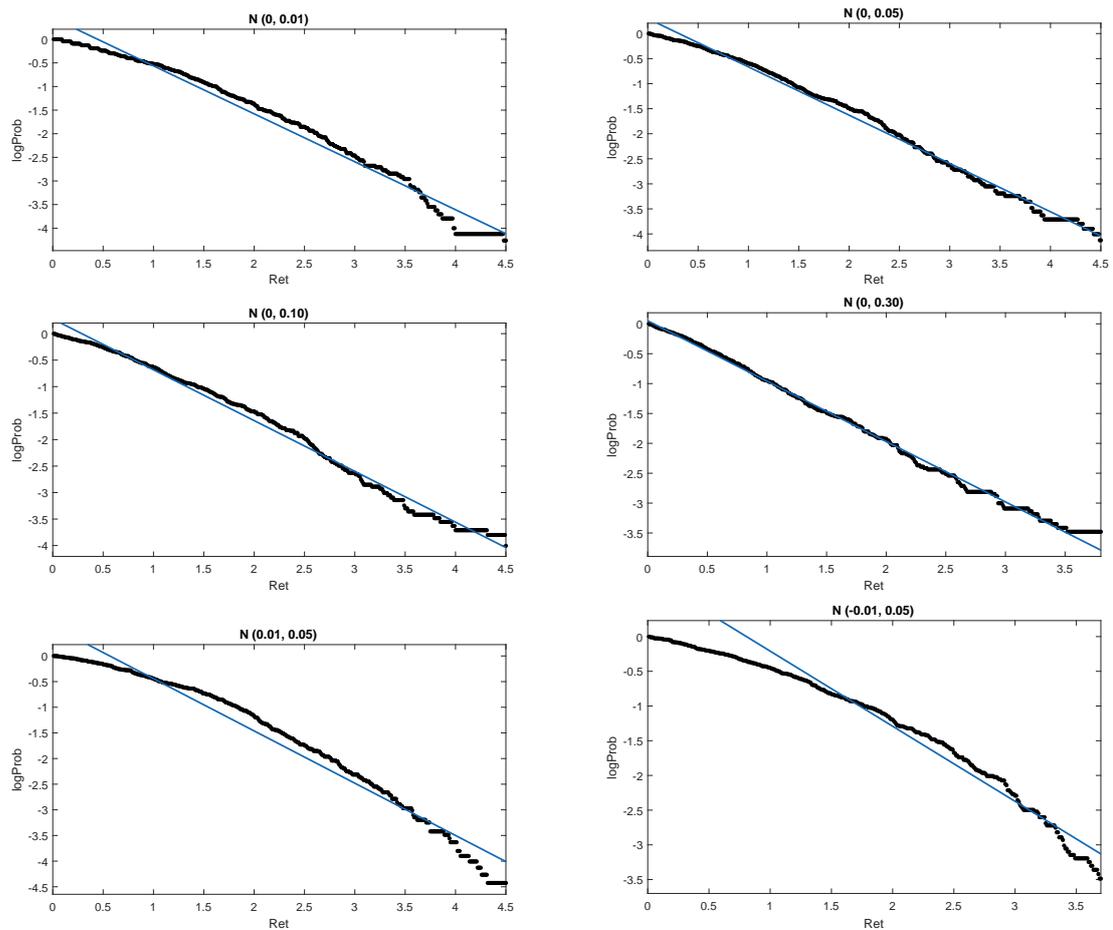


Figura 16 – Exemplos de ajustes lineares na escala semi-log.

Nas Figuras 16 e 17, temos exemplos de ajustes lineares realizados nas escalas semi-log e log-log, respectivamente. Analisando os resultados obtidos, observamos que as séries de dados na escala semi-log foram ajustadas em toda a sua extensão, exceto para os mercados com tendência de baixa. Em contrapartida, vemos que na escala log-log somente os dados que estão na cauda da distribuição tiveram uma melhor aproximação, cooperando com os resultados esperados da literatura para as séries reais [47, 49].

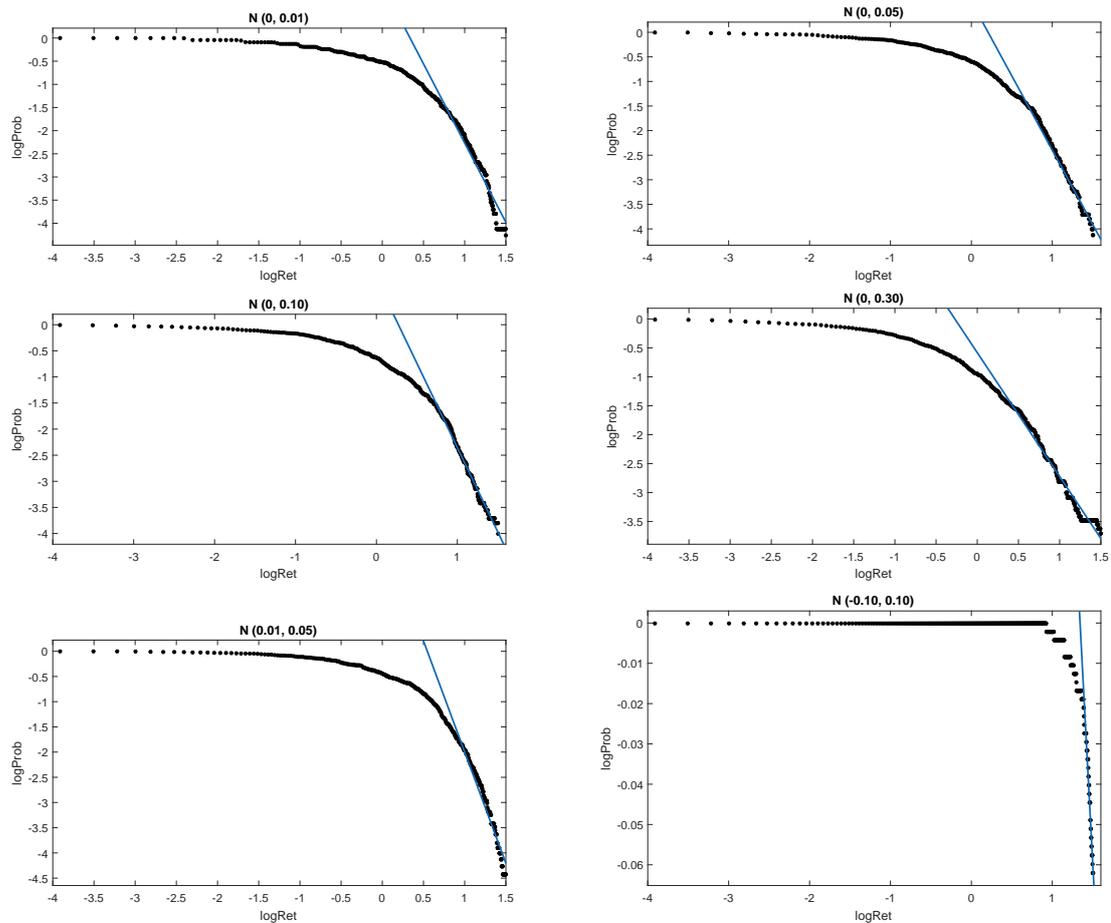


Figura 17 – Exemplos de ajustes lineares na escala log-log.

Ao realizarmos o ajuste dos dados pela função exponencial e pela lei de potência sem mudanças nas escalas, notamos que mais uma vez, se considerarmos toda a série nessa modelagem, a melhor aproximação ocorrerá ao utilizarmos a função exponencial. Com ela, perdemos informação somente nos extremos da distribuição. Porém, considerando somente a cauda da função de distribuição dos dados, conforme o comportamento previsto para séries de mercados reais em [17], a lei de potência apresenta um melhor ajuste a fdp nesse local, corroborando com os trabalhos observados na literatura [26], onde a lei de potência melhor modela os eventos extremos. Veja exemplos na Figuras 18 e 19.

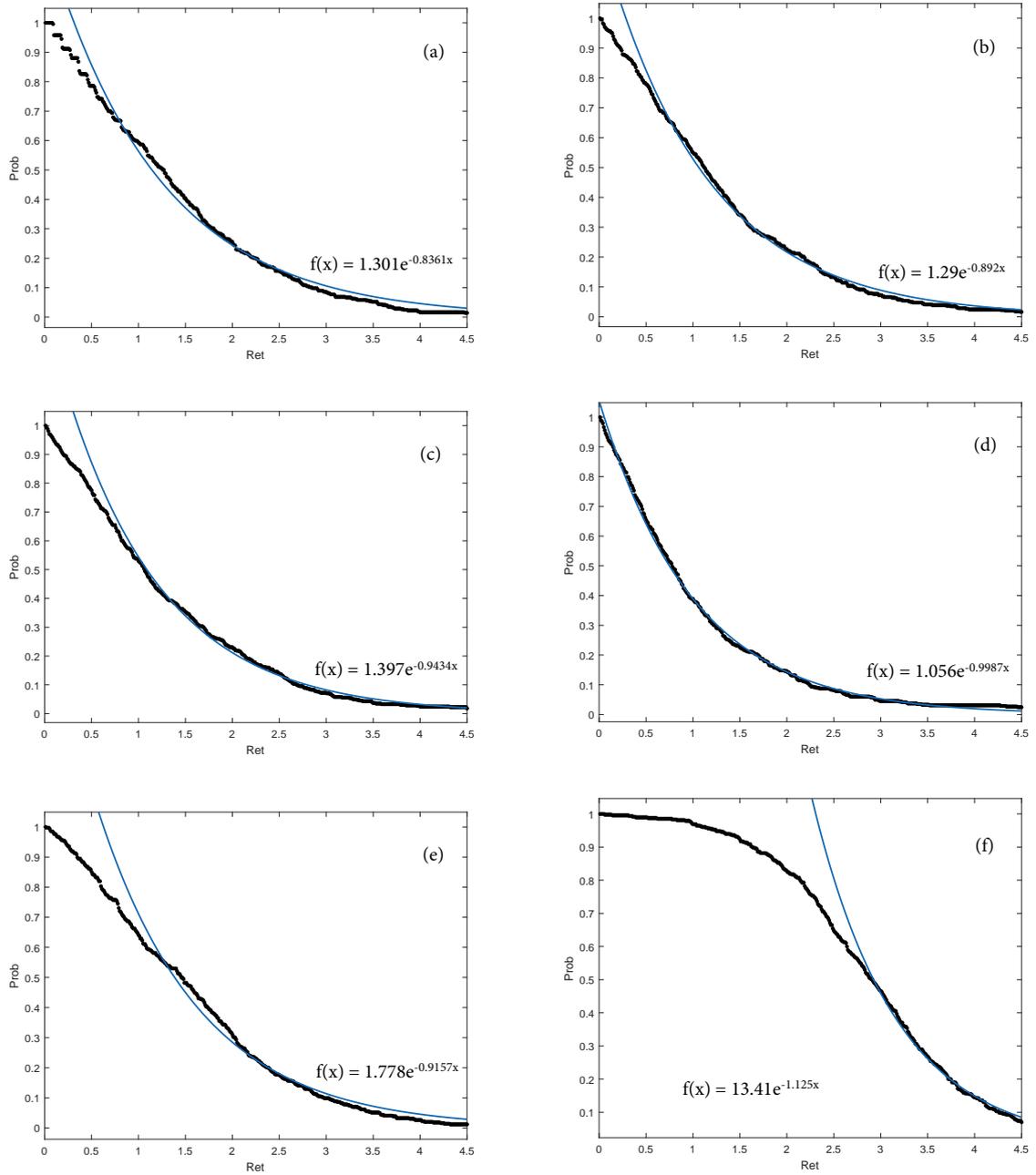


Figura 18 – Exemplos de ajustes pela exponencial sem transformação na escala para as simulações com *drifts*: $N(0, 0.01)$ (a), $N(0, 0.05)$ (b), $N(0, 0.10)$ (c), $N(0, 0.30)$ (d), $N(0.01, 0.05)$ (e) e $N(-0.10, 0.01)$ (f).

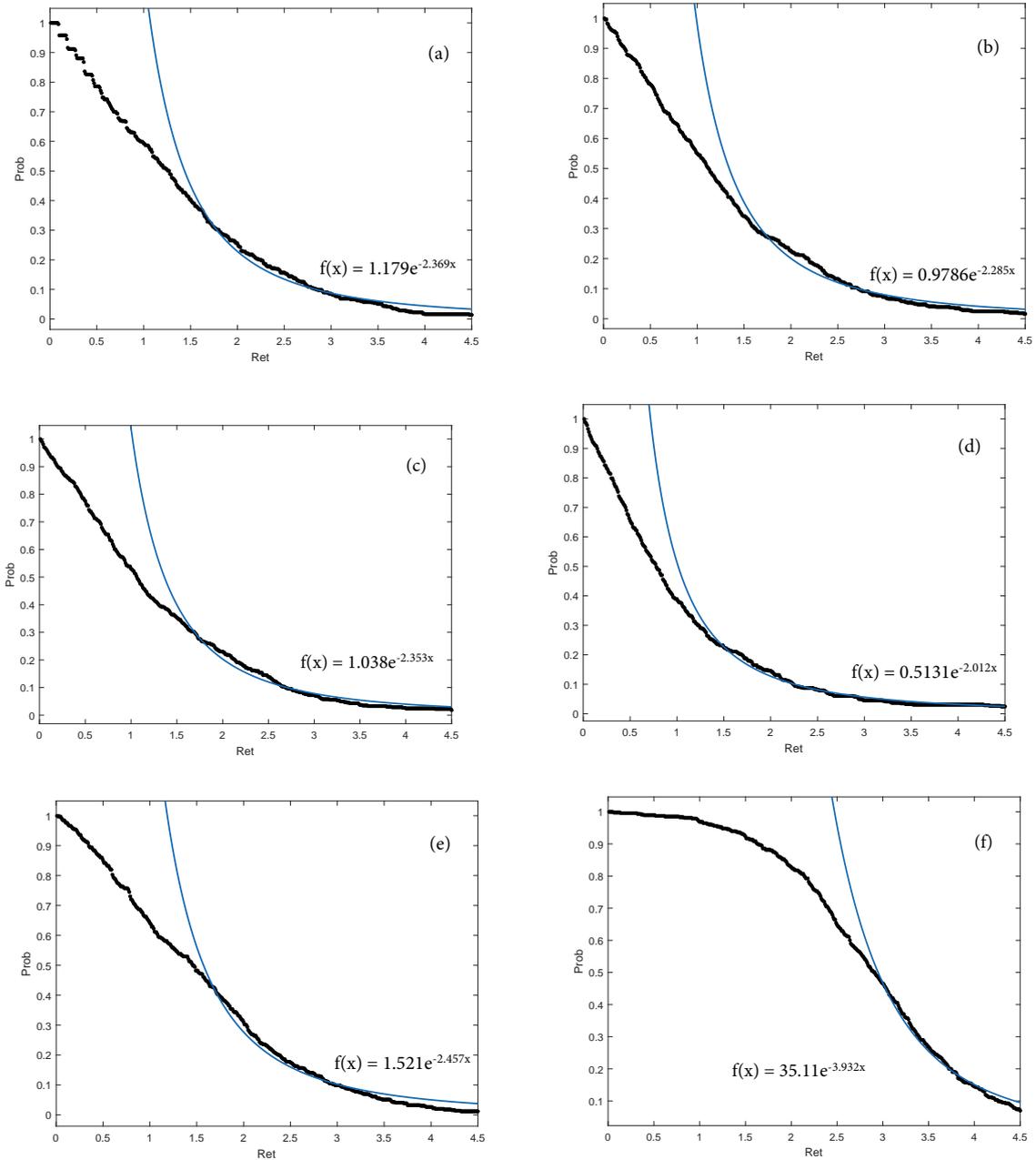


Figura 19 – Exemplos de ajustes pela lei de potência sem transformação na escala para as simulações com *drifts*: $N(0, 0.01)$ (a), $N(0, 0.05)$ (b), $N(0, 0.10)$ (c), $N(0, 0.30)$ (d), $N(0.01, 0.05)$ (e) e $N(-0.10, 0.01)$ (f).

