

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ADMINISTRAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO**

**Alexandre Lerch Franco**

**Aplicação da Análise de Componentes Independentes  
em Estudo de Eventos em Finanças**

**Porto Alegre  
2008**

**Alexandre Lerch Franco**

**Aplicação da Análise de Componentes Independentes  
em Estudo de Eventos em Finanças**

Tese apresentada ao Programa de Pós-Graduação em Administração da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do grau de Doutor em Administração.

Orientador: Prof. Gilberto de Oliveira Kloeckner, Ph.D.

**Porto Alegre**

**2008**

**Alexandre Lerch Franco**

**Aplicação da Análise de Componentes Independentes  
em Estudo de Eventos em Finanças**

Conceito Final:

Tese aprovada em 05 de dezembro de 2008.

BANCA EXAMINADORA:

---

Prof. Dr. Oscar Claudino Galli – PPGA/EA/UFRGS

---

Prof. Dr. Paulo Sérgio Ceretta – UFSM

---

Prof. Dr. Newton Carneiro Affonso da Costa Júnior – UFSC

---

Orientador – Prof. Gilberto de Oliveira Kloeckner, Ph.D – PPGA/EA/UFRGS

*À minha esposa Luciana  
pela motivação e ajuda em todos os sentidos*

## **AGRADECIMENTOS**

Ao prof. Gilberto de Oliveira Kloeckner, Ph.D., pelo estímulo e motivação para seguir em frente, pelo exemplo de professor, pelo carisma e pela amizade construída ao longo desta jornada.

À minha querida esposa Luciana, pela motivação e pelo tempo em que estive ausente.

Aos meus pais, pelo tempo em que estive ausente e pela educação que me deram.

Ao prof. Dr. Oscar Claudino Galli, pelo apoio e pela disponibilidade.

Aos demais professores de finanças do PPGA que, de alguma forma, contribuíram para a solidificação de meu conhecimento científico em finanças.

À Universidade Federal do Rio Grande do Sul e, em especial, aos professores e funcionários do PPGA, por propiciar um ambiente agradável e de completa infra-estrutura para o desenvolvimento desta tese.

Aos demais familiares, colegas de trabalho, amigos e às demais pessoas que de alguma forma fizeram a diferença nesta longa caminhada.

Muito obrigado!

*“A informação de que você dispõe não é a informação que você deseja.  
A informação que você deseja não é a informação que você necessita.  
A informação que você necessita não é a informação que você consegue obter.  
A informação que você consegue obter custa mais do que você deseja pagar.”*  
(Anônimo)

*“The market is like a beautiful woman: endlessly fascinating,  
endlessly complex, always changing, always mystifying.”*  
Adam Smith

## RESUMO

Nas últimas duas décadas, estudos empíricos em finanças têm utilizado o método de estudo de eventos para detectar retornos anormais no entorno de eventos que, teoricamente, deveriam ser incorporados instantaneamente no preço dos títulos. O método de estudo de eventos, a partir da década de 90, com a massificação das planilhas eletrônicas e dos pacotes estatísticos, se popularizou no meio acadêmico brasileiro, sendo um dos principais métodos de pesquisa em finanças com ênfase em mercado de capitais ou finanças corporativas. Apesar da eficácia do método em detectar a anormalidade dos retornos, comprovada em diversos estudos empíricos, acredita-se que o método seja pouco eficiente em determinar a verdadeira amplitude do retorno anormal, uma vez que são necessários pressupostos estatísticos e argumentos econômico-financeiros que podem não ser válidos. O fato de que cada modelo apresenta um desempenho diferente de captura dos retornos anormais contribui com a tese de que os modelos utilizados atualmente não conseguem filtrar totalmente o retorno anormal da série normal. Portanto, este estudo teve como objetivo principal testar a aplicabilidade do método de Análise de Componentes Independentes - ICA - em detectar retornos anormais em séries temporais e comparar o seu desempenho com os modelos geradores de retornos anormais mais utilizados em testes empíricos. Com este objetivo, foram realizadas milhares de simulações envolvendo parâmetros semelhantes aos do mercado de ações brasileiro, com o uso de algoritmos de simulação elaborados exclusivamente para esta finalidade. Os resultados sugerem que o método ICA é capaz de detectar anormalidades em séries temporais, fornecendo, desta forma, a descoberta do real impacto do retorno anormal nos elementos da amostra, necessitando apenas de uma modelagem prévia em função do tamanho da amostra e sua variância.

Palavras-chave: estudo de evento; análise de componentes independentes; métodos de pesquisa em finanças; simulação.

# **Independent Component Analysis Application on Events Study in Finance**

## **ABSTRACT**

In the last two decades financial empiric studies have used the event study method to detect abnormal return on events that in theory should be instantly incorporated on securities price. This method became popular to Brazilian academic environment through the intensification usage of electronic worksheet and statistic packages in the 90's turning into one of the main research methods for financial studies with emphasis on stock market and corporative financing. Despite the efficiency of the method in detecting abnormalities it's believed that it's least effective on establishing the real amplitude of the abnormal return considering that statistics presupposed and economic and financial arguments may not be valid. The fact that each model shows a different performance on capturing abnormal returns contributes to the idea that today's models can't completely filter the abnormal return on a normal series. Therefore this study has as a main objective to test the applicability of the Independent Component Analysis method – ICA – in detecting abnormal returns in time series and comparing its performance against abnormal return generating models more used on empiric tests. With this objective, thousands of simulations involving parameters similar to the Brazilian stock market with the usage of simulation algorithms elaborated exclusively for this purpose. The results suggest that ICA method is capable of detecting abnormalities in time series supplying in this form a discovery on the real impact of abnormal return on sample elements needing only a previous molding due to the size of its sample and variance.

Key-words: event study; independent component analysis; financial research method; simulation.



## LISTA DE ILUSTRAÇÕES

Figura 2.1 – Reação de preços a novas informações .....	31
Figura 2.2 – Fases de um Estudo de Evento .....	42
Figura 2.3 – Linha do Tempo de um Estudo de Evento .....	43
Figura 2.4 – Distribuição de frequência dos retornos pela forma logarítmica .....	45
Figura 3.1 – Representação sistemática do processo de ICA.....	67
Figura 3.2 – Distribuição de duas variáveis gaussianas independentes .....	70
Figura 4.1 – Resultados das simulações, percentual médio de <i>outliers</i> corretamente identificados em 1000 simulações com 100 repetições cada, variando algoritmo e padrão de <i>outlier</i> . Os desvios-padrão dos percentuais estão em parênteses. ....	98
Figura 5.1 – 83 séries normais aleatórias com média e desvios idênticos aos das séries da amostra original.....	108
Figura 5.2 – 84 séries misturadas sendo 83 séries normais e uma contendo a anormalidade de 0,90% .....	109
Figura 5.3 – Componentes encontradas em uma das simulações com anormalidade de 0,90% .....	110
Figura 5.4 – Primeira componente independente detectada pelo algoritmo <i>FastICA</i> .....	111
Figura 5.5– Componentes encontradas em uma das simulações com anormalidade de 4,50% .....	115
Figura 5.6– Nona componente independente detectada com a nova estrutura teórica.....	116
Figura 5.7– Componentes detectadas na dimensão de amostra 5 x 100.....	117
Figura 5.8– Primeira componente detectada na dimensão de amostra 5 x 100.....	118

## LISTA DE TABELAS

Tabela 2.1 – Modelos de Estudo de Eventos utilizados em encontros da ANPAD .....	59
Tabela 2.2 – Estudos de Evento em ENANPAD por ano .....	59
Tabela 2.3 – Artigos nos encontros da ANPAD que utilizaram estudo de evento .....	61
Tabela 2.4 – Estudos de Evento, por ano e periódico. Para cada periódico, todos os artigos que continham um estudo de evento estão incluídos. Artigos de <i>survey</i> e metodológicos estão excluídos. ....	64
Tabela 3.1 – Trabalhos publicados sobre Análise de Componentes Independentes 1994-2007.	78
Tabela 3.2 – Trabalhos publicados na área de Finanças com o uso do método ICA .....	81
Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA .....	83
Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA .....	84
(continuação) .....	84
Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA .....	85
(continuação) .....	85
Tabela 4.1 – Relação de ações de maior liquidez na Bovespa em 2006 .....	90
Tabela 5.1 – Nível de anormalidade detectada para uma simulação - teste efetuado no software estatístico R® .....	101
Tabela 5.2 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - teste t para os modelos tradicionais de retorno anormal .....	103
Tabela 5.3 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções dos modelos AR(1) para 4.47 e 10 desvios .....	105
Tabela 5.4 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções dos modelos AR(1) para 15 e 20 desvios .....	106
Tabela 5.5 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções para séries normais – Baragona e Battaglia .....	107
Tabela 5.6 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das 1000 simulações pós-inclusão da anormalidade para 8 e 12 desvios .....	113

<b>Tabela 5.7–</b> Freqüência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das 1000 simulações pós-inclusão da anormalidade para 11 e 12 desvios .....	<b>114</b>
<b>Tabela 5.8–</b> Resumo dos testes com o método ICA – nível de anormalidade detectada em 100% das mil simulações e a respectiva faixa de detecção em desvios-padrão .....	<b>119</b>

# SUMÁRIO

<b>INTRODUÇÃO</b> .....	<b>14</b>
<b>1 PROBLEMA DE PESQUISA E OBJETIVOS</b> .....	<b>18</b>
1.1 O PROBLEMA DE PESQUISA .....	18
1.2 OBJETIVOS .....	23
1.2.1 Objetivo principal .....	24
1.2.2 Objetivos específicos .....	24
1.3 JUSTIFICATIVA E CONTRIBUIÇÕES DO ESTUDO .....	24
1.4 ORGANIZAÇÃO DO TRABALHO .....	27
<b>2 MÉTODOS DE PESQUISA EM FINANÇAS</b> .....	<b>29</b>
2.1 HIPÓTESE DE MERCADO EFICIENTE .....	29
2.2 EFICIÊNCIA INFORMACIONAL .....	35
2.3 TESTES DE EFICIÊNCIA DE MERCADO .....	38
2.4 TESTES DE PREVISIBILIDADE DE RETORNOS .....	39
2.5 TESTES DE ESTUDO DE EVENTOS .....	41
2.5.1 Definição do evento .....	42
2.5.2 Seleção da amostra .....	43
2.5.3 Escolha do método .....	44
2.5.4 Modelos para determinação dos retornos anormais .....	46
2.5.5 Modelos estatísticos .....	46
2.5.6 Modelos econômico-financeiros .....	51
2.6 TESTES DE INFORMAÇÃO PRIVADA .....	56
2.7 UTILIZAÇÃO DO MÉTODO DE ESTUDO DE EVENTO EM FINANÇAS .....	57
2.7.1 Estudos de eventos no mercado brasileiro .....	57
2.7.2 Estudos de eventos em outros mercados .....	62
<b>3 ANÁLISE DE COMPONENTES INDEPENDENTES</b> .....	<b>65</b>
3.1 DEFINIÇÃO .....	65
3.2 O MÉTODO ICA .....	66
3.2.1 Restrições do Método ICA .....	68
3.2.1.1 As fontes devem ser independentes .....	69
3.2.1.2 As fontes devem apresentar distribuições não gaussianas .....	70
3.2.1.3 A matriz de mistura deve ser quadrada .....	71
3.2.1.4 Os sinais devem ser estacionários .....	71
3.2.2 Determinação das componentes por maximização da não-gaussianidade .....	72
3.2.2.1 Medidas para a gaussianidade .....	73
3.3 O ALGORITMO FASTICA .....	74
3.3.1 Propriedades do algoritmo <i>FastICA</i> .....	77
3.4 APLICABILIDADE DO MÉTODO ICA .....	78
3.4.1 Uso do método ICA em Finanças .....	79
3.4.2 Uso do método ICA em outras áreas do conhecimento .....	82

<b>4 MÉTODO.....</b>	<b>87</b>
4.1 DESCRIÇÃO DO EXPERIMENTO.....	87
4.1.1 Etapas do experimento.....	88
4.1.1.1 Primeira etapa: criação do banco de dados.....	88
4.1.1.2 Segunda etapa: medindo o desempenho dos modelos tradicionais.....	91
4.1.1.3 Terceira etapa: construindo o experimento com o método ICA.....	94
4.1.1.3.1 Testes efetuados na base de dados .....	98
<b>5 RESULTADOS.....</b>	<b>101</b>
5.1 RESULTADOS DOS MODELOS TRADICIONAIS.....	101
5.2 RESULTADOS DOS TESTES COM O MÉTODO ICA .....	104
5.2.1 Teste do ICA com modelos auto-regressivos.....	104
5.2.2 Teste do ICA em séries aleatórias gaussianas - Fase I.....	106
5.2.3 Teste do ICA em séries aleatórias gaussianas - Fase II.....	112
5.2.4 Teste do ICA para dimensões menores de dados .....	116
5.2.4 Considerações sobre os resultados dos testes com o método ICA.....	118
<b>6 DISCUSSÕES, CONCLUSÕES E CONSIDERAÇÕES FINAIS .....</b>	<b>120</b>
6.1 DISCUSSÕES SOBRE OS RESULTADOS.....	120
6.2 CONCLUSÕES.....	123
6.3 LIMITAÇÕES DA TESE.....	125
6.4 SUGESTÕES PARA PESQUISAS FUTURAS.....	126
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>128</b>
<b>ANEXO A – Algoritmos elaborados no ambiente Matlab® .....</b>	<b>142</b>
<b>ANEXO B – CV Lattes resumido .....</b>	<b>153</b>

## INTRODUÇÃO

Inicialmente, o conceito de mercado eficiente esteve ligado ao conceito de *random walk*<sup>1</sup>, que sustenta a hipótese de que os preços passados não carregam nenhuma informação sobre os preços futuros.

De acordo com Paulos (2004), a Hipótese de Mercado Eficiente (HME) remonta à tese de Eugene Fama, de 1964, ao trabalho do economista Paul Samuelson, ganhador do Prêmio Nobel, e a outros estudiosos da década de 60.

Sua origem, contudo, deve-se ao trabalho, escrito em 1900, do matemático Louis Bachelier, aluno do matemático francês Henri Poincaré, que somente cinquenta anos depois teve a sua tese descoberta. A idéia central de Bachelier foi: “A expectativa matemática do especulador é zero.”

Apesar da idade de Bachelier na época, a matemática que ele desenvolveu para explicar os preços das opções sobre os títulos do governo francês anteciparam em cinco anos a descoberta de Einstein sobre o movimento dos elétrons que, por sua vez, forneceu a base da teoria da marcha aleatória em finanças.

Para os estatísticos, a eficiência de mercado consiste no caminho aleatório de preços. Para os investidores, a eficiência de mercado significa a impossibilidade de arbitragem a custo zero, ou seja, um mercado é dito eficiente quando nenhuma estratégia for superior à estratégia ingênua de comprar e reter títulos (*buy and hold*), tornando o mercado de capitais um *fair game*.

---

<sup>1</sup> *random walk* é um processo estocástico de estado discreto e tempo discreto, i.e., é o equivalente à uma caminhada aleatória, sob a ótica do estatístico, significa a independência das variações de preços, ou seja, as variações nos preços são perturbações caóticas de média zero. Os primeiros autores a estudarem o caminho aleatório das séries de preços de títulos foram os estatísticos Kendall (1953) e Roberts (1959).

Neste ponto, surge um paradoxo curioso envolvendo a eficiência informacional dos mercados financeiros. Se for verdade, que os preços dos títulos contém toda a informação disponível, seja ela pública ou não, significa que os preços não contém nenhuma informação nova. Nesse sentido, o paradoxo aparece: como explicar o equilíbrio de um mercado onde os preços nada informam?

Segundo Sá (2001, p. 349) o mercado eficiente como uma teoria de equilíbrio informacional, com o passar do tempo, acumulou mais dúvidas do que convicções. O conceito está tão mergulhado em contradições teóricas e anomalias empíricas que sua aparente simplicidade esconde uma enorme complexidade:

A economia dos mercados financeiros é incapaz de gerar riquezas mas é o lugar de destino da riqueza gerada na economia real. A riqueza nova dos ativos reais se transporta aos títulos em *cashflows* e ganhos de capital necessariamente. Se o mercado for eficiente, entretanto, a assimilação da informação pelos preços deles remove toda a memória e oblitera o caminho da poupança do investidor na direção da nova riqueza.

Dessa forma, quem recebe a nova riqueza da economia real em um mercado eficiente o faz fortuitamente, pelo acaso de ter o papel certo, não pela procura investigada do papel certo. O sentido de “eficiência” restringe-se, assim, à propriedade de o mercado inibir as transferências de riqueza entre as pessoas.

Segundo o autor (Sá, 2001), existe abundante evidência de que agentes financeiros informados são capazes de captar a informação do enriquecimento dos ativos reais a tempo de proporcionar ganhos anormais consistentemente na compra e venda de títulos, embora isso contradiga a eficiência dos mercados.

O paradoxo da circularidade torna mais complexa a discussão de eficiência de mercado, podendo, inclusive, ser uma das explicações para os ciclos de aceleração e desaceleração dos mercados de capitais. Ao imaginar um mercado ineficiente, os investidores buscarão os retornos anormais sempre que estiverem disponíveis, tornando o mercado, dessa forma, cada vez

mais eficiente. Sendo o mercado eficiente, a busca por retornos anormais não seria possível, fazendo que com que todos os investidores não reagissem frente aos eventos, surgindo, então, oportunidades de retorno anormal, tornando o mercado ineficiente.

Várias são as versões da Hipótese de Eficiência de Mercado (HEM), dependendo das informações que se presumem já estejam incorporadas no preço dos títulos.

Fama (1970), categorizou três formas de eficiência, diferenciando-as quanto ao nível de disponibilidade das informações.

A forma mais fraca sugere que todas as informações sobre os preços dos títulos no passado já estão incorporadas no preço atual, hipótese que destrói as regras e padrões da escola conhecida como Análise Técnica, que usa apenas as informações dos preços passados para projetar preços futuros.

A forma mais forte sugere que nenhuma pessoa com informações privilegiadas (*insiders*) conseguiria obter retornos superiores no mercado utilizando-se de tais informações para acumular riqueza com o prejuízo dos desinformados.

Entre essas duas, Fama (1970) categorizou como sendo a forma semiforte (*semi-strong form*), a hipótese de que os preços refletem não apenas o seu histórico (forma fraca), como também todas as informações disponíveis publicamente (balanços patrimoniais, demonstrativos financeiros, entre outras). Na década de 90, Fama (1991) propôs a alteração no nome das formas de eficiência informacional e os testes de hipótese da forma semiforte foram denominados como **estudos de evento**.

Bernstein (1997, p. 198), relembra o ensaio filosófico sobre as probabilidades, de Laplace, que estava convencido de que não existe algo como sorte, ou acaso:

Os eventos presentes estão ligados aos eventos precedentes por um vínculo baseado no princípio óbvio de que uma coisa não pode ocorrer sem uma causa que a produza... Todos os eventos, mesmo aque-



les que, devido à insignificância, não parecem seguir as grandes leis da natureza, resultam delas tão necessariamente como as revoluções do Sol.

A mais flagrante contradição empírica de um mercado eficiente é a aparente previsibilidade de perdas e ganhos anormais quando da publicação dos lucros das empresas.

Segundo Sá (2001), existem evidências de que os participantes do mercado incorporam a informação que está chegando aos preços e, ao mesmo tempo, subestimam, ou superestimam, o ajuste (i.e., não o fazem por completo) ou não haveria ganhos ou perdas anormais logo adiante.

Portanto, este estudo apresenta os conceitos que circundam a Hipótese de Eficiência de Mercado, com mais ênfase no método de teste da forma semiforte, denominado **estudo de evento**.

A questão problemática que justifica o assunto principal desta tese está descrita no próximo Capítulo, assim como as justificativas de verificação de um modelo alternativo.

Um dos objetivos deste estudo é aprimorar a capacidade dos investidores informados em detectar ineficiências no mercado de capitais, buscando para isso novos métodos para realizar os testes de ineficiências de mercado, contribuindo para a transferência de riqueza, enquanto for possível.

O processo de comunicação entre a economia real e a financeira sendo imperfeito, faz com que seja rentável investigar as informações que ainda não foram processadas. O canal de comunicação, sendo imperfeito, fomenta o ruído que permite o mercado existir.

Segundo Black (1986), na formação do preço dos ativos financeiros pode-se teorizar que o papel dos agentes informados é impedir que preço e valor divirjam muito, enquanto o dos desinformados é comprar e vender no ruído.

## 1 PROBLEMA DE PESQUISA E OBJETIVOS

Neste tópico são apresentados o problema de pesquisa e os objetivos do estudo. Resumidamente, alguns conceitos necessários são apresentados e discutidos. No próximo capítulo, estes conceitos são reapresentados de forma mais clara e abrangente.

### 1.1 O PROBLEMA DE PESQUISA

O método denominado “estudo de evento”, em finanças, descreve as etapas de um teste empírico onde são estudados os impactos de um evento econômico-financeiro no comportamento dos títulos que compõem a amostra dos elementos que mantém, pelo menos, uma coisa em comum, o evento em estudo.

Trabalhos empíricos, por exemplo, no Brasil, envolvendo o uso de estudos de evento são utilizados para verificar o efeito nos retornos dos títulos de eventos como: o impacto de mudanças societárias; empresas em concordata; reação a *stock splits* e *stock dividends*; reações racionais frente mudanças no cenário econômico; reação à fusões e aquisições; o efeito das decisões de investimento e financiamento; lançamento de *ADR's*; reações dos investidores à bonificações, entre outros. Estes são alguns dos assuntos pesquisados no mercado brasileiro, desde 1997, e publicados em encontros da Anpad.<sup>2</sup>

Segundo Dennis e McConnell (1986), a análise de um evento por este método, pressupõe que muitas das informações relevantes envolvendo o evento em questão devem se refletir nos preços dos títulos, em períodos próximos ao evento.

---

<sup>2</sup> Um quadro resumo dos trabalhos no Brasil e no exterior é apresentado no capítulo 2.

Brown e Warner (1980) afirmam que o mais importante, em um estudo de evento, não é simplesmente identificar o retorno anormal, mas sim verificar a extensão e o tamanho dessa anormalidade, em comparação a um modelo de equilíbrio de determinação de retornos normais.

O problema consiste em modelar o retorno normal, de forma que apenas o retorno anormal seja isolado da série. Vários modelos foram elaborados para estimar o retorno normal dos títulos. Brown e Warner (1980), apresentam duas categorias: os modelos estatísticos e os modelos econômicos.<sup>3</sup>

Os modelos estatísticos se dividem em:

- a) modelo ajustado à média;
- b) modelo ajustado ao mercado; e
- c) modelo ajustado ao risco e ao mercado.

Os modelos estatísticos são os mais utilizados. No Brasil, a preferência dos pesquisadores é pelo modelo ajustado ao risco e ao mercado, conforme pesquisa realizada por Soares, Rostagno e Soares (2002). Nos Estados Unidos, por exemplo, de acordo com Brown e Warner (1980), a preferência é pelo modelo ajustado ao mercado.

O modelo ajustado à média encontra muita resistência por parte dos pesquisadores, uma vez que carrega a premissa de que os retornos, uma vez afastados pelo efeito de um evento, tendem a retornar à média.

Bernstein (1997), lembra questões pertinentes: “Quando as observações afastam-se da média do passado, quais as probabilidades de regressarem àquela média no futuro? E se regressarem, pararão na média ou a ultrapassarão?”

Bernstein (1997, p. 144) ressalva que, mesmo que a marcha aleatória (*random walk*) seja uma descrição válida da realidade do mercado de ações,

---

<sup>3</sup> Os modelos e suas características são apresentados em suas formas algébricas no capítulo 2. A preocupação nesta etapa do trabalho foi a de apresentar a situação problemática que envolve a utilização dos modelos existentes.

mesmo que as mudanças nos preços tenham o comportamento equivalente à uma distribuição normal perfeita, a média será algo diferente de zero:

A tendência ascendente não deve surpreender. A riqueza dos detentores de ações ordinárias aumentou a longo prazo, à medida que a economia e as receitas e os lucros das empresas cresceram. Como mais ações subiram do que caíram de preço, a mudança média nos preços das ações deve resultar em mais de zero.

Os modelos ajustados ao mercado, ajustados ao mercado e ao risco e o modelo econômico do CAPM, conforme descrito por Kloeckner (1995), são semelhantes e consistentes, apenas com pressupostos diferenciados. No modelo de mercado, presume-se que todas as ações têm um risco sistemático ( $\beta$ ) igual à um e um intercepto ( $\alpha$ ) igual à zero.

Portanto, as críticas de um modelo provocam conseqüências na aceitação de outro. Os modelos de retorno ajustado ao mercado e o ajustado ao risco e ao mercado encontram críticas pela definição do retorno que deve ser definido como retorno de mercado. Essa mesma crítica é enfrentada pelos modelos econômicos, que necessitam de uma *proxy* que corresponda a carteira teórica do mercado.

Procianoy e Verdi (2006), por exemplo, utilizaram o retorno do índice da Fundação Getúlio Vargas (FGV-100) como retorno de mercado para medir o retorno anormal nas ações adicionadas e excluídas dos índices IBOVESPA, IBRX50 e IBRX100. Ao medir o mesmo evento no FGV-100, os autores utilizaram o retorno do IBOVESPA, como retorno de mercado.

Penteado e Famá (2002), estudaram a (in)adequação do índice Ibovespa como parâmetro para medidas de risco, principalmente na composição do modelo CAPM – *Capital Assets Pricing Model*, conforme proposto e estabelecido por Sharpe (1964). Segundo os autores, o portfólio que seria o “ativo-padrão” sugerido por Sharpe, deveria ser aquele em que cada ativo teria participação proporcional ao seu valor de mercado, e não de liquidez, conforme preconiza o índice Ibovespa.

O fato de o índice Ibovespa ser constituído por empresas ponderadas pela liquidez provoca um desvio ao modelo proposto por Sharpe. Penteado e Famá (2002, p. 51), ao notarem que a inclinação da reta de regressão da evolução do preço das ações aumenta quando usado um índice onde a participação de cada ação se dá na proporção do valor de mercado da respectiva empresa em relação ao valor de mercado total da carteira, sugerem que há indícios de subavaliações no mercado brasileiro:

A utilização do custo assim calculado, como fator de decisão para a escolha entre projetos alternativos, pode ter levado à aceitação de alternativas que – se o custo houvesse sido estimado conforme prescrito por Sharpe – poderiam ter sido rejeitadas.

Ampliando o assunto, tornam-se pertinentes os resultados obtidos por Halfeld e Silva (1999) ao estudarem a eficiência do mercado na América Latina pela composição de carteiras de ações construídas com base apenas em tradicionais indicadores da análise fundamentalista. Contrariamente ao previsto pelo CAPM, os autores encontraram indícios de que o beta não consegue explicar o retorno futuro das ações. No período inteiro da análise (janeiro de 1992 a junho de 1998), o beta não foi capaz de explicar os retornos das ações. A evidência mais forte foi no período que antecedeu a crise do México, onde o beta funcionava ao contrário, ou seja, as ações mais arriscadas ofereciam retornos médios menores. Os próprios autores explicam, conforme descrito por Halfeld e Silva (1999) em Haugen (2000, p. 173):

Esta anomalia pode ser explicada pela grande concentração do Ibovespa em poucos papéis de grande liquidez. No Brasil, ações com alto beta são as de empresas grandes e líquidas que, no período da análise, tiveram um desempenho ruim. Como consequência, chega-se a um resultado estranho: ações pouco arriscadas, em termos de beta, oferecem retornos mais altos.

Os modelos econômicos CAPM e APT – *Arbitrage Pricing Theory*, segundo Camargos (2002), por obedecerem, além dos pressupostos estatísticos, restrições econômicas, deveriam indicar medidas mais precisas de retornos anormais.

De acordo com Copeland et al. (2005, p. 180), o CAPM pode ser visto como um caso particular de APT, quando os retornos dos ativos são descritos por uma distribuição normal multivariada.