Dado que os principais fatos estilizados da economia foram observados nas séries do mercado artificial, os resultados encontrados nos experimentos corroboram em linhas gerais de comportamento com a metodologia proposta.

4.1.2 Análises dos Parâmetros dos Agentes Financeiros no Ambiente Artificial

Como nosso mercado simulado foi implementado com o algoritmo PSO, dizemos que a melhor partícula do enxame é aquela que dispõe de mais riqueza, levando-se em consideração o volume total de ações e de capital que ela possui. Da mesma forma, dizemos que a pior partícula é aquela que é detentora de menos riqueza. Nesse trabalho, essas partículas são também chamadas de agente mais rico e agente mais pobre, respectivamente.

Para a melhor partícula no caso dos RW's, à medida que mudamos os valores do desvio-padrão aumentando gradativamente de 0.01, 0.05 a 0.10, os limiares de venda (δ) e de compra (λ) foram sofrendo inversão e passou-se a comprar caro e vender barato as ações, conforme Figura 20.

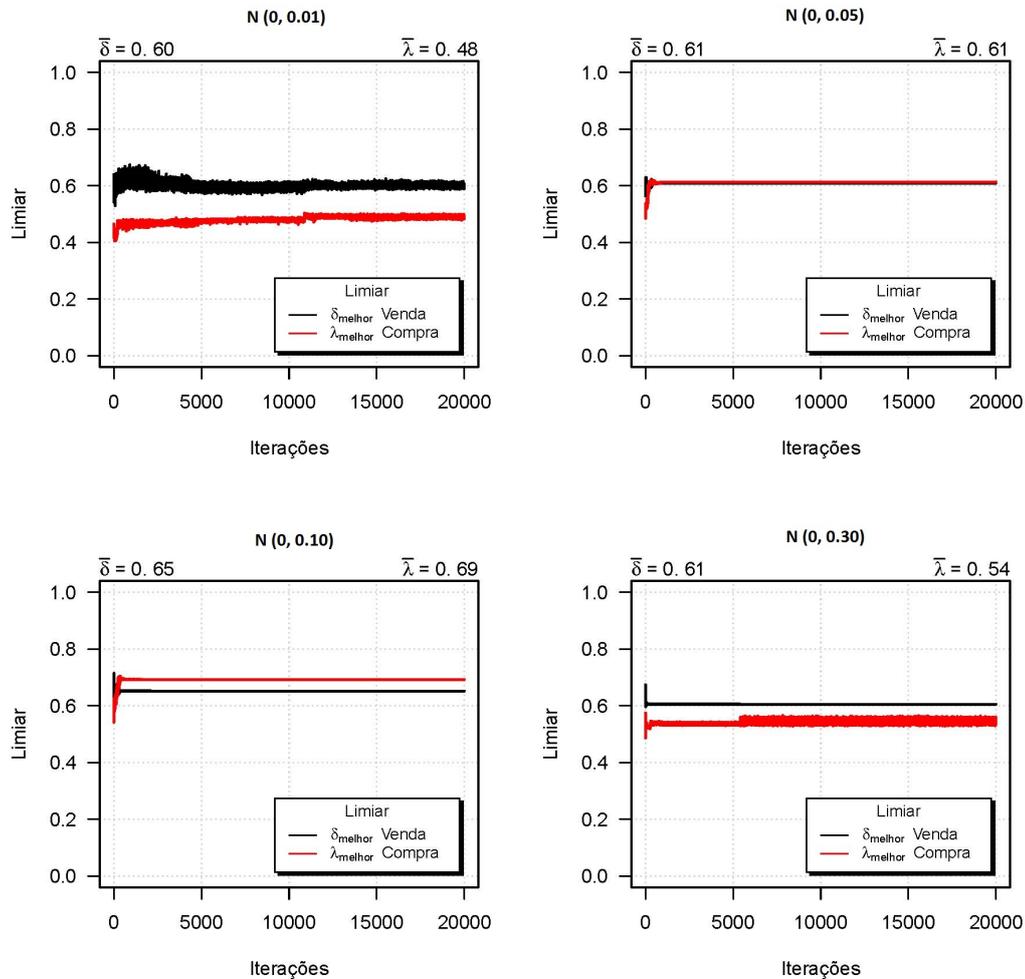


Figura 20 – Ilustração dos limiares de venda e de compra para a melhor partícula nos RW's.

Porém, ao assumirmos $\sigma = 0.30$, o processo voltou a ter o comportamento desejado. Isso sugere, que utilizar $\sigma = 0.10$ ou $\sigma = 0.30$ por exemplo, podem representar os casos em que as ações foram compradas em momentos de baixa e vendidas em momentos de alta no mercado. Para a pior partícula dos RW's, somente quando $\sigma = 0.01$ e $\sigma = 0.30$ as ações negociadas no ambiente artificial eram compradas pelo menor preço e vendidas pelo maior. Ver Figura 21.

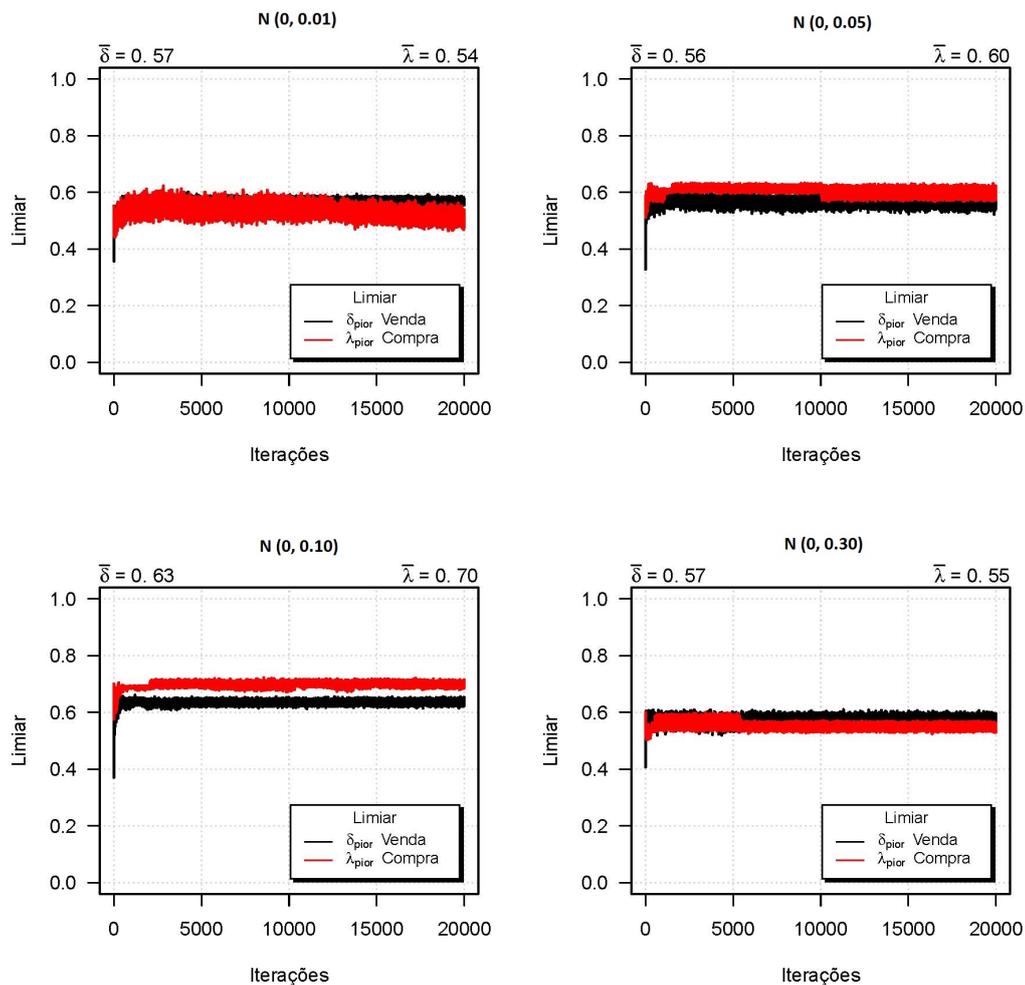


Figura 21 – Ilustração dos limiares de venda e de compra para a pior partícula nos RW's.

Ao analisarmos a segunda rodada de simulações, observamos que quando a média da distribuição utilizada assume valores negativos, representando um mercado com tendência de baixa, a melhor partícula apresenta o limiar de venda maior do que o limiar de compra em todos os casos. Indicando que estamos vendendo caro e comprando barato as ações no mercado, veja exemplos na Figura 22.

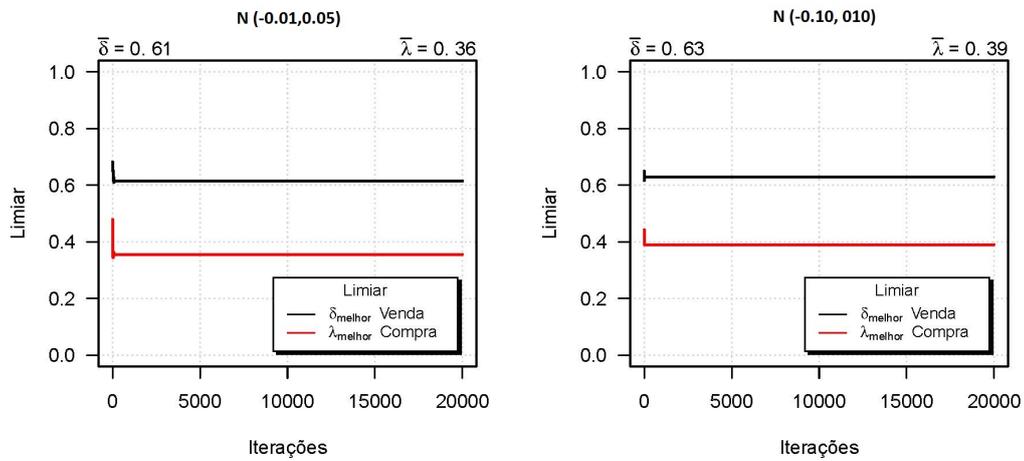


Figura 22 – Ilustração de exemplos dos limiares de venda e de compra para a melhor partícula nas simulações com *drifts* $N(-0.01, 0.05)$ e $N(-0.10, 0.10)$.

Por outro lado, conforme exibido na Figura 23, nos deparamos com limiares de venda menores do que os limiares de compra para as piores partículas nessas simulações. Sugerindo que as ações sofreram desvalorização no momento de venda e valorização no processo de compra.

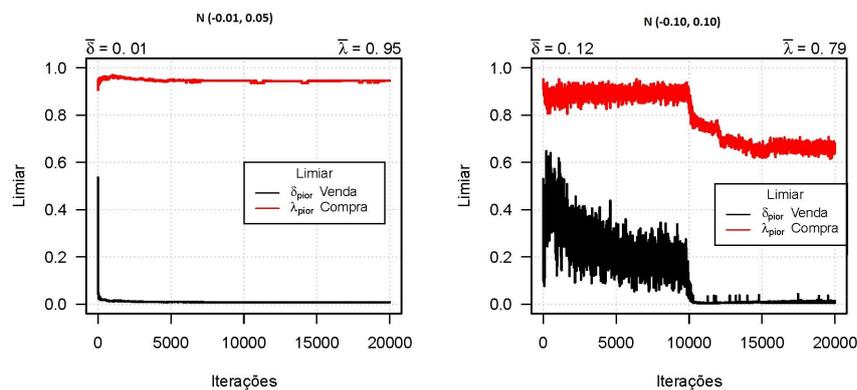


Figura 23 – Ilustração de exemplos dos limiares de venda e de compra para a pior partícula nas simulações com *drifts* $N(-0.01, 0.05)$ e $N(-0.10, 0.10)$.

Se por um lado, conforme exemplos na Figura 25, para os mercados com tendência de alta, a melhor partícula apresenta os limiares de compra inferiores aos limiares de venda para todos os *drifts* da segunda rodada de simulações. Por outro, para a pior partícula, os limiares obtidos estão muito próximos. O de venda é sempre maior do que o de compra,

conforme exemplos na Figura 24. Embora que em alguns casos estes limiares são iguais (*drift* $N(0.01, 0.01)$, *drift* $N(0.01, 0.10)$, *drift* $N(0.05, 0.10)$).

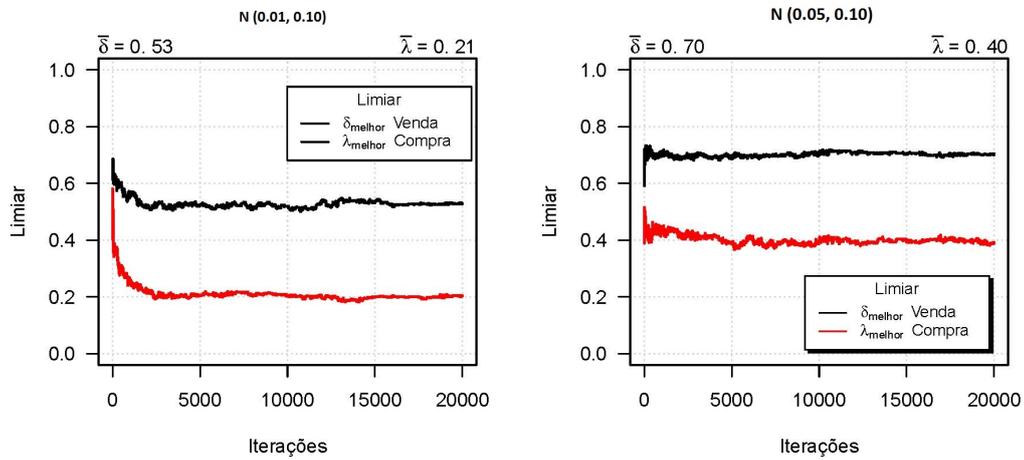


Figura 24 – Ilustração de exemplos dos limiares de venda e de compra para a pior partícula nas simulações com tendência de alta nos mercados.

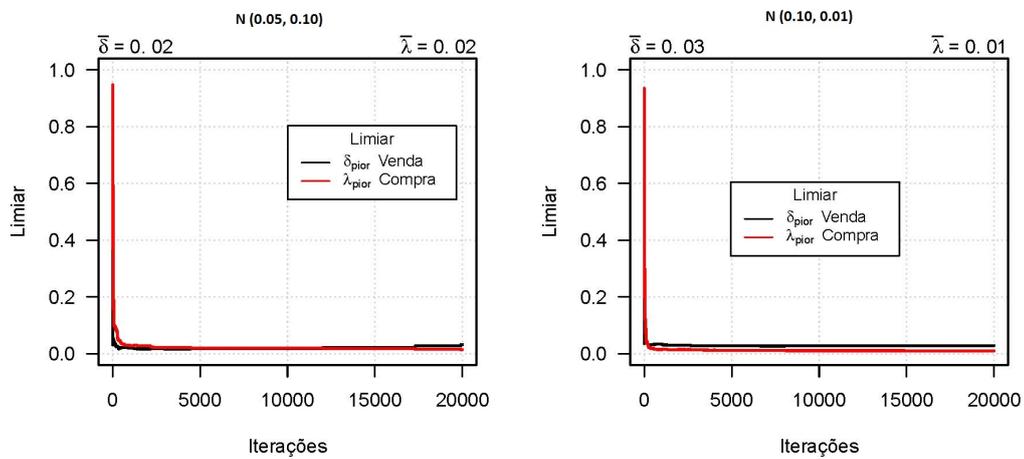


Figura 25 – Ilustração de exemplos dos limiares de venda e de compra para a melhor partícula nas simulações com tendência de alta nos mercados.

A Tabela 10 expõe os resultados do teste de Jarque-Bera para os limiares.

Tabela 10 – Teste de Normalidade de Jarque-Bera para os limiares δ e λ em cada RW.

σ	Estatística		p -valor		Hipótese Aceita	
	δ	λ	δ	λ	δ	λ
0.01	12772	26779	<0.0001*	<0.0001*	H_1	H_1
0.05	20527000	14728000	<0.0001*	<0.0001*	H_1	H_1
0.10	117310000	14002000	<0.0001*	<0.0001*	H_1	H_1
0.30	7622400000	308.18	<0.0001*	<0.0001*	H_1	H_1

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

De acordo com a Tabela 10, pode-se dizer que as séries de dados referentes aos limiares δ e λ não encontram-se normalmente distribuídos. Sendo assim, para comparação das variâncias e verificação quanto a homogeneidade entre os limiares de venda (δ) e de compra (λ) da primeira rodada de simulações, é indicada a realização do teste de Levene. As Tabelas 11 a 16 exibem os resultados obtidos com o teste de Levene.

Tabela 11 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.05$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.01] - \lambda[0.01]$	19.521	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.05] - \lambda[0.05]$	133.55	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.01] - \delta[0.05]$	18044	< 0.0001*	H_1 (Var. Heterogêneas)
$\lambda[0.01] - \lambda[0.05]$	12647	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.01] - \lambda[0.05]$	9739.7	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.05] - \lambda[0.01]$	26239	< 0.0001*	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Tabela 12 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.10$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.01] - \lambda[0.01]$	19.521	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.10] - \lambda[0.10]$	183.28	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.01] - \delta[0.10]$	18423	< 0.0001*	H_1 (Var. Heterogêneas)
$\lambda[0.01] - \lambda[0.10]$	10780	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.01] - \lambda[0.10]$	8452.3	< 0.0001*	H_1 (Var. Heterogêneas)
$\delta[0.10] - \lambda[0.01]$	26817	< 0.0001*	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Tabela 13 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.01$ vs $\sigma = 0.30$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.01] - \lambda[0.01]$	19.521	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.30]$	41010	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.01] - \delta[0.30]$	19002	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\lambda[0.01] - \lambda[0.30]$	2011.8	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.01] - \lambda[0.30]$	1197.8	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.01]$	27844	$< 0.0001^*$	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Tabela 14 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.05$ vs $\sigma = 0.10$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.05] - \lambda[0.05]$	133.55	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.10] - \lambda[0.10]$	183.28	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.05] - \delta[0.10]$	20.069	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\lambda[0.05] - \lambda[0.10]$	2.6403	0.1042	H_0 (Var. Homogêneas)
$\delta[0.05] - \lambda[0.10]$	152.14	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.10] - \lambda[0.05]$	166.98	$< 0.0001^*$	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Tabela 15 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.05$ vs $\sigma = 0.30$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.05] - \lambda[0.05]$	133.55	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.30]$	41010	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.05] - \delta[0.30]$	85.097	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\lambda[0.05] - \lambda[0.30]$	8784.2	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.05] - \lambda[0.30]$	35732	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.05]$	196.3	$< 0.0001^*$	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Tabela 16 – Teste de homogeneidade das variâncias (Levene) com δ vs λ para $\sigma = 0.10$ vs $\sigma = 0.30$ nos RWs.

Parâmetro	Estatística	p -valor	Hipótese Aceita
$\delta[0.10] - \lambda[0.10]$	183.28	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.30]$	41010	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.10] - \delta[0.30]$	15.441	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\lambda[0.10] - \lambda[0.30]$	8784.2	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.10] - \lambda[0.30]$	37227	$< 0.0001^*$	H_1 (Var. Heterogêneas)
$\delta[0.30] - \lambda[0.10]$	196.3	$< 0.0001^*$	H_1 (Var. Heterogêneas)

Nota: O símbolo * indica teste significativo ao nível de significância de 5%.

Através do teste de Levene foi verificada a não homogeneidade de variância das amostras para a versão de *Random Walks*, exceto quando comparamos os limiares de compra (λ) para $\sigma = 0.05$ e $\sigma = 0.10$. Nesse caso, não temos evidências para rejeitar a hipótese de igualdade de variâncias nesses *drifts*. O que caracteriza dispersão semelhante nos dados dessas séries.

A Tabela 17 mostra os resultados dos testes de Mann-Kendall (*Rank Test*) e de Levene para δ e λ em cada variação de σ no RW.

Tabela 17 – Teste de tendência e homogeneidade para as médias dos limiares δ e λ de acordo com os RWs.

σ	Limiaries	Teste	Estatística	p -valor	Hipótese Aceita
$\sigma = 0.01$	δ	Rank	0.76716	0.443	H_0 (Sem tendência)
	λ	Rank	-0.51739	0.6049	H_0 (Sem tendência)
	$\delta \times \lambda$	Levene	1.3554	0.2491	H_0 (Var. homogêneas)
$\sigma = 0.05$	δ	Rank	-0.66012	0.5092	H_0 (Sem tendência)
	λ	Rank	-1.7306	0.0835	H_0 (Sem tendência)
	$\delta \times \lambda$	Levene	2.0105	0.1616	H_0 (Var. homogêneas)
$\sigma = 0.10$	δ	Rank	1.124	0.261	H_0 (Sem tendência)
	λ	Rank	1.088	0.2765	H_0 (Sem tendência)
	$\delta \times \lambda$	Levene	0.81346	0.3708	H_0 (Var. homogêneas)
$\sigma = 0.30$	δ	Rank	-0.303	0.7617	H_0 (Sem tendência)
	λ	Rank	0.7315	0.4645	H_0 (Sem tendência)
	$\delta \times \lambda$	Levene	0.0139	0.9066	H_0 (Var. homogêneas)

Com essa tabela, observamos estatisticamente que todas as séries dos limiares não apresentam tendência. Além disso, ao não rejeitarmos a hipótese nula no teste de variância, podemos assumir igualdade de variâncias entre elas.

Em geral, foi observado que apesar do agente mais pobre possuir a quantidade de capital maior que a quantidade do agente mais rico, como acontece o inverso com a quantidade de ações, isto é, o agente mais rico detêm mais ações do que o menos rico, então há uma compensação e a quantidade de riqueza medida pelo *fitness* da partícula, torna-se aproximadamente igual para ambos os agentes em todos os casos de *drifts* com tendência de baixa no mercado (os *drifts* negativos). Observe a ilustração na Figura 26.

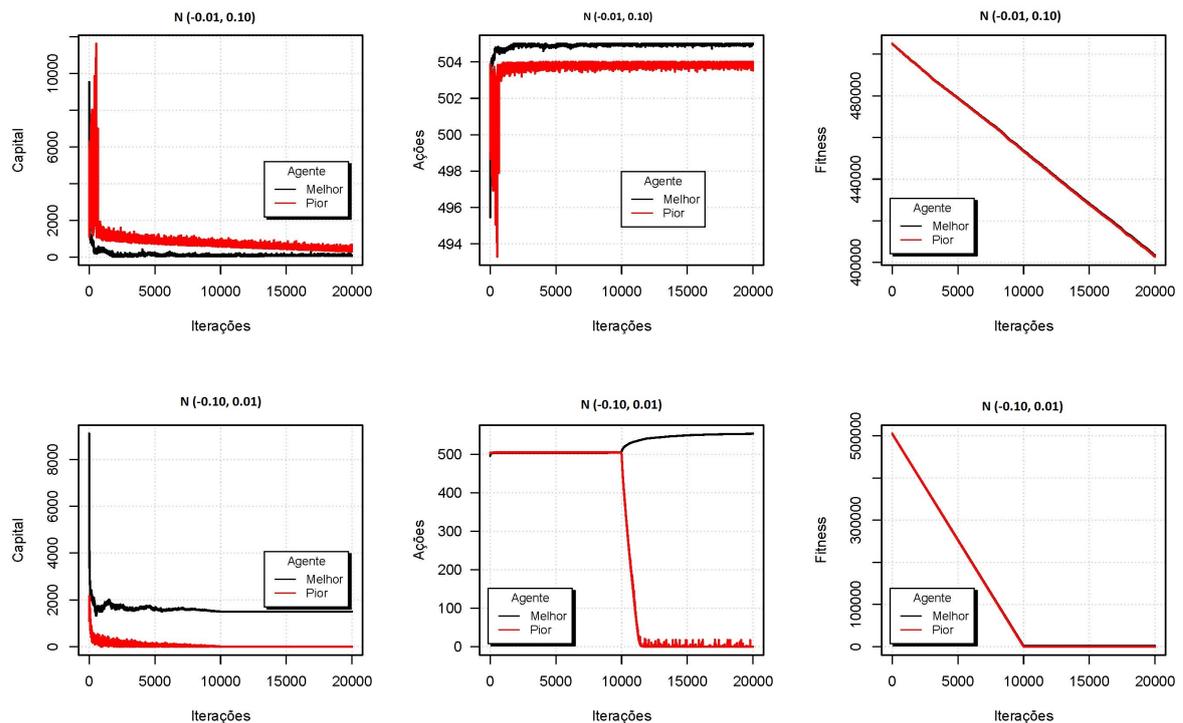


Figura 26 – Ilustração de exemplos do capital, da quantidade de ações e do *fitness* para dois mercados com tendência de baixa.

Para os casos que simulam mercado em alta (com *drift* positivo), todos eles apresentaram comportamento similar. Em contrapartida, a quantidade de capital do agente menos rico é superior a quantidade daquele que possui maior *fitness* (a melhor partícula). Veja Figura 27.

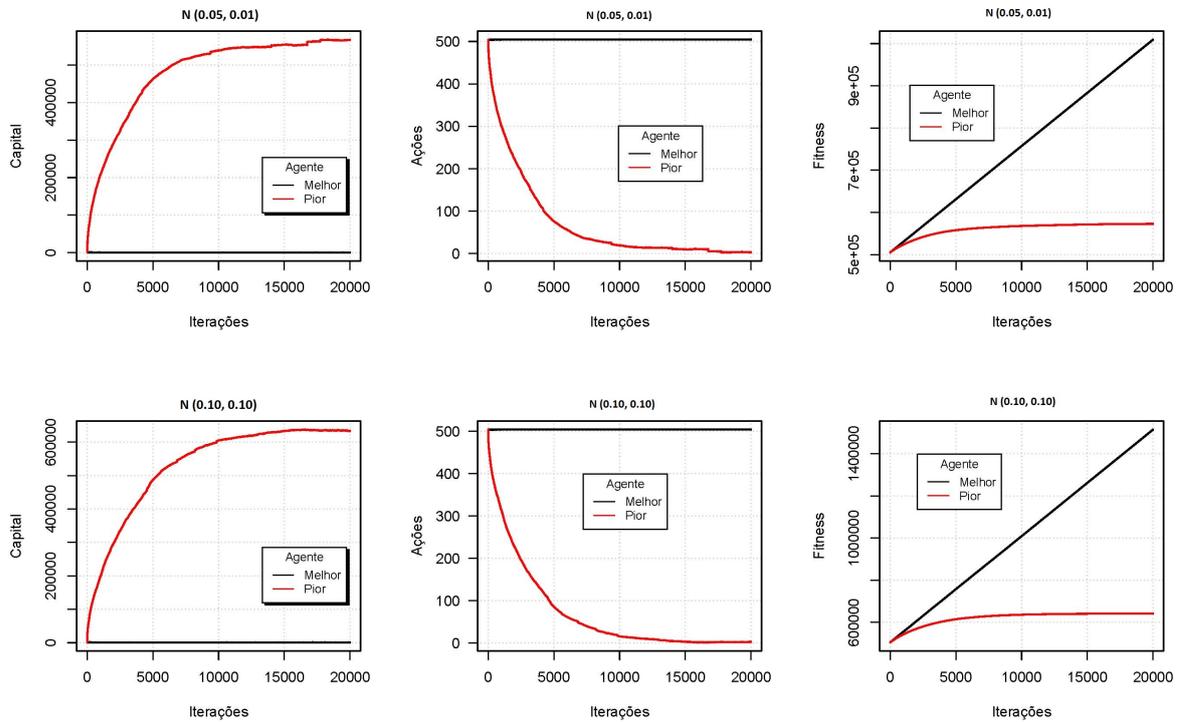


Figura 27 – Ilustração de exemplos do capital, quantidade de ações e *fitness* para dois mercados com tendência de alta.

No caso dos RW's, a quantidade de capital do agente mais pobre é sempre maior. Porém, como ele tem menos ações do que o agente mais rico e essas possuem seus preços sofrendo valorização a cada rodada de negociação, então a riqueza total do agente mais rico é sempre maior, conforme exemplificado na Figura 28. Isso sugere, que essa seja uma boa representação do comportamento de um ativo com boa liquidez, onde as negociações acontecem e os bons preços de mercado desses ativos auxiliam no seu processo de venda, tornando-se mais fácil de ser realizado.