A adoção do modelo CAPM, por exemplo, implica a validação de, pelo menos, três pressupostos<sup>4</sup>: (1) que todos os investidores utilizam a otimização de carteiras de Markowitz; (2) que o índice representativo de mercado seja uma das carteiras eficientes e (3) que o único fator que diferencia o retorno dos títulos é a sensibilidade destes às mudanças no retorno de mercado, medido pelo “beta”. Ainda, Markowitz (1952) presumiu que as probabilidades dos retornos da carteira se comportariam como uma distribuição normal, o que nem sempre ocorre.

Soares, Rostagno e Soares (2002) ao medirem o desempenho de três modelos de estudo de eventos: (1) ajustado ao mercado; (2) ajustado ao risco e ao mercado e (3) o CAPM, verificaram que, apesar do CAPM ser o modelo mais sofisticado, seu resultado foi inferior ao apresentado pelo modelo de retorno ajustado ao risco e ao mercado.

A falta de padronização na utilização dos fatores e dos procedimentos pode ocasionar resultados diferentes para uma mesma base de dados. Apesar das taxas CDI e SELIC serem semelhantes, não existe uma padronização para a definição da taxa livre de risco a ser adotada quando da utilização do modelo CAPM. Soares, Rostagno e Soares (2002), por exemplo, utilizaram a variação diária do CDI como taxa livre de risco. Por outro lado, autores como Lima e Terra (2006) utilizam a taxa SELIC como taxa livre de risco.

McGovern (2005, p.110) esclarece que testar a Hipótese de Eficiência de Mercado usando um modelo de precificação de ativos como medida padrão para o retorno justo de uma ação compreende, na realidade, testar a HEM “em associação à” algum modelo de precificação de ativos. O autor sugere que os resultados dos testes geram um dilema. Existem os pesquisadores que, ao encontrar anormalidades no mercado, concluem que o mercado é ineficiente e que os investidores são irracionais, pois se fossem racionais te-

riam se beneficiado desta anormalidade até o ponto de sua extinção. Por outro lado, os críticos desta hipótese e, por consequência, defensores da HEM apresentam outra interpretação sobre as anormalidades detectadas. Estes sugerem que o mercado é eficiente e que os investidores são racionais. As anormalidades só aparecem porque são geradas por modelos de precificação inadequados. Portanto, a fonte da anormalidade são os modelos de precificação de ativos e não a hipótese de eficiência de mercado.

Em face dessas críticas e dificuldades de modelagem do retorno normal para as séries em estudo, surgem questões pertinentes: (1) existe uma metodologia capaz de detectar os retornos anormais sem precisar de pressupostos como: retorno à média, equilíbrio de mercado, nível de risco linear, além dos pressupostos econômicos de taxas livres de risco, retornos de portfólio de mercado, mantendo intacta a série original? (2) existe uma metodologia capaz de detectar o retorno anormal sem precisar utilizar o modelo residual da diferença do retorno real frente a um modelo estabelecido que possa não significar a realidade se o evento não ocorresse?

## 1.2 OBJETIVOS

De acordo com o explicitado no problema de pesquisa, esta tese busca testar a aplicabilidade de uma nova metodologia para a análise de eventos em finanças. O objetivo principal e os específicos são explicitados na sequência.

---

<sup>4</sup> Sem mencionar as premissas básicas de validação estrutural do CAPM: que exista uma taxa livre de risco e que o retorno de mercado seja superior à essa taxa.

### 1.2.1 Objetivo principal

O objetivo da presente tese é testar a capacidade do método de Análise de Componentes Independentes - ICA<sup>5</sup> em capturar retornos anormais e comparar o seu desempenho com os modelos mais utilizados atualmente.

### 1.2.2 Objetivos específicos

Complementando o objetivo principal, os objetivos específicos são os seguintes:

a) Analisar o desempenho dos modelos mais utilizados atualmente em estudos de eventos, replicando a técnica adaptada utilizada por Kloeckner (1995) e repetida em Soares, Rostagno e Soares (2002);

b) Testar a capacidade do método de Análise de Componentes Independentes em capturar os retornos anormais na mesma base simulada para os modelos atuais;

c) Identificar a potencialidade do método alternativo, frente aos resultados atingidos nos últimos objetivos.

## 1.3 JUSTIFICATIVA E CONTRIBUIÇÕES DO ESTUDO

O teste de eficiência da forma semiforte, renomeado como **estudo de evento** por Fama (1991), analisados, testados e utilizados por Brown e Warner (1980 e 1985), MacKinlay (1997), Campbell, Lo e MacKinlay (1997), entre

---

<sup>5</sup> ICA – *Independent Component Analysis*



outros, permanece da mesma forma como foi concebido. A metodologia originalmente concebida há quase quatro décadas continua sendo utilizada para mensurar a velocidade do ajustamento dos preços dos títulos ao redor de uma data específica, quando da divulgação de informações relevantes, tais como: pagamento de dividendos, anúncios trimestrais, fusões e aquisições, entre outras.

Apesar da aparente eficácia do método, apresentada em Kloeckner (1995) em séries com retornos mensais, posteriormente confirmada por Soares, Rostagno e Soares (2002) em séries com retornos diários, em detectar a anormalidade dos retornos, acredita-se que o método seja pouco eficiente em determinar a verdadeira amplitude do retorno anormal, uma vez que são necessários pressupostos estatísticos e argumentos econômico-financeiros que podem não ser válidos.

A detecção do retorno anormal com o uso dos métodos atuais foi comprovada; a questão é que se ambas as metodologias fossem capazes de identificar a amplitude deste retorno, os resultados obtidos deveriam ser idênticos. O fato de que cada modelo apresenta um desempenho diferente de captura dos retornos anormais (aos níveis de significância de 1% e 5%) contribui com a tese de que os modelos utilizados atualmente não conseguem filtrar totalmente o retorno anormal da série normal, ou seja, parte do retorno anormal, em alguns modelos, continua dentro da série normal, o que faz com que os resultados de cada modelo sejam diferentes. Portanto, a análise precisa da amplitude da anormalidade fica prejudicada.

A análise de componentes independentes é uma técnica que busca identificar os fatores independentes comuns às séries de uma amostra. A aparente similaridade com a análise de componentes principais é mera coincidência de títulos. Na análise de componentes principais, o principal objetivo é reduzir o número de fatores, ou variáveis explicativas, classificando-as de acordo com o poder de explicação da variável dependente, ou seja, o objetivo é determinar as variáveis que são responsáveis pela maior variância possível

dos dados. Diferente disso, a análise de componentes independentes busca a identificação de fatores, ou *sources*, estatisticamente independentes<sup>6</sup>.

Portanto, a análise de componentes independentes (ICA) pode ser considerada como uma técnica de separação cega de fontes (*blind source separation*) pois não é necessária a modelagem ou um pré-conhecimento das fontes existentes nos dados misturados<sup>7</sup>.

Em finanças, a análise de componentes independentes é utilizada para identificar mecanismos desconhecidos que sejam responsáveis pela movimentação não aleatória dos retornos dos títulos. Apesar de ser uma técnica desenvolvida há mais de uma década, poucos são os estudos de sua aplicabilidade em finanças. Testes já foram realizados em diversas áreas do conhecimento: separação de sons, de imagens, de sinais de telecomunicações, de frequências de rádio, sinais de satélite, entre outros<sup>8</sup>.

A principal característica da análise de componentes independentes é que a informação obtida sobre cada fator, pelo fato de serem independentes, não estará diluída em mais de um fator, ou seja, se existe um fator comum entre as séries analisadas (e.g., retornos de títulos), ele estará totalmente reconstruído em apenas um fator, o que difere da análise de componentes principais que pode apresentar resultados correlacionados.

A lógica que justifica este estudo parte do princípio que, se uma amostra de retornos de títulos representar o agrupamento de diferentes empresas, de diferentes setores, em diferentes períodos, apresentando apenas o evento em estudo como fator comum entre as séries, espera-se que, com o uso da análise de componentes independentes seja possível detectar este fator comum a todas as empresas sem a necessidade de criar modelos de retornos normais baseados em pressupostos econômico-financeiros.

---

<sup>6</sup> uma fonte é considerada independente quando não existe nenhuma informação sobre o comportamento dessa fonte em outras fontes encontradas.

<sup>7</sup> Dados misturados, na técnica de separação cega de fontes, pode ser descrito como uma combinação linear de séries temporais ou uma sobreposição de imagens, onde exista a necessidade de separação.

<sup>8</sup> No Capítulo 3 são apresentados os principais testes de aplicabilidade do método ICA.

Dessa forma, espera-se obter informações mais puras, ou mais realistas sobre a dimensão do impacto do evento nas séries, uma vez que não será mais necessária a modelagem e estimação do que seria o suposto retorno normal das séries.

Teorias como regressão à média, aproximação linear de risco e retorno e os pressupostos do CAPM, APT, entre outros, não seriam mais necessárias. O evento, uma vez existente, seria detectado e estudado em sua forma mais clara, pois ele seria completamente separado das séries em estudo.

Portanto, espera-se que os resultados deste estudo contribuam para o reconhecimento do impacto dos eventos que afetam os retornos dos títulos, contribuindo para a evolução dos estudos em finanças sobre o efeito das informações públicas nos retornos dos títulos.

#### 1.4 ORGANIZAÇÃO DO TRABALHO

Além da breve contextualização, este capítulo trata do problema de pesquisa, dos objetivos da tese, da justificativa, das contribuições e da organização do estudo.

O Capítulo 2 aborda os principais conceitos teóricos sobre a Hipótese de Eficiência de Mercado, as formas de medição da eficiência informacional e as evidências empíricas de estudos de eventos no Brasil e no exterior.

O Capítulo 3 apresenta o método de Análise de Componentes Independentes, suas características, propriedades e aplicabilidades.

O Capítulo 4 contempla o método e a descrição do experimento, reportando-se à definição da amostra e etapas do procedimento metodológico.

O Capítulo 5 apresenta os resultados de cada etapa do experimento.

E, por último, no Capítulo 6 são apresentadas: a discussão sobre os resultados obtidos, as conclusões desta tese, as limitações do estudo e as recomendações para estudos futuros.

## 2 MÉTODOS DE PESQUISA EM FINANÇAS

Objetivando enriquecer a análise em estudo e sustentar a discussão sobre as técnicas atualmente utilizadas nos testes de Hipótese de Eficiência de Mercado (HEM), este capítulo apresenta os conceitos por trás da HEM e apresenta as principais pesquisas empíricas em busca de retornos anormais no Brasil e no exterior. Privilegia-se, neste tópico, as pesquisas que utilizaram como método o estudo de evento, ou seja, pesquisas de eficiência de mercado da forma semi-forte.

### 2.1 HIPÓTESE DE MERCADO EFICIENTE

Até os anos de 1960, segundo Sá (2001, p.342), as bases de dados de preços de ativos financeiros das diversas bolsas americanas não desfrutavam de aprovação acadêmica como fontes de pesquisa:

A teoria financeira das empresas não tinha uma direção definida, os economistas não acreditavam que os mercados financeiros pudessem ser descritos racionalmente e os dados de bolsas de valores carregavam a suspeita da contaminação de arbitrariedades contábeis.

O inicial desinteresse dos economistas e a vasta riqueza dos números oscilando nas bolsas de valores chamaram a atenção de estatísticos que, inicialmente, concluíram que os preços dos títulos “caminhavam aleatoriamente”.

Os economistas logo descobriram que os preços dos títulos não podiam caminhar aleatoriamente porque, por exemplo, os preços de um contrato futuro tinham um encontro marcado na data de vencimento, caso contrário,

surgiria um problema de arbitragem. Neste sentido, o preço de uma ação estaria predestinado a encontrar o seu valor intrínseco em um certo momento.

Segundo Sá (2001), a correção estava em que não eram os preços que caminhavam aleatoriamente, e sim os retornos percentuais de períodos sucessivos. O retorno de um determinado período seria igual ao do período anterior somado à uma variação aleatória simétrica de média zero.

Desta forma, segundo Sá (2001, p. 343), o rendimento dos títulos mantinha-se na média constante da classe de risco do título:

No mercado eficiente, por conseqüência, os títulos de mesma classe de risco são substituíveis entre si porque nenhum tem, no longo prazo, uma taxa de crescimento de preços diferente da expectativa de rentabilidade mínima desejada pelos investidores para a classe de risco, igual ao custo de capital daquela classe.

Conseqüentemente, a teoria do mercado eficiente se tornou uma das mais estudadas por pesquisadores. Inúmeros trabalhos empíricos constatam, em determinados momentos, a ineficiência dos mercados, gerando retornos anormais para os investidores.

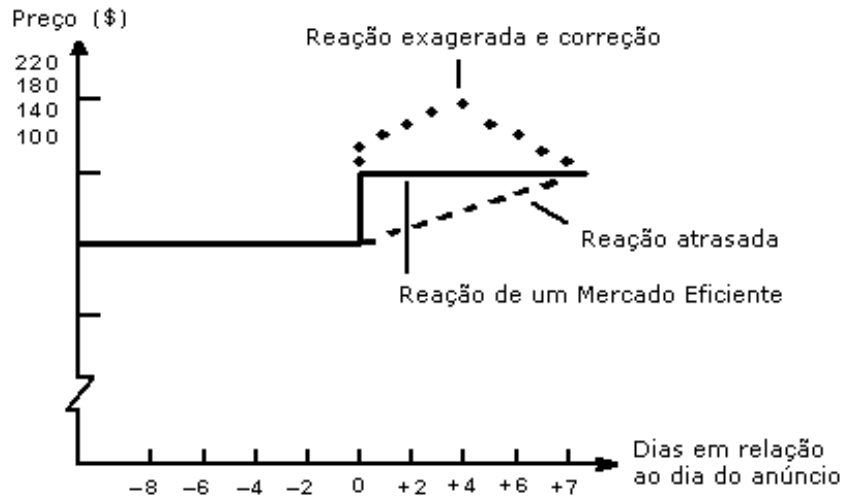
Segundo Ross, Westerfield e Jaffe (2002, p. 279), “um mercado eficiente de capitais é aquele no qual os preços das ações refletem completamente as informações disponíveis”. Portanto, a hipótese de mercado eficiente apresenta implicações para investidores e empresas:

Como a informação se reflete imediatamente nos preços, os investidores só devem esperar obter uma taxa normal de retorno. O conhecimento da informação, ao ser divulgada, não traz vantagem alguma a um investidor. O preço ajusta-se antes que se possa negociar a ação com o uso da informação.

As empresas devem esperar receber o valor justo dos títulos que vendem. A palavra *justo* significa que o preço que recebe pelos títulos emitidos é igual ao seu valor presente. Portanto, em mercados eficientes de capitais não existem oportunidades valiosas de financiamento decorrentes de enganar os investidores.

As possíveis reações de preços a novas informações, segundo Ross *et al.*(2002) podem ser: (1) ajuste imediato, caso este que explica um merca-

do eficiente; (2) reação retardada ou (3) reação exagerada, como pode ser visto na Figura 2.1.



**Figura 2.1 – Reação de preços a novas informações**

Fonte: adaptado de Ross *et. al.* (2002, p. 281)

Em outras palavras, em um mercado eficiente, não deveria existir a possibilidade de transferência de riqueza dos dirigentes das empresas aos acionistas, e vice-versa.

Porém, se os mercados de capitais, conforme enfatiza Sá (2001, p. 353), “são mecanismos direcionadores da poupança para os melhores investimentos de ativos reais é porque não são informacionalmente eficientes.”

O conceito de mercado absolutamente eficiente ou ineficiente permanece no campo teórico abstrato. A natureza complexa dos mercados de capitais impossibilita que uma afirmação seja verdadeira e outra, conseqüentemente, seja falsa. Paulos (2004, p. 190), explica a hipótese “paradoxal” do mercado eficiente afirmando que a crença predominante em uma hipótese engendra sua própria falsidade:

Se a vasta maioria dos investidores acreditar na Hipótese do Mercado Vagaroso, todos estarão convictos de que realmente vale a pena identificar tendências e analisar empresas, e, assim agindo, constroem um mercado eficiente. Portanto, se a maioria dos investidores

acreditar que a Hipótese do Mercado Vagaroso é verdadeira, suas próprias iniciativas converterão em realidade a Hipótese do Mercado Eficiente. Concluimos que, se a Hipótese do Mercado Eficiente for falsa, isso significa que a maioria dos investidores não acha que a Hipótese do Mercado Vagaroso seja verdadeira, isto é, se a Hipótese do Mercado Eficiente for falsa, a maioria dos investidores acredita que ela (HME) é verdadeira.

Portanto, a verdade ou falsidade da Hipótese do Mercado Eficiente não é imutável, e depende fundamentalmente das crenças dos investidores.

Paulos (2004) conclui que, a maioria dos investidores não acredita na HEM, motivo pelo qual ele afirma crer que ela se sustenta, mas apenas com alguma precariedade e em boa parte do tempo, mas não o tempo todo.

Por conseqüência, trabalhos empíricos sobre eficiência de mercado têm contaminado de informações, muitas vezes dúbias, o conhecimento de pesquisadores e investidores em todo o mundo. Jensen (1978), já afirmava naquela época que, se fosse feita uma análise do volume de estudos sobre finanças, a hipótese de eficiência de mercado seria a mais testada e fundamentada de todas as ciências sociais.

De acordo com Brito (1981), existem três níveis possíveis de eficiência de mercado:

a) eficiência alocacional: diz respeito ao equilíbrio perfeito entre poupadores e tomadores de aplicações financeiras, operacionalizada por intermediários financeiros que com habilidade desenvolvem títulos impedindo que uma das partes imponha preços sobre a outra.

b) eficiência operacional: diz respeito à forma com que as instituições financeiras assessoram seus clientes em relação à circulação e disponibilização dos títulos no mercado financeiro.

c) eficiência informacional: diz respeito à velocidade com que as informações relevantes sobre os agentes do mercado (empresa, mercado, instituições, etc) são refletidas nos preços dos títulos. Diz-se que um mercado é



“totalmente eficiente” quando os preços refletem integralmente as informações, sejam elas públicas ou privadas.

Tendo em vista o objetivo desta tese, são apresentadas neste Capítulo apenas as formas de medição de eficiência informacional, propostas por Fama (1970, 1991).

Para Paxson e Wood (2001), a expressão eficiência de mercado é utilizada para explicar a relação existente entre informação e preço de mercado. Segundo os autores, embora haja relatos de testes de eficiência de mercado realizados em 1900, somente em 1953, Maurice Kendall deu forma a essa idéia. Entretanto, a definição formal de eficiência de mercado, em seu nível informacional, foi dada por Fama em 1970.

Da mesma forma, Downes e Goodman (1993) descrevem que segundo a teoria de mercado eficiente, os preços de mercado refletem o conhecimento e as expectativas de todos os investidores. Seus adeptos consideram inútil procurar ações subavaliadas ou fazer previsões de movimentos de mercado. Qualquer novo desenvolvimento se reflete no preço das ações de uma companhia, tornando impossível vencer o mercado. Essa mesma teoria, objeto de ruidosas disputas, afirma que um investidor que arremessasse dardos na lista de ações de um jornal teria as mesmas chances de desempenho no mercado que qualquer investidor profissional.

Portanto, a hipótese de eficiência de mercado, metodologia sugerida por Roberts (1959), ganhou fama com a publicação de um artigo escrito por Fama (1970) no *Journal of Finance*. Segundo o autor, um mercado no qual os preços dos ativos sempre reflitam completamente todas as informações disponíveis é chamado de eficiente.

Fama constatou, através de testes empíricos, que o preço das ações mudavam aleatoriamente. Todos os testes de eventos revelaram que era impossível uma estratégia de negociação gerar lucros com base em novas informações divulgadas pela empresa. Nesse mesmo sentido, Brealey e Myers (2002) afirmam que em mercados eficientes a compra ou venda de qualquer

título pelo preço vigente no mercado nunca será uma transação com valor presente líquido positivo.

De acordo com Fama (1970), em um mercado eficiente, os preços dos ativos flutuam em torno do seu valor intrínseco, pelo qual novas informações ocasionam rápidas mudanças nesse valor. Após o ajuste pela nova informação, os preços voltariam a flutuar aleatoriamente, o conhecido *random walk*. Os preços de um ativo descrevem um caminho aleatório se as mudanças de preço são independentes e identicamente distribuídas em relação aos preços anteriores.

Quando os preços descrevem um caminho aleatório, a melhor previsão para o preço futuro de um ativo é o seu preço atual, pois este preço representa todas as informações disponíveis sobre o ativo.

Paxson e Wood (2001) afirmam que desde a sua concepção, o conceito de mercado eficiente tem gerado muitas controvérsias na literatura financeira e econômica. A hipótese de mercado eficiente relaciona-se com a questão do preço de equilíbrio dos ativos financeiros, como consequência da concorrência em tais mercados. A implicação mais importante talvez seja que o preço de qualquer título reflete o seu verdadeiro valor, de forma racional; dessa maneira, em um mercado eficiente os investidores acreditam que os títulos que eles adquirirem encontram-se precificados de maneira justa.

O mercado eficiente, porém, não pressupõe que os preços de mercado estejam sempre iguais ao valor real. Conforme Damodaran (1997), os desvios em torno do valor de mercado e do valor venal devem ser aleatórios. Dessa forma, nenhum grupo de investidores seria capaz de, consistentemente, encontrar ações subavaliadas ou superavaliadas utilizando qualquer metodologia de avaliação.

As principais implicações da hipótese de mercado eficiente, segundo Damodaran (1997), são as seguintes:

a) Em um mercado eficiente, a pesquisa e a avaliação do patrimônio líquido seriam uma tarefa cara que não traria benefícios. As probabilidades

de se encontrar uma ação subvalorizada seriam sempre 50:50, refletindo a aleatoriedade dos erros de precificação.

b) Em um mercado eficiente, uma estratégia de diversificação aleatória do perfil de ações ou de indexação aleatória ao mercado, com pouco ou nenhum custo de informações e custos de execução mínimos, seria superior a qualquer outra estratégia que gerasse mais informações e maiores custos de execução.<sup>9</sup>

c) Em um mercado eficiente, uma estratégia de se reduzir ao máximo as negociações, ou seja, criar uma carteira e não alterá-la a menos que houvesse necessidade de caixa, seria superior a qualquer estratégia de investimento.

Por outro lado, Ross, Westerfield e Jaffe (2002) afirmam que mesmo os mais ardorosos defensores da hipótese do mercado eficiente, admitem algum grau de ineficiência ou a existência de operações dos chamados *insiders*<sup>10</sup> e não se surpreenderiam se fosse verificado que os mercados são ineficientes na sua forma mais forte.

## 2.2 EFICIÊNCIA INFORMACIONAL

De acordo com Shannon e Weaver (1975), informação é um processo de seleção efetuado por algum agente, entre eventos passíveis de ocorrer em um dado ambiente. Na origem da informação encontram-se, de um lado, sinais físicos emanados de um objeto ou ambiente qualquer, na forma de vibrações sonoras, radiações elétricas ou luminosas, moléculas odoríferas etc; e, de outro lado, um agente (ou sujeito) capaz de extrair algum sentido, ou orientação, ou significado desses sinais.

---

<sup>9</sup> Ou seja, uma estratégia ingênua de investimento superaria qualquer estratégia mais cara.

Por este motivo, para que ocorra informação haverá sempre necessidade de interação (ou comunicação) entre um sujeito e um objeto, ou sujeito a sujeito.

Segundo Dantas (2003, p.13), a era atual pode ser caracterizada como “a era da informação”. O autor define a entrada desta nova era como sendo a terceira revolução industrial:

A denominada “terceira revolução industrial” ou “revolução da informação” deve o seu impulso inicial à Segunda Grande Guerra, quando vultosos recursos financeiros e equipes científicas da mais alta qualificação foram mobilizados para desenvolver máquinas capazes, tanto de efetuar rapidamente cálculos muito complexos, quanto, principalmente, de processar a informação: reunir dados de diferentes origens, postos em forma numérica (binária) e associá-los conforme algoritmos pré-definidos.

Com a massificação dos computadores, catalisada pelo uso intenso e progressivo da internet, a informação passou a ter um valor cada vez maior. Antes da revolução digital da década de 90, a informação que vinha escrita em uma notícia nos jornais era absorvida gradualmente pela sociedade. Hoje, a informação é disponibilizada, muitas vezes, durante o evento. Ou seja, a queda das torres gêmeas foi um evento assistido por toda a população do planeta; bastava-se ter acesso à internet ou televisão. Pelo fato de termos acesso cada vez mais rápido à informação, a sociedade passou a valorizá-la com mais frequência, em busca de conseqüências do tipo causa e efeito para decidir os rumos da vida e do trabalho de cada pessoa.

A pesquisa de Jones (2000) é um exemplo de valoração da informação nesta nova era. O autor encontrou retornos anormais positivos no mercado britânico – um dos mercados mais conservadores – para anúncios de investimento em pesquisa e desenvolvimento. Na segmentação por tamanho da empresa, as maiores empresas apresentaram retornos anormais significativos nos anúncios de pesquisa e desenvolvimento.

---

<sup>10</sup> A melhor tradução para o termo *insider*, no contexto de finanças é “pessoa bem informada”, que seria equivalente aquela pessoa detentora de informações privilegiadas, em sua maioria, ainda não divulgadas ao público.

O fato de a informação de investimentos em P&D ser relevante, mostra que o mercado está, de uma certa forma, apostando no futuro da empresa, pois este tipo de investimento não gera fluxo de caixa instantâneo. É um sinal de valoração de uma informação, sem a devida necessidade de geração de caixa<sup>11</sup> associada ao evento. E essa mais-valia determinada pelo mercado é gerada no momento da transformação de uma informação privada em pública, objetivo do anúncio.

Bromwich (1992) diferencia informação pública de uma informação privada da seguinte forma: uma informação só pode ser definida como pública quando esta é apresentada da mesma forma para todos os participantes do mercado em um mesmo instante. Esta característica de informação pública resulta no sentido de que ninguém estará disposto a pagar por esta informação, pois este investimento não poderá ser recuperado como gerador de retornos diferenciados.

Segundo Leite e Sanvicente (1990), a ocorrência de um evento envolvendo as ações de uma empresa terá “conteúdo informacional”, ou “valor” como informação se nos disser algo que ainda não sabemos. Portanto, teoricamente, uma informação pública não deveria ter valor. Porém, o momento em que a informação privada torna-se pública (publicação da reavaliação de ativos, por exemplo) pode gerar retornos diferenciados se o valor dessa informação for superior à expectativa do mercado. Estima-se, portanto, que o mercado seja um agente determinante na valoração das informações fornecidas pelas empresas. E a principal característica é a maneira pela qual esta informação está sendo utilizada pelo mercado.

Colaborando com esta tese, Hopwood (1973), considera mais importante que a utilização da informação contábil, na avaliação da performance gerencial, objeto do seu estudo, a maneira como essa informação será utilizada.

Portanto, verificar o comportamento do mercado a uma nova informação, quando esta se torna pública, contribui para a identificação da percep-

---

<sup>11</sup> Tem uma geração de valor, mas não de caixa.

ção de valor que o mercado tinha e, conseqüentemente, do valor atingido pela utilidade da informação.

### 2.3 TESTES DE EFICIÊNCIA DE MERCADO

Inicialmente, testes de eficiência de mercado eram aplicados sem arcabouço teórico. Fama (1970), trouxe uma nova visão para a eficiência dos mercados. O autor disponibilizou, através da Hipótese de Eficiência de Mercado (HEM), um pensamento crítico sobre a dúvida de que os mercados podem, eventualmente, apresentar alguma ineficiência.

O autor classificou em três os níveis de eficiência de um mercado:

a) Eficiência Fraca: nenhum investidor poderia obter retornos em excesso mediante a análise dos preços históricos. A partir desta afirmação deduz-se que a análise técnica não teria utilidade.

b) Eficiência Semi-Forte: nenhum investidor poderia obter retornos em excesso baseado em informações públicas (demonstrações financeiras, notícias), ou seja, os preços rapidamente se ajustariam às novas informações. A partir desta afirmação, a análise fundamentalista não teria utilidade.

c) Eficiência Forte: nenhum investidor poderia obter retornos anormais usando qualquer informação, seja ela pública ou privada, ou seja, os preços refletem todas as informações acerca dos ativos.