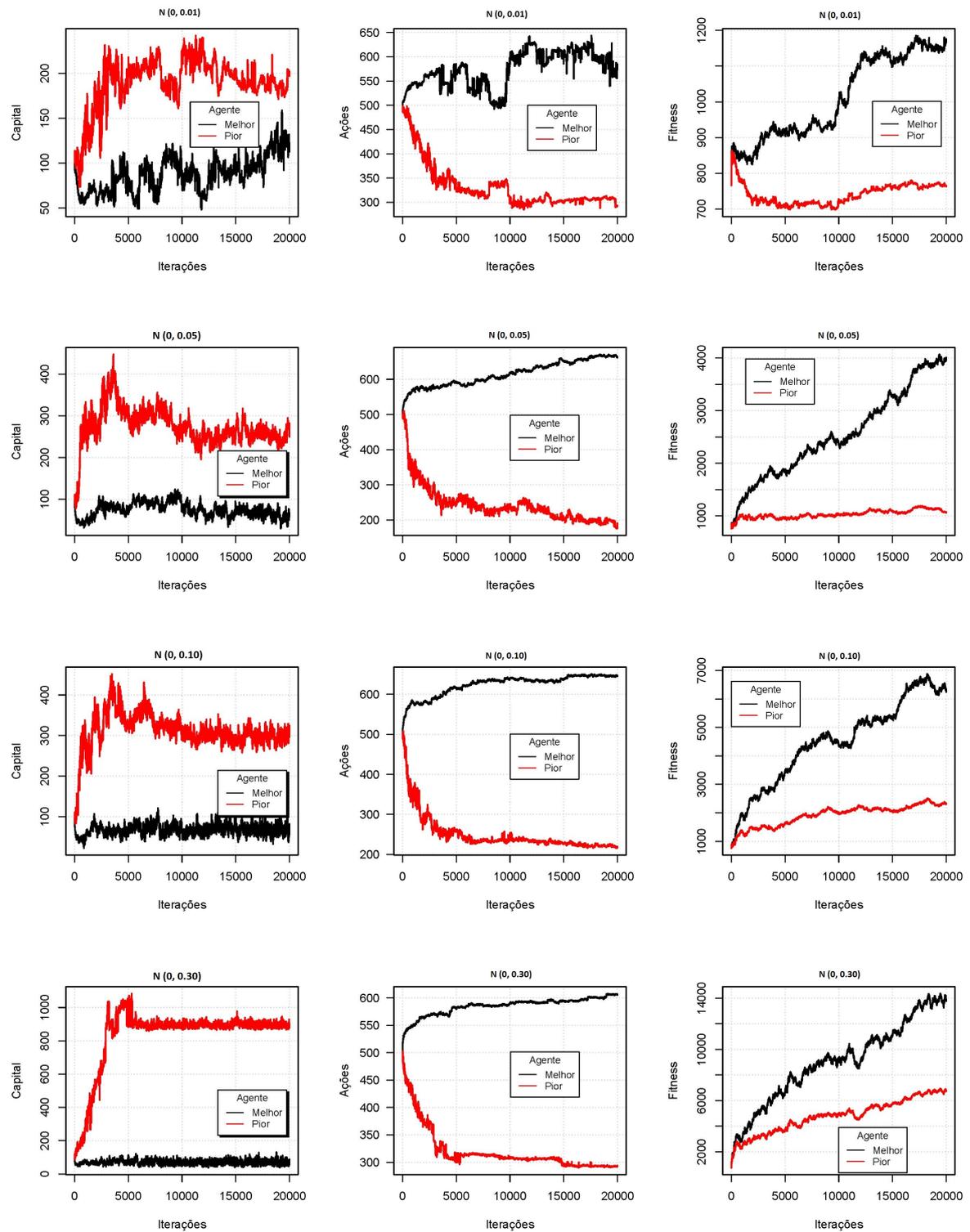


Figura 28 – Ilustração da evolução do capital, da quantidade de ações e do *fitness* para os mercados com RW.

Observando as Tabelas 18 e 19 e analisando a distribuição de riqueza entre os agentes financeiros na última rodada de negociações, através do Coeficiente de Gini, percebemos que seus valores são bem próximos (parecidos), com exceção dos casos em que simulamos mercados com tendência de baixa, a partir de $\theta = -0.05$ e $\theta = -0.10$. Neles, o valor correspondente ao coeficiente aumentou bastante, indicando um grande desequilíbrio na distribuição de riqueza. Nos demais casos há uma boa regularidade na forma como ela foi distribuída e não observa-se concentração da riqueza com apenas um único agente.

Tabela 18 – Valores médios para δ , λ , *fitness* e índice de Gini (\bar{g}) em relação aos melhores e piores agentes em 30 simulações de cada *drift* negativo.

drift	$\bar{\delta}_{melhor}$	$\bar{\lambda}_{melhor}$	\bar{Fit}_{melhor}	$\bar{\delta}_{pior}$	$\bar{\lambda}_{pior}$	\bar{Fit}_{pior}	\bar{g}
N(-0.01,0.01)	0.0283	0.0145	496144.9300	0.0063	0.9460	403870.4433	0.0024
N(-0.01,0.05)	0.0267	0.0155	495547.9500	0.0082	0.9456	403106.0167	0.0024
N(-0.01,0.10)	0.0301	0.0151	494143.2167	0.0061	0.9378	402801.3733	0.0024
N(-0.05,0.01)	0.0165	0.0214	460137.9100	0.0241	0.9137	401.2105	0.9202
N(-0.05,0.05)	0.0324	0.0172	455637.6067	0.0221	0.9000	1968.1346	0.7592
N(-0.05,0.10)	0.0190	0.0206	457088.4100	0.0179	0.8870	3582.0061	0.7215
N(-0.10,0.01)	0.0281	0.0154	420479.1900	0.0087	0.5837	0.0141	0.9758
N(-0.10,0.05)	0.0216	0.0166	422485.5733	0.0092	0.5845	0.0158	0.9769
N(-0.10,0.10)	0.0209	0.0165	418105.4400	0.0063	0.6556	0.0154	0.9744

Tabela 19 – Valores médios para δ , λ , *fitness* e índice de Gini (\bar{g}) em relação aos melhores e piores agentes em 30 simulações de cada *drift* positivo.

drift	$\bar{\delta}_{melhor}$	$\bar{\lambda}_{melhor}$	\bar{Fit}_{melhor}	$\bar{\delta}_{pior}$	$\bar{\lambda}_{pior}$	\bar{Fit}_{pior}	\bar{g}
N(0.01,0.01)	0.5621	0.2116	606143.8900	0.0234	0.0155	515348.6300	0.0015
N(0.01,0.05)	0.5459	0.1886	606497.2567	0.0299	0.0133	515521.7833	0.0015
N(0.01,0.10)	0.5259	0.2039	606956.5533	0.0368	0.0142	517422.2500	0.0015
N(0.05,0.01)	0.7178	0.4190	1009484.0333	0.0372	0.0127	572912.5933	0.0039
N(0.05,0.05)	0.7073	0.3663	1009689.3000	0.0289	0.0140	574272.4833	0.0039
N(0.05,0.10)	0.7027	0.3890	1008534.3033	0.0319	0.0143	569402.5800	0.0039
N(0.10,0.01)	0.7864	0.4840	1512954.2667	0.0284	0.0106	641523.6300	0.0052
N(0.10,0.05)	0.7668	0.4357	1513982.3333	0.0319	0.0151	641269.3833	0.0052
N(0.10,0.10)	0.7649	0.4603	1514467.6333	0.0362	0.0128	642182.2400	0.0052

Nota-se que para mercados em baixa, os limiares dos melhores agentes são muito pequenos em comparação aos limiares do mercado em alta. Isso sugere que no mercado em baixa, os melhores agentes são mais conservadores. Já os piores agentes, tendem a comprar mais caro e vender barato.

Por outro lado, quando o mercado está em alta, os melhores agentes vendem caro e compram barato. Além disso, os piores agentes apresentam os limiares de venda e compra muito baixos.

Com essas observações, aparentemente, a média usada na geração dos preços das ações negociadas no mercado influencia no modo como a riqueza é distribuída entre os indivíduos que participam do mesmo. Na Figura 29, ilustramos essa distribuição para um caso em que há concentração de riqueza e para um caso em que não há.

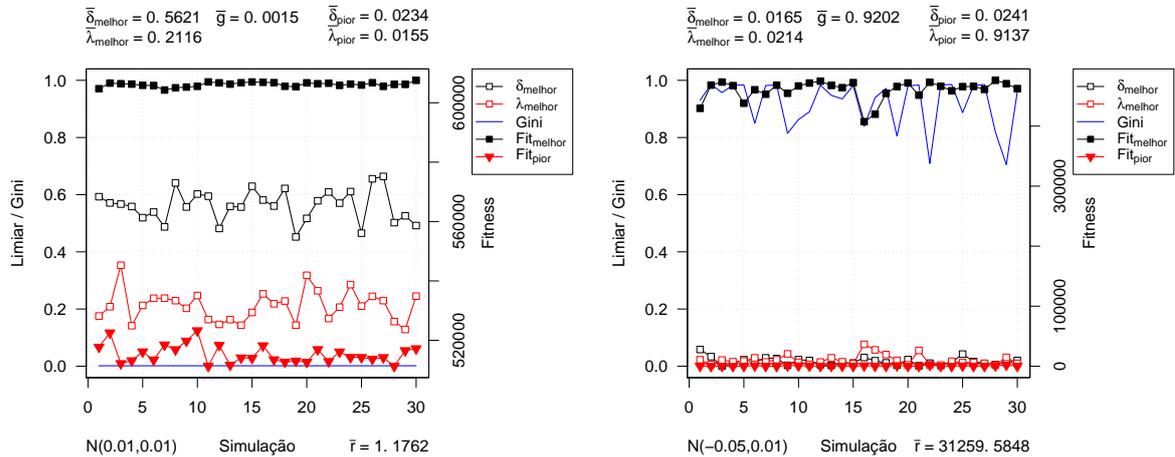


Figura 29 – Ilustração de exemplos da distribuição de riqueza com o Coeficiente de Gini.

A fim de ilustrarmos a movimentação da melhor partícula no espaço de busca em função dos seus limiares de venda (δ) e de compra (λ) a cada rodada de negociações, geramos o plot 3D apresentado a seguir. As esferas azuis indicam a posição inicial na movimentação, enquanto as esferas vermelhas indicam o posição final.

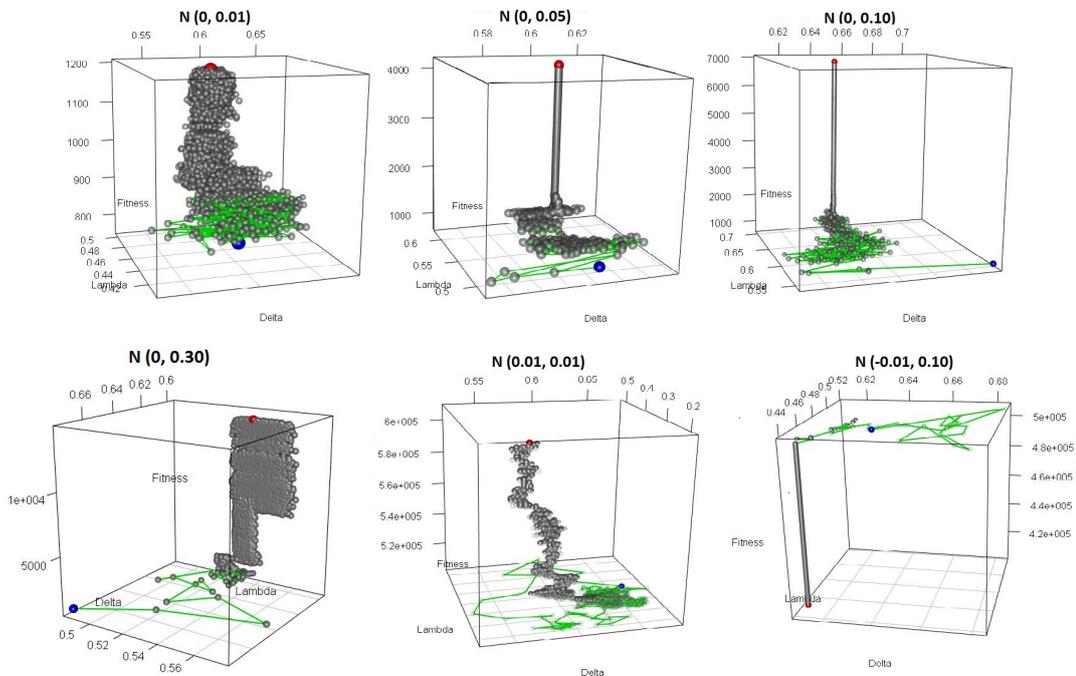


Figura 30 – Plot 3D com a movimentação da melhor partícula em função dos limiares de venda (δ), de compra (δ) e da quantidade de riqueza (*fitness*) no caso de alguns mercados simulados.

4.2 Aplicação em Mercados Reais

Obtivemos as séries de retornos dos preços das ações para o mercado real de acordo com a Equação 4. Essas séries são ilustradas na Figura 31. Em seguida, demos início à verificação dos fatos estilizados da economia mais comuns na literatura.

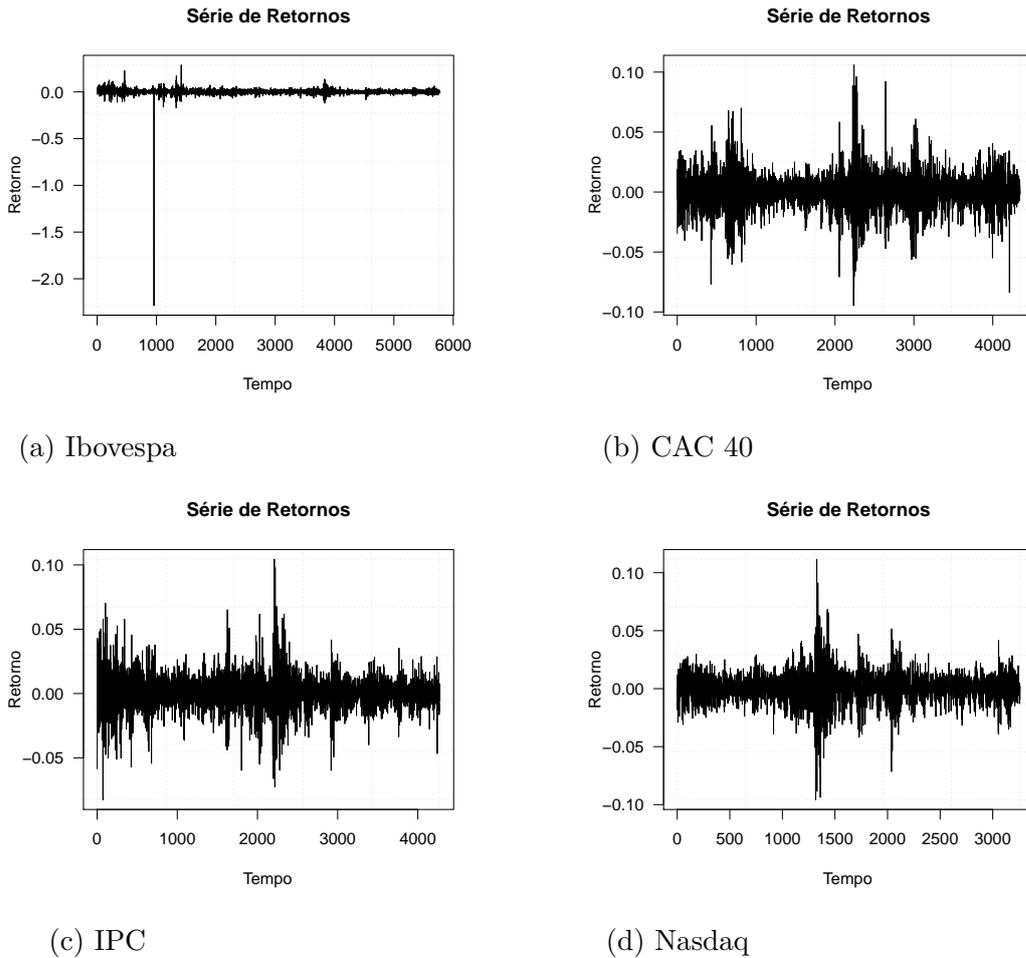
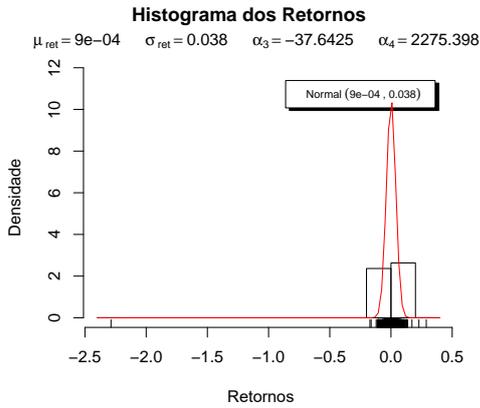
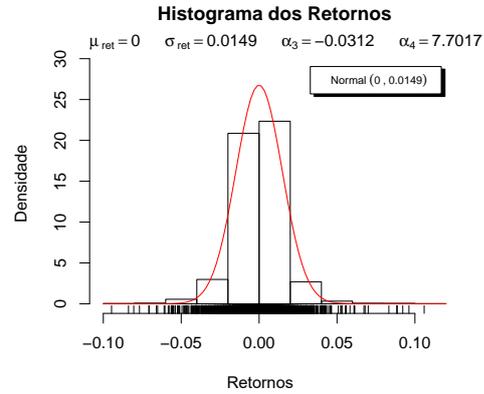


Figura 31 – Séries de retornos dos preços dos mercados reais.

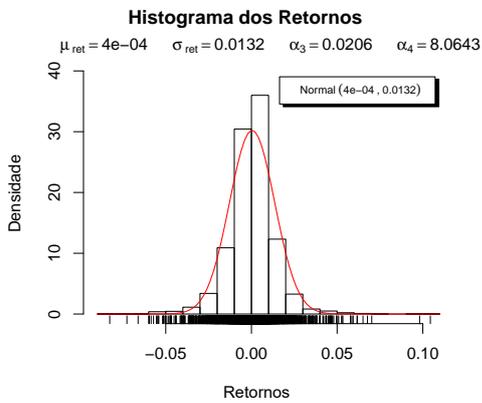
Como pode-se observar na Figura 32, com o terceiro e o quarto momentos amostrais, α_3 e α_4 respectivamente, todas as séries de dados apresentam excesso de curtose e assimetria. Sugerindo que as mesmas não poderiam ser modeladas por uma distribuição gaussiana. As séries apresentam distribuição leptocúrtica, bem como o verificado no trabalho de [15].



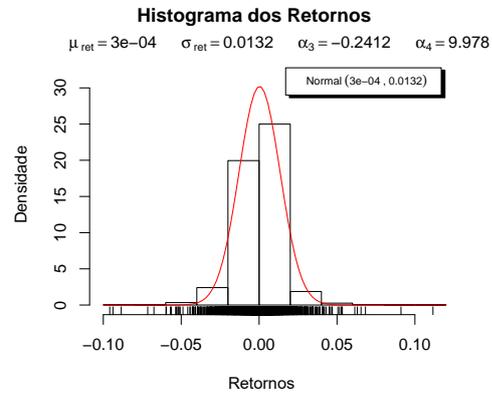
(a) Ibovespa



(b) CAC 40



(c) IPC



(d) Nasdaq

Figura 32 – Histograma das séries de retornos dos preços dos mercados reais.

Através da função de autocorrelação dos retornos, comprovamos a independência nas séries. Observa-se nas ilustrações da Figura 33 que os valores da FAC são nulos ou muito próximos de zero para as pequenas *lags*, caracterizando ausência de autocorrelação linear no conjunto de dados.

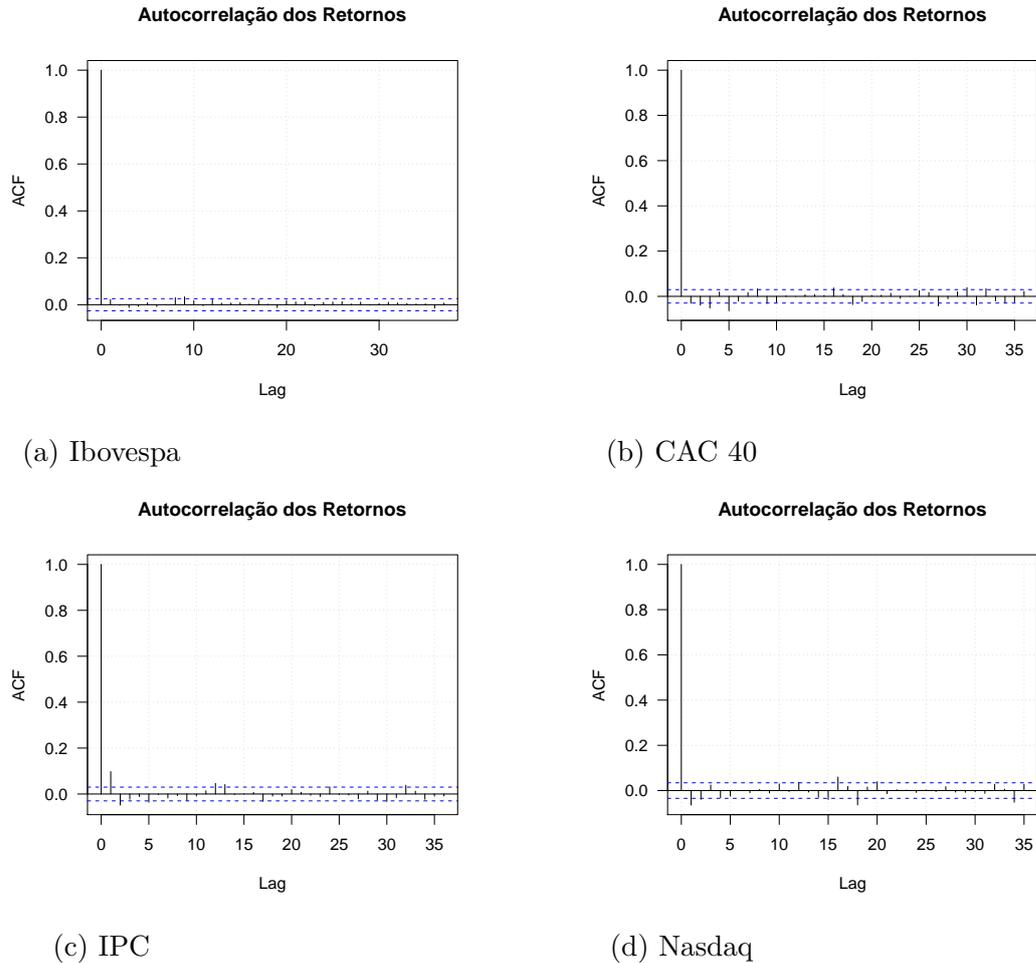
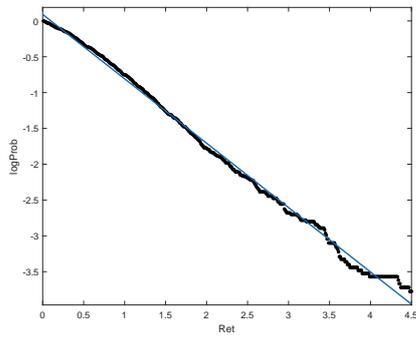
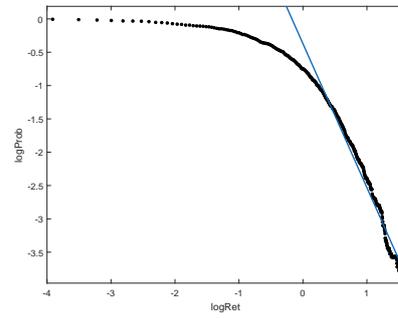


Figura 33 – Autocorrelações das séries de retornos dos preços dos mercados reais.

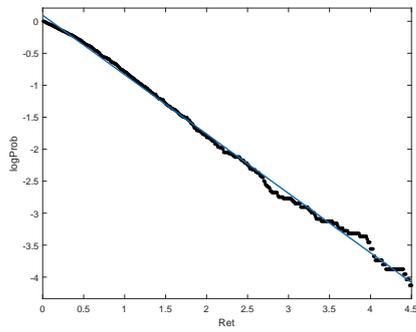
Com as funções linear, exponencial e lei de potência, realizamos o ajuste das distribuições com as séries dos retornos normalizadas pela Equação 5. Na Figura 34, observamos que a modelagem pela função linear na escala semi-log, correspondente ao ajuste pela função exponencial, obteve um ajuste com toda a série de dados. Por outro lado, ao utilizarmos a escala log-log na modelagem, que corresponde ao ajuste pela lei de potência, se considerarmos somente os dados que formam a cauda da distribuição, a aproximação apresentará melhores resultados, corroborando com o verificado em [17, 45]. A linha azul equivale ao ajuste realizado, enquanto os pontos pretos representam os dados utilizados.



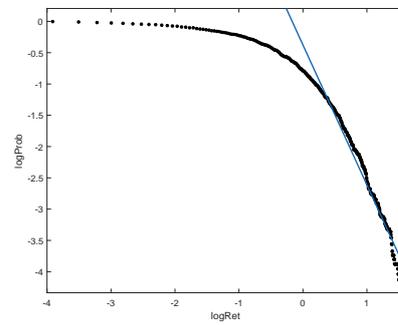
(a)



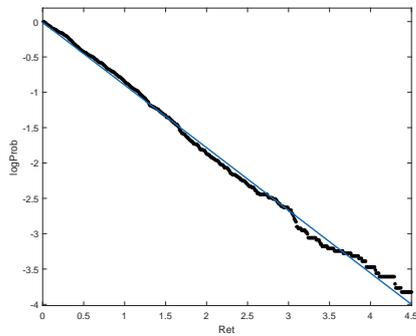
(b)



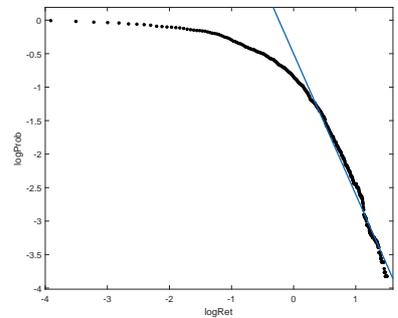
(c)



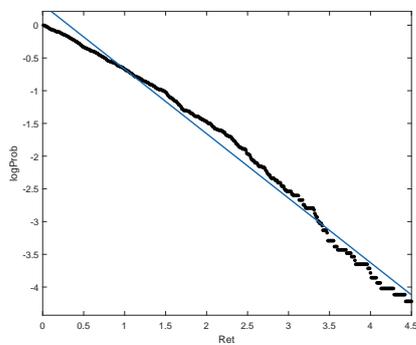
(d)



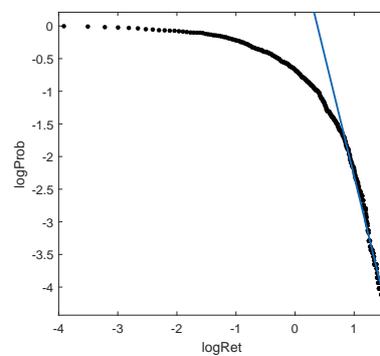
(e)



(f)



(g)



(h)

Figura 34 – Ajustes lineares nas escalas semi-log (CAC 40 (a), Ibovespa (c), IPC (e) e Nasdaq (g)) e log-log (CAC 40 (b), Ibovespa (d), IPC (f) e Nasdaq (h)).

Conforme pode-se observar na Figura 35, em analogia ao realizado nos mercados simulados, ao efetuarmos a aproximação pela lei de potência e pela exponencial sem transformação nas escalas, a curva da exponencial aderiu melhor à curva gerada pelo conjunto de dados. Já a lei de potência, exibiu melhor modelagem somente à cauda da fdp.