Posteriormente, Fama (1991) reclassificou os testes de eficiência de mercado para possível aplicação em diferentes mercados e contribuindo para a confirmação ou rejeição das hipóteses de eficiência. Portanto, os testes de eficiência foram redefinidos da seguinte forma:

a) **Testes de Previsibilidade de Retornos**: baseados na aleatoriedade temporal das taxas de rentabilidade das ações. Verifica a hipótese de eficiência de mercado na forma fraca.

b) **Testes de Estudos de Evento**: baseados na análise dos ajustes de preços a informações novas no mercado e sua velocidade de ajuste. Verifica a hipótese de eficiência de mercado na forma semi-forte.

c) **Testes de Informação Privada**: baseados na verificação de obtenção de retornos maiores por parte de *insiders*, utilizando critérios de definição diversos. Verifica a hipótese de eficiência de mercado na forma forte.

Na seqüência, são detalhados os testes de eficiência propostos por Fama (1991) e as evidências empíricas de ineficiência dos mercados na forma semi-forte, ou seja, através de estudos de eventos, objeto do presente estudo.

## 2.4 TESTES DE PREVISIBILIDADE DE RETORNOS

Os testes de previsibilidade de retornos, originalmente denominados testes de eficiência na forma fraca, foram os primeiros testes a surgirem na teoria econômica. Os primeiros pesquisadores a testarem a “caminhada aleatória”, ou *random walk*, foram Kendall (1953) e Roberts (1959). Kendall (1953), após ter efetuado um teste com 22 séries de ações e commodities do Reino Unido, verificou que os dados possuíam correlação serial próxima de zero, ou seja, seguiam um caminho sem rumo.

Roberts (1959), realizou um teste cego da seguinte forma: simulou uma série de retornos de títulos com números aleatórios e outra com retornos efetivamente realizados, e concluiu que não era possível diferenciar uma série da outra.

Estes são exemplos de testes de correlação serial, que objetivam identificar se existe alguma correlação entre os dados que possa ser traduzida como informação relevante para dados futuros.

A expressão básica de um teste de correlação serial, de acordo com Elton e Gruber (1995), pode ser descrita da seguinte forma:

$$r_t = a + br_{t-1-T} + \varepsilon_t \quad (2.1)$$

Onde, o termo  $a$  significa o retorno esperado não correlacionado com o retorno anterior, o termo  $b$  significa a forma de correlação do retorno atual com o retorno anterior. Se  $T=0$  trata-se do relacionamento entre o retorno de hoje e o retorno de ontem.  $\varepsilon$  é o erro aleatório, incorpora a variabilidade do retorno não relacionado com o anterior.

Outros testes que analisam a independência dos retornos são descritos por Fama (1965). O teste de corrida de sinal busca identificar a direção ou a tendência da série temporal. Segundo Fama (1965), existem três corridas possíveis para configurar uma tendência: positiva, negativa ou sem mudanças.

Os testes de estratégia, ou testes de regra de filtro se assemelham muito à análise do comportamento da série ao que os analistas técnicos (grafistas) chamam de suportes e resistências. Neste teste, um patamar de preços é definido e, enquanto não for atingido não devem ser tomadas ações de compra ou venda do respectivo título. Menos formal que os testes de correlação serial e os de corrida de sinal, esta técnica testa o desempenho de estratégias<sup>12</sup> de compra e venda de títulos em comparação ao resultado da técnica passiva do *buy and hold*.

---

<sup>12</sup> Atualmente existem inúmeros indicadores técnicos que servem de matéria-prima para a confecção de *trading systems*. Os analistas técnicos utilizam estes pacotes de indicadores para identificar momentos de compra e venda de títulos, em sua maioria no curto prazo.



## 2.5 TESTES DE ESTUDO DE EVENTOS

Os estudos de eventos, originalmente denominados como testes para analisar a eficiência semi-forte de um mercado, têm como pressuposto básico que nenhum investidor pode obter retornos extraordinários (anormais) utilizando-se de informações públicas.

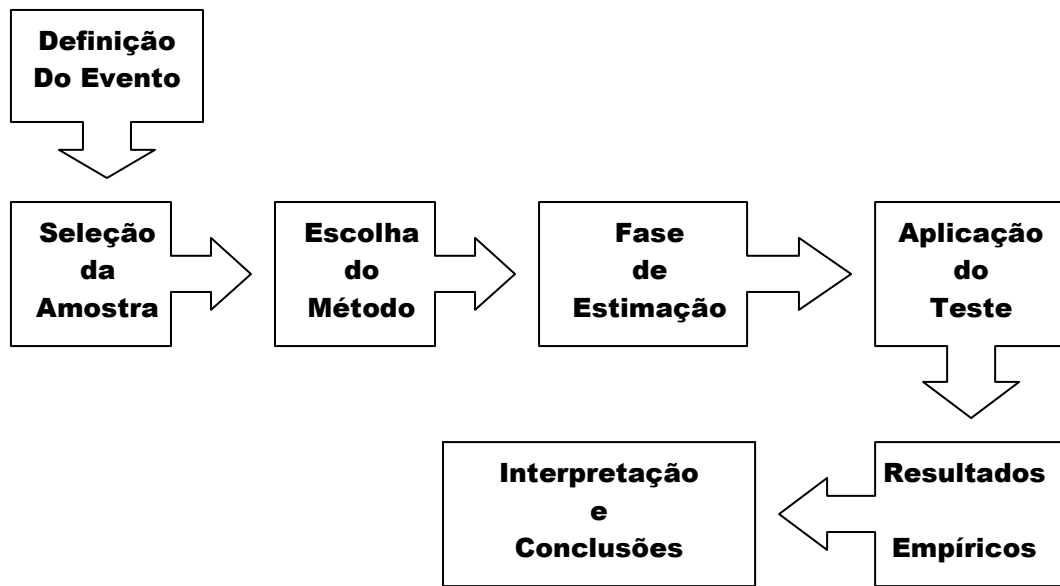
Estes testes se preocupam mais em medir a velocidade de ajuste dos preços nos dias próximos ao dia da publicação (anúncio), ao invés de medir a intensidade<sup>13</sup> deste ajuste, tendo em vista as limitações do método atual de detecção de retornos anormais.

Segundo Paxson e Wood (2001), o teste denominado “estudo de evento” descreve uma metodologia empírica que objetiva estudar o impacto das informações públicas nos retornos dos títulos. Uma vez parametrizado o comportamento normal da amostra, calcula-se o retorno anormal, seja ele superior ou inferior, que, em tese, seria uma consequência do evento em estudo.

De acordo com MacKinlay (1997) e Campbell, Lo e MacKinlay (1997), um estudo de evento engloba uma seqüência de fases que podem ser descritas conforme Figura 2.2.

---

<sup>13</sup> A significância dos ajustes é medida, a amplitude comparativa é que fica prejudicada.



**Figura 2.2 – Fases de um Estudo de Evento**

As fases descritas na Figura 2.2, assim como as suas características, são detalhadas na seqüência.

### 2.5.1 Definição do evento

Uma vez definido o evento a ser estudado, marca-se como data zero a data em que a notícia sobre o evento se tornou pública. Ao redor desta data, são coletados os preços de cada elemento da amostra. O período entre os preços pode variar, sendo mais comum a utilização de preços mensais, como utilizado por Kloeckner (1995), ou cotações diárias, conforme Soares, Ros-tagno e Soares (2002).

A estratégia de coleta de preços ao redor da data zero (data do evento), visa identificar duas possíveis situações: (1) se houve indícios de vazamento de informações, ou seja, movimentações no mercado antes do anúncio; ou (2) medir a velocidade de ajuste dos preços frente à nova informação,

pressupondo que o evento em estudo seja relevante e, por este motivo, ocorra um ajuste de preços nos títulos.

A data zero (data do evento) é o ponto de alinhamento dos elementos da amostra. Todas as empresas são alinhadas por este ponto. São, portanto, criadas três janelas de períodos, conforme ilustrado na Figura 2.3: (1) a janela de estimação que é utilizada para calcular os parâmetros do modelo de determinação de retornos anormais; (2) a janela do evento que deve compreender os dias ao redor do evento que deverão ser analisados; e (3) a janela pós-evento, ou janela de comparação, utilizada para medir se houve alteração significativa nos preços dos títulos pós-evento<sup>14</sup>.

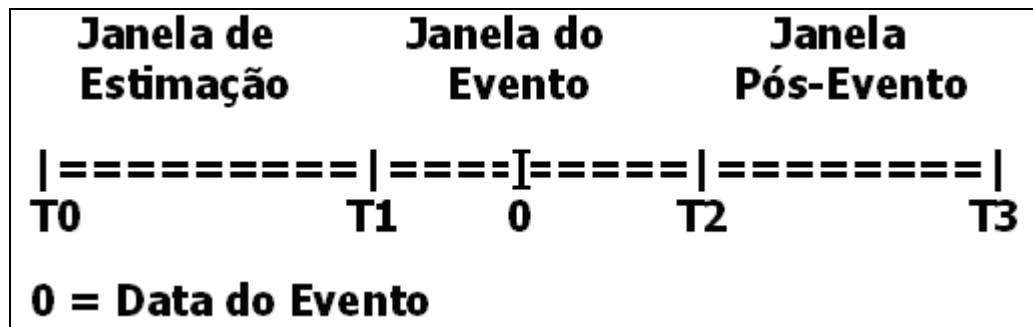


Figura 2.3 – Linha do Tempo de um Estudo de Evento

### 2.5.2 Seleção da amostra

Após a definição do evento que será estudado, faz-se a coleta de dados dos elementos que irão compor a amostra. Nesta fase, torna-se necessária a realização de um “filtro”, por parte do pesquisador, para a seleção da amostra. As ações que irão compor a amostra precisam ter em comum o anúncio do evento. Acontece que o evento pode ter impacto diferenciado entre as empresas, sendo que para empresas grandes, uma aquisição, fusão ou

<sup>14</sup> Um dos testes estatísticos utilizados para medir a diferença entre a média das janelas é o teste *One-Way ANOVA*, um teste de análise de variâncias que determina se as médias, e outras estatísticas, de duas ou mais populações são iguais.

alienação (por exemplo), pode não ocasionar retornos anormais significativos, o que para uma empresa muito pequena seria evidente.

Um outro aspecto que deve ser levado em conta quando da coleta de dados, diz respeito ao tipo de empresa que irá compor a base amostral. Empresas multinacionais, muitas vezes, com o uso de artifícios contábeis em função da necessidade de conversão da contabilidade brasileira para os padrões internacionais, ajustam os resultados da subsidiária para reduzir a carga tributária, fato este que leva o impacto do evento para as ações do grupo no exterior. Outra questão é a utilização de empresas do setor bancário que diferem de outros setores da economia, com controles específicos e enquadramentos contábeis diferenciados.

Portanto, os critérios que foram utilizados na seleção das ações que compõem a amostra devem ser detalhados e justificados, para que os resultados tenham sustentação teórico-empírica. Em geral, em pesquisas empíricas, as decisões dos pesquisadores sobre o delineamento do método podem ser mais importantes que os resultados estatísticos propriamente ditos.

### 2.5.3 Escolha do método

Para definir o método de cálculo do retorno anormal, primeiramente define-se como retorno anormal a diferença entre o retorno real (observado) e o suposto retorno normal, construído com base em critérios e pressupostos estatísticos e econômico-financeiros, como será visto a seguir.

O retorno anormal de um título  $i$  em uma data  $t$ , de acordo com Carmegos (2002), pode ser descrito da seguinte forma:

$$RA_{it} = R_{it} - E(R_{it} | X_t) \quad (2.2)$$

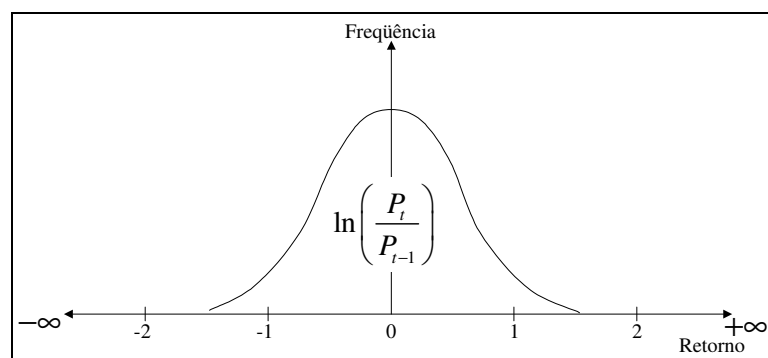
onde,  $RA_{it}$ ,  $R_{it}$  e  $E(R_{it}|X_t)$  são, respectivamente, o retorno anormal, o retorno observado do título (efetivo) e o retorno normal teórico do ativo  $i$  no período  $t$ , com base nas informações  $X$ , condicionantes para o modelo de geração de retornos normais. Supondo que  $E(R_{it}|X_t) = E(R_{it})$ , tem-se:

$$RA_{it} = R_{it} - E(R_{it}) \quad (2.3)$$

Para a determinação do retorno por período, Soares, Rostagno e Soares (2002) justificam a utilização do retorno na forma logarítmica por resultar em uma distribuição mais próxima da distribuição normal, possibilitando a utilização de testes estatísticos paramétricos, como o teste t para a análise dos resultados. Segundo os autores, a distribuição de freqüência dos retornos pela forma logarítmica resultaria em uma distribuição normal semelhante à descrita na Figura 2.4.

O preço do título, em uma capitalização contínua, pode ser medido da seguinte forma:

$$R = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (2.4)$$



**Figura 2.4 – Distribuição de freqüência dos retornos pela forma logarítmica**

Fonte: Soares, Rostagno e Soares (2002, p. 6)

Apesar de demonstrado o benefício da utilização do logaritmo do retorno dos títulos, não existem estudos que comprovem o seu benefício em

termos de detecção do retorno anormal, em detrimento da utilização dos retornos da forma discreta. Trabalhos envolvendo ambas as capitalizações são validados com estatísticas não-paramétricas, ou paramétricas quando satisfeitos os seus pressupostos, sem a perda de sua integridade.

#### **2.5.4 Modelos para determinação dos retornos anormais**

De acordo com Brown e Warner (1980, 1985), MacKinlay (1997) e Campbell, Lo e MacKinlay (1997), os modelos para detecção dos retornos anormais podem ser divididos em duas categorias: (1) modelos com base em pressupostos estatísticos e (2) modelos com base em pressupostos econômico-financeiros.

#### **2.5.5 Modelos estatísticos**

Os modelos estatísticos não dependem de pressupostos econômico-financeiros. Porém, algumas premissas devem ser satisfeitas, as quais são descritas a seguir:

##### **a) Modelo de retorno ajustado à média**

Neste modelo, utiliza-se apenas a média dos retornos passados como base para analisar o comportamento dos retornos no entorno do evento em estudo.

O retorno anormal seria a diferença entre o retorno observado e uma constante, e pode ser descrito da seguinte forma:

$$RA_{it} = R_{it} - \overline{R}_i \quad (2.4)$$

onde,  $\overline{R}_i$  significa a média aritmética dos retornos passados da ação  $i$ .

Pesquisadores que utilizam este modelo, acreditam que exista uma tendência de regressão à média em que os valores, uma vez deslocados em função de um evento, tendem à retornar ao seu curso natural.

Segundo Sá (2001, p. 350), a justificativa para esta tese é que empresas não vão sistematicamente mal ou bem, senão existiriam apenas as muito ruins e as muito boas:

O mercado de títulos é mais volátil que os fundamentos econômicos das empresas porque irracionalmente extrapola a média para regressar a ela. Na regressão à média, as séries históricas de preços contém, pois, a memória do regresso, contraditando que o mercado seja eficiente.

Porém, o autor, em contraposição, afirma que nunca se sabe quando a cotação vai regressar à média, tampouco se a média ficou constante ao longo dos ciclos, de modo que o fenômeno da regressão não traduz de que forma ele pode ser usado para prever o futuro.

A probabilidade de erro neste modelo é proporcional ao tamanho do período utilizado como parâmetro para a média (constante), pois a tendência de um ativo financeiro é, no longo prazo, superar os investimentos sem risco oferecendo retornos maiores, similares à classe de risco em que se encontra. Este motivo já seria suficiente para justificar uma média futura superior à média anterior.

## b) Modelo de retorno ajustado ao mercado

Neste modelo, o *benchmarking* para a determinação do retorno anormal é um indicador de retorno do mercado onde se situa a ação.

Algebricamente, o retorno anormal é medido pela diferença do retorno observado com o retorno de mercado, e pode ser descrito assim:

$$RA_{it} = R_{it} - R_{mt} \quad (2.5)$$

onde  $R_{mt}$  representa o retorno do portfólio de mercado observado no mesmo período, em geral, representado pelo índice Ibovespa ou FGV-100.

Assume-se, com este modelo que: (1) os retornos são iguais para todos os ativos e equivalentes ao retorno do mercado e (2) que o retorno do mercado pode ser descrito como uma soma ponderada em relação ao volume financeiro negociado.

Como visto anteriormente, a hipótese de que o índice Ibovespa seja adequado para expressar o retorno de mercado é confrontada por diversos autores, assim como Penteado e Famá (2002) que entendem que a forma de cálculo<sup>15</sup> do índice não corresponderia ao que preconiza a teoria proposta por Sharpe (1964) quando da composição do modelo CAPM.

Uma variação na determinação do retorno anormal descrito em 2.5, proposta por Brown e Warner (1985), utilizada inicialmente por Leal (1988/1989) e, posteriormente por Bueno, Braga e Almeida (2000) é calcular o valor do IC – Índice de Comparação, que consiste em calcular a razão entre a variação de preços de um ativo e a variação de preços de um índice representativo de mercado.

---

<sup>15</sup> A participação das empresas no índice Ibovespa é ponderada pela liquidez, sendo que a proposta teórica de Sharpe era de que um índice de retorno de mercado deveria ser aquele em que cada ativo tenha participação proporcional ao seu valor de mercado.



A análise pode ser efetuada para dias anteriores ou posteriores ao anúncio/evento, apenas alterando a posição dos valores dos preços da data zero de numeradores para denominadores.

Um exemplo algébrico deste modelo, utilizado para detectar utilização de informações privadas, pode ser descrito da seguinte forma:

$$LnIC = Ln \left( \frac{\frac{P_0}{P_t}}{\frac{I_0}{I_t}} \right) \quad (2.6)$$

onde,

$P_0$  é o preço da ação na data do evento;

$P_t$  é o preço da ação  $t$  dias/pregões antes do anúncio/evento;

$I_0$  é a cotação do índice de mercado na data do evento;

$I_t$  é a cotação do índice de mercado  $t$  dias/pregões antes do evento.

Em estudos de eventos realizados nos Estados Unidos, segundo Brown e Warner (1980), o modelo mais utilizado para o cálculo dos retornos anormais foi o modelo ajustado ao mercado.

Testes realizados por Kloeckner (1995) e Soares, Rostagno e Soares (2002) confirmaram as conclusões obtidas por Brown e Warner (1980 e 1985) que concluíram que não há melhorias significativas na utilização de metodologias mais sofisticadas do que o modelo ajustado ao mercado. Todos os modelos que serão descritos na seqüência obtiveram resultados semelhantes na capacidade de detecção de retornos anormais sendo a sua utilização uma opção (e não necessidade) do pesquisador.

### c) Modelo de retorno ajustado ao risco e ao mercado

O pressuposto básico do modelo ajustado ao risco e ao mercado é que exista uma relação linear entre o retorno da ação e o retorno de mercado.

Segundo Brown e Warner (1980), o modelo pode ser descrito da seguinte forma:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad (2.7)$$

sendo,  $E(\varepsilon_{it}) = 0$  tem-se:

$$E(R_{it}) = \alpha_i + \beta_i E(R_{mt}) \quad (2.8)$$

onde,

$R_{it}$  é o retorno observado da ação  $i$  no instante  $t$ ;

$R_{mt}$  é o retorno observado da carteira de mercado no instante  $t$ ;

$\alpha_i$  e  $\beta_i$  são parâmetros *OLS – Ordinary Least Squares* – estimados pela regressão linear.

Segundo Campbell, Lo e Mackinlay (1997), admitindo-se que os retornos da ação e da carteira de mercado tenham distribuição normal, é possível estimar os parâmetros  $\alpha_i$  e  $\beta_i$  da seguinte forma:

$$\beta_i = \frac{Cov(R_{it}, R_{mt})}{Var(R_{mt})} \quad (2.9)$$

$$\alpha_i = E(R_{it}) - \beta_i E(R_{mt}) \quad (2.10)$$

Segundo os autores, os parâmetros devem ser calculados fora da janela do evento, para que a modelagem não seja influenciada pelo evento.

Logo, o retorno anormal de um ativo pode ser obtido pela diferença entre o retorno observado do ativo e a relação linear entre o ativo e a carteira de mercado:

$$RA_{it} = R_{it} - (\alpha_i + \beta_i R_{mt}) \quad (2.11)$$

### 2.5.6 Modelos econômico-financeiros

A esperança dos pesquisadores ao utilizarem os modelos econômico-financeiros é que, uma vez incluídos os pressupostos teóricos da Teoria Financeira nos modelos estatísticos, resultaria em uma modelagem mais robusta e, conseqüentemente, mais explicativa dos possíveis retornos normais, evidenciando anormalidades em dimensões mais precisas e realistas.

#### a) Modelo CAPM

O modelo CAPM – *Capital Asset Pricing Model*, proposto e estabelecido por Sharpe (1964), Lintner (1965) e Mossin (1966) foi desenvolvido a partir da contribuição de Harry Markowitz<sup>16</sup> (1952) que em seu artigo intitulado *Portfolio Selection*, publicado no *Journal of Finance*, ao utilizar as técnicas de programação quadrática para combinar ações em uma carteira, verificou que

---

<sup>16</sup> Markowitz (1952) desenvolveu um modelo capaz de otimizar uma carteira de investimentos fornecendo (1) a maior taxa de retorno esperada frente à um determinado nível de risco ou (2) o menor nível de risco para uma determinada taxa de retorno esperada, investimentos localizados na chamada “fronteira eficiente”.

é possível aumentar o retorno e, ao mesmo tempo, reduzir a variância da carteira, ou seja, era descoberta a fronteira eficiente de uma carteira de títulos. A partir deste momento, risco e variância se tornaram inseparáveis.

Anos mais tarde, William Sharpe (1963) contribuiu com Markowitz reduzindo a necessidade de cálculo de covariâncias entre os papéis individuais. Sua solução foi estimar a variância de cada papel em relação ao mercado como um todo.

Dessa simplificação, junto com os trabalhos de Treynor (1961) e, posteriormente complementados por Lintner (1965) e Mossin (1966), surgiu o modelo do CAPM que expressa o valor teórico dos ativos se todos os investidores seguissem à risca as recomendações de Markowitz para a formação de suas carteiras.

Segundo Copeland, Weston e Shastri (2005), “a relação de equilíbrio entre risco e retorno proposta pelo modelo CAPM, utiliza as seguintes premissas:

- a) Os investidores são indivíduos avessos ao risco, os quais buscam maximizar a utilidade esperada de sua riqueza no final do período do investimento;
- b) Os investidores são tomadores de preço que têm expectativa homogênea sobre o retorno dos ativos, os quais assumem distribuição normal;
- c) Existe um ativo livre de risco, o qual se pode emprestar e tomar emprestado sem limites de quantias;
- d) A quantia de ativos é fixa. Todos os ativos são divisíveis e negociáveis no mercado;
- e) Os ativos do mercado estão acessíveis a todos, as informações não tem custos e estão disponíveis a todos os investidores. Os

custos e taxas de captação são iguais aos custos e taxas de aplicação;

- f) Não existem imperfeições de mercado, tais como: impostos, regulações e restrições para vendas de curto prazo”;

Conforme Haugen (2000), o modelo do CAPM parte do pressuposto de que todos os investidores utilizam a ferramenta de otimização de Markowitz para montar carteiras. O modelo sugere ainda: (1) que o índice representativo de mercado é uma das carteiras eficientes de Markowitz e (2) que o único fator responsável pelas diferenças de retorno entre os títulos é a sensibilidade destes às oscilações do mercado (medido pelo “beta”).

Logo, o CAPM é um modelo que relaciona o retorno de um título com o retorno não diversificável (retorno de mercado), considerando esta relação o único fator de diferenciação entre os retornos. Algebricamente, o modelo pode ser descrito da seguinte forma:

$$E(R_i) = R_f + \beta_i [E(R_m) - R_f] \quad (2.12)$$

onde,

$E(R_i)$  é o retorno esperado do título  $i$ ;

$E(R_m)$  é o retorno esperado da carteira de mercado;

$R_f$  é a taxa livre de risco;

$\beta_i$  é a relação de risco entre o título  $i$  e a carteira de mercado.

Dessa forma, o retorno anormal medido com o uso do modelo CAPM pode ser descrito da seguinte forma:

$$RA_{it} = R_{it} - \{R_f + \beta_i [E(R_m) - R_f]\} \quad (2.13)$$

Kloekner (1995, p. 263), propôs uma forma alternativa de medir o retorno anormal com as premissas do CAPM, forma equivalente à utilizada por

Black (1972). Segundo Kloeckner (1995, p. 263), alternativamente, poderia ser utilizada a versão original do CAPM e, no lugar do  $R_{zt}$  utilizar-se-ia a taxa livre de risco  $R_f$  :

$$RA_{it} = R_{it} - [R_{zt}(1 - \beta_i) + \beta_i R_{mt}] \quad (2.14)$$

“sendo  $R_{zt}$  o retorno no período  $t$  de um portfólio de variância mínima, cujos retornos não são correlacionados com os retornos do portfólio de mercado.”

## b) Modelo APT

O modelo APT – *Arbitrage Pricing Theory*, proposto por Ross (1976), se difere do modelo CAPM por diversos motivos, sendo o principal o fato de que o modelo APT pressupõe que os retornos dos títulos são gerados por uma série de fatores setoriais e/ou gerais de mercado (i.e., há correlação entre os retornos de dois títulos quando são influenciados pelo mesmo fator). Outra diferença é o fato de o modelo CAPM supor que os retornos dos ativos têm distribuição normal, hipótese não necessária no modelo APT.

Na versão mais simplificada, o conceito de arbitragem, conforme cita Neftci (1996), significa tomar posições simultâneas em títulos diferentes que, sem risco, garanta um retorno superior ao que seria possível adquirir investindo em um ativo livre de risco. Se essa estratégia existe, diz-se que há uma oportunidade de arbitragem.

O princípio por trás da teoria da arbitragem é que seria impossível montar duas carteiras totalmente livres de risco com retornos diferenciados, pois essa diferença seria rapidamente eliminada pelo processo de arbitragem.

A teoria de arbitragem é mais robusta do que a teoria do CAPM, sendo algumas das justificativas descritas abaixo:

- a) Não é assumido nenhum pressuposto sobre a distribuição normal dos retornos dos ativos;
- b) Não é feita nenhuma suposição sobre a função utilidade individual do investidor;
- c) A teoria permite que a taxa de retorno de equilíbrio dos ativos seja dependente de vários fatores e não apenas do risco do mercado;
- d) Qualquer cesta de ativos pode ser considerada, sem a necessidade de envolver o universo dos ativos existentes;
- e) Não há necessidade de se estabelecer um portfólio de mercado;
- f) O APT é facilmente estendido para análise multiperíodo.

Em suma, a Teoria da Arbitragem (APT) estabelece uma análise fatorial dos elementos que mais influenciam no retorno do ativo. É criada uma função linear e seus resultados são testados empiricamente.

A fórmula básica do modelo é a seguinte:

$$RA_{it} = R_{it} - [E(R_i) + \beta_{i1}F_1 + \beta_{i2}F_2 + \dots + \beta_{in}F_n + \varepsilon_i] \quad (2.15)$$

onde,

$RA_{it}$  é o retorno anormal do título  $i$ , no período  $t$ ;

$R_{it}$  é o retorno do título  $i$ , no período  $t$ ;

$E(R_i)$  é o retorno esperado do título  $i$ ;

$\beta_{i1}$  representa a relação linear entre o título  $i$  e o fator 1 ( $F_1$ ), e assim sucessivamente para os demais betas e fatores.

Portanto, no modelo APT, a relação entre o retorno esperado do título e as sensibilidades aos fatores deve ser linear, podendo ser positiva ou negativa.

De acordo com Soares et al. (2002), a teoria não especifica quais são esses fatores, apenas afirma que podem ser fatores de âmbito setorial ou macroeconômico que estão sendo responsáveis por uma parte do risco que não pode ser anulada pela diversificação.