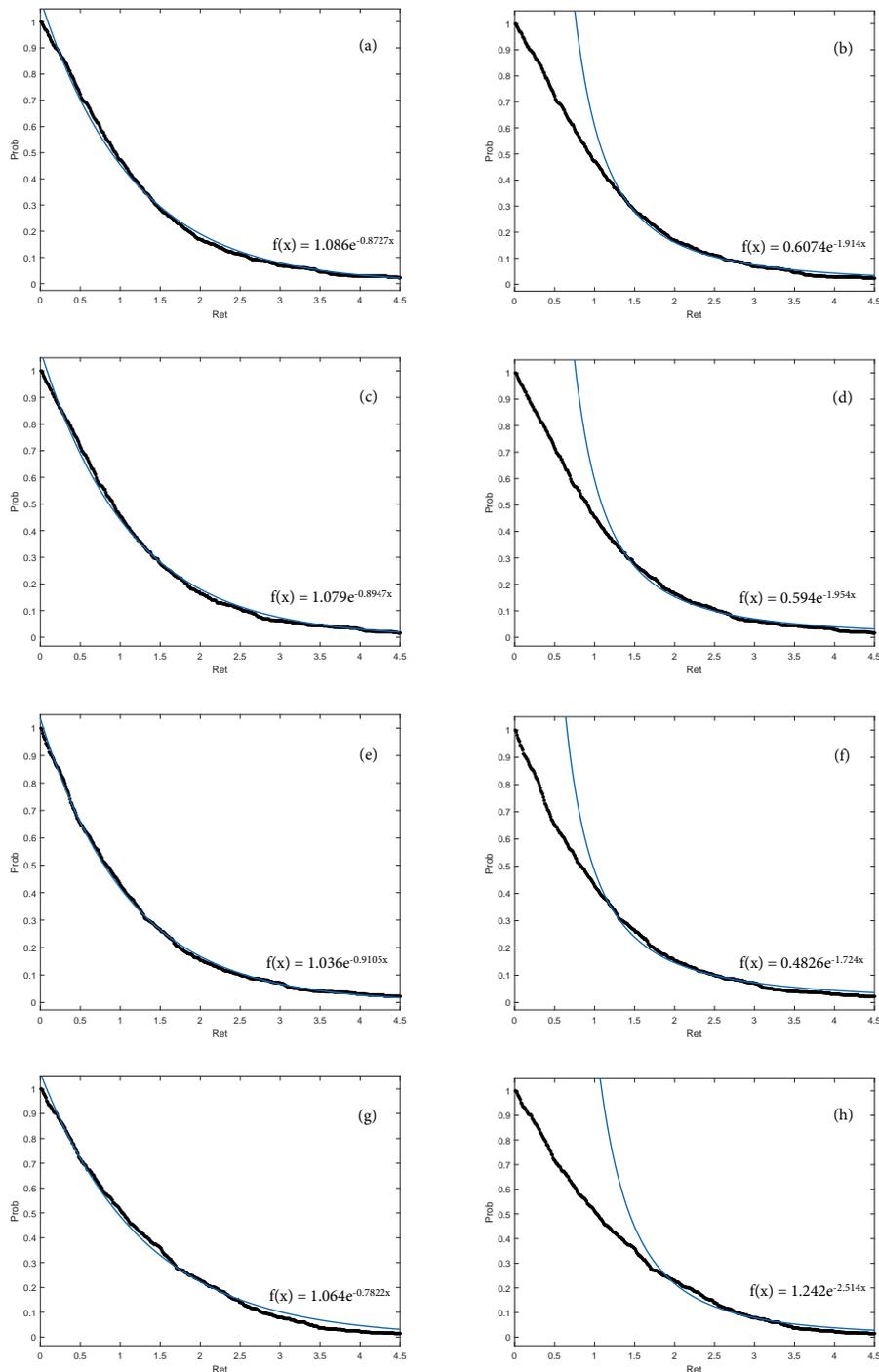


Figura 35 – Ajustes pela função exponencial (CAC 40 (a), Ibovespa (c), IPC (e) e Nasdaq (g)) e lei de potência (CAC 40 (b), Ibovespa (d), IPC (f) e Nasdaq (h)).

Ao analisarmos os limiares de compra e de venda gerados nos experimentos com as séries reais, notamos que tanto para a melhor partícula como para a pior, os valores médios são bem próximos (Figura 36). Indicando que as partículas encontravam-se bem posicionadas nas simulações e não há grandes divergências entre os comportamentos dos piores e dos melhores agentes financeiros nas iterações do algoritmo.

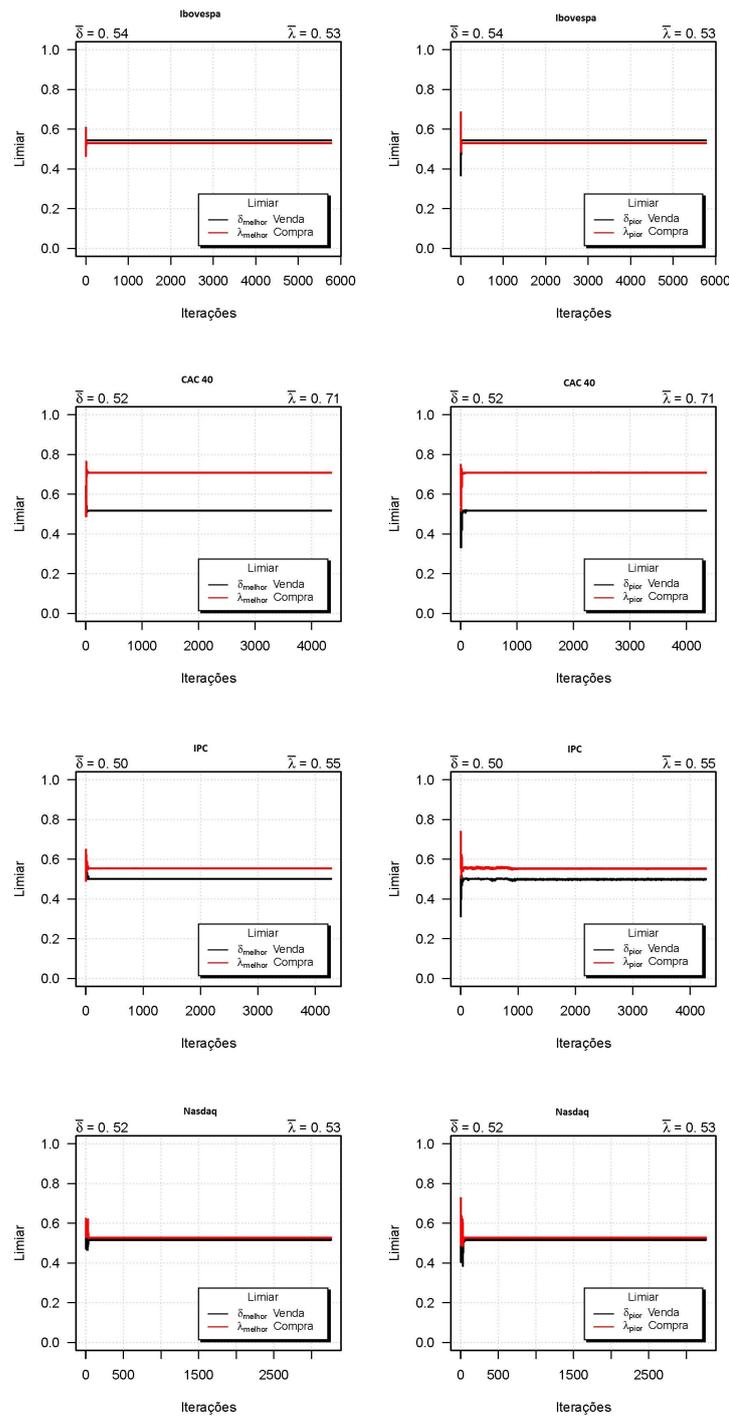


Figura 36 – Limiar de venda (δ) e de compra (λ) das melhores partículas nas iterações.

Observa-se ainda, que somente para o mercado do índice Bovespa o limiar de venda foi maior do que o limiar de compra. Nesse caso, esse fato sugere que para o período analisado, as ações foram compradas em momentos de baixa no mercado e estiveram sendo vendidas em momentos de alta, ou seja, quando houve valorização nos valores das mesmas.

Para os demais casos, apesar das pequenas oscilações observadas nos preços dos ativos nos mercados, o limiar de compra se sobrepôs. Com isso, em geral, as ações estavam sendo compradas em momentos de alta e vendidas em períodos de baixa nos mercados, sugerindo desvalorização em seus preços.

Podemos observar nas Figuras 37 e 38 que para todos os mercados reais, a quantidade de ações da melhor partícula foi superior à quantidade da pior. No caso do índice Bovespa, o volume de capital do melhor e do pior agente financeiro foi bem parecido, enquanto para os demais mercados, a pior partícula concentrou maior quantidade de capital. Ao analisarmos o montante de riqueza dessas partículas, observamos que conforme esperado, a melhor partícula concentra mais riqueza e tanto ela como a pior apresentaram valores próximos.

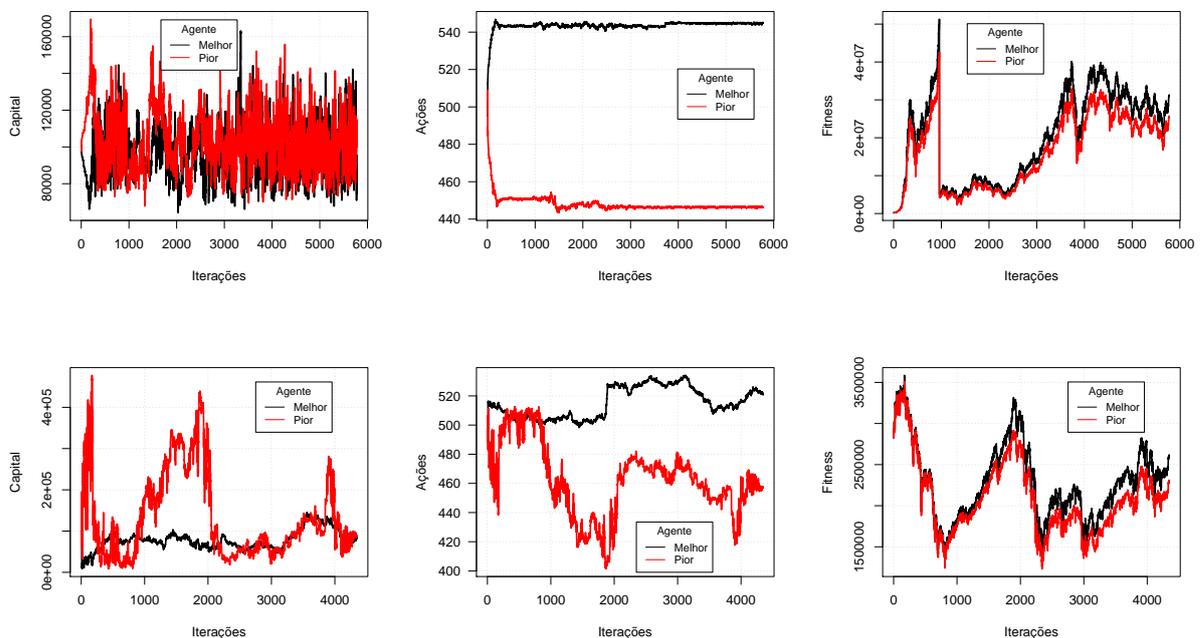


Figura 37 – Ilustração da quantidade de capital das séries de retornos dos preços dos índices Ibovespa e CAC 40, respectivamente.

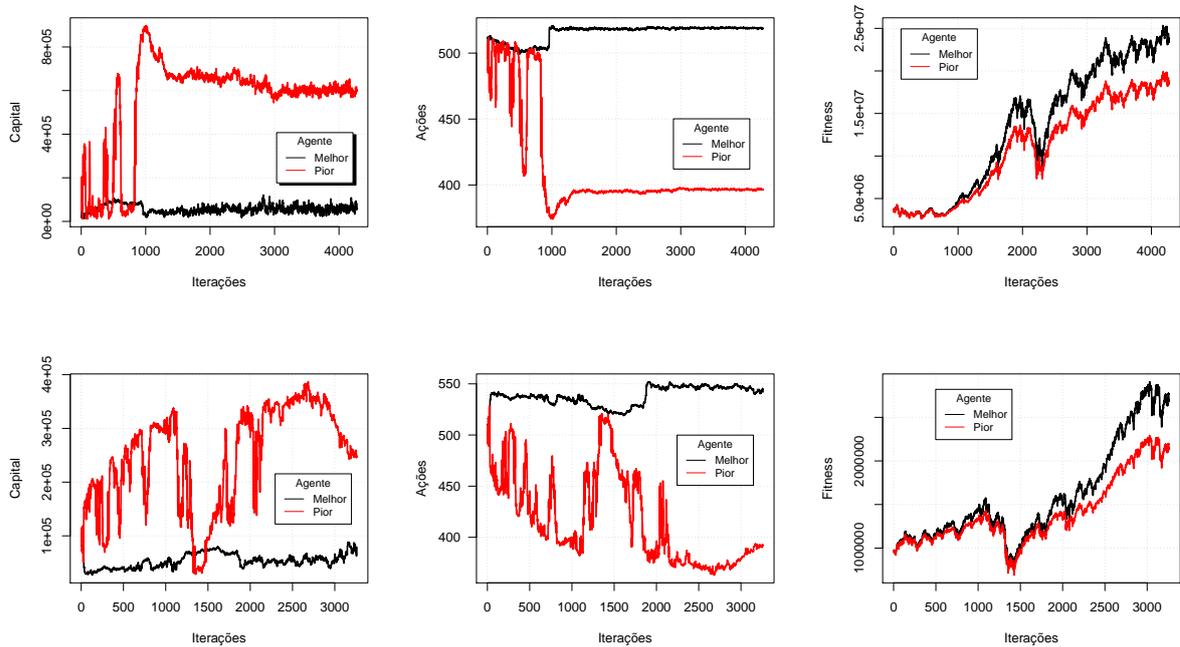


Figura 38 – Ilustração da quantidade de capital das séries de retornos dos preços dos índices IPC e Nasdaq, respectivamente.

A Tabela 20 exhibe os resultados obtidos ao mensurarmos a distribuição de riqueza (\bar{g}) entre os agentes financeiros através do coeficiente de Gini. Para todos os mercados reais, os resultados apontam para uma distribuição de riqueza bem igualitária entre os agentes. Ilustramos na Figura 39, os valores encontrados para essa medida, para os limiares de venda e de compra da melhor e da pior partícula, além da quantidade de riqueza (*fitness*) dos agentes para a última rodada de negociações no PSO.