## 2.6 TESTES DE INFORMAÇÃO PRIVADA

Os testes de informação privada, originalmente denominados testes de Hipótese de Eficiência de Mercado da forma forte, partem da premissa que os preços dos títulos já refletem todas as informações sobre os mesmos, seja ela pública ou privada (i.e., um investidor de posse de uma informação privada<sup>17</sup> não conseguiria auferir ganhos superiores aos de outros investidores que não tivessem tal informação, uma vez que o mercado, sendo eficiente, anularia tal possibilidade).

Estes testes, pelo fato da dificuldade em se obter informações privadas, são os menos testados e, pelo fato de que os testes de HEM estejam em níveis seqüenciais de acesso às informações (a hipótese de eficiência fraca só precisa da cotação histórica do título), provocam uma anulação da necessidade de testes mais fortes, uma vez reconhecida a ineficiência em um nível inferior. Em outras palavras, a rejeição da eficiência na forma semi-forte implica, conseqüentemente, na rejeição da eficiência na forma forte.

Uma dificuldade nos testes de informação privada é a falta de conhecimento dessas informações, pelo fato de não estarem publicamente disponíveis. Os principais estudos deste nível de eficiência se concentram em estu-

---

<sup>17</sup> Não pública



dar o desempenho de recomendações de analistas de mercado, administradores de carteira e gestores de fundos de investimento que possivelmente seriam *insiders*.

Uma metodologia alternativa para o estudo da forma forte baseia-se na utilização do método de estudo de evento calculando o retorno anormal “antes” do evento se tornar público. A evidência de retorno anormal antes da informação se tornar pública sugere que este retorno deve-se ao movimento de investidores que detêm informações privilegiadas e, por este motivo, privadas, sendo este retorno adquirido de forma ilegal, sujeito a multas e penalidades por parte dos órgãos reguladores, no caso brasileiro a CVM – Comissão de Valores Mobiliários.

## 2.7 UTILIZAÇÃO DO MÉTODO DE ESTUDO DE EVENTO EM FINANÇAS

Nesta etapa do contexto teórico, são apresentados os testes de Hipótese de Eficiência de Mercado da forma semi-forte, ou seja, estudos empíricos que utilizaram o método de estudo de evento, utilizados para verificar possíveis ajustes nos preços dos títulos antes, durante e depois dos eventos se tornarem públicos.

### 2.7.1 Estudos de eventos no mercado brasileiro

Durante duas décadas, pelo menos, estudos empíricos em finanças têm utilizado o método de estudo de eventos para detectar retornos anormais no entorno de eventos que, teoricamente, deveriam ser incorporados instantaneamente no preço dos títulos.

O método de estudo de eventos, a partir da década de 90, com a massificação das planilhas eletrônicas e dos pacotes estatísticos, se popularizou no meio acadêmico brasileiro, sendo um dos principais métodos de pesquisa em finanças com ênfase em mercado de capitais ou finanças corporativas.

A utilização do método de estudos de eventos tornou-se comum na última década do século XX, pesquisadores, como Lamb (1993), Kelm (1996), Schiehl (1996b), Navarro (1996), Prux (1998), utilizaram o método de estudo de evento em suas dissertações de mestrado ao estudar o impacto no retorno dos títulos de eventos como: variação no ativo permanente, alterações nos demonstrativos contábeis, resultados econômico-financeiros, entre outros.

Estudos empíricos como os de Leal (1988/1989), Leal e Amaral (1990) e Leite e Sanvicente (1991) foram um dos primeiros a utilizar o método de estudo de evento no Brasil. Leal (1988/1989) e Leal e Amaral (1990) estudaram o efeito da emissão pública de ações e utilizaram o modelo de mercado adaptado em Índice de Comparação (IC) com dados diários. Leite e Sanvicente (1991), usando o modelo de mercado, testaram o efeito da publicação das demonstrações contábeis. Bertucci, Amaral e Garcia (2002), estudaram o efeito da subscrição de ações no mercado brasileiro. Os autores utilizaram o modelo de mercado e cotações diárias.

Para ilustrar a importância e disseminação do método, foram levantados os trabalhos empíricos apresentados nos Enanpad's (Encontros da ANPAD) em um período de 10 anos, entre 1996 e 2006, que utilizaram o estudo de evento como método para análise de retornos anormais.

A Tabela 2.1 apresenta a quantidade de trabalhos neste período em função dos modelos para cálculo dos retornos anormais utilizados:

Verifica-se a predominância do uso do modelo ajustado ao mercado, mantendo coerência com a preferência em estudos americanos. Quase a metade dos estudos utilizou o modelo de retorno ajustado ao mercado e, em segundo lugar, o modelo ajustado ao risco e ao mercado. O modelo CAPM, talvez pela difícil decisão de escolher um ativo que possa representar o ativo

livre de risco (e.g., Selic, C-Bonds, CDI), foi o modelo ajustado ao mercado menos utilizado pelos autores.

**Tabela 2.1 – Modelos de Estudo de Eventos utilizados em encontros da ANPAD**

<b>Modelo</b>	<b>Número de trabalhos (1996-2006)</b>
Retorno ajustado à média	2
Retorno ajustado ao mercado	13
Método do Índice de Comparação	4
Retorno ajustado ao risco e ao mercado	7
Retorno ajustado ao mercado - CAPM	3

O maior número de estudos de eventos em finanças apresentados em Enanpad's ocorreu em 2004, com a apresentação de sete estudos empíricos, como pode ser visto na Tabela 2.2.

A Tabela 2.3 apresenta a relação completa dos trabalhos apresentados nos encontros da ANPAD, no período de 1996 a 2006. Analisando a Tabela 2.3, verifica-se que o evento mais estudado é o impacto nas ações quando da publicação de demonstrações contábeis, anuais e trimestrais, seguido dos eventos de fusões e aquisições, governança corporativa e emissão de ADR's.

**Tabela 2.2 – Estudos de Evento em ENANPAD por ano**

<b>Ano</b>	<b>Estudos de Evento</b>
1996	1
1997	1
1998	2
1999	1
2000	2
2001	3
2002	1
2003	2
2004	7
2005	3
2006	3

De uma forma mais diferenciada e específica, as alterações em estrutura de capital, política de dividendos e procedimentos acionários (e.g., subscrição, recompra, desdobramentos, bonificações, entre outros) são estudadas com frequência. Verifica-se, também, o estudo de ações de *insiders*, não apenas como parte mas como objetivo fim da pesquisa. No levantamento efetuado, constatou-se que alguns autores mencionam o método do Índice de Comparação como sendo o modelo de mercado, não especificando que trata-se de uma forma alternativa de cálculo do retorno do modelo de mercado.

Em três estudos, os autores utilizaram dois modelos de cálculo de retorno normal – o modelo ajustado ao mercado e, para a mesma amostra, o modelo ajustado ao risco e ao mercado. A utilização de dois modelos pode representar uma insegurança por parte dos pesquisadores sobre a eficácia dos modelos em capturar os retornos anormais. A questão que deve ser estudada, quando da utilização de mais de um modelo seria o que fazer caso os dois modelos apresentassem resultados opostos?

Conforme descrito anteriormente, os trabalhos de Kloeckner (1995) e posteriormente de Soares et al. (2002) comprovaram que ambos os modelos são eficazes na detecção de retornos anormais, não havendo sinalização de resultados diferenciados que favoreça um em detrimento de outro.

**Tabela 2.3 – Artigos nos encontros da ANPAD que utilizaram estudo de evento**

<b>Autores (ano)</b>	<b>Evento</b>	<b>Cotação</b>	<b>Método</b>
Schiehl (1996)	demonstrações contábeis	mensal	Retorno ajustado ao mercado
Navarro (1997)	política de dividendos	semanal	Retorno ajustado ao mercado
Herrera e Procianoy (1998)	anúncio de concordata	diária	Ajustado ao Mercado + Risco e ao Mercado
Vieira e Procianoy (1998)	desdobramentos e dividendos em ações	diária	Retorno ajustado ao mercado
Cordeiro, Perobelli e Arbex (1999)	eventos macroeconômicos	diária	Retorno ajustado à média
Bueno, Braga e Almeida (2000)	fusões e aquisições	diária	Método do Índice de Comparação
Perobelli e Ness Jr. (2000)	lucros trimestrais	diária	Retorno ajustado ao mercado
Holthausen e Galli (2001)	lançamento de ADR's	diária	Retorno ajustado ao mercado
Procianoy e Antunes (2001)	demonstrações contábeis	diária	Retorno ajustado ao mercado
Vieira e Procianoy (2001)	bonificações	diária	Retorno ajustado ao risco e ao mercado
Novis Neto e Saito (2002)	pagamento de dividendos	diária	Retorno ajustado ao mercado
Comerlato e Terra (2003)	apresentação aos analistas	diária	Retorno ajustado ao mercado - CAPM
Gabrielli e Saito (2003)	recompra de ações	diária	Retorno ajustado ao mercado
Sarlo Neto, Loss e Nossa (2004)	demonstrações contábeis (ON versus PN)	mensal	Retorno ajustado ao risco e ao mercado
Lima e Terra (2004)	informações contábeis trimestrais e anuais	diária	Retorno ajustado ao mercado - CAPM
Brito, Batistella e Famá (2004)	fusões e aquisições no setor bancário	diária	Retorno ajustado ao risco e ao mercado
Medeiros e Matsumoto (2004)	emissão pública de ações e <i>insiders</i>	diária	Retorno ajustado ao risco e ao mercado
Firmino, Santos e Matsumoto (2004)	política de dividendos	diária	Retorno ajustado ao mercado*
Garcia, Sato e Caselani (2004)	lançamento de ADR's	diária	Ajustado ao Mercado + Risco e ao Mercado
Lima e Terra (2004)	informações financeiras versus governança corporativa	diária	Retorno ajustado ao mercado - CAPM
Markoski e Moreno (2005)	classificação de risco	diária	Retorno ajustado à média
Camargos e Barbosa (2005)	fusões e aquisições	diária	Método do Índice de Comparação
Patrocínio, Kayo e Kimura (2005)	criação de valor em fusões e aquisições	diária	Retorno ajustado ao mercado
Camargos e Barbosa (2006)	lançamento de ADR's	diária	Método do Índice de Comparação
Camargos e Barbosa (2006)	governança corporativa	diária	Método do Índice de Comparação
Rochman e Eid. Jr. (2006)	governança corporativa versus ações de <i>insiders</i>	diária	Ajustado ao Mercado + Risco e ao Mercado

\*adaptado

### 2.7.2 Estudos de eventos em outros mercados

Conforme descrito em Soares et al. (2002) e, segundo Brown e Warner (1980), o modelo ajustado ao mercado é o modelo mais utilizado para o cálculo dos retornos anormais em estudos de evento realizados nos Estados Unidos. Dois motivos estariam corroborando com esta escolha: (1) é o modelo de maior simplicidade de cálculo e (2) apresenta resultados similares aos modelos mais sofisticados, como pode ser visto nos trabalhos de Kloeckner (1995) e Soares et al. (2002).

Autores como Chan et al. (1995) estudaram a reação do mercado americano a 447 anúncios de decisões de investimentos entre 1978 e 1990. Os autores segmentaram sua amostra segundo diferentes motivos: expansão do negócio, redução de custos, redução da capacidade produtiva, consolidação de instalações, outras razões e, finalmente, sem razão explícita. O retorno anormal do mercado aos acontecimentos foi mensurado a partir de um estudo de evento, variando significativamente em função do tipo de investimento anunciado – positivamente em se tratando de investimentos que elevam o fluxo de caixa futuro esperado da empresa.

Da mesma forma, a reação do mercado inglês a anúncios de investimentos diversos foi examinada por Jones (2000). O estudo compreendeu 563 ocorrências entre 1991 e 1996, utilizando um estudo de evento com retornos ajustados ao mercado. O autor encontrou retornos anormais positivos para o anúncio de aquisição de ativos imobilizados.

Examinando o mercado espanhol com uma amostra de 80 empresas não financeiras, no período de 1990 a 1997, Del Brio et al. (2002) encontraram evidências significativas de reação positiva do mercado a realizações de investimento, aumento no nível de caixa e no volume de dívida. Porém, encontraram evidências de reação negativa a aumentos no nível de capital próprio na estrutura da empresa.

Uma pesquisa nos principais periódicos de finanças que compreendeu o período entre 1974 e 2000 foi efetuada por Khotari e Warner (2006). Os autores encontraram mais de 500 estudos de evento nos principais periódicos de administração, finanças, economia e análise quantitativa, sendo o auge atingido em 1983, com 38 artigos utilizando o método de estudo de evento. O periódico com maior frequência foi o *Journal of Finance*, com 212 artigos, seguido do *Journal of Financial Economics*, com 207 artigos.

A Tabela 2.4 a seguir apresenta um resumo da pesquisa e evidencia que, de acordo com a previsão dos autores, a tendência é que a utilização do método de estudos de evento continue crescendo.

No mesmo artigo, Khotari e Warner (2006) mencionam, colaborando com a presente tese, que apesar da popularidade do método, a questão da definição do modelo de retornos esperados continua sem solução. Utilizando a expressão de Fama (1998, p. 291 apud Khotari e Warner, 2006), os autores concluem:

On a somewhat depressing note, Fama (1998, p. 291) concludes that “all models for expected returns are incomplete descriptions of the systematic patterns in average returns,” which can lead to spurious indications of abnormal performance in an event study.

Por este motivo, entende-se que a metodologia de estudo de eventos pode ser melhorada. Com a massificação da informática, com a velocidade atual dos processadores, dos sistemas computacionais existentes e com a tecnologia de processamento de dados utilizada por diversas áreas do conhecimento, tem-se o ambiente propício para inovações no que diz respeito à metodologia de detecção dos retornos anormais.

No próximo capítulo são apresentadas as características e aplicabilidade da técnica de separação cega de fontes, conhecida como Análise de Componentes Independentes, que pode contribuir para o avanço tecnológico nas pesquisas em finanças.

**Tabela 2.4 – Estudos de Evento, por ano e periódico. Para cada periódico, todos os artigos que continham um estudo de evento estão incluídos. Artigos de *survey* e metodológicos estão excluídos.**

<b>Year</b>	<b><i>Journal of Business</i></b>	<b><i>Journal of Finance</i></b>	<b><i>Journal of Financial Economics</i></b>	<b><i>Journal of Financial and Quant. Analysis</i></b>	<b><i>Review of Financial Studies</i></b>	<b>Grand Total</b>
1974	2		2	1		5
1975		2	2	1		5
1976		5	1	1		7
1977		5	5	1		11
1978	1	5	4	1		11
1979		7		2		9
1980	3	4	2	2		11
1981	1	3	4	2		10
1982	1	6	2	1		10
1983	2	14	18	4		38
1984		5	5	1		11
1985	2	4	7	2		15
1986	2	7	14	4		27
1987		7	18	1		26
1988	1	4	7	5	1	18
1989		11	11	1	1	24
1990	5	17	7	6	2	37
1991	5	17	2	4	1	29
1992	4	13	9	4	1	31
1993	5	7	5	5	3	25
1994	1	10	10	5		26
1995	1	8	14	11	2	36
1996	1	7	10	5	3	26
1997	3	8	12	3		26
1998	1	14	11	3		29
1999	1	7	12	1	4	25
2000	2	15	13	5	2	37
<b>Totals</b>	<b>44</b>	<b>212</b>	<b>207</b>	<b>82</b>	<b>20</b>	<b>565</b>

Fonte: Khotari e Warner (2006, p. 50)



### 3 ANÁLISE DE COMPONENTES INDEPENDENTES

Neste capítulo, é apresentado o método de Análise de Componentes Independentes: suas características, aplicabilidades - tanto em finanças quanto em outras áreas do conhecimento - e suas limitações.

#### 3.1 DEFINIÇÃO

A técnica denominada Análise de Componentes Independentes (ICA<sup>18</sup>) tem como objetivo identificar e separar fontes<sup>19</sup> estatisticamente independentes, a partir de uma determinada mistura de fontes originais.

A Análise de Componentes Independentes, ou *blind source separation*, é muito utilizada em tratamento de sinais de áudio, na separação de fontes sonoras, o chamado *cocktail-party problem*.

Conforme descrito em Leite (2005), o problema chamado de *cocktail-party problem* pode ser exemplificado da seguinte forma:

- Suponha uma sala com três microfones instalados em locais diferentes, onde estejam três pessoas conversando simultaneamente.

- Cada microfone gravará uma amostra da conversa durante um período  $t$ , que podem ser expressas como  $x_1(t)$ ,  $x_2(t)$  e  $x_3(t)$ . Essas amostras são, portanto, misturas das falas, que podem ser expressas como  $s_1(t)$ ,  $s_2(t)$  e  $s_3(t)$ .

---

<sup>18</sup> a partir deste capítulo, será utilizada a sigla ICA (Independent Component Analysis) como referência para Análise de Componentes Independentes.

<sup>19</sup> Fontes podem ser descritas como elementos independentes que foram misturados.

A expressão matemática resultante da captura de cada microfone, tendo em vista a distância deste com a conversa gravada, pode ser descrita algebricamente da seguinte forma:

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \quad (3.1)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \quad (3.2)$$

$$x_3(t) = a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t) \quad (3.3)$$

onde os pesos  $a_{ij}$  correspondem à capacidade de captura de cada microfone e da distância destes à conversa gravada.

Ao admitir-se que não haja atrasos (*delay*) ou outros fatores adicionais nas três misturas, e, se os parâmetros  $a_{ij}$  fossem conhecidos, o problema se resumiria na solução de uma equação linear. Porém, o problema é que nem as fontes e nem os pesos das misturas são conhecidos.

Partindo destes pressupostos, a técnica de ICA visa identificar e separar as três fontes originais. Vários algoritmos foram elaborados para automatizar a busca de fontes com o uso do método ICA. Os mais utilizados são: JADE, infomax e Fastlca.

Baragona e Battaglia (2007) testaram os três algoritmos nomeados acima. O Fastlca obteve o melhor desempenho na determinação de outliers em séries temporais multivariadas.<sup>20</sup>

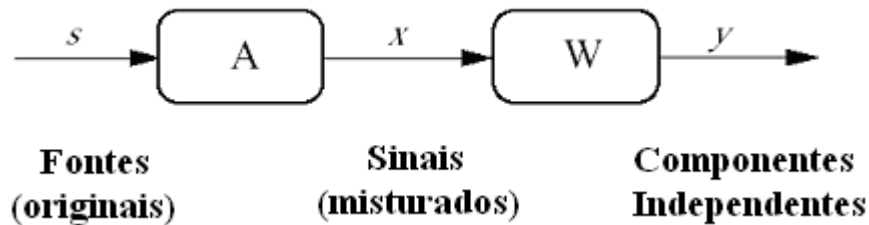
## 3.2 O MÉTODO ICA

O método de Análise de Componentes Independentes (ICA), consiste em determinar vetores estatisticamente independentes ( $y$ ) chamados de

---

<sup>20</sup> Esse experimento será mais comentado no próximo capítulo.

componentes independentes, de uma mistura composta por vetores ( $x$ ), misturados de fontes originais ( $s$ ).



**Figura 3.1 – Representação sistemática do processo de ICA**  
 Fonte: adaptado de Back and Weigend (1997, p. 3)

A Figura 3.1 apresenta a sistemática por trás do método ICA. Partindo do pressuposto da existência de fontes originais ( $s$ ) que foram misturadas através de uma matriz  $A$ , originando os sinais observados ( $x$ ). Uma vez descoberta a matriz  $W$  que anula a mistura (*demixing*) torna-se possível converter os sinais misturados ( $x$ ) nas respectivas fontes originais ( $y$ ).

Conforme descrito em Jutten e Héroult (1991); Comon (1994); Hyvärinen e Oja (2000) e Leite (2005), o modelo de mistura consiste em  $n$  dados observáveis  $x_1, \dots, x_n$  que resultam de uma combinação linear de  $N$  variáveis aleatórias  $s_1, \dots, s_n$  independentes entre si:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n, \forall i = 1, \dots, N \quad (3.4)$$

No modelo acima, apenas as variáveis aleatórias  $x_i$  são conhecidas. A partir delas, serão estimados os coeficientes da mistura  $a_{ij}$  e as componentes independentes  $s_j$ .

Como multiplicação de matrizes, o modelo assume a seguinte forma:

$$x = A.s \quad (3.5)$$

O modelo, segundo Hyvärinen e Oja (2000), pode ser escrito, também, da seguinte forma:

$$x = \sum_{i=1}^n a_i s_i \quad (3.6)$$

Segundo os autores, o modelo ICA é generativo<sup>21</sup>, o que significa que ele descreve como os dados observados são gerados a partir de um processo de mistura das suas componentes.

Após estimar a matrix de mistura (A) torna-se possível obter os componentes independentes usando a sua matriz inversa (W):

$$s = W.x \quad (3.7)$$

A técnica utilizada pelo ICA para encontrar as componentes independentes está relacionada com o método de separação cega de fontes (*BSS – Blind Source Separation*), pois nada é conhecido sobre a matriz de mistura e poucas são as considerações acerca das fontes estimadas.

### 3.2.1 Restrições do Método ICA

Para que o método seja eficaz, algumas restrições nas variáveis são necessárias. Estas seriam, portanto, as pressuposições necessárias para garantir a convergência dos sinais observados em fontes originais.<sup>22</sup>

---

<sup>21</sup> Generativo: que se refere à geração

<sup>22</sup> No capítulo seguinte serão comentadas as pressuposições e os seus impactos em uma amostra de séries temporais.

### 3.2.1.1 As fontes devem ser independentes

Por definição, o método pressupõe que as fontes sejam independentes. Variáveis aleatórias são consideradas independentes quando a informação de uma variável não fornece qualquer informação sobre a outra.

Uma evidência fraca de independência é a ausência de correlação entre duas variáveis. Duas variáveis aleatórias  $y_1$  e  $y_2$  são ditas não correlacionadas se a sua covariância for zero:

$$E\{y_1 y_2\} - E\{y_1\}E\{y_2\} = 0 \quad (3.8)$$

Se duas variáveis são independentes, pode-se afirmar que elas são não correlacionadas. Porém, conforme explicam Hyvärinen e Oja (2000), ausência de correlação não implica em independência.

Leite (2005, p. 29), explica que se duas variáveis aleatórias são independentes, então elas são descorrelacionadas, mas o inverso não é sempre verdade. A autora exemplifica, supondo duas variáveis aleatórias  $x$  e  $y$  tais que  $x$  é uniforme entre  $-1$  e  $1$  e  $y = x^2$ . Dessa forma,  $x$  e  $y$  são variáveis estatisticamente dependentes, o que pode ser confirmado pela expressão:

$$E\{x^2 \cdot y^2\} = E\{x^6\} = \frac{1}{7 \cdot 2^6} \neq \frac{1}{15 \cdot 2^6} = E\{x^2\}E\{x^4\} = E\{x^2\}E\{y^2\} \quad (3.9)$$

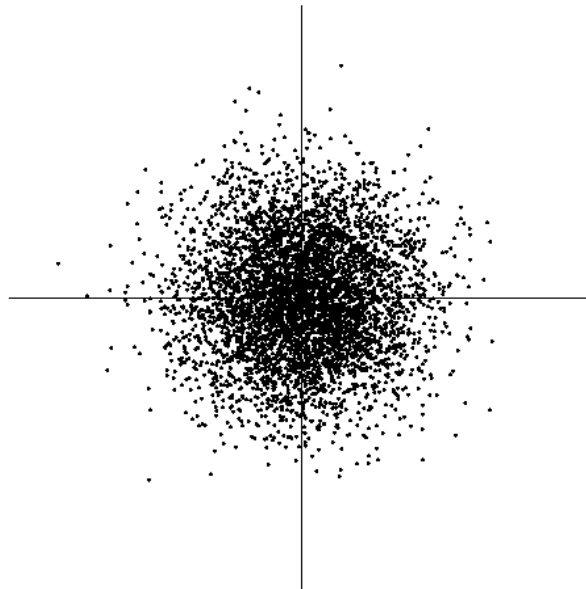
Porém, tem-se que:

$$\text{cov}(x, y) = E\{x^3\} - E\{x\}E\{x^2\} = 0, \quad (3.10)$$

o que demonstra que  $x$  e  $y$  são descorrelacionadas, apesar de dependentes entre si.

### 3.2.1.2 As fontes devem apresentar distribuições não gaussianas

Outra restrição, segundo Hyvärinen e Oja (2000), é que as componentes independentes tenham distribuições não gaussianas. O motivo deve-se ao fato de que a distribuição de duas variáveis gaussianas gera uma figura de densidade completamente simétrica, conforme pode ser visto na Figura 3.2:



**Figura 3.2 – Distribuição de duas variáveis gaussianas independentes**

Fonte: Hyvärinen e Oja (2000, p. 26)

A distribuição da Figura 3.2, sendo isotrópica<sup>23</sup>, não sofre alteração quando submetida a rotações, tornando-se impossível a estimação da matriz de mistura das duas componentes originais.

---

<sup>23</sup> Isotrópico: diz-se de, ou corpo cujas propriedades físicas são idênticas em todas as direções.

### 3.2.1.3 A matriz de mistura deve ser quadrada

Assumir que a matriz de mistura ( $A$ ) seja quadrada significa impor que o número de componentes independentes seja igual ao número de sinais observados.

Quando o número de sinais supera o número de fontes a serem encontradas, pode-se limitar a busca das fontes para que não sejam encontrados ruídos (*noise*). Este processo é definido como sendo de *Minimização da Informação Mútua*.

Porém, se existirem mais fontes que sinais, algumas não poderão ser separadas da mistura, comprometendo todo o resultado, uma vez que a sua permanência prejudica a identificação de todas as componentes independentes.

### 3.2.1.4 Os sinais devem ser estacionários

Segundo Back e Weigend (1997) este é um pressuposto de praticamente todas as modelagens de séries temporais, não apenas do método ICA.

Um sinal  $x(t)$  é considerado estacionário se o seu valor esperado for constante (i.e., ao se remover a média, o valor esperado assume valor nulo:  $E[x(t)] = 0$ ).

### 3.2.2 Determinação das componentes por maximização da não-gaussianidade

A técnica empregada para determinar as componentes independentes compreende a medição e interpretação da gaussianidade dos sinais, na tentativa de maximizar a não-gaussianidade das variáveis como medida de separação dos sinais em fontes.

Partindo do pressuposto que os sinais observados são compostos por uma mistura de fontes originais:

$$x = A.s \quad (3.11)$$

As componentes independentes podem ser determinadas pelo inverso da matriz de mistura:

$$s = A^{-1}.x \quad (3.12)$$

Conforme descrito em Leite (2005), supondo uma combinação linear qualquer de  $x_i$  dada por  $y = b^t.x$ . Como  $x = A.s$ , pode-se escrever  $y = b^t.x$ , ou seja,  $y = b^t.A.s$ :

$$y = \underbrace{[b_1 b_2 \dots b_n]}_{\mathbf{b}^t} \cdot \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & & a_{nn} \end{bmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}}_{\mathbf{S}} \quad (3.13)$$

O vetor  $b^t$ , portanto, pode ser descrito como uma isca, para capturar a situação de menor gaussianidade de  $y$ . Se  $b^t$  for uma linha da inversa da matriz  $A$ , então  $b^t$  resultará em um vetor-linha com um dos elementos igual a 1 e os demais iguais a zero, e  $y$  será uma das fontes originais.



Como apenas os sinais observados são conhecidos, o vetor  $b^t$  não pode ser determinado em um primeiro momento. Torna-se necessário um estimador que forneça uma aproximação do vetor  $b^t$ .

Pelo teorema do valor central, descrito por Leite (2005, p. 34):

...a soma de variáveis aleatórias independentes tende a uma distribuição gaussiana, supondo que as  $n$  componentes tenham distribuições idênticas. Assim, a soma de duas variáveis aleatórias é mais gaussiana que as variáveis originais. Partindo desse pressuposto,  $y$  será mais gaussiano do que qualquer componente independente  $s_i$  e será o menos gaussiano possível quando, de fato, se igualar a uma das  $s_i$ .

Esse teorema sugere que se deve variar o vetor  $b^t$  e observar a distribuição de  $y$ . Ao se encontrar o vetor  $b^t$  que maximize a não-gaussianidade de  $y$ , ter-se-á encontrado uma componente independente.

Segundo Leite (2005), o espaço de otimização do vetor  $b^t$  tem dois máximos locais,  $s_i$  e  $-s_i$ , fato este que provoca a perda do sinal verdadeiro da componente.

### 3.2.2.1 Medidas para a gaussianidade

Conforme visto no subitem anterior, torna-se necessário encontrar uma forma de medir a gaussianidade da distribuição de  $y$ . Devido a sua simplicidade teórica e computacional, a principal medida de gaussianidade utilizada nos algoritmos ICA é a curtose.

A lógica por trás dessa medida se sustenta na hipótese de que variáveis perfeitamente gaussianas apresentam excesso de curtose igual a zero, e que para a grande maioria das variáveis não gaussianas, o excesso de curtose é diferente de zero.

Outra forma de medir a gaussianidade da distribuição de  $y$  é chamada de negentropia (*negentropy*).

A entropia de uma variável aleatória está relacionada à quantidade de informação que essa variável carrega. Quanto mais imprevisível for a variável, maior será a entropia e menor será a quantidade de informação.

Uma variável gaussiana possui mais entropia entre todas as variáveis aleatórias de mesma variância que não sejam gaussianas (i.e., uma distribuição gaussiana é menos estruturada que qualquer outra distribuição).

Para obter uma medida de não-gaussianidade utiliza-se uma versão normalizada de entropia diferencial, chamada de negentropia. Estatisticamente, a negentropia é a medida ótima de não-gaussianidade.

A utilização da negentropia, por ser uma medida de não-gaussianidade mais robusta, exige maiores recursos computacionais e, por este motivo, poucos algoritmos o utilizam.

Uma alternativa intermediária é a aproximação da negentropia, usando pressupostos básicos de primeira ordem: média zero e variância igual a um (i.e., variáveis padronizadas). Com essa simplificação nos cálculos a necessidade computacional é reduzida e o efeito é uma aproximação da real negentropia, atingindo-se uma medida de não-gaussianidade mais robusta que a análise de curtose.<sup>24</sup>

### 3.3 O ALGORITMO FASTICA

O algoritmo *FastICA* foi desenvolvido para separar as fontes originais de uma mistura utilizando o procedimento de aproximação da negentropia que tem como idéia principal o ajuste da sistemática de mistura de modo que

as densidades de probabilidades da cada uma das estimativas sejam as mais distantes possíveis de uma variável gaussiana, procedimento este usualmente chamado de maximização da não-gaussianidade.

Hyvärinen e Oja (2000) desenvolveram o algoritmo *FastICA* para ser utilizado nos programas estatísticos e matemáticos R® e Matlab® com o objetivo de facilitar e otimizar os cálculos do procedimento de estimação da não-gaussianidade das variáveis.

O procedimento de maximização da não-gaussianidade está associado ao Teorema do Limite Central<sup>25</sup> que, em linhas gerais, estabelece que o resultado da soma de variáveis aleatórias resulta em uma distribuição mais gaussiana que qualquer variável unitária pertencente ao conjunto da mistura.

Como visto anteriormente, o algoritmo *FastICA* utiliza o conceito de negentropia, i.e., uma medida de entropia normalizada em relação a uma variável gaussiana. Este procedimento se sustenta em uma teoria da informação que diz que em um conjunto de variáveis aleatórias de mesma variância, a variável gaussiana é aquela que possui a maior entropia.

A negentropia de uma variável aleatória  $x$  é dada por:

$$J(x) = H(x_g) - H(x) \quad (3.14)$$

onde  $x_g$  corresponde a uma variável aleatória gaussiana com a mesma variância que  $x$ . Logo, a negentropia será nula se  $x$  for uma variável gaussiana.

O problema recai na estimação de entropia das variáveis. Hyvärinen e Oja (2000) utilizam uma aproximação baseada nos chamados momentos polinomiais, o que gera uma estimativa da negentropia:

$$J(x) = \alpha (E\{G(y)\} - E\{G(v)\})^2 \quad (3.15)$$

---

<sup>24</sup> O algoritmo utilizado neste experimento utiliza o método de aproximação da negentropia para medida de gaussianidade das variáveis observadas.

<sup>25</sup> Exemplo: dados não viciados têm distribuição de probabilidade retangular (todos com probabilidade igual a 1/6). Porém, a distribuição de probabilidade da soma dos valores obtidos em várias jogadas tende a forma gaussiana.

onde  $G$  é uma função não-linear, não-quadrática,  $\alpha$  é uma constante e  $v$  é uma variável aleatória gaussiana de média zero e variância unitária. Como conseqüência da utilização dessa variável gaussiana normalizada, torna-se necessário restringir a potência de cada uma das estimativas das fontes, assumindo, por exemplo, que  $E\{y_i\} = E\{w_i^T x\} = 1$ .

Para descrever como é feita a maximização da não-gaussianidade pelo algoritmo *FastICA*, são descritos a seguir os passos para recuperar uma fonte, i.e., o ajuste de uma das linhas da matriz  $W$ , descrita como sendo  $w_i^T$ , de modo que  $y_i = w_i^T x$  resulte em uma estimativa satisfatória de uma fonte original. O problema de otimização pode ser descrito da seguinte forma:

$$\tilde{w}_i \arg \max_{w_i} (E\{G(y_i)\} - E\{G(v)\})^2 \quad (3.16)$$

com a restrição que  $E\{y_i\} = E\{w_i^T x\} = 1$ .

Utilizando o método de Lagrange, é possível verificar que esse problema de otimização é resolvido quando a seguinte condição é satisfeita:

$$E\{xG'(w_i^T x)\} + \beta w_i = 0 \quad (3.17)$$

onde  $\beta$  é uma constante que depende do valor ótimo de  $\tilde{w}_i$ .

Aplicando-se o método de Newton para a resolução de (3.17), uma vez que as misturas já estarão branqueadas<sup>26</sup>, faz-se o giro seqüencial conforme etapas descritas abaixo:

Etapa 1: Definir um vetor  $w$  inicial (e.g. aleatório);

$$\text{Etapa 2: } w_i = E\{xG'(w_i^T x)\} - E\{G'(w_i^T x)\}w_i \quad (3.18)$$

$$\text{Etapa 3: } w_i = w_i / \|w_i\| \quad (3.19)$$

<sup>26</sup> Se duas variáveis aleatórias são descorrelacionadas e têm variâncias iguais a 1, então elas são chamadas de brancas. O branqueamento é efetuado através de uma transformação linear dos dados.

Etapa 4: se não convergir, retorne à etapa 2.

Para extrair  $N$  fontes com a regra de ajustes descrita acima é necessário executá-la para  $N$  vetores  $w_i$ . Logo, para que não seja recuperada sempre a mesma fonte, i.e., convergir sempre para um mesmo resultado ótimo, foi inserido no final do algoritmo uma etapa para garantir a ortogonalidade dos vetores obtidos. Para isso, Hyvärinen e Oja (2000) incluíram o método de ortogonalização de Gram-Schmidt ao final de cada ajuste. Esse procedimento recebeu o nome de *deflation*, que, a cada iteração, retira a contribuição dos vetores já estimados. Dessa forma, podem ser estimadas  $N$  fontes originais, desde que a mistura seja composta por  $N$  sinais observáveis.

### 3.3.1 Propriedades do algoritmo *FastICA*

As diferenças do algoritmo *FastICA* frente aos outros métodos automatizados para Análise de Componentes Independentes, segundo Hyvärinen e Oja (2000, p. 16) são as seguintes:

1. A convergência é cúbica (ou ao menos quadrática), fato que o diferencia e o torna mais rápido que outros algoritmos onde a convergência é linear;
2. O algoritmo apresenta uma plataforma amigável (*userfriendly*), pois não existem passos ou parâmetros para escolha durante o processo;
3. O algoritmo encontra precisamente os componentes independentes de (praticamente) toda distribuição não-gaussiana usando qualquer não-linearidade  $g$ ;
4. O algoritmo pode ser otimizado pela escolha conveniente da não-linearidade  $g$ ;

5. Os componentes independentes podem ser estimados um a um. Isso reduz o tempo de processamento em casos onde apenas um componente deve ser encontrado;
6. O algoritmo *FastICA* tem a maior parte das vantagens de um algoritmo neural: é semelhante, distribuído, computacionalmente simples e requer pouco espaço na memória.

### 3.4 APLICABILIDADE DO MÉTODO ICA

Desde 2000, ano em que o algoritmo *FastICA* se tornou acessível no ambiente Matlab, as pesquisas sobre a aplicabilidade do método ICA não pararam de crescer.

Por ser um método ainda em desenvolvimento e evolução, a maioria dos artigos internacionais referem-se à adaptação de alguma fase ou etapa de algum algoritmo em desenvolvimento.

A Tabela 3.1 apresenta uma síntese dos trabalhos apresentados que tinham como foco principal a utilização ou o desenvolvimento do método de Análise de Componentes Independentes, desde 1994.

**Tabela 3.1 – Trabalhos publicados sobre Análise de Componentes Independentes 1994-2007**

<b>Tipo</b>	<b>%</b>	<b>Quantidade</b>
Método	32,35%	22
Medicina	23,53%	16
Finanças	14,71%	10
Biologia	11,76%	8
Eng. Elétrica	10,29%	7
Outros	2,94%	2
Computação	1,47%	1
Geofísica	1,47%	1
Física	1,47%	1
<b>Total</b>	<b>100,00%</b>	<b>68</b>

Como pode ser visto na Tabela 3.1, aproximadamente 1/3 dos artigos referem-se a: inovações no método, aprimoramento de algoritmos e novos procedimentos de estimação das fontes. A medicina, por ser a ciência mais antiga, vem em segundo lugar, seguida pela área de finanças que a partir do trabalho de Back e Weigend (1997) vem ampliando as áreas de aplicação do método.

Em 2000, já relatavam os autores Hyvärinen e Oja (2000) das aplicações mais usuais:

- Separação de sinais neurológicos obtidos pela técnica chamada de Magnetoencefalografia (MEG);
- Busca de fatores escondidos em séries financeiras<sup>27</sup>;
- Redução de manchas e outros ruídos em imagens naturais;
- Separação de sinais de telefonia móvel (CDMA) para eliminar a interferência entre as comunicações.

Portanto, aplicações do método ICA podem ser encontradas em diferentes áreas do conhecimento, tais como: processamento de áudio e imagem, processamento de sinais biomédicos, telecomunicações e finanças.

### 3.4.1 Uso do método ICA em Finanças

A primeira<sup>28</sup> aplicação do método ICA para analisar a estrutura de retorno de ações foi de Back e Weigend (1997). Os autores utilizaram o algoritmo *JADE (Joint Approximate Diagonalization of Eigenmatrices)* de dois estágios e aplicaram o método em uma base de dados de retornos diários de 28 grandes empresas registradas na bolsa de Tóquio, entre 1986 e 1989. Os au-

<sup>27</sup> Esta aplicabilidade em séries financeiras será detalhada no próximo subitem.

<sup>28</sup> Os próprios autores reconhecem pois colocaram no título.

tores, após reduzirem para apenas quatro componentes independentes (dos 28 possíveis), reconstruíram a série do Banco de Tóquio-Mitsubishi e puderam constatar os momentos de alta frequência, i.e., movimentos bruscos responsáveis pela movimentação das ações. Os autores concluíram que o método é útil para analisar e entender os mecanismos que movem as séries temporais em finanças.

Chan e Cha (2001) testaram a aplicabilidade do método ICA em um modelo fatorial, i.e., os autores buscaram encontrar os fatores mais relevantes para uma amostra de séries temporais. Como a ausência de correlação entre os fatores é uma necessidade no modelo fatorial, os autores testaram a utilização do método ICA em uma amostra de retornos diários de 22 ações que compõem o índice Hang Seng em Hong Kong, entre 1992 e 2000. Os autores encontraram oito componentes independentes que, por definição, são não-correlacionados entre si.

A busca por fatores independentes não foi a única aplicabilidade do método ICA em finanças. Autores mais recentes, descritos na Tabela 3.2 utilizaram a técnica como estratégia de investimento.

Yen (2004), por exemplo, em sua tese de doutorado, testou o desempenho de um portfólio de ações construído por componentes independentes. Segundo o autor, as componentes independentes seguem um processo chamado variância gama (VG), que seria uma evolução do movimento browniano.

Conforme pesquisa efetuada, as principais áreas pesquisadas em finanças com o uso do método de Análise de Componentes Independentes foram:

- Busca por fatores desconhecidos, responsáveis pela movimentação do retorno de ativos financeiros;
- Busca por fatores independentes para modelagem fatorial;
- Análise do desempenho de retorno de uma carteira de ações com base nas componentes independentes encontradas;



- Administração do risco de carteiras de ações.

A Tabela 3.2 apresenta a relação completa dos trabalhos encontrados, na área de finanças que utilizaram o método ICA.

**Tabela 3.2 – Trabalhos publicados na área de Finanças com o uso do método ICA**

<b>Autores</b>	<b>Objeto de Estudo</b>	<b>Tipo de Publicação</b>
BACK, Andrew D. and WEIGEND, Andreas S. (1997)	Administração de Risco de Carteira de Títulos	Paper: International Journal of Neural Systems
MALAROIU, Simona; KIVILUOLO, Kimmo and OJA, Erkki (1999)	previsão de séries temporais	Paper: Helsinki University of Technology
CHAN, Lai-Wan and CHA, Siu-Ming (2001)	Gestão de Carteiras	Paper: International Symposium on Independent Component Analysis and Blind Signal Separation
FU, Tak-chung et al. (2001)	Retorno de Investimentos	Paper: Hong Kong Polytechnic University
CHAN, Lai-Wan (2002)	Previsão de desempenho	Paper: International Joint Conference on Neural Networks
CHIU, Kai-Chun and XU, Lei (2003)	Gestão de Carteiras	Paper: International Workshop on Computational Intelligence in Economics and Finance
YEN, Ju-Yi J. (2004)	Retorno de Investimentos	Tese: Faculty of the Graduate School of the University of Maryland
CHEN, Ying; HÄRDLE, Wolfgang and SPOKOINY, Vladimir (2005)	Análise de Risco	Paper: Economic Risk Review
CHEN, Ying; HÄRDLE, Wolfgang and SPOKOINY, Vladimir (2005)	Administração de Risco de Carteira de Títulos (ações Alemãs)	Discussion Paper: Center for Applied Statistics and Economics
MADAN, Dilip B. and YEN, Ju-Yi J. (2005)	Retorno de Investimentos	Paper: Department of Finance, University of Maryland

Como pôde ser visto na Tabela 3.1 a área de finanças representa menos de 15% dos trabalhos publicados com o uso do método ICA. Uma das

razões seria a falta de visão dos pesquisadores do uso da ferramenta para outras finalidades. A área de finanças apresenta um vasto ambiente para novas aplicações. A Tabela 3.2 evidencia que o mercado de capitais e principalmente a gestão do risco dominam o uso do método.

### **3.4.2 Uso do método ICA em outras áreas do conhecimento**

De um total de 68 trabalhos encontrados, mais da metade se referem a adaptações e evoluções do método de análise e para soluções biomédicas, principalmente na área de exames de imagens naturais ou vídeos.

O artigo sobre o método, publicado há mais tempo, é o de Comon (1994) que propôs o primeiro algoritmo computacional, introduzindo o conceito como sendo uma extensão à usual análise de componentes principais.

Mais recentemente, os artigos sobre o método ICA baseiam-se na análise de desempenho dos algoritmos otimizadores e das redes neurais artificiais construídas com os mesmos. Mutihac e Hulle (2003), por exemplo, testaram o desempenho de seis algoritmos em um experimento. O *FastICA* obteve o melhor desempenho, i.e., foi o algoritmo que obteve a melhor aproximação das fontes originais simuladas.

Na área da biomedicina, os assuntos primordiais referiam-se à adaptação da técnica na análise de imagens de exames para auxiliar no diagnóstico de doenças. Atualmente, as pesquisas buscam encontrar fatores que auxiliem na prevenção de infartos do miocárdio e na detecção de diversos tipos de câncer (os quais não são possíveis o diagnóstico por visualização humana, por exemplo).

Na engenharia elétrica, as pesquisas iniciais tratavam da separação de ruídos em sinais de áudio. Atualmente, além das ondas sonoras, ondas eletromagnéticas e de radiofrequência são os objetos de separação.

Com o avanço da tecnologia, a engenharia elétrica está cada vez mais sendo utilizada no campo da medicina, biologia, geofísica, entre outras. Um exemplo é a dissertação de mestrado de Leite (2005) na área de engenharia elétrica com o objetivo de aplicar a análise de componentes independentes na identificação de regiões lesionadas em mamografias, i.e., um possível início de processo cancerígeno.

Na área da tecnologia digital, principalmente na área de sistemas de computação, os campos de aplicação do método ICA vão desde o reconhecimento de voz, face, vasos sanguíneos, impressão digital e até a interpretação da íris ocular.

A Tabela 3.3 apresenta uma compilação dos principais trabalhos nas diversas áreas do conhecimento e a sua evolução ao longo da última década.

**Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA**

<b>Autores</b>	<b>Área do Conhecimento</b>	<b>Objeto de Estudo</b>
COMON, Pierre (1994)	Método	proposta de algoritmo
HATEREN, J. H. van and SCHAAF, A. (1997)	Biologia	Diagnóstico por ICA
BARTLETT, Marian Stewart and SEJNOWSKI, Terrence J. (1997)	Método	análise de imagens
CICHOCKI, A.; DOUGLAS, S.C. and AMARI, S.(1998)	Eng. Elétrica	Ruídos de Sinais
JUNG, Tzyy-Ping et al. (1998)	Medicina	separação de sinais
McKEOWN, Martin J. et al. (1998)	Medicina	Infarto do miocárdio
HUVÄRINEN, Aapo. (1998)	Método	análise de imagens
GIANNALOPOULOS, Xavier (1998)	Método	desempenho de algoritmos
HATEREN, J. H. van and RUDERMAN, D. L. (1998)	método	tratamento de imagem
JUNG, Tzyy-Ping et al. (1999)	Medicina	separação de sinais
HANSEN, Lars Kai, LARSEN, Jan and KOLENDA (1999)	Método	Multimídia - Fala, sons e imagens
HYVÄRINEN, Aapo. (1999)	Método	apresentação do método ICA
CICHOCKI, Andrzej et al (1999)	Método	separação de imagens
HYVÄRINEN, Aapo; SÄRELÄ, Jaakko and VIGÁRIO, Ricardo (1999)	Método	proposta de novo método

**Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA****(continuação)**

<b>Autores</b>	<b>Área do Conhecimento</b>	<b>Objeto de Estudo</b>
CASEY, Michael A. and WESTNER, Alex. (2000)	Eng. Elétrica	Separação de áudio
MAKEIG, Scott et al. (2000)	Física	separação de sinais
McKEOWN, Martin J. (2000)	Medicina	Diagnóstico por ICA
MORITZ et al (2000)	Medicina	Diagnóstico por ICA
JUNG, Tzyy-Ping et al. (2000)	Medicina	Diagnóstico por ICA
MORITZ, Chad H. et al. (2000)	Medicina	Diagnóstico por ICA
HYVÄRINEN, Aapo and OJA, Erkki (2000)	Método	proposta de algoritmo
LEE, Te-Won; GIROLAMI, M.; BELL, A. J. and SEJNOWSKI, T. J. (2000)	Método	convergência de teorias de BSS
LAPPALAINEN, Harri and HONKELA, Antti (2000)	Método	desenvolvimento de método de análise
ERIKSSON, Jan; KARVANEN, Juha and KOIVUNEN, Visa. (2000)	Método	separação de sinais
BINGHAM, Ella and HYVÄRINEN, Aapo (2000)	Método	proposta de algoritmo
LIEBERMEISTER, Wolfram (2001)	Biologia	Diagnóstico por ICA
CALHOUN, V. D. et al. (2001)	Medicina	neurodiagnóstico por ICA
LEE, Jae Sung et al. (2001)	Medicina	Infarto do miocárdio
BROWN, Glen D.; YAMADA, Satoshi and SEJNOWSKI (2001)	Medicina	neurodiagnóstico por ICA
CHEUNG, Yiu-ming and XU, Lei. (2001)	Método	classificação de IC's
CARDOSO, Jean-François and COMON, Pierre (2001)	Método	modelagem
CARDOSO, Jean-François (2001)	Método	modelagem
BACK, Kyungim et al. (2002)	Computação	Identificação de Face
YU, Dan; SATTAR, Farook and MA, Kai-Kuang (2002)	Eng. Elétrica	separação de sinais
HARITPOULOS, Michel; YIN, Hujun and ALLINSON, Nigel M. (2002)	Eng. Elétrica	Ruídos de Som
HAVRAN, C. et al. (2002)	Eng. Elétrica	Identificação de Face

**Tabela 3.3 – Trabalhos publicados em outras áreas com o uso do método ICA****(continuação)**

<b>Autores</b>	<b>Área do Conhecimento</b>	<b>Objeto de Estudo</b>
AIRES, Filipe; ROSSOW, William B. and CHÉDIN, Alain (2002)	Geofísica	Análise de série cronológica de Geofísica
CARROLL, Timothy J. et al. (2002)	Medicina	Diagnóstico por ICA
QUIGLEY, Michelle A. et al. (2002)	Medicina	Diagnóstico por ICA
MARTOGLIO, Ann-Marie et al (2002)	Medicina	diagnóstico de câncer de ovário
FIORI, Simone (2002)	Método	modelagem de redes neurais
STONE, James V. (2002)	Método	aplicabilidades do método ICA
HYVÄRINEN, Aapo and OJA, Erkki (2002)	Método	proposta de novo algoritmo
LEE, Su-In and BATZOGLOU, Serafim. (2003)	Biologia	Diagnóstico por ICA
LIAO et al (2003)	Biologia	Diagnóstico por ICA
McKEOWN, Martin J.; HANSEN, Lars Kai and SEJNOWSKI, Terrence J. (2003)	Biologia	Diagnóstico por ICA
ANEMÜLLER, Jörn; SEJNOWSKI, Terrence J. and MAKEIG, Scott. (2003)	Biologia	Diagnóstico por ICA
DRAGHICI, Sorin et al (2003)	Medicina	diagnóstico de HIV
KANO, Manabu; HASEBE, Shinji and HASHIMOTO, Iori. (2003)	Método	aplicação em controle estatístico de processos
JUNG, Tzzy-Ping et al. (2003)	Método	separação de sinais
MUTIHAC, Radu and HULLE, Marc M. Van (2003)	Método	comparativo de algoritmos de redes neurais
ZAMBONI, Nicola and SAUER, Uwe (2004)	Biologia	Diagnóstico por ICA
SCHOLZ, M. et al. (2004)	Biologia	Diagnóstico por ICA
DEBENER, Stefan et al. (2004)	Medicina	diagnóstico neurológico por ICA
OJA, E. (2005)	Método	Apliação de algoritmos
LEONG, W. Y. and HOMER, J. (2006)	Eng. Elétrica	Eliminar ruído
CHINDARO, S.; SIRLANTZIS, K. and FAIRHURST, M. C. (2006)	Eng. Elétrica	Classificação de cores
HUANG, De-Shuang and ZHENG, Chun-Hou. (2006)	Medicina	diagnóstico de câncer

Verifica-se na Tabela 3.3 que a predominância dos estudos foi a aplicabilidade do método em vários campos da medicina, com ênfase na utilização do método para leitura e/ou interpretação de exames, contribuindo para o correto diagnóstico da doença. Como existem inúmeras doenças, muitas inclusive ainda sem cura, o método de Análise de Componentes Independentes

encontra um ambiente propício para evolução. Por este motivo, muitos trabalhos de melhorias de algoritmos têm como objetivo ampliar a aplicabilidade do método nos campos das ciências biológicas (medicina e biologia).

Com o objetivo de testar a aplicabilidade do método de Análise de Componentes Independentes e, por conseqüência, ampliar a participação da área de finanças nos trabalhos com o método, no próximo capítulo é efetuado um experimento por simulação e os resultados são descritos no capítulo seguinte.

## 4 MÉTODO

Neste capítulo, é apresentado o método que foi utilizado para realizar o experimento, as suas etapas e o seu desenvolvimento.

O método utilizado no estudo é o de teste por simulação, o que caracteriza o estudo como sendo de caráter experimental.

A escolha por este procedimento deve-se ao fato de que não existem estudos sobre a aplicabilidade do método proposto em estudo de eventos em finanças, tornando os resultados de uma possível aplicação prática sem poder de comprovação ou generalização de sua eficácia.

### 4.1 DESCRIÇÃO DO EXPERIMENTO

O objetivo deste estudo é testar a aplicabilidade do método de Análise de Componentes Independentes na análise de estudo de eventos em finanças (*event study*, comparando, na medida do possível, com os resultados dos métodos tradicionais de estudo de eventos, descritos no capítulo 2.

Portanto, as duas primeiras etapas do experimento assemelham-se ao método aplicado por Kloeckner (1995) e Soares, et al. (2002).

O experimento foi dividido em três etapas principais:

Etapa 1: criação de um banco de dados: coleta e preparação;

Etapa 2: análise do desempenho dos modelos tradicionais de detecção de retornos anormais, por simulação;

Etapa 3: análise do desempenho do método ICA na detecção de retornos anormais, por simulação.

#### **4.1.1 Etapas do experimento**

A primeira etapa do experimento consistiu na construção do banco de dados que seria utilizado para comparar os resultados dos diferentes modelos de detecção de anormalidades. A segunda etapa consistiu em medir o desempenho dos três modelos geradores de retornos anormais mais utilizados em estudos de eventos. A terceira e última etapa consistiu em estudar o desempenho do método ICA em detectar retornos anormais.

A seguir, são apresentados os passos de cada etapa do processo.

##### **4.1.1.1 Primeira etapa: criação do banco de dados**

Da mesma forma como foi realizado no trabalho de Soares, et al. (2002), a seleção da amostra envolveu os passos descritos a seguir.

Inicialmente, foram selecionadas no banco de dados Economática® as 100 ações de empresas negociadas na bolsa de valores de São Paulo (Bovespa) de maior liquidez para o ano de 2006.

Este procedimento, em busca de quantidade significativa de ações influentes na Bovespa, teve como objetivo tornar os resultados representativos do mercado analisado.

Segundo Van Horne (1997), a expressão “liquidez” pode ser descrita como “a facilidade com que a ação é negociada e está relacionada com o tempo em que o ativo leva para ser negociado.”



Foi utilizado o indicador de liquidez, calculado pela Economática®, segundo a fórmula abaixo:

$$LQ_{it} = \sqrt{\frac{v_{it}}{V_t} \cdot \frac{n_{it}}{N_t} \cdot \frac{S_{it}}{S_t}} \cdot 100 \quad (4.1)$$

onde  $LQ_{it}$  é o índice de liquidez da ação  $i$  no período  $t$ ,  $S$  é número total de pregões da amostra;  $N$  é o número de negócios registrados nestes pregões.  $V$  é o volume financeiro gerado nos pregões. Os numeradores das frações referem-se aos indicadores específicos da ação  $i$ .

O período para determinação do índice de liquidez compreendeu aquele a partir do dia 2 de janeiro de 2006 até o dia 28 de dezembro do mesmo ano. Inicialmente, foram classificadas as ações em função do índice de liquidez, do maior ao menor, e foram selecionadas as primeiras 100 (cem) ações de maior liquidez. Na seqüência, foram eliminadas da amostra as ações que apresentaram uma ou mais ausências de negociação (em um ou mais pregões) no período. Por fim, a amostra ficou composta por 83 ações. No trabalho de Soares, et al. (2002), esse procedimento resultou em uma amostra com 69 ações, fato que comprova que, entre 2002 e 2006, houve um aumento de liquidez geral na bolsa, i.e. o volume financeiro de negociações foi crescente e, por conseqüência, mais negócios foram realizados com as empresas de maior liquidez.