Tabela 20 – Coeficiente de Gini para os mercados reais

Mercado	$\bar{\delta}_{melhor}$	$\bar{\lambda}_{melhor}$	\overline{Fit}_{melhor}	$\bar{\delta}_{pior}$	$\bar{\lambda}_{pior}$	\overline{Fit}_{pior}	\bar{g}
Bovespa	0.5431	0.5292	31412135.8667	0.5431	0.5292	1.2570	0.0173
CAC40-Franca	0.5167	0.7082	2796226.4000	0.5170	0.7082	1.2168	0.0145
IPC-Mexico	0.5011	0.5536	23744817.8667	0.4998	0.5527	18700031.4333	0.0082
Nasdaq	0.5164	0.5267	2681465.7333	0.5164	0.5267	2131066.7333	0.0139

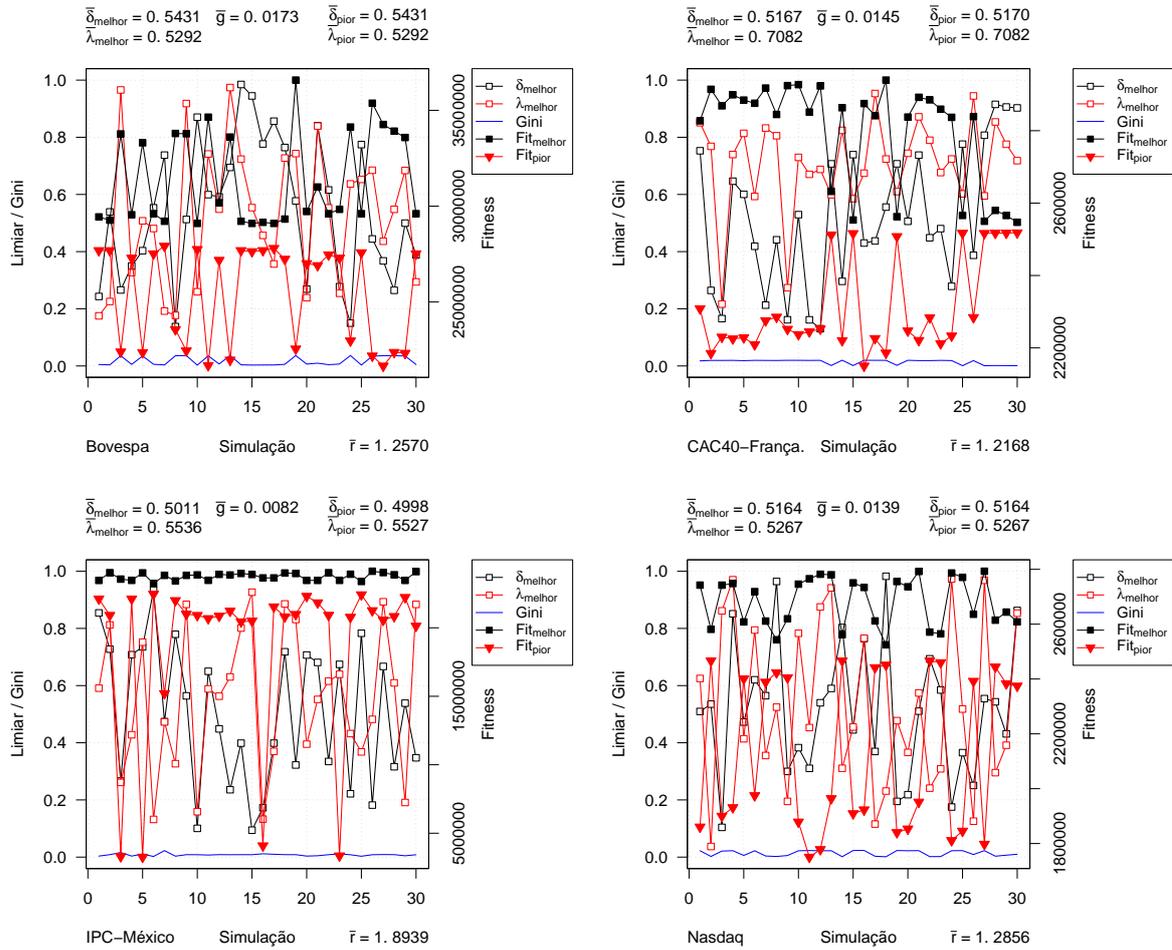
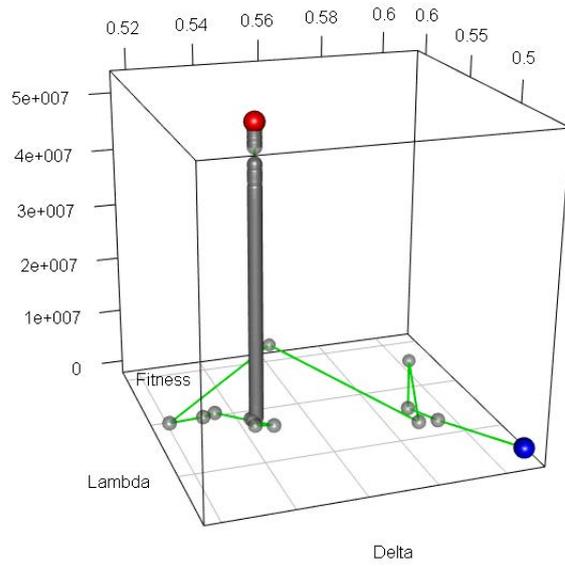
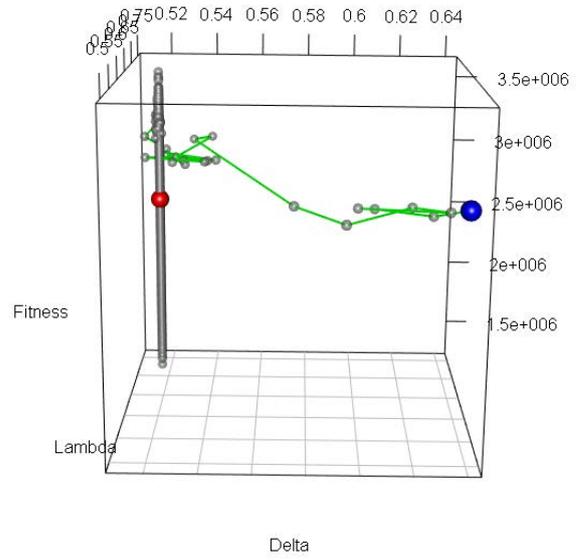


Figura 39 – Coeficiente de Gini da quantidade de *fitness* (riqueza) nos mercados reais.

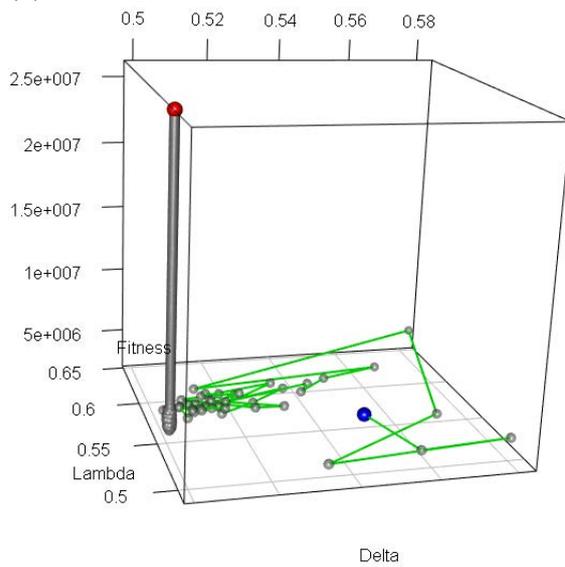
Como mostra na Figura 40, ilustramos o espaço de busca do algoritmo indicando a movimentação da melhor partícula para as séries reais adotadas nesse trabalho. Em que delta representa o limiar de venda (δ), lambda o limiar de compra (λ) e *fitness* a quantidade de riqueza dessas partículas. As esferas azuis indicam a posição inicial na movimentação, já as esferas vermelhas apontam a posição final.



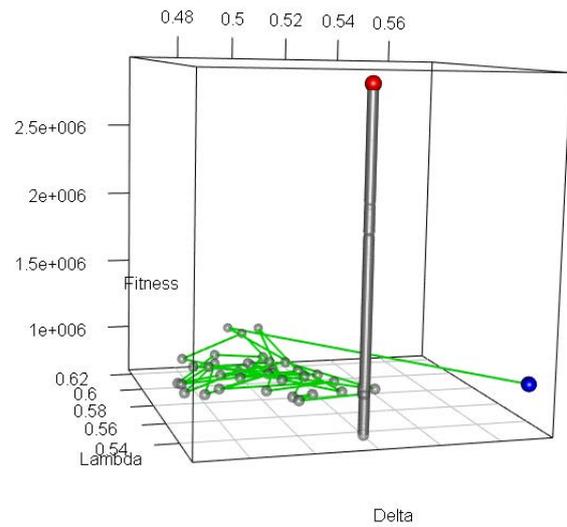
(a) Ibovespa



(b) CAC 40



(c) IPC



(d) Nasdaq

Figura 40 – Plot 3D com a movimentação da melhor partícula em função dos limiares de venda (δ), de compra (δ) e da quantidade de riqueza (*fitness*) no caso dos mercados reais.

5 Conclusões

Nosso estudo teve como objetivo simular um mercado financeiro em três situações gerais e realizar a modelagem observando as principais abordagens utilizadas nas séries financeiras. A primeira rodada de experimentos consistiu em mercados do tipo *Random Walk* (RW), enquanto a segunda simulou mercados em baixa, quando há desvalorização nos preços das ações seguindo um *drift* com distribuição gaussiana de média negativa, e mercados em alta, onde há valorização nos preços das ações de acordo com um *drift* com distribuição gaussiana de média positiva. Posteriormente, utilizamos séries de preços reais e investigamos o comportamento dos mercados gerados.

Os mercados financeiros foram construídos tendo como principal característica a interação entre as partículas que o compõem. Para implementação do ambiente artificial, utilizamos o algoritmo *Particle Swarm Optimization* (PSO), que vem sendo bastante aplicado a diversos problemas de busca e otimização, permitindo a troca de informações entre as partículas.

As séries de preços geradas foram analisadas e através das suas respectivas séries de retornos comprovamos a existência de fatos estilizados da economia em nossos dados, tais como assimetria, excesso de curtose e ausência de autocorrelação linear. Dados os fatos estilizados dos mercados artificiais, ao realizarmos comparações com os observados nos mercados reais podemos validar o ambiente artificial construído.

A partir da normalização das séries de preços obtidas, modelamos suas fdp's pelo ajuste linear nas escalas semi-log e log-log e pelas funções exponencial e lei de potência. Verificamos que no ajuste linear, a escala semi-log, correspondente a um ajuste exponencial, obteve melhores resultados aos comparados à escala log-log, que corresponde ao ajuste pela lei de potência. Ao investigarmos as séries normalizadas sem transformações nas escalas, a função exponencial aderiu melhor aos dados do que a lei de potência. Observou-se que, conforme mencionado em diversos trabalhos na literatura, a aproximação pela função exponencial obtém resultados consideráveis ao realizarmos o ajuste em toda a série de dados, enquanto a aproximação realizada pela lei de potência concentra seu melhor ajuste somente aos dados que compõem a cauda da distribuição. Esse fato foi verificado tanto para as séries simuladas, como para as séries reais.

Além disso, estudamos os limiares de venda (δ) e de compra (λ) que integram o ambiente artificial. Notamos que, em geral, o algoritmo nos forneceu o resultado desejado:

através de limiares de venda maiores do que limiares de compra, visualizamos mercados em que os ativos estariam sendo comprados em períodos de baixa e vendidos em períodos de alta. Por outro lado, quando obtivemos limiares de compra maiores do que os limiares de venda, tínhamos exemplos de mercados que sugerem que os ativos foram comprados em momentos de alta e vendidos em momentos de baixa, isto é, foram comprados quando os mesmos estavam com seus preços sofrendo valorização e vendidos quando houve desvalorização dessas séries nos mercados.

Com um conjunto de regras extremamente simples conseguimos obter as mesmas características de um mercado real, reproduzindo sua dinâmica. Essas regras consistem na utilização dos limiares de venda (δ_i) e de compra (λ_i) como suporte às negociações. A análise dos limiares consiste numa forma que pode ser utilizada para auxiliar na tomada de decisão de investimentos nos mercados financeiros estudados, onde as mesmas baseiam-se em regras simples de comportamento dos agentes.

No mercado Ibovespa, o agente financeiro que obteve sucesso com riqueza total superior a dos demais agentes possui, em média, limiar de venda igual a 54% e de compra igual a 53% em relação ao mercado. Isso sugere, que para obtermos sucesso ao investir no índice Bovespa os agentes financeiros devem ser mais conservadores, limitando-se a comprar as ações somente em períodos de baixa e vendendo-as somente em períodos de alta. Portanto, deve ter seus limiares de venda e de compra próximo aos gerados com as simulações.

Porém, nos demais mercados estudados (Nasdaq, IPC e CAC 40) os agentes mostraram-se mais agressivos, aceitando vender as ações em períodos de baixa e comprar em períodos de alta. Com limiares de venda variando de 50% a 52% em relação ao mercado e limiares de compra podendo chegar até 71%, como no caso do índice CAC 40.

Considerando ainda os limiares de venda e de compra, investigamos o comportamento dos melhores e dos piores agentes no mercado, de acordo com seus valores de *fitness*. Através do Coeficiente de Gini mensuramos a distribuição de riqueza entre eles. Verificamos que a mesma encontra-se bem distribuída entre os participantes do mercado na última rodada de negociações, exceto nos casos em que a média do *drift* usado na geração dos preços era de -0.05 e -0.10 . Nesses casos, o valor do Coeficiente de Gini aumentou bastante e nos permitiu concluir que havia concentração de riqueza entre alguns poucos agentes no grupo, não condizendo com a realidade do mercado real. Nos mercados gerados com as séries reais, todos apresentaram uma distribuição de riqueza mais igualitária e tanto o melhor agente, quanto o pior tiveram sua distribuição de riqueza total com valores bem próximos.

5.1 Trabalhos Futuros

Como trabalho futuro, desejamos aperfeiçoar o algoritmo PSO utilizado para as simulações e construir um método de previsão adequado que auxilie na identificação de eventos extremos e, conseqüentemente, na tomada de decisão dos participantes do mercado.

Pode-se utilizar medidas de risco, como o *Value-at-Risk* (VaR) para mensurar a pior perda esperada num mercado, durante um determinado intervalo de tempo e considerando um dado nível de confiança. Com o objetivo de auxiliar na avaliação da distribuição de ganhos ou perdas dos agentes financeiros.

Referências Bibliográficas

- 1 UERJ, F. de E. . **Inteligência de Enxame: PSO**. Disponível em: <<http://www.eng.uerj.br/~nadia/pso.pdf>> .
- 2 BANDA, N.; ENGELBRECHT, A.; ROBINSON, P. Continuous emotion recognition using a particle swarm optimized narx neural network. In: IEEE. **Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on**. [S.l.], 2015. p. 380–386.
- 3 BANK.COM, W. **The World Bank**. Disponível em: <<http://www.worldbank.org/>> .
- 4 BARROS, R. P. d. O.; FOGUEL, M. N. O.; ULYSSEA, G. O. Desigualdade de renda no brasil: uma análise da queda recente. Instituto de Pesquisa Econômica Aplicada (IPEA), v. 1, 2006.
- 5 BARROS, T. d. S. Análise técnica e fundamentalista: Ensaio sobre os métodos de análise. **Revista de Administração e Negócios da Amazônia**, v. 7, n. 2, p. 39–63, 2015.
- 6 BOX, G. E.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2008.
- 7 MORETTIN, P. A.; TOLOI, C. M. de C. **Modelos para previsão de séries temporais**. [S.l.]: Instituto de matemática pura e aplicada, 1981. v. 1.
- 8 KWAPIEŃ, J.; DROŹDŹ, S. Physical approach to complex systems. **Physics Reports**, Elsevier, v. 515, n. 3, p. 115–226, 2012.
- 9 KARAZMODEH, M.; NASIRI, S.; HASHEMI, S. M. Stock price forecasting using support vector machines and improved particle swarm optimization. **Journal of Automation and Control Engineering**, v. 1, n. 2, p. 173–176, 2013.
- 10 MIRAMONTES, O.; VOLKE, K. **Fronteras de la física en el siglo XXI**. [S.l.]: CopIt ArXives, 2013.
- 11 YANG, X.-S. et al. **Swarm intelligence and bio-inspired computation: theory and applications**. [S.l.]: Newnes, 2013.
- 12 MENEZES, S. R. L. de; FREITAS, R. S.; PARPINELLI, R. S. Mineração em grandes massas de dados utilizando hadoop mapreduce e algoritmos bio-inspirados: Uma revisão sistemática mining of massive data bases using hadoop mapreduce and bio-inspired algorithms: A systematic review. 2016.
- 13 PRADO, J. R. do; SARAMAGO, S. F. Otimização por colônia de partículas. **FAMAT em Revista**, p. 87, 2005.