Quanto aos diferentes tipos de ações (ON, PN, etc.), foram selecionadas as ações de maior liquidez para cada empresa, apresentando apenas um tipo de ação por empresa, retirando, portanto, uma possível correlação entre as ações da amostra. A Tabela 4.1 apresenta a descrição das 83 ações de maior liquidez em 2006, que compõem o experimento. Para o cálculo do retorno diário das ações utilizou-se a capitalização contínua. A fórmula utilizada no cálculo dos retornos foi a seguinte:

$$R = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (4.2)$$

Tabela 4.1 – Relação de ações de maior liquidez na Bovespa em 2006

Num	Ação/Empresa	Num	Ação/Empresa
1	Petrobras PN	42	CPFL Energia ON
2	Vale Rio Doce PNA	43	Loj Americanas PN
3	Bradesco PN	44	Lojas Renner ON
4	Usiminas PNA	45	Klabin PN
5	Telemar-Tele NL Par PN	46	Souza Cruz ON
6	Bco Itau Hold Finan PN	47	Celesc PNB
7	Gerdau PN	48	Tractebel ON
8	Sid Nacional ON	49	Duratex PN
9	Cemig PN	50	Localiza ON
10	Net PN	51	Telemig Celul Part PN
11	Unibanco UnN1	52	Transmissao Paulist PN
12	Cosan ON	53	Contax PN
13	Braskem PNA	54	Ipiranga Pet PN
14	Itausa PN	55	Nossa Caixa ON
15	Vivo Part PN	56	Suzano Papel PNA
16	Brasil ON	57	Dasa ON
17	Bradespar PN	58	Energias BR ON
18	Eletronbras ON	59	Telesp PN
19	TAM PN	60	Universo Online PN
20	ALL America Latina UNT N2	61	Comgas PNA
21	Ambev PN	62	Porto Seguro ON
22	Copel PNB	63	Acesita PN
23	Tim Participacoes PN	64	Unipar PNB
24	CCR Rodovias ON	65	Confab PN
25	Embratel Part PN	66	lochp-Maxion PN
26	Natura ON	67	Randon Part PN
27	Brasil Telecom PN	68	OHL Brasil ON
28	Sadia PN	69	Weg PN
29	Gerdau Met PN	70	AES Tiete PN
30	Gol PN	71	Guararapes ON
31	Submarino ON	72	Suzano Petroquim PN
32	Aracruz PNB	73	Grendene ON
33	Cyrela Realty ON	74	Ultrapar PN
34	Embraer ON	75	Paranapanema PN
35	Pao de Acucar PN	76	Marcopolo PN
36	Votorantim C P PN	77	Saraiva Livr PN
37	Sabesp ON	78	Fosfertil PN
38	Telemar Norte Leste PNA	79	Eternit ON
39	Brasil T Par PN	80	Copesul ON
40	Rossi Resid ON	81	Ideiasnet ON
41	Light ON	82	Coteminas PN
		83	Forjas Taurus PN

A amostra, finalmente, ficou composta por uma matriz com dimensão 83 x 245 (83 ações e 245 retornos diários). Dessa amostra, obteve-se a média e o desvio-padrão de cada ação.

A etapa seguinte consistiu em simular os resultados obtidos com a aplicação dos modelos tradicionais de detecção de retornos anormais. Este procedimento é descrito a seguir.

#### **4.1.1.2 Segunda etapa: medindo o desempenho dos modelos tradicionais**

Para esta etapa, foram utilizados os pacotes matemático-estatísticos R® (*freeware version*) e a plataforma do Matlab®, versão 7.4.0.287 (R2007a) para rodar as simulações dos modelos testados.

O procedimento foi semelhante ao utilizado por Soares, et al. (2002), porém com uma capacidade de simulação maior, uma vez que no trabalho dos autores foi utilizado o software Excel® e foram efetuadas apenas 28 simulações para cada modelo testado. Utilizando o R® e o Matlab® foi possível efetuar mais de 1000 simulações para cada modelo testado.

Os modelos testados foram: (1) o modelo de retorno ajustado ao mercado; (2) o modelo de retorno ajustado ao risco e ao mercado e (3) o modelo de retorno ajustado ao modelo CAPM.

As séries simuladas foram obtidas com a geração de séries aleatórias para cada ação, mantendo-se os parâmetros das séries originais (média e desvio-padrão). O objetivo deste procedimento foi o de eliminar a influência de qualquer evento que pudesse influenciar o retorno das ações durante o período analisado, evitando com isso a contaminação do estudo.

Para cada simulação, uma nova amostra de 83 séries aleatórias era gerada.

O índice representativo de mercado foi obtido com a média dos retornos diários das 83 ações que compõem a amostra inicial. Dessa forma, a

primeira observação do índice representativo de mercado corresponde à média aritmética da primeira observação das 83 ações, e assim por diante.

Para o modelo de retorno ajustado ao modelo CAPM, a taxa livre de risco adotada foi a variação média diária do CDI no ano de 2006.

A etapa seguinte do experimento consistiu em introduzir, artificialmente, diferentes níveis de anormalidade na última observação, de uma forma diferente daquela que foi realizada por Soares, et al. (2002). Os autores, em seu trabalho, adicionaram a anormalidade na última observação (245), mas mantiveram a observação anterior, i.e., criaram uma observação de número 246. Entende-se que este procedimento pode anular a eficiência da simulação, uma vez que o passo da observação 245 para a 246 seja apenas a anormalidade introduzida, sem que qualquer movimentação natural tenha ocorrido.

Portanto, no presente experimento, foi mantida a dimensão original da matriz de dados (83 x 245) e a última observação foi “contaminada” com a anormalidade, i.e., o retorno de número 245 foi reescrito com o acréscimo da anormalidade. A fórmula usada foi a seguinte:

$$Rca = [(1 + Rsa).(1 + An)] - 1 \quad (4.3)$$

onde,  $Rsa$  é o retorno de número 245 (última observação de cada série) gerado aleatoriamente,  $An$  é a anormalidade introduzida e  $Rca$  é o retorno 245 com o efeito da anormalidade.

Para cada simulação de séries aleatórias, 201 anormalidades foram testadas, i.e., foram testadas simulações contendo de 0,00% à 2,00%, variando dois pontos percentuais para cada simulação.

Logo, foram efetuadas 1000 simulações para os três modelos e, em cada simulação, foram testadas a detecção das 201 anormalidades, i.e., para esta etapa do experimento, mais de 201 mil testes estatísticos foram efetuados.

Da mesma forma como foi utilizado por Kloeckner (1995, p. 268), a técnica estatística empregada para avaliar a capacidade dos modelos foi o teste  $t$  de Student.

Com base no teorema, descrito em Kloeckner (1995, p. 269) que diz que “a soma de variáveis randômicas independentes e padronizadas distribui-se de acordo com uma normal com média  $m$ , igual à soma dos valores esperados das variáveis, e variância  $n_s$ , igual ao número de variáveis na amostra”, foram efetuados os seguintes procedimentos:

1- após a determinação dos retornos anormais, gerados em cada modelo testado, determinou-se o desvio-padrão das ações que compõem cada série aleatória com o objetivo de padronizar os retornos anormais, usando a equação abaixo:

$$ZAR_i = \frac{AR_i}{s_i} \quad (4.4)$$

onde  $ZAR_i$  é o retorno anormal padronizado para a hipótese nula igual a zero,  $s_i$  é o desvio-padrão da ação  $i$  e  $AR_i$  é o retorno anormal da ação  $i$ , na simulação testada.

Utilizando o teorema descrito anteriormente, o valor da estatística  $t$  é obtido da seguinte forma:

$$t = \frac{\sum_{i=1}^n ZAR_i}{\sqrt{n}} \quad (4.5)$$

Em anexo é apresentado o algoritmo elaborado na plataforma Matlab® e os resultados obtidos, tanto no pacote R® quanto no Matlab® são descritos no próximo capítulo.

#### 4.1.1.3 Terceira etapa: construindo o experimento com o método ICA

Nesta etapa, o principal objetivo é testar a aplicabilidade do método de Análise de Componentes Independentes em substituição ou complemento dos modelos geradores de retornos anormais usualmente utilizados. Portanto, para fins de comparação dos resultados, foi utilizada a mesma base de dados da etapa anterior.

O procedimento metodológico desta etapa se baseia no estudo realizado por Baragona e Battaglia (2007), onde os autores testaram a eficiência do método ICA em detectar *outliers*<sup>29</sup> em séries temporais.

Os autores utilizaram o método de simulação para analisar a capacidade do método ICA em detectar *outliers* e testaram o desempenho de três algoritmos: FastICA, JADE e Infomax.

Em seu artigo, Baragona e Battaglia (2007) simularam a contaminação de processos auto-regressivos de primeira ordem, de acordo com o modelo de mistura descrito a seguir.

Considere  $y_t = (y_{1t}, \dots, y_{ht})'$ ,  $t = 1, \dots, n$  como sendo um conjunto de variáveis observadas e  $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{ht})'$  como sendo uma seqüência de vetores gaussianos independentes identicamente distribuídos com média  $(0, 0, \dots, 0)'$  e matriz de variância – covariância  $\Sigma_\varepsilon = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_h^2\}$ . Suponha que as séries temporais observadas são obtidas por uma transformação linear de  $k$  variáveis não observáveis  $x_t = (x_{1t}, \dots, x_{kt})'$  com  $k \leq h$  ( $h =$  número de fatores), e um vetor *outlier* de magnitude  $\omega$  e padrão  $\Delta_t$  ( $\Delta_t = 1$  quando o elemento for um *outlier* e  $\Delta_t = 0$  quando não for), acrescido de um ruído serial representado por  $\{\varepsilon_t\}$ , resultando na seguinte expressão:

<sup>29</sup> *Outliers* podem ser descritos como observações discordantes em séries temporais que podem enviesar as estimativas dos mínimos quadrados dos parâmetros ARMA, por exemplo.

$$y_t = Ax_t + \omega\Delta_t + \varepsilon_t \quad (4.6)$$

Na simulação realizada por Baragona e Battaglia (2007), o modelo (4.6) foi considerado com  $h = 5$  séries temporais observadas e  $k = 4$  fatores ocultos (ou componentes independentes). Inicialmente, os autores geraram séries temporais contendo fatores não-observáveis  $x_{it}, i = 1, \dots, k; t = 1, 2, \dots, n$ ; correspondente a modelos auto-regressivos de primeira ordem AR(1) que, de acordo com Pindyck e Rubinfeld (2004), podem ser escritos da seguinte forma:

$$x_{it} = \phi_i x_{i,t-1} + a_{it} \quad (4.7)$$

onde  $\{a_{it}\}$  são ruídos brancos (de variância unitária),  $i = 1, \dots, k$ .

Os autores geraram 1.000 (mil) séries temporais com dimensão  $n=100$ . Os parâmetros dos modelos auto-regressivos foram todos iguais a 0,7 (valor absoluto) com sinais alternados, sendo  $(\phi_1, \phi_2, \phi_3, \phi_4) = (0,7; -0,7; 0,7; -0,7)$ .

O ruído  $\{\varepsilon_t\}$  foi assumido como tendo média igual a zero e diagonal da matriz de variância – covariância com entradas na diagonal principal todas iguais a 0,04. Na seqüência, foi considerado de perturbação de magnitude

$\omega = (3,3,3,3,3)$  e uma matriz de mistura  $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0,5 & 1 & 0 & 0 \\ 0 & 0,5 & 1 & 0 \\ 0 & 0 & 0,5 & 1 \\ 0 & 0 & 0 & 0,5 \end{pmatrix}$ , e procederam-se

as mil simulações para as séries temporais observadas  $y_t$ , conforme descrito na equação (4.6).

Os autores explicam que a magnitude do *outlier* (3) é praticamente a mesma magnitude da variância, tanto das séries observadas, quanto dos fatores a serem encontrados.

A idéia por trás do experimento é a de identificar o sinal *outlier* como sendo uma fonte desconhecida, que estaria contaminando os quatro modelos auto-regressivos de primeira ordem.

Três diferentes padrões de *outliers* foram introduzidos no experimento:

- 1- Outlier no final da série: foi introduzido o *outlier* na última observação ( $t=100$ ) pois, segundo os autores, métodos usuais desenvolvidos para detectar *outliers* em séries temporais não são capazes de detectar *outliers* na última (ou na primeira) observação;
- 2- Outlier em três períodos seqüenciais (*outlier patch*): foram introduzidos três *outliers*, nas observações  $t= 51, 52$  e  $53$ , sendo um em cada observação;
- 3- Outlier de mudança de nível (*level shift*): o terceiro teste foi introduzir uma mudança de nível na observação  $t=51$  (*i.e.*,  $\Delta_t = 1$  para  $t \geq 51$ , ou zero, caso contrário).

Para implementar o método ICA, os autores utilizaram três algoritmos: FastICA<sup>30</sup>, JADE e Infomax.

Para a detecção dos *outliers*, os autores utilizaram o valor de  $\mu_i \pm 4,47\sigma_i$  como limite para determinar se a componente extraída do método ICA continha o *outlier* introduzido, sendo  $\mu_i$  e  $\sigma_i$ , respectivamente, a média e o desvio-padrão da componente  $i$  extraída.

Este procedimento de detecção utiliza o Teorema da Desigualdade de Chebyshev para identificar a existência de *outliers*. A aplicabilidade, neste caso, é relevante pois a utilização do método não exige o conhecimento prévio da distribuição de probabilidade da variável estudada, no caso a distribuição da componente independente encontrada.

---

<sup>30</sup> O algoritmo FastICA foi testado na plataforma Matlab® e foi adotada a versão padrão do algoritmo.



De acordo com o Teorema da Desigualdade de Chebyshev, conforme descrito em Spiegel (2001, p. 94), sendo  $X$  uma variável aleatória (discreta ou contínua) tendo média  $\mu$  e variância  $\sigma^2$ , as quais são finitas, se  $\varepsilon$  for um número positivo qualquer,

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad (4.8)$$

ou, com  $\varepsilon = k\sigma$ ,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad (4.9)$$

Portanto, ao utilizar o limite de  $\mu_i \pm 4,47\sigma_i$  significa que a probabilidade de  $X$  diferir de sua média por mais de 4,47 desvios-padrão é menor ou igual a 0,05 (5%), e a probabilidade de  $X$  estar dentro de 4,47 desvios-padrão de sua média é maior ou igual a 0,95 (95%).

Os resultados encontrados por Baragona e Battaglia (2007) indicam que os três algoritmos testados (FastICA, JADE e Infomax) apresentaram desempenhos similares, i.e., a dificuldade de detecção em um algoritmo foi semelhante em outro, não havendo diferenças significativas na escolha de um algoritmo em detrimento de outro.

Os resultados apontam que os três algoritmos conseguiram detectar o *outlier* na última observação em praticamente 100% das simulações, o de quebra de estrutura (*level shift*) foi identificado em aproximadamente 98% nos algoritmos JADE e FastICA sendo que no algoritmo Infomax em apenas 88% das simulações. A maior dificuldade de detecção foi dos três *outliers* em seqüência, que só foram detectados em aproximadamente 40% das simulações. Os resultados das simulações são apresentados na Figura 4.1.

Algorithm	Outlier $t=100$	Patch $t = 51, 52, 53$			Level shift
		$\geq 1$	all	Ave. no.	
JADE	100 (0.0)	80.01 (1.394)	42.25 (1.382)	2.37 (0.607)	98.83 (0.343)
infomax	99.84 (0.111)	73.44 (1.483)	34.25 (1.449)	2.19 (0.690)	88.16 (1.06)
FastICA	100 (0.0)	79.64 (1.464)	42.03 (1.397)	2.37 (0.611)	98.49 (0.419)

Figura 4.1 – Resultados das simulações, percentual médio de *outliers* corretamente identificados em 1000 simulações com 100 repetições cada, variando algoritmo e padrão de *outlier*. Os desvios-padrão dos percentuais estão em parênteses.

Fonte: Baragona e Battaglia (2007, p. 16)

#### 4.1.1.3.1 Testes efetuados na base de dados

Foram efetuados diferentes testes na base de dados, objetivando identificar o comportamento do algoritmo FastICA e o seu desempenho na detecção dos retornos anormais inseridos<sup>31</sup>.

Em um primeiro momento, as simulações para detecção de *outlier* na última observação, efetuadas por Baragona e Battaglia (2007), foram replicadas, porém, com algumas adaptações.

Primeiramente, foram realizadas 1000 simulações onde a estratégia para contaminar as séries originais se baseou na estrutura teórica de Baragona e Battaglia (2007). Para os autores, a inclusão de um *outlier* em uma série temporal pode ser o resultado de uma mistura de séries originais com a série contendo o *outlier*, i.e., valores com zero absoluto e apenas o último elemento com o valor do *outlier*.

<sup>31</sup> No próximo capítulo (Resultados) são detalhadas as dificuldades encontradas em cada teste efetuado e as suas conseqüências no desenvolvimento do estudo.

Portanto, na primeira fase do teste, foram geradas 83 séries do tipo AR (1) e uma série contendo a anormalidade crescente. As 84 séries, após geradas, foram misturadas, utilizando a mesma matriz de mistura dos autores.

Em um segundo momento, contribuindo para uma situação mais realista, a estratégia adotada para inserir a anormalidade nas séries foi semelhante ao experimento com os modelos tradicionais. As últimas observações foram reescritas com a anormalidade inserida, variando de 0,00% à 0,90% em intervalos de 0,10%, porém, manteve-se ainda a estrutura das séries, i.e., modelos auto-regressivos de primeira ordem gerados com variáveis gaussianas aleatórias.

Após essa primeira fase de testes, foi testado o desempenho de detecção do retorno anormal em séries gaussianas aleatórias, com média e desvios-padrão equivalentes às ações da base de dados, i.e., na mesma estrutura de amostra dos testes realizados nos modelos geradores de retornos anormais. Porém, utilizando-se ainda da estrutura teórica de Baragona e Battaglia (2007), que considera as  $n$  séries observáveis como sendo uma mistura de  $n-1$  séries de fontes originais, combinadas linearmente com um vetor de anormalidade.

A diferença desta interpretação teórica de Baragona e Battaglia (2007), para as simulações efetuadas por Kloeckner (1995) e Soares et al. (2002) é que nos dois últimos os autores incluem o evento anormal nas séries originais, mantendo o mesmo número de séries, sendo que Baragona e Battaglia (2007) consideram a anormalidade (e.g., *outlier*) como sendo um vetor a ser misturado em uma combinação linear (e.g., aleatória) com as séries originais, resultando em  $n+1$  séries observáveis.

As simulações realizadas por Baragona e Battaglia (2007) não refletem a realidade, pois, a princípio, é impossível determinar ou afirmar que a anormalidade, representada por um vetor, influencie as séries de uma forma equivalente a uma mistura ou combinação vetorial.

Na realidade, não existirão  $n+1$  séries observáveis, pois a anormalidade não pode ser descrita como uma série original. O número de séries originais, quando da aplicação do método ICA em séries contaminadas, é reduzido, i.e., a representação de uma série original é forçada a ser substituída pela série (vetor) que represente a anormalidade inserida artificialmente.

Este foi o motivo que influenciou este estudo em testar o desempenho do algoritmo FastICA em séries gaussianas. Para que as fontes possam ser encontradas pelo método ICA, conforme visto no capítulo 3, as séries observáveis não podem ser gaussianas perfeitas. Com a inclusão da anormalidade crescente na última observação, as séries tornavam-se cada vez menos gaussianas, favorecendo a detecção e separação do vetor de anormalidade das séries originais, mesmo com a perda de uma série original.

Em resumo, os testes efetuados com o uso do método ICA foram:

- 1- aplicação da estrutura teórica de Baragona e Battaglia (2007): 83 séries do tipo AR (1), misturadas com uma série contendo a anormalidade na última observação e as demais observações nulas;
- 2- aplicação da estrutura teórica de Baragona e Battaglia (2007) sem a geração dos modelos auto-regressivos, apenas com séries aleatórias gaussianas;
- 3- aplicação da estrutura teórica realizada nos modelos tradicionais, conforme descrito em 4.1.1.2.

Os resultados descritos no próximo capítulo esclarecem e apontam algumas características específicas do método ICA.

## 5 RESULTADOS

Os resultados dos testes realizados são apresentados neste capítulo e estão divididos em função das etapas de medição da *performance* dos modelos geradores de retornos anormais e do método ICA.

A discussão sobre os resultados é apresentada no próximo capítulo.

### 5.1 RESULTADOS DOS MODELOS TRADICIONAIS

Em um primeiro teste, foram efetuadas simulações crescentes, de 0% até o nível significativo do teste *t* de Student, em intervalos de 0,01%, utilizando-se o ambiente do software estatístico R®. A Tabela 5.1 mostra a partir de que nível de anormalidade introduzida os modelos foram capazes de detectar os retornos anormais, apresentando os resultados para 5% e 1% de significância.

**Tabela 5.1 – Nível de anormalidade detectada para uma simulação - teste efetuado no software estatístico R®**

Modelos de Retorno Anormal	teste t	
	5% de sig.	1% de sig.
Ajustado ao mercado	0,71%	>1%
Ajustado ao risco e ao mercado	0,67%	0,98%
CAPM	0,66%	0,97%

O modelo de retorno CAPM, dentre os três testados, se mostrou mais sensível, detectando anormalidades a partir de 0,66% e a partir de 0,97%, significantes 5% e 1%, respectivamente.

Os resultados foram semelhantes aos encontrados por Soares et al. (2002), i.e., os três modelos apresentaram resultados similares, apenas com a diferença que no trabalho dos autores o modelo mais sensível foi o ajustado ao risco e ao mercado.

Após essa simulação, foi elaborado um algoritmo no programa Matlab® para capturar as frequências das milhares de simulações possíveis, a fim de que os resultados encontrados sejam mais robustos<sup>32</sup>.

Foram efetuadas, portanto, 1000 simulações sendo acrescentado em cada simulação, na última observação, um retorno anormal variando de 0,00% (para testar o erro tipo I, i.e., rejeição da hipótese nula quando a mesma é verdadeira) até 1,00%, variando a cada 0,01%, o que resultou em 101 mil simulações para cada modelo testado.

Com a aplicação do teste *t*, descrito no item 4.1.1.2, foram acumuladas as frequências no nível de anormalidade que cada modelo iniciava a detecção, i.e., tornava o resultado do teste significativo a 5% e a 1%.

A Tabela 5.2 apresenta os resultados das simulações para as significâncias de 5% e 1%, respectivamente.

A Tabela 5.2 foi comprimida para uma melhor visualização. Os níveis de anormalidade entre 0,05% e 0,30%, e entre 0,70% e 2,00% não foram apresentados porque permaneceram constantes, sem alteração.

Ao analisar a Tabela 5.2 verifica-se que em nenhuma das simulações ocorreu detecção ao nível de 0,00%. Portanto, confirma-se a não incidência dos modelos no erro tipo I, i.e., rejeição da hipótese nula quando a mesma é verdadeira.

Os modelos ajustados ao mercado e ajustados ao risco e ao mercado praticamente produziram resultados iguais. Ambos detectaram 100% das simulações com anormalidade igual ou superior a 0,47% e 0,65%, significantes ao nível de 5% e 1%, respectivamente.

---

<sup>32</sup> O anexo A apresenta os algoritmos utilizados no presente estudo, escritos no ambiente Matlab®.

**Tabela 5.2 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - teste t para os modelos tradicionais de retorno anormal**

Anormalidade	teste t => 5% de sig.			teste t => 1% de sig.		
	MERCADO	RISCO	CAPM	MERCADO	RISCO	CAPM
0,00%	0	0	0	0	0	0
0,01%	0	0	0	0	0	0
0,02%	0	0	0	0	0	0
0,03%	0	0	0	0	0	0
0,04%	0	0	0	0	0	0
0,05%	0	0	0	0	0	0
...	...	...	...	...	...	...
0,30%	0	0	0	0	0	0
0,31%	0	0	0	0	0	0
0,32%	0	0	0	0	0	0
0,33%	0	0	0	0	0	0
0,34%	0	0	0	0	0	0
0,35%	0	10	10	0	0	0
0,36%	0	10	10	0	0	0
0,37%	10	20	30	0	0	0
0,38%	30	30	30	0	0	0
0,39%	30	30	40	0	0	0
0,40%	40	40	40	0	0	0
0,41%	40	40	50	0	0	0
0,42%	60	50	60	0	0	0
0,43%	70	60	70	0	0	0
0,44%	70	70	70	0	0	0
0,45%	70	70	70	0	0	0
<b>0,46%</b>	90	80	<b>100</b>	0	0	0
<b>0,47%</b>	<b>100</b>	<b>100</b>	100	0	0	0
0,48%	100	100	100	0	0	0
0,49%	100	100	100	0	0	0
0,50%	100	100	100	0	0	0
0,51%	100	100	100	0	0	0
0,52%	100	100	100	0	0	0
0,53%	100	100	100	0	10	10
0,54%	100	100	100	0	20	20
0,55%	100	100	100	10	30	30
0,56%	100	100	100	30	30	30
0,57%	100	100	100	40	30	40
0,58%	100	100	100	40	40	40
0,59%	100	100	100	40	40	50
0,60%	100	100	100	60	50	60
0,61%	100	100	100	70	60	70
0,62%	100	100	100	70	70	70
0,63%	100	100	100	70	70	70
<b>0,64%</b>	100	100	100	90	80	<b>100</b>
<b>0,65%</b>	100	100	100	<b>100</b>	<b>100</b>	100
0,66%	100	100	100	100	100	100
0,67%	100	100	100	100	100	100
0,68%	100	100	100	100	100	100
0,69%	100	100	100	100	100	100
0,70%	100	100	100	100	100	100
...	...	...	...	...	...	...
2,00%	100	100	100	100	100	100

De qualquer forma, os resultados obtidos são equivalentes aos encontrados por Kloeckner (1995, p. 270) e Soares et al. (2002): é indiferente, para o pesquisador, optar por um ou outro modelo gerador de retornos de controle uma vez que os resultados apresentados são muito similares.

## 5.2 RESULTADOS DOS TESTES COM O MÉTODO ICA

A estratégia nesta etapa do experimento foi a de gradualmente trazer para a mesma estrutura do teste dos modelos tradicionais, iniciando pelo teste da estrutura teórica de Baragona e Battaglia (2007).

### 5.2.1 Teste do ICA com modelos auto-regressivos

Inicialmente, para testar a capacidade de detecção do método ICA, foi utilizada a estrutura teórica de Baragona e Battaglia (2007), i.e., foram geradas 1000 séries de modelos auto-regressivos de primeira ordem gerados com variáveis gaussianas aleatórias para cada nível de anormalidade testado.

Os níveis de anormalidade testados compreenderam, neste primeiro teste, os valores de 0,00% até 0,90%, com variação de 0,10%, i.e., foram testadas 1000 simulações em 10 níveis de anormalidade cada, totalizando 10mil simulações.

O método utilizado para verificar se o método ICA conseguiu detectar a série com a anormalidade foi o mesmo do experimento de Baragona e Battaglia (2007), o Teorema<sup>33</sup> da Desigualdade de Chebyshev, que foi testado para

---

<sup>33</sup> A partir deste ponto o Teorema da Desigualdade de Chebyshev, por questões de facilidade de leitura, será descrito como o teste de Chebyshev.



4,47 desvios e 10 desvios da média, representando significância estatística de 5% e 1%, respectivamente.

Os resultados obtidos não foram satisfatórios. A Tabela 5.3 apresenta o percentual de detecção das séries com a anormalidade introduzida. O método ICA conseguiu identificar em 100% das simulações as anormalidades a partir de 0,80%, sendo que não houve diferença em função do número de desvios-padrão do teste de Chebyshev.

Verifica-se que, nos testes efetuados, 40% das séries sem anormalidade foram classificadas como tendo anormalidade, o que evidencia o erro tipo I. Ao ampliar o número de desvios-padrão do teste de Chebyshev verifica-se a ausência de ocorrências do erro tipo I, e o mais importante, sem a perda da capacidade de detecção dos retornos anormais.

**Tabela 5.3 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções dos modelos AR(1) para 4.47 e 10 desvios**

Anormalidade	frequência detecção	
	5% sig.	1% sig.
0,00%	40	0
0,10%	20	0
0,20%	20	0
0,30%	10	0
0,40%	70	0
0,50%	40	0
0,60%	20	0
0,70%	80	40
0,80%	100	100
0,90%	100	100

Para resolver o problema do erro tipo I, foram ampliadas as faixas de detecção da anormalidade para níveis cada vez maiores. Após inúmeras tentativas, verificou-se que nos níveis de 15 e 20 desvios-padrão em torno da média, em mais de 10.000 simulações não ocorrem erros do tipo I, sendo esta faixa de detecção a mais segura. Apesar da distância das faixas serem muito superiores ao definido pelo teste de Chebyshev (4,47 e 10 desvios) verifica-se que a capacidade de detecção permaneceu próxima da que era atingida nos padrões originais.

Portanto, as simulações que resultaram na Tabela 5.3 foram repetidas, porém com níveis de detecção de 15 e 20 desvios-padrão em torno da média como sendo o “fator de detecção” da anormalidade.

A Tabela 5.4 apresenta os resultados das simulações com as novas faixas de 15 e 20 desvios. Verifica-se que, desta forma, o erro tipo I foi totalmente eliminado, o que evidencia que esta faixa de classificação torna os resultados mais robustos. Verifica-se que, nesta faixa, o método ICA conseguiu identificar 100% de anormalidades a partir de 0,90% de magnitude, i.e., uma diferença de apenas 0,10% em relação ao atingido pelos parâmetros do teste de Chebyshev. Testes posteriores com faixas superiores a 30 desvios apresentaram resultados semelhantes aos da Tabela 5.4.

**Tabela 5.4 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções dos modelos AR(1) para 15 e 20 desvios**

Anormalidade	frequência detecção	
	15 desvios	20 desvios
0,00%	0	0
0,10%	0	0
0,20%	0	0
0,30%	0	0
0,40%	1	0
0,50%	1	0
0,60%	0	0
0,70%	18	18
0,80%	99	99
0,90%	100	100

### 5.2.2 Teste do ICA em séries aleatórias gaussianas - Fase I

Uma vez encontrados os níveis seguros para detectar as anormalidades partiu-se para a segunda fase do experimento: testar a capacidade de detecção do método ICA em séries aleatórias gaussianas sem a necessidade de serem geradas por modelos auto-regressivos, ou seja, da forma como foram efetuados os testes nos modelos tradicionais.

Em um primeiro momento, desta segunda fase do experimento, manteve-se a estrutura teórica de Baragona e Battaglia (2007), que simula a contaminação das séries, seja por *outlier* ou por retorno anormal, como sendo a mistura de  $n$  séries originais com uma série contendo o evento.

Os resultados das 10 mil simulações para as séries aleatórias gaussianas, de mesma média e desvio-padrão das séries originais, na estrutura teórica de Baragona e Battaglia (2007), são apresentados na Tabela 5.5.

**Tabela 5.5 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das detecções para séries normais – Baragona e Battaglia**

Anormalidade	frequência detecção	
	15 desvios	20 desvios
0,00%	0	0
0,10%	0	0
0,20%	0	0
0,30%	0	0
0,40%	0	0
0,50%	0	0
0,60%	0	0
0,70%	100	100
0,80%	100	100
0,90%	100	100

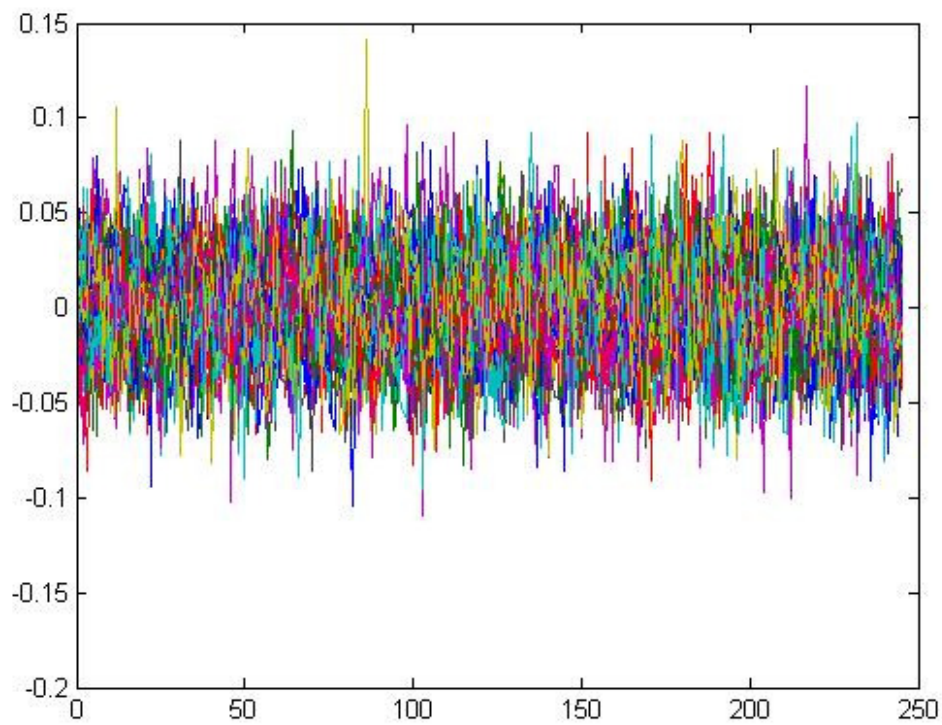
Na Tabela 5.5 verifica-se que o modelo de simulação aleatório sem modelagem auto-regressiva apresenta maior eficiência na detecção dos retornos anormais, tendo sido detectadas 100% das anormalidades a partir de 0,70%.

Para exemplificar como é feita a detecção da anormalidade pelo método ICA, a seguir são apresentadas as plotagens das séries originais, séries contaminadas antes do método ICA, e dos componentes principais encontrados, com ênfase na apresentação da componente que representa o vetor de contaminação.

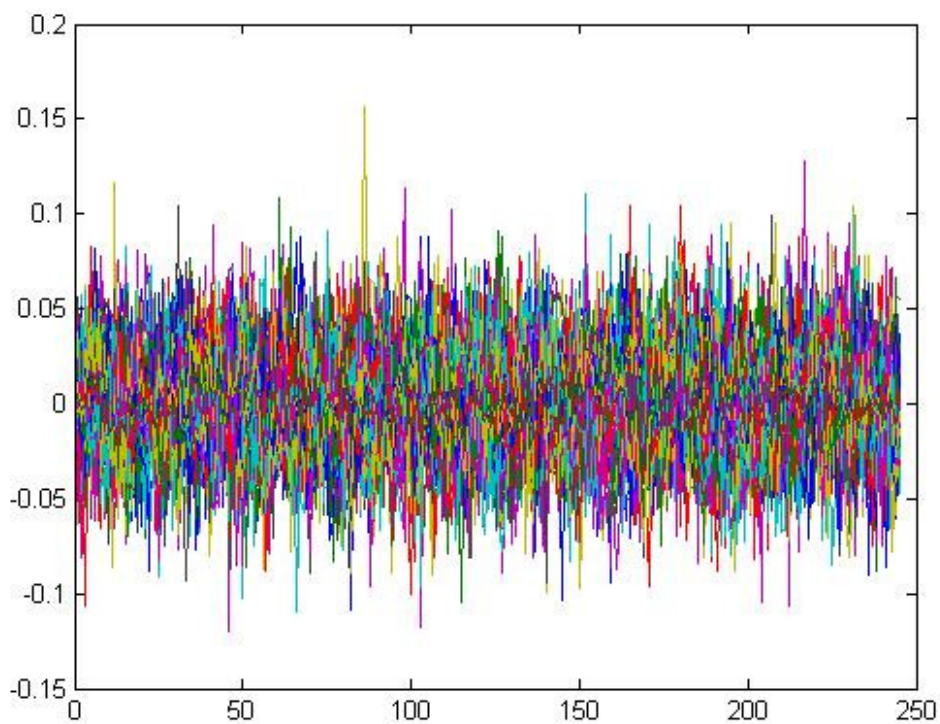
A Figura 5.1 apresenta as 83 séries originais aleatórias, de mesma média e desvio-padrão das séries utilizadas na fase de teste dos modelos tradicionais.

A Figura 5.2 apresenta as 84 séries misturadas, ou seja, as 83 séries originais da Figura 5.1 misturadas com uma série apresentando o retorno anormal na última observação e as demais com valor zero.

Observa-se, pela Figura 5.2, que é praticamente impossível identificar a combinação linear efetuada nas 83 séries com a inclusão do vetor com a normalidade de 0,90%, sendo necessária a utilização do método ICA para separação das séries.

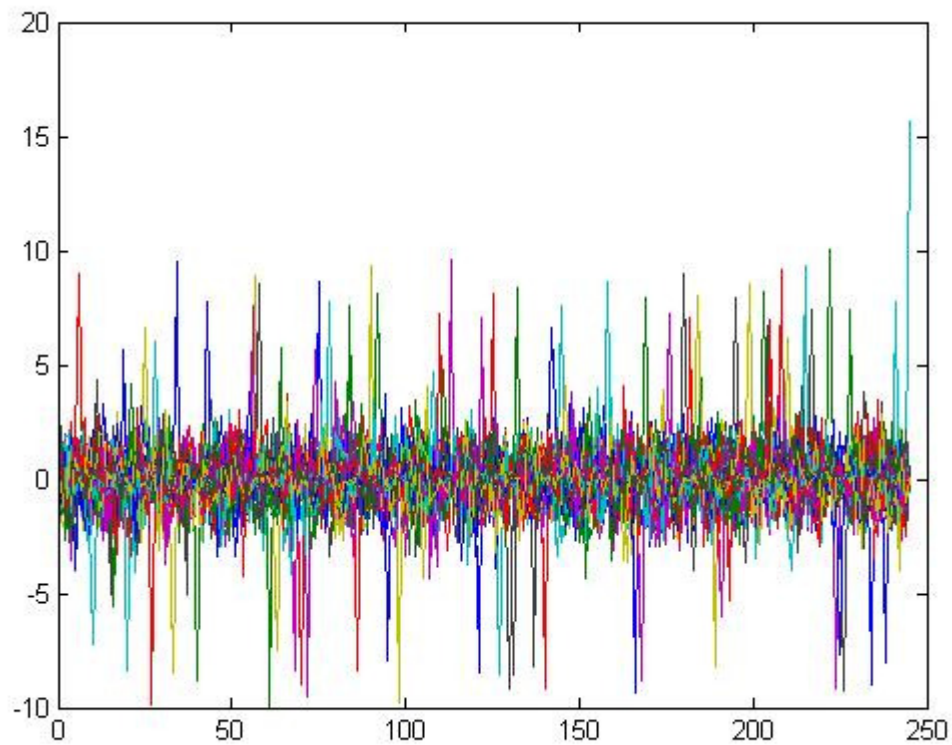


**Figura 5.1 – 83 séries normais aleatórias com média e desvios idênticos aos das séries da amostra original**



**Figura 5.2 – 84 séries misturadas sendo 83 séries normais e uma contendo a anomalia de 0,90%**

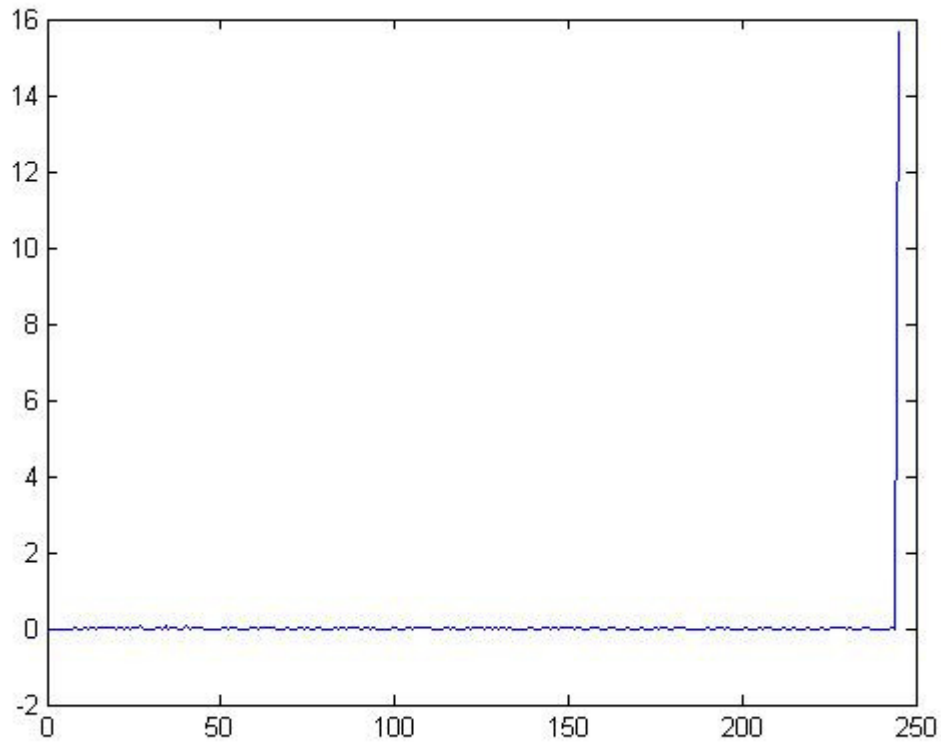
A Figura 5.3 apresenta as componentes independentes encontradas, após a aplicação do método ICA na amostra da Figura 5.2.



**Figura 5.3 – Componentes encontradas em uma das simulações com anormalidade de 0,90%**

Ao analisar a Figura 5.3 fica visível a componente da última observação como sendo a mais expressiva da amostra. Observa-se que as demais componentes se situaram em sua grande maioria no intervalo entre 5 e -5 desvios-padrão em torno da média, com algumas séries apresentando pontos bem próximos de 10 desvios, o que provoca o erro tipo I. Em uma análise visual, apenas a componente mais expressiva está a mais de 15 desvios da média.

Em uma análise mais precisa, a Figura 5.4 apresenta a plotagem componente de número um detectada, ou seja, a componente de menor gaussianidade dentre todas as encontradas pelo algoritmo *FastICA*.



**Figura 5.4 – Primeira componente independente detectada pelo algoritmo *FastICA***

Na Figura 5.4, pode-se verificar a característica da série que apresenta a anormalidade inserida. O método ICA detectou esta série e classificou-a como sendo a de menor gaussianidade entre as componentes encontradas, ou seja, a primeira componente encontrada, na simulação com anormalidade de 0,90%, é a componente que continha apenas a anormalidade, i.e., o vetor com 0,009 na última observação e 0,000 nas demais.

Portanto, os resultados desta segunda fase foram satisfatórios. O método ICA conseguiu detectar todas as anormalidades superiores a 0,70% nas milhares de simulações efetuadas com a modelagem teórica de Baragona e Battaglia (2007).

### 5.2.3 Teste do ICA em séries aleatórias gaussianas - Fase II

Em uma terceira e última fase, foram efetuadas simulações nas séries originais com a inclusão da anormalidade diretamente nas séries. Isto significa que nesta terceira fase está sendo considerada a possibilidade de existir um número maior de componentes independentes, ao invés de limitá-los em número equivalente ao de séries originais, o que contraria os princípios de Análise de Componentes Independentes, mas que torna mais realista a situação em estudo.

Na terceira fase do experimento, ao invés de misturar as séries originais simuladas<sup>34</sup> com um vetor contendo a anormalidade, as séries simuladas foram reescritas, tendo sido incluída na última observação o retorno anormal, da mesma forma como foi realizado nos testes descritos em 4.1.1.2, com o uso da equação 4.3.

Os resultados preliminares indicavam que o método ICA era incapaz de detectar as anormalidades inseridas na última observação. Nenhuma das anormalidades, de 0,10% a 0,90% foram detectadas, utilizando a faixa de teste de 15 e 20 desvios-padrão.

Tentativas de ampliação do retorno anormal evidenciaram que a faixa de 15 e 20 desvios era grande demais, em relação à magnitude das anormalidades, ou seja, as anormalidades não eram classificadas após a detecção.

Em um segundo momento, simulações com aumento de anormalidade e redução da faixa de classificação foram sendo efetuadas, até que se obtiveram os seguintes resultados para acréscimos de 0,50% a cada simulação de 0,00% a 4,5% em uma faixa de 8 e 12 desvios, descritos na Tabela 5.6.

---

<sup>34</sup> As séries simuladas continuam sendo séries aleatórias gaussianas com mesma média e desvio-padrão das séries originais de mercado, mesmo padrão utilizado nos testes dos modelos tradicionais.



**Tabela 5.6 – Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das 1000 simulações pós-inclusão da anormalidade para 8 e 12 desvios**

Anormalidade	frequência detecção	
	8 desvios	12 desvios
0,00%	4	0
0,50%	22	0
1,00%	51	0
1,50%	63	0
2,00%	67	0
2,50%	100	0
3,00%	100	29
3,50%	100	87
4,00%	100	100
4,50%	100	100

A Tabela 5.6 apresenta três descobertas em função desta nova estrutura teórica de apenas reescrever a última observação de cada série: (1) a capacidade do método ICA em detectar anormalidades; (2) a perda de eficiência do método ICA e (3) a mudança necessária de faixa para classificação das séries.

Infelizmente, por este método, com os resultados obtidos, verifica-se uma perda significativa de eficiência na detecção de anormalidades. Verifica-se também, ao visualizar a Tabela 5.6 que 4% das simulações provocaram o erro tipo I quando foi utilizado o parâmetro de 8 desvios-padrão como medida de corte para detectar as anormalidades. Porém, verifica-se que para 12 desvios, apesar da perda de eficiência, o erro tipo I não ocorre.

Para corrigir os erros do tipo I, novas simulações foram realizadas. Os resultados são apresentados na Tabela 5.7.

A Tabela 5.7 revela que os 11 desvios proporcionam a eliminação do erro tipo I e provocam uma perda de eficiência de detecção de 2,50% para 3,50%. Em outras simulações efetuadas, verificou-se que, para cada redução ou ampliação em 1 desvio-padrão na faixa de classificação, a eficiência era ampliada ou reduzida em meio ponto percentual.

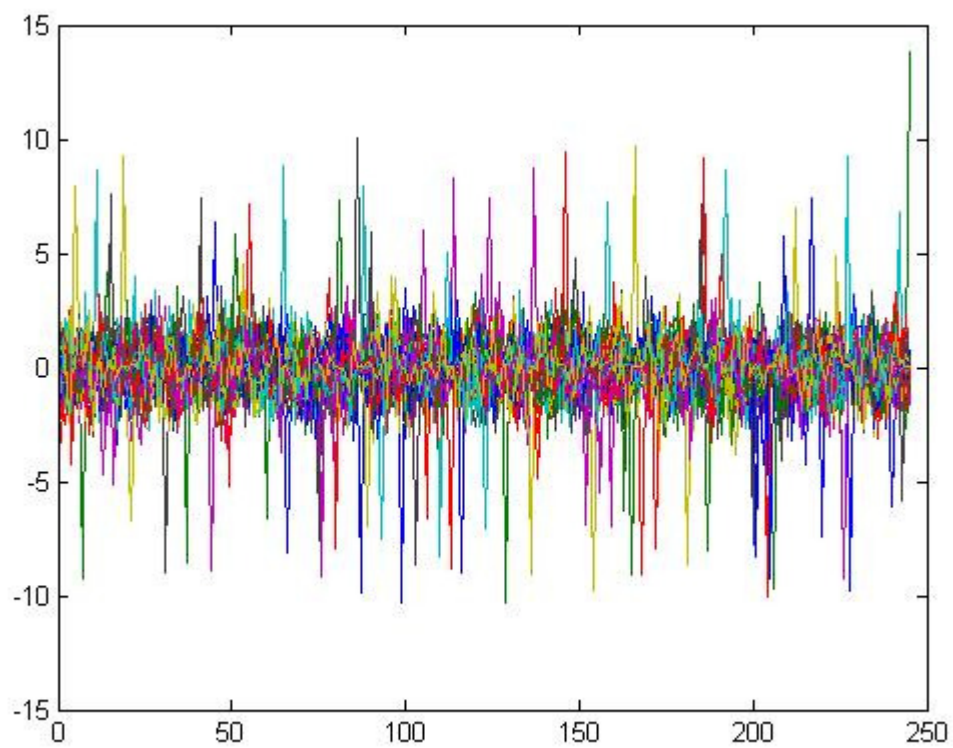
**Tabela 5.7– Frequência de detecção de mil simulações para cada nível de anormalidade introduzida - resultado das 1000 simulações pós-inclusão da anormalidade para 11 e 12 desvios**

Anormalidade	frequência detecção	
	11 desvios	12 desvios
0,00%	0	0
0,50%	0	0
1,00%	0	0
1,50%	0	0
2,00%	10	0
2,50%	44	0
3,00%	92	36
3,50%	<b>100</b>	80
4,00%	100	<b>100</b>
4,50%	100	100

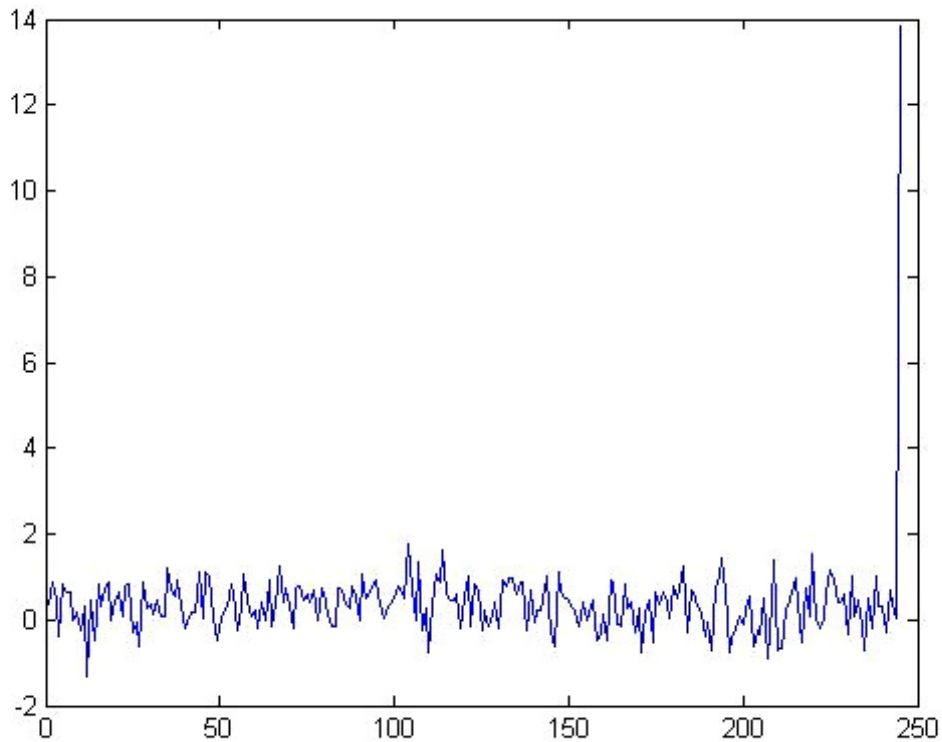
Porém, apenas a partir de 11 desvios é que o método ICA não implicava em rejeição da hipótese nula quando ela era verdadeira.

A Figura 5.5 mostra as componentes independentes encontradas em uma das 1000 simulações desta terceira fase do experimento. Como pode ser visto, as componentes apresentam amplitude próxima a 10 desvios, sendo que a amplitude máxima da componente com a anormalidade de 4,50% é de quase 15 desvios. Por este motivo é que no início não era possível identificar as anormalidades, pois elas nunca se situavam entre 15 e 20 desvios.

A perda de eficiência do método ICA, nesta nova estrutura teórica, é visível quando se efetua a análise individual da componente independente que apresenta a anormalidade inserida. A Figura 5.6 apresenta a componente com a anormalidade de 4,50%. Em uma rápida comparação com a Figura 5.4 é visível a perturbação maior da componente encontrada com esta nova estrutura teórica.



**Figura 5.5– Componentes encontradas em uma das simulações com anormalidade de 4,50%**



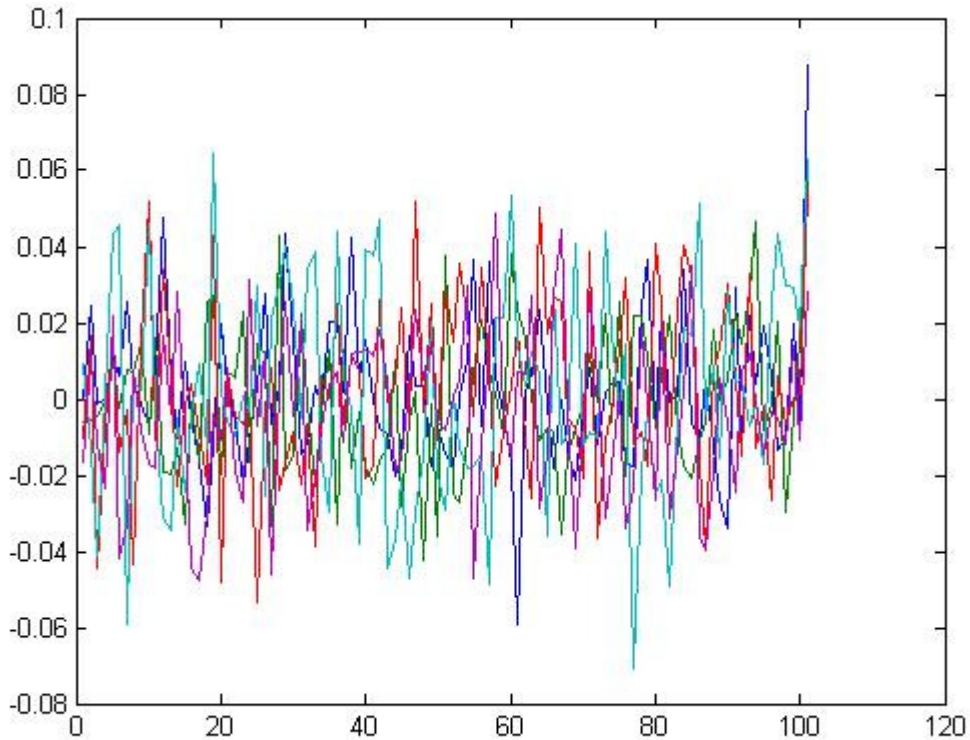
**Figura 5.6– Nona componente independente detectada com a nova estrutura teórica**

Em função da volatilidade presente nas demais observações, as quais não foram explicadas pelas demais componentes e, por conseqüência, permaneceram na componente com a anormalidade, a componente da Figura 5.6 só foi encontrada, por exemplo em uma simulação, na nona posição, ou seja, na simulação foram encontradas séries menos gaussianas que a que continha o conteúdo explicativo da anormalidade inserida.

#### **5.2.4 Teste do ICA para dimensões menores de dados**

Em outro teste, uma amostragem de dimensões semelhantes ao que foi simulado por Baragona e Battaglia (2007) foi testada. Foram definidas as dimensões da amostra em cinco séries com 100 observações e foi testada a eficiência na detecção das anormalidades apenas reescrevendo-se a última observação com a anormalidade.

Os resultados sugerem que existe uma relação direta entre o número de séries e a magnitude da componente independente relacionada com a anormalidade. A Figura 5.7 apresenta as componentes encontradas em uma das simulações com a dimensão definida em 5 x 100.