- 14 XIONG, T. et al. Forecasting interval time series using a fully complex-valued rbf neural network with dpso and pso algorithms. **Information Sciences**, Elsevier, v. 305, p. 77–92, 2015.
- 15 CONT, R. Empirical properties of asset returns: stylized facts and statistical issues. Taylor & Francis, 2001.
- 16 BARROS, R. P. d.; HENRIQUES, R.; MENDONÇA, R. A estabilidade inaceitável: desigualdade e pobreza no brasil. Instituto de Pesquisa Econômica Aplicada (Ipea), 2001.
- 17 MATIA, K. et al. Scale-dependent price fluctuations for the indian stock market. **EPL (Europhysics Letters)**, IOP Publishing, v. 66, n. 6, p. 909–914, 2004.
- 18 GLERIA, I.; MATSUSHITA, R.; SILVA, S. D. Scaling power laws in the sao paulo stock exchange. **Economics Bulletin**, v. 7, n. 3, 2002.
- 19 MANTEGNA, R. N. Lévy walks and enhanced diffusion in milan stock exchange. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 179, n. 2, p. 232–242, 1991.
- 20 LEVY, M.; SOLOMON, S. New evidence for the power-law distribution of wealth. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 242, n. 1, p. 90–94, 1997.
- 21 MENESES, A.; MARIANO, F. **Mercado Financeiro**. [S.l.]: Elsevier, 2011.
- 22 SOBRINHO, V.; BERTOLO, B. Matemática financeira. Atlas, 1997.
- 23 MORAIS, M. A evolução da bolsa de valores como instrumento do mercado de capitais. 1989.
- 24 BM&FBOVESPA. **Instituto Educacional BOVESPA**. Disponível em: <http://lojavirtual.bmf.com.br/LojaIE/portal/pages/pdf/Apostila_PQO_Cap_05_V2.pdf>.
- 25 FALCÃO, E. Divulgação em demonstrações financeiras de companhias abertas. **Caderno de estudos**, SciELO Brasil, n. 12, p. 01–13, 1995.
- 26 NETO, P. S. de M. et al. Market volatility modeling for short time window. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 390, n. 20, p. 3444–3453, 2011.
- 27 MANTEGNA, R. N.; STANLEY, H. E. An introduction to econophysics: correlation and complexity in finance. **Cambridge, UK: Cambridge University**, 2000.
- 28 NETO, M. et al. **Um método para análise de mercados de ações utilizando séries temporais de índices financeiros**. Tese (Doutorado) — Universidade Federal de Pernambuco, 2012.
- 29 BONABEAU, E.; DORIGO, M.; THERAULAZ, G. **Swarm intelligence: from natural to artificial systems**. [S.l.]: Oxford university press, 1999.

- 30 KENNEDY, J. Review of engelbrecht's fundamentals of computational swarm intelligence. **Genetic Programming and Evolvable Machines**, Springer, v. 8, n. 1, p. 107–109, 2007.
- 31 EBERHART, R. C.; SHI, Y. **Computational intelligence: concepts to implementations**. [S.l.]: Elsevier, 2011.
- 32 JAMES, K.; RUSSELL, E. Particle swarm optimization. In: **Proceedings of 1995 IEEE International Conference on Neural Networks**. [S.l.: s.n.], 1995. p. 1942–1948.
- 33 EBERHART, R. C.; KENNEDY, J. A new optimizer using particle swarm theory. In: NEW YORK, NY. **Proceedings of the sixth international symposium on micro machine and human science**. [S.l.], 1995. v. 1, p. 39–43.
- 34 CLERC, M. **Particle swarm optimization**. [S.l.]: John Wiley & Sons, 2010. v. 93.
- 35 EBERHART, R. C.; SHI, Y.; KENNEDY, J. **Swarm intelligence**. [S.l.]: Elsevier, 2001.
- 36 COHEN, S. C.; CASTRO, L. N. de. Data clustering with particle swarms. In: IEEE. **2006 IEEE International Conference on Evolutionary Computation**. [S.l.], 2006. p. 1792–1798.
- 37 CHENG, S.; SHI, Y.; QIN, Q. Experimental study on boundary constraints handling in particle swarm optimization: From population diversity perspective. 2013.
- 38 YANG, X.-S. **Engineering optimization: an introduction with metaheuristic applications**. [S.l.]: John Wiley & Sons, 2010.
- 39 PULIDO, M.; MELIN, P.; CASTILLO, O. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the mexican stock exchange. **Information Sciences**, Elsevier, v. 280, p. 188–204, 2014.
- 40 MAJHI, R. et al. Prediction of s&p 500 and djia stock indices using particle swarm optimization technique. In: IEEE. **Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence)**. **IEEE Congress on**. [S.l.], 2008. p. 1276–1282.
- 41 HEGAZY, O.; SOLIMAN, O. S.; SALAM, M. A. Comparative study between fpa, ba, mcs, abc, and pso algorithms in training and optimizing of ls-svm for stock market prediction. **International Journal of Advanced Computer Research**, International Journal of Advanced Computer Research, v. 5, n. 18, p. 35, 2015.
- 42 WANG, F.; PHILIP, L.; CHEUNG, D. W. Complex stock trading strategy based on particle swarm optimization. In: IEEE. **2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)**. [S.l.], 2012. p. 1–6.
- 43 NENORTAITE, J.; SIMUTIS, R. Stocks' trading system based on the particle swarm optimization algorithm. In: SPRINGER. **International Conference on Computational Science**. [S.l.], 2004. p. 843–850.

- 44 NENORTAITE, J.; SIMUTIS, R. Adapting particle swarm optimization to stock markets. In: IEEE. **5th International Conference on Intelligent Systems Design and Applications (ISDA'05)**. [S.l.], 2005. p. 520–525.
- 45 WANG, F. et al. Return intervals approach to financial fluctuations. In: SPRINGER. **International Conference on Complex Sciences**. [S.l.], 2009. p. 3–27.
- 46 PLEROU, V. et al. On the origin of power-law fluctuations in stock prices. **Quantitative Finance**, Taylor & Francis, v. 4, n. 1, p. C11–C15, 2004.
- 47 YAMASAKI, K. et al. Scaling and memory in volatility return intervals in financial markets. **Proceedings of the National Academy of Sciences of the United States of America**, National Acad Sciences, v. 102, n. 26, p. 9424–9428, 2005.
- 48 GABAIX, X. et al. **Institutional investors and stock market volatility**. [S.l.], 2005.
- 49 PLEROU, V. et al. Econophysics: financial time series from a statistical physics point of view. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 279, n. 1, p. 443–456, 2000.
- 50 BAI, J.; NG, S. Tests for skewness, kurtosis, and normality for time series data. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 23, n. 1, p. 49–60, 2005.
- 51 GOPIKRISHNAN, P. et al. Scaling of the distribution of fluctuations of financial market indices. **Physical Review E**, APS, v. 60, n. 5, p. 5305, 1999.
- 52 BOUCHAUD, J.-P.; POTTERS, M. **Theory of financial risk and derivative pricing: from statistical physics to risk management**. [S.l.]: Cambridge university press, 2003.
- 53 CURTO, J. D. Distribuições de pareto estáveis: Aplicação aos índices psi20, dax e djia. _____. **Métodos quantitativos**, v. 3, p. 107–129, 2003.
- 54 GUERGELOT, M. A.; ARTUSO, A. R. Distribuição de lévy aplicado aos retornos do índice bovespa e das ações da vale e petrobrás. **Revista CIATEC-UPF**, v. 6, n. 2, p. 48–65, 2014.
- 55 WEGMAN, E. J. Time series analysis: Theory, data analysis and computation. **Center for Computational Statistics, George Mason University**, 1996.
- 56 GUJARATI, D. N.; PORTER, D. C. **Econometría (Quinta edición)**. [S.l.]: México: McGRAW-HILL/INTERAMERICANA EDITORES, SA DE CV, 2010.
- 57 FAMA, E. F. The behavior of stock-market prices. **The Journal of Business**, v. 38, n. 1, p. 34–105, 1965.
- 58 FAMA, E. F. Random walks in stock market prices. **Financial analysts journal**, CFA Institute, v. 51, n. 1, p. 75–80, 1995.

- 59 SOARES, S. S. D. Distribuição de renda no Brasil de 1976 a 2004 com ênfase no período entre 2001 e 2004. Instituto de Pesquisa Econômica Aplicada (Ipea), 2006.
- 60 GIRALDES, M. d. R. Equidade em áreas sócio-econômicas com impacto na saúde em países da união européia. **Cadernos de Saúde Pública**, SciELO Public Health, v. 17, n. 3, p. 533–544, 2001.
- 61 BAGOLIN, I. P.; JÚNIOR, S. d. S. P. A desigualdade da distribuição da educação e crescimento no Brasil: índice de gini e anos de escolaridade. **Estudos do CEPE, Santa Cruz do Sul-RS**, v. 18, p. 7–31, 2003.
- 62 JUNIOR, C.; MOREIRA, L. Regional concentration of gross value in the domestic production of pinion in Paraná state. **Ciência Florestal**, SciELO Brasil, v. 26, n. 3, p. 853–861, 2016.
- 63 AMATO, C. M. **Ecologia de Populações de *Ocotea porosa* (Nees) Barroso em Áreas Submetidas a Diferentes Graus de Perturbação**. 2008. Dissertação (Mestrado).
- 64 MANDELROT, B. The variation of certain speculative prices. **The Journal of Business**, v. 36, n. 4, p. 394–419, 1963.
- 65 MANTEGNA, R. N.; STANLEY, H. E. et al. Scaling behaviour in the dynamics of an economic index. **Nature**, v. 376, n. 6535, p. 46–49, 1995.
- 66 CONT, R.; POTTERS, M.; BOUCHAUD, J.-P. Scaling in stock market data: stable laws and beyond. In: **Scale invariance and beyond**. [S.l.]: Springer, 1997. p. 75–85.
- 67 QUEIROS, S. M. D.; ANTENEODO, C.; TSALLIS, C. Power-law distributions in economics: a nonextensive statistical approach. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **SPIE Third International Symposium on Fluctuations and Noise**. [S.l.], 2005. p. 151–164.
- 68 GABAIX, X. et al. A unified econophysics explanation for the power-law exponents of stock market activity. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 382, n. 1, p. 81–88, 2007.
- 69 SILVA, A. C.; YAKOVENKO, V. M. Stochastic volatility of financial markets as the fluctuating rate of trading: An empirical study. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 382, n. 1, p. 278–285, 2007.
- 70 MCCAULEY, J. L.; GUNARATNE, G. H. An empirical model of volatility of returns and option pricing. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 329, n. 1, p. 178–198, 2003.
- 71 SILVA, A. C.; PRANGE, R. E.; YAKOVENKO, V. M. Exponential distribution of financial returns at mesoscopic time lags: a new stylized fact. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 344, n. 1, p. 227–235, 2004.
- 72 KELLY, B.; JIANG, H. Tail risk and asset prices. **Review of Financial Studies**, Soc Financial Studies, v. 27, n. 10, p. 2841–2871, 2014.

- 73 LIMA, N. F. D. **Modelagem de Mercados Inspirada em Gases Ideais e Teoria da Colisão**. 2012. Dissertação (Mestrado).
- 74 GRESNIGT, F.; KOLE, E.; FRANSES, P. H. Interpreting financial market crashes as earthquakes: A new early warning system for medium term crashes. **Journal of Banking & Finance**, Elsevier, v. 56, p. 123–139, 2015.
- 75 BRZEZINSKI, M. Do wealth distributions follow power laws? evidence from ‘rich lists’. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 406, p. 155–162, 2014.
- 76 DRĂGULESCU, A.; YAKOVENKO, V. M. Exponential and power-law probability distributions of wealth and income in the united kingdom and the united states. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 299, n. 1, p. 213–221, 2001.
- 77 CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. Power-law distributions in empirical data. **SIAM review**, SIAM, v. 51, n. 4, p. 661–703, 2009.
- 78 (IBOVESPA) índice B. **índice Bovespa (Ibovespa)**. Disponível em: <http://www.bmfbovespa.com.br/pt_br/produtos/indices/indices-amplos/indice-bovespa-ibovespa.htm>.
- 79 INVESTIDORES, B. R. com. **A BM&FBOVESPA - Perfil e Histórico**. Disponível em: <<http://ri.bmfbovespa.com.br/static/ptb/perfil-historico.asp?idioma=ptb>>.
- 80 CAPELLE-BLANCARD, G.; CHAUDHURY, M. Efficiency tests of the french index (cac 40) options market. In: **EFMA 2002 London Meetings**. [S.l.: s.n.], 2001.
- 81 DEVILLE, L. Time to efficiency of the french cac 40 index options market. In: **EFMA 2004 Basel Meetings Paper**. [S.l.: s.n.], 2003.
- 82 EXCHANGE, M. S. **Mexican Stock Exchange. Information, news, market index, stock quotes, and global markets**. Disponível em: <<http://www.stockexchange.com.mx/>>.
- 83 NASDAQ.COM. **National Association of Securities Dealers Automatic Quotation System (Nasdaq) Definition**. Disponível em: <<http://www.nasdaq.com/investing/glossary/n/national-association-of-securities-dealers-automatic-quotation-system>>.
- 84 NASDAQ.COM. **Common stock Definition**. Disponível em: <<http://www.nasdaq.com/investing/glossary/c/common-stock>>.
- 85 MANN, H. B. Nonparametric tests against trend. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 245–259, 1945.
- 86 KENDALL, M. G. **Rank correlation methods**. 4th. ed. [S.l.]: London: Charles Griffin, 1970.
- 87 FONTELLES, M. Bioestatística aplicada à pesquisa experimental. **São Paulo: Livraria da Física**, 2012.

- 88 BROWN, M. B.; FORSYTHE, A. B. Robust tests for the equality of variances. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 69, n. 346, p. 364–367, 1974.
- 89 THADEWALD, T.; BÜNING, H. Jarque–bera test and its competitors for testing normality—a power comparison. **Journal of Applied Statistics**, Taylor & Francis, v. 34, n. 1, p. 87–105, 2007.
- 90 NEWBOLD, P.; CARLSON, W. L.; THORNE, B. M. **Statistics for Business and Economics**. 8th. ed. [S.l.]: Pearson, 2013.
- 91 BARROS, R. P. d.; MENDONÇA, R. S. P. d. Os determinantes da desigualdade no brasil. Instituto de Pesquisa Econômica Aplicada (Ipea), 1995.
- 92 MILLER, S. J. The method of least squares. **Mathematics Department Brown University**, p. 1–7, 2006.
- 93 GAVIN, H. P. **The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems**. 2016.
- 94 YUAN, Y.-x. A review of trust region algorithms for optimization. In: **ICIAM**. [S.l.: s.n.], 2000. v. 99, p. 271–282.