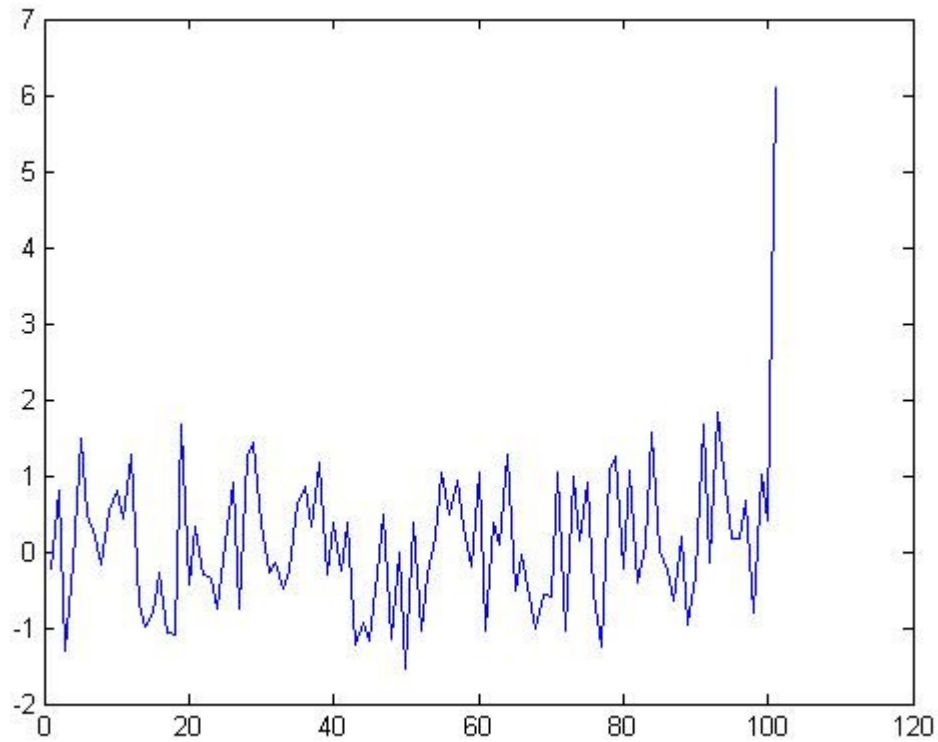


**Figura 5.7– Componentes detectadas na dimensão de amostra 5 x 100**

A presença da componente com a anormalidade na última observação é visível na Figura 5.7, após a detecção das componentes pelo ICA. Porém, verifica-se que a sua amplitude não é muito diferente das demais componentes, fato este que pode prejudicar o processo de detecção pelo teste de Chebyshev.

A Figura 5.8 apresenta a primeira componente detectada na amostragem de dimensão 5 por 100. Ao analisá-la, observa-se que o retorno anormal apresentou magnitude equivalente a pouco mais de 6 desvios-padrão, o que prejudica a sua detecção na faixa entre 15 e 20 desvios e colabora com a hipótese de que a quantidade de variáveis da amostra influencia a amplitude

da componente encontrada. Foram utilizados os mesmos parâmetros de média e desvio-padrão que foram utilizados nos testes anteriores sendo que a magnitude da anormalidade foi menor.



**Figura 5.8– Primeira componente detectada na dimensão de amostra 5 x 100**

Neste último experimento, a perda de eficiência também foi demonstrada, tendo sido detectadas anormalidades apenas de magnitude superior a 4,00%.

#### **5.2.4 Considerações sobre os resultados dos testes com o método ICA**

O experimento abrangeu quatro testes de detecção com o uso do método ICA. O primeiro teste consistiu na replicação da estrutura teórica de Ba-

ragona e Battaglia (2007), com a geração de modelos auto-regressivos de primeira ordem. O segundo teste, objetivando modelar a simulação da mesma forma como foi realizada para os modelos tradicionais de estudo de evento, consistiu em gerar séries aleatórias gaussianas com mesma média e desvio-padrão dos dados originais. Porém, em uma primeira fase, manteve-se a estrutura teórica de Baragona e Battaglia (2007), i.e., foi incluída uma série na base de dados, sendo esta a série contendo o *outlier*. O terceiro teste foi semelhante ao teste anterior, porém, ao invés de incluir uma nova série na base de dados, todas as séries foram reescritas sendo que na última observação foi incluído o *outlier*, da mesma forma como foi feito em Soares et al. (2002). O quarto e último teste foi analisar o desempenho do método ICA em base de dados de dimensões menores, por exemplo, cinco séries com 100 dados de retornos diários. Um resumo dos resultados obtidos em cada um dos testes está descrito na Tabela 5.8.

**Tabela 5.8– Resumo dos testes com o método ICA – nível de anormalidade detectada em 100% das mil simulações e a respectiva faixa de detecção em desvios-padrão**

Resultado	modelos AR (1)	gaussianas com $n+1$ componentes	gaussianas com $n$ componentes	dimensão reduzida com $n$ componentes
% de anormalidade detectada em 100% dos casos	0,90%	0,70%	3,50%	4%
faixa de detecção (em desvios-padrão)	15	15	11	6

Apesar da funcionalidade do método ICA em todas as estruturas teóricas testadas, apenas a estrutura de Baragona e Battaglia (2007) apresentou resultados equivalentes aos testes efetuados nos modelos tradicionais. A perda de eficiência na detecção nas outras hipóteses teóricas é significativamente elevada o que faz com que a necessidade de reflexão sobre o assunto seja alterada: substituir a ênfase nos ajustes empíricos para aplicação do método pela validação da estrutura teórica por trás da problemática é um tema que será debatido no próximo capítulo.

## 6 DISCUSSÕES, CONCLUSÕES E CONSIDERAÇÕES FINAIS

O principal objetivo desta tese foi testar a capacidade do método de Análise de Componentes Independentes - ICA em detectar retornos anormais em séries temporais financeiras e comparar o seu desempenho com os modelos mais utilizados atualmente.

A seguir, são apresentadas as discussões sobre os resultados obtidos com o estudo, as conclusões que podem ser extraídas e as considerações finais.

### 6.1 DISCUSSÕES SOBRE OS RESULTADOS

De acordo com os resultados obtidos pelas simulações, é latente a problemática da questão da inserção da anormalidade nas séries temporais.

Baragona e Battaglia (2007), ao testarem o desempenho do método ICA em detectar *outliers* em séries temporais, utilizaram uma estrutura teórica composta por cinco séries temporais misturadas, de um total de quatro séries originais oriundas de modelos auto-regressivos de primeira ordem e um vetor de perturbação que, dentre outras versões, foi apresentado como um vetor de zeros sendo a última observação com variância 150% superior à variância média das séries originais.<sup>35</sup>

Utilizando a mesma estrutura de geração dos modelos auto-regressivos e da matriz de mistura, o experimento efetuado refletiu a concepção teórica por trás da lógica dos autores.

---

<sup>35</sup> Apesar de o método utilizado por Baragona e Battaglia (2007) não ser o objetivo principal deste estudo, uma das críticas é justamente sobre a amplitude do *outlier* em teste em relação à variância das séries originais, o que pode ter contribuído para o bom desempenho do método ICA.



O método ICA é, conforme visto no capítulo 3, amplamente utilizado para separar fontes, ou seja, identificar as variáveis independentes e separá-las, desde que a mistura tenha número igual ou superior ao número de componentes independentes. Nas simulações efetuadas, onde foram incluídos vetores que correspondiam às anormalidades, os resultados são favoráveis ao método ICA.

Em comparação com o desempenho dos modelos tradicionais, que detectaram anormalidades a partir de 0,65%, significantes a 1%, o método ICA, na concepção teórica de Baragona e Battaglia (2007), apresentou desempenho semelhante.

Inicialmente, com o teste dos modelos auto-regressivos de primeira ordem, os resultados sugeriam uma eficiência na detecção de anormalidades superiores a 0,90%. Na tentativa com séries aleatórias gaussianas os resultados foram semelhantes aos dos modelos tradicionais, sendo que a componente independente com a anormalidade começou a ser identificada a partir de 0,70%, sem a necessidade de modelagem ou de simplificações teóricas.

A perda de eficiência nos testes com os modelos auto-regressivos pode ser explicada pelo fato de que o método ICA busca separar as componentes independentes pelo seu nível de gaussianidade, i.e., os modelos auto-regressivos são, por definição, menos gaussianos que uma série aleatória gaussiana. Por este motivo, o método ICA tem maior dificuldade em detectar as séries com o *outlier*, pois as demais séries podem estar com um nível de gaussianidade próximo do limite de decisão de captura do algoritmo.

A questão teórica por trás desses resultados é que não pode ser interpretada como realista, pelo menos em um primeiro momento. Os resultados de eficiência foram positivos quando misturadas as séries originais com uma série composta pela anormalidade, o que resulta em um número equivalente de componentes independentes<sup>36</sup> e séries observáveis.

Porém, nos testes efetuados com os modelos tradicionais, ao invés de combinar linearmente as séries originais com um vetor semelhante, a anor-

malidade foi introduzida nas séries, reescrevendo a última observação, mantendo o mesmo número de séries originais antes da contaminação.

A diferença de eficiência observada nos resultados dos testes diz respeito ao fato de que, nas primeiras duas fases do experimento, o número de componentes independentes era igual ao número de séries observáveis. Na terceira e última fase do experimento, pelo fato de a anormalidade ter sido incluída nas séries originais, o número de séries observáveis tornou-se inferior ao número de componentes independentes, pois o objetivo era separar das séries originais a anormalidade inserida, na forma de componente independente de menor gaussianidade.

Acredita-se que esta perda de eficiência é equivalente ao que ocorre nas tentativas de limpeza de fotos naturais envelhecidas, as quais apresentam focos de umidade, mofo e sujeira que não representam padrões para serem enquadrados como uma imagem independente sobreposta. Os resultados da separação de duas imagens sobrepostas é bem superior aos obtidos na limpeza digital de fotos envelhecidas.

Foram efetuadas tentativas de detecção, retirando uma série após a mistura com o vetor contendo a anormalidade, objetivando ficar com 83 séries observáveis e 84 possíveis componentes independentes. Os resultados mostraram a incapacidade do método ICA em detectar as anormalidades, por maiores que sejam, se uma das séries for excluída da amostra.

Após a aplicação dos testes, verificou-se que a variância da base de dados é um fator relevante no desempenho do método ICA. Séries com menor variância podem ser parametrizadas para detecção de anormalidades em menores níveis de desvios-padrão. Outro fator relevante é o número de séries da amostra. No último teste, observou-se que a anormalidade foi detectada a apenas seis desvios-padrão da média. Um fator relevante é o fato de que este último teste utilizou a variância das cinco ações mais líquidas da Bovespa, provavelmente as de menor variância em relação às demais.

---

<sup>36</sup> Se o vetor contendo a anormalidade for considerado como sendo uma componente independente.

## 6.2 CONCLUSÕES

Os resultados obtidos nas simulações sugerem que, uma vez interpretada a contaminação das séries de retornos como sendo uma mistura de vetores sendo um deles o retorno anormal, o método ICA pode ser considerado como eficiente na detecção sendo o seu desempenho similar aos modelos geradores de retornos anormais tradicionais, com algumas vantagens:

- 1- Ausência de necessidade de modelagem estatística ou teórica<sup>37</sup>;
- 2- Ausência de necessidade de número mínimo de séries<sup>38</sup> (tamanho da amostra) por não utilizar testes estatísticos paramétricos;
- 3- É identificada a real magnitude do retorno anormal, sem perdas em função de modelagem ou simplificações matemáticas;
- 4- Uma vez detectada a anormalidade, torna-se viável a reconstrução das séries originais sem o efeito do retorno anormal descoberto, identificando o impacto real desta nos retornos normais das ações.

Por outro lado, se o efeito dos retornos anormais for teoricamente interpretado como uma ampliação ou redução nos retornos normais em uma data específica, da forma como foi simulada a terceira fase do experimento, os resultados sugerem que o método ICA é menos eficiente que os modelos geradores de retornos anormais tradicionais.

Em tese, uma mistura, seja de imagens, sons ou retornos diários, por exemplo, tem na sua concepção as raízes originais das variáveis misturadas que, em sua essência, são diferentes. Nessa lógica, o método ICA funciona, pois a sua eficácia é atingida ao separar componentes independentes, i.e.,

---

<sup>37</sup> O modelo não precisa de pressupostos como o CAPM, por exemplo. (beta, retorno de mercado, taxa livre de risco, entre outros.)

<sup>38</sup> Baragona e Battaglia (2007) simularam com apenas cinco séries temporais.

quanto mais diferenças entre as séries forem identificadas melhores serão os resultados de sua aplicação.

Em uma simulação preliminar, foi testada a capacidade de detecção em uma amostra com vinte séries aleatórias gaussianas sem a inclusão de anormalidade, o que resultaria em um erro do tipo I. Os resultados não acusaram a presença de erro tipo I, corroborando com a hipótese de que uma amostra maior pode ser prejudicial para uma possível detecção de anormalidade que não existe, i.e., semelhanças entre séries podem provocar o reconhecimento de um padrão que, na verdade, não passariam de supostas coincidências temporais.

As conclusões que podem ser extraídas pelos resultados das simulações efetuadas são descritas a seguir:

1. A capacidade de detecção de retornos anormais com o uso do método ICA, por um lado, evita o embasamento da solução em pressupostos teóricos e/ou estatísticos, mas, por outro lado, exige uma análise prévia do comportamento das séries antes da sua aplicação.
2. Conforme descrito no capítulo 5, as dimensões da amostra, assim como a sua variância, são fatores que interferem na magnitude de apresentação da componente individual que carregará a informação do retorno anormal.
3. Portanto, antes de utilizar o método ICA, o pesquisador, de posse da amostra, deverá efetuar um teste simulado com séries aleatórias gaussianas de mesma variância dos elementos em estudo para identificar o nível de precisão atingível pelo método e a magnitude atingida pela anormalidade.
4. De posse dessas informações, o pesquisador poderia proceder a análise e, uma vez identificada a anormalidade, esta poderá ser reconstruída no padrão original, descobrindo-se a real magnitude do impacto da anormalidade nos retornos das ações.

5. Uma vez detectada a anormalidade, o pesquisador poderá reconstruí-la ou, se preferir, poderá reconstruir as séries em estudo sem o efeito desta componente, identificando caso a caso o real impacto do retorno anormal no desempenho das ações.

Apesar das dificuldades encontradas e das necessidades de parametrização para conseguir detectar as anormalidades, acredita-se que o método ICA seja uma ferramenta útil.

Contribuindo com a tese de McGovern (2005), de que atualmente todos os testes de eficiência de mercado são, na realidade, testes conjuntos de várias hipóteses, por exemplo de aceitação de modelos de precificação de ativos, o método ICA colabora com a versão dos que acreditam ser o mercado eficiente, os investidores racionais e que as anormalidades detectadas atualmente são geradas por modelos de precificação inadequados. O método ICA, por não precisar de modelos de precificação como *benchmark* para detectar as anormalidades, contribui para a aceitação de que o mercado realmente possa ser ineficiente e, por conseqüência, existam investidores irracionais, caso ocorra a detecção de retornos anormais.

Portanto, se o mercado financeiro espelhar a estrutura teórica de Baragona e Battaglia (2007), com uma modelagem prévia, o método ICA é capaz de identificar a amplitude real do retorno anormal, contribuindo para o desenvolvimento da área de finanças no que diz respeito à interpretação dos fenômenos que interferem no desempenho dos títulos, fornecendo, inclusive, a possibilidade de classificação dos eventos em função do nível de anormalidade sofrida.

### 6.3 LIMITAÇÕES DA TESE

A principal limitação deste estudo diz respeito ao fato de que foram efetuadas simulações, sem a possibilidade de testar o desempenho do método

ICA em uma amostra que contivesse um evento real. A inclusão do evento na amostra pode não ser fidedigna com a realidade.

Outra limitação refere-se ao fato de que foram testadas séries temporais simuladas com mesma média e desvio-padrão de retornos reais do ano de 2006, i.e., resultados diferentes podem ser encontrados se ocorrerem alterações significativas na estrutura da amostra, principalmente no que se refere à variância. Amostras com menor variância poderão resultar em detecções em níveis menores de anormalidade, uma vez que foram testadas anormalidades com dimensões absolutas.

#### 6.4 SUGESTÕES PARA PESQUISAS FUTURAS

O aspecto metodológico da mensuração do retorno anormal em um estudo de eventos é um dos menos discutidos em estudos de finanças, porém, é a metodologia mais utilizada. Em função disso, em que pese a necessidade de entender ou explicar fenômenos responsáveis pela movimentação no retorno das ações, a busca por uma metodologia cada vez mais isenta de subjetivismos ou de necessidade de embasamento teórico, seja estatístico ou econômico-financeiro, deve ser o principal objeto de pesquisa.

Para futuras pesquisas, recomenda-se a aplicação do método ICA em um caso real de retorno anormal e as diferenças que podem ser encontradas quando da aplicação dos métodos tradicionais, na busca pela real amplitude da anormalidade.

Algoritmos novos estão sendo elaborados, alguns com a capacidade de detectar um número maior de componentes independentes em relação ao número de misturas apresentadas, o que pode contribuir para melhorar o desempenho do método ICA em estudo de eventos.

Acredita-se que uma das lacunas aqui apresentadas, a definição teórica do impacto do retorno anormal nas ações, possa ser o objeto principal em um estudo mais aprofundado. Entender de que forma o retorno anormal se incorpora aos retornos normais das ações, talvez seja a resposta para que o método ICA se torne mais eficaz, contribuindo para a detecção e mensuração de todo e qualquer efeito positivo ou negativo nos retornos das ações.

## REFERÊNCIAS BIBLIOGRÁFICAS

AIRES, Filipe; ROSSOW, William B. and CHÉDIN, Alain. Rotation of EOFs by the independent component analysis: toward a solution of the mixing problem in the decomposition of geophysical time series. **Journal of the atmospheric sciences**, USA, 1, January. 2002. v. 59, p. 111-123.

ANEMÜLLER, Jörn; SEJNOWSKI, Terrence J. and MAKEIG, Scott. Complex independent component analysis of frequency-domain electroencephalographic data. **Neural networks**, n. 16, p. 1311-1323, 2003.

BACK, Andrew D. and WEIGEND, Andreas S. A First Application of Independent Component Analysis to Extracting Structure from Stock Returns, **International Journal of Neural Systems**, v. 8, n.5, Oct. 1997. Special issue on Data Mining in Finance.

BACK, Kyungim et al. **PCA vs. ICA: a comparison on the FERET data set**. Colorado State University – Department of Computer Science, 2002. Disponível em: <[http://www.cs.colostate.edu/~draper/publications/baek\\_cvrip02.pdf](http://www.cs.colostate.edu/~draper/publications/baek_cvrip02.pdf)>. Acesso em: 05 out. 2007.

BARAGONA, Roberto; BATTAGLIA, Francesco. Outliers detection in multivariate time series by independent component analysis. **Neural Computation**. Vol. 19, n.º 7, p. 1962-1984, julho/2007.

BARTLETT, Marian Stewart and SEJNOWSKI, Terrence J. Independent components of face images: a representation for face recognition. In: SYMPOSIUM ON NEURAL COMPUTATION PROCEEDINGS, n. 4, 1997. **Joint Symposium on Neural Computation Proceedings**, 1997, p. 3-10.

BERNSTEIN, Peter L. **Desafio aos Deuses**: a fascinante história do risco; tradução Ivo Korytowski. Rio de Janeiro: Campus, 1997.

BERTUCCI, L. A.; AMARAL, H. F.; GARCIA, F. G. A estratégia de financiamento com subscrições de ações: o caso de empresas brasileiras. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO DOS CURSOS DE GRADUAÇÃO EM ADMINISTRAÇÃO, XXIII ANGRAD, 2002, Rio de Janeiro. **Anais...**Rio de Janeiro: Angrad, ago., 2002, 9 p.

BINGHAM, Ella and HYVÄRINEN, Aapo. A fast fixed-point algorithm for independent component analysis of complex valued signals. **International Journal of Neural Systems**, 2000. v. 10, n. 1, p. 1-8.

BLACK, Fisher. Capital market equilibrium with restricted borrowing. **Journal of Business**. v. 45, n. 3, jul., 1972.

\_\_\_\_\_. Noise. **Journal of Finance** vol. 41, p. 529-543, 1986.



BREALEY, Richard A. e MYERS, Stewart C. **Principles of Corporate Finance**. 7<sup>th</sup> ed. New York: McGraw-Hill, 2002.

BRITO, G. A. S.; BATISTELLA, F. D.; FAMA, R.. Fusões e Aquisições no Setor Bancário: Avaliação Empírica do Efeito Sobre o Valor das Ações. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII ENANPAD, 2004, Curitiba (PR). **Anais...**, Rio de Janeiro: ANPAD, 2004.

BRITO, Ney Roberto O. de. **Mercado de capitais e a estrutura empresarial brasileira**. Rio de Janeiro: Guanabara, 1981.

BROMWICH, Michael. **Financial reporting, information and capital markets**. London: Pitman Publishing, 1992.

BROWN, Glen D.; YAMADA, Satoshi and SEJNOWSKI, Terrence. Independent component analysis at the neural cocktail party. **Trends in neurosciences**, v. 24, n. 1, p. 54-63, January 2001.

BROWN, S. J.; WARNER, J. B. Measuring security price performance. **Journal of Financial Economics**. v.8, n.3, p.205-258, set., 1980.

\_\_\_\_\_. Using daily stock returns: the case of event studies. **Journal of Financial Economics**. v.14, p. 3-31, set., 1985.

BUENO, A. F.; BRAGA, R. F. R.; ALMEIDA, R. J. Pesquisa sobre a eficiência informacional no Mercado brasileiro nos casos de fusões e aquisições. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIV ENANPAD, 2000, Florianópolis (SC). **Anais...** Rio de Janeiro: ANPAD, set., 2000, 12 p.

CALHOUN, V. D. et al. fMRI activation in a visual-perception task: network of areas detected using the general linear model and independent component analysis. **Ideal library – neuroimage**, n.14, p. 1080-1088, 2001.

CAMARGOS, Marcos Antônio de. **Eficiência informacional do mercado de capitais brasileiro pós-Plano Real**: um estudo de evento dos anúncios de fusões e aquisições. 147p. Dissertação (Mestrado em Administração) – Centro de Pós-Graduação e Pesquisa em Administração, Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2002.

\_\_\_\_\_; BARBOSA, F. V.; Análise Empírica do Retorno Acionário Anormal e da Reação do Mercado de Capitais Brasileiro aos Anúncios de Fusões e Aquisições Ocorridos entre 1994 e 2001. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIX ENANPAD, 2005, Brasília (DF). **Anais...**, Rio de Janeiro: ANPAD, 2005.

\_\_\_\_\_. Emissão de ADRs, Retorno Acionário Anormal e o Comportamento das Ações no Mercado Doméstico: Evidências Empíricas. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXX ENANPAD, 2006, Salvador (BA). **Anais...** Rio de Janeiro: ANPAD, 2006.

\_\_\_\_\_. Evidências Empíricas do Impacto da Adesão aos Níveis Diferenciados de Governança corporativa sobre o Comportamento das Ações na Bovespa.. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXX ENANPAD, 2006, Salvador (BA). **Anais...** Rio de Janeiro: ANPAD, 2006b.

CAMPBELL, J. Y.; LO, A. W.; MACKINLAY, A. C. **The econometrics of financial markets**. 2<sup>th</sup> ed. New Jersey: Princeton University Press, 1997. 611 p.

CARDOSO, Jean-François and COMON, Pierre. **Tensor-based independent component analysis**. École Nationale Supérieure des Télécommunications, France, 2001. Disponível em: <<http://www.tsi.enst.fr/~cardoso/Papers.PDF/eusipco90.pdf>>. Acesso em: 05 out. 2007.

\_\_\_\_\_. **The three easy routes to independent component analysis contrast and geometry**. École Nationale Supérieure des Télécommunications, France 2005, Disponível em: <<http://www.tsi.enst.fr/~cardoso/Papers.PDF/ica01.pdf>>. Acesso em: 05 out. 2007.

CARROLL, Timothy J. et al. Confounding effect of large vessels on MR perfusion images analyzed with independent component analysis. **AJNR Am J Neuroradiol**, Denver, n. 23, p. 1007-1012, June/July 2002.

CASEY, Michael A. and WESTNER, Alex. Separation of mixed audio sources by independent subspace analysis. In: INTERNATIONAL COMPUTER MUSIC CONFERENCE, Berlin, august 2000. **Anais...** Berlin: ICMC, 2000.

CHAN, Lai-Wan and CHA, Siu-Ming. Selection of independent factor model in finance. In: INTERNATIONAL CONFERENCE ON INDEPENDENT COMPONENT ANALYSIS AND BLIND SIGNAL SEPARATION, n.3, 2001, California, USA. **Anais...USA: International Workshop on ICA**, 2001.

CHAN, Lai-Wan. The prediction performance of independent factor models. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2002, Hawaii, USA. **Anais...** USA: IEEE, 2002.

CHAN, S., GAU, G., WANG, K. Stock Market Reaction to Capital Investment Decisions: Evidence from Business Relocations. **Journal of Financial and Quantitative Analysis**, v. 30, n. 1, 1995.

CHEN, Ying; HÄRDLE, Wolfgang and SPOKOINY, Vladimir. GHICA – Risk analysis with GH distributions and independent components. **Economic risk**, 2005. Humboldt-Universität zu Berlin. Disponível em: <<http://sfb649.wiwi.hu-berlin.de>>. Acesso em: 17 out. 2007.

\_\_\_\_\_. Portfolio value at risk based on independent component analysis. **Economic risk**, 2005. Humboldt-Universität zu Berlin. Disponível em: <<http://sfb649.wiwi.hu-berlin.de>>. Acesso em: 17 out. 2007.

CHEUNG, Yiu-ming and XU, Lei. Independent component ordering in ICA time series analysis. **Elsevier – neurocomputing**, n. 41, p. 145-152, 2001.

CHINDARO, S.; SIRLANTZIS, K. and FAIRHURST, M. C. ICA-based multi-colour space texture classification system. **Electronics letters**, v. 42, n. 21, October 2006. Disponível em: <<http://web.ebscohost.com/ehost/pdf?vid=2&hid=12&sid=4afe2664-9986-46c3-a0a6-3b66400da67e%40sessionmgr108>>. Acesso em: 18 set. 2007.

CHIU, Kai-Chun and XU, Lei. Stock forecasting by arch driven gaussian tfa and alternative mixture-of-experts models. In: INTERNATIONAL WORKSHOP ON COMPUTATIONAL INTELLIGENCE IN ECONOMICS AND FINANCE, 3. 2003, North Carolina, USA. **Anais...**North Carolina, USA: CIEF, 2003. p. 1096-1099.

CICHOCKI, A.; DOUGLAS, S.C. and AMARI, S. Robust techniques for independent component analysis (ICA) with noisy data. **Elsevier – neurocomputing**, n. 22, p. 113-129, July 1998.

CICHOCKI, Andrzej et al. Neural networks for blind separation with unknown number of sources. **Elsevier – Neurocomputing**, n. 24, p. 55-93, 1999.

COMERLATO, Glória M. B. ; TERRA, P. R. S. . Um Estudo sobre a Relevância das Informações Geradas nas Apresentações das Companhias Abertas para a Associação Brasileira dos Analistas do Mercado de Capitais. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVII ENANPAD, 2003, Atibaia. **Anais...**, Rio de Janeiro: ANPAD, 2003.

COMON, P. Independent component analysis – a new concept? **Signal Processing**. n. ° 36, p. 287-314, 1994.

COPELAND, Tom; WESTON, J. Fred; SHASTRI, Kuldeep. **Financial theory and corporate policy**. 4<sup>th</sup> ed. USA: Addison-Wesley, 2005.

CORDEIRO, F.; PEROBELLI, F.; ARBEX, M. Expectativas racionais e eficiência informacional: análise do mercado acionário brasileiro num cenário de regras no período 1997-1999. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIII ENANPAD, 1999, Foz do Iguaçu (PR). **Anais...**, Rio de Janeiro: ANPAD, 1999.

DAMODARAM, A. **Avaliação de Investimentos**: ferramentas e técnicas para a determinação do valor de qualquer ativo. Rio de Janeiro: Qualitymark Ed., 1997.

DANTAS, Marcos. Informação e trabalho no capitalismo contemporâneo. **Lua Nova**. n. 60 (2003) Disponível em <<http://www.scielo.br>>. Acessado em: 01 de junho de 2006.

DEBENER, Stefan et al. **Independent components of the auditory novelty odd-ball**. University Hospital Hamburgo-Eppendorf, Germany, 2004. Disponível em: <<http://www.sccn.ucsd.edu/papers/BRES-D-04-05245.pdf>>. Acesso em: 05 out. 2007.

DEL BRIO, E., MIGUEL, A., PINDADO, J. **Investment and Firm Value: an Analysis Using Panel Data**. SSRN: disponível em <<http://www.ssrn.com>>. Acessado em: 22 de maio de 2006.

DENNIS, D. K.; McCONNELL, J. J. Corporate mergers and security returns. **Journal of financial economics**, v. 16, n. 2, p. 143-87, Jun., 1986.

DOWNES, John e GOODMAN, Jordan Elliott. **Dicionário de Termos Financeiros e de Investimentos**. Tradução: Ana Rocha Tradutores Associados. São Paulo: Nobel, 1993.

DRAGHICI, Sorin et al. Mining HIV dynamics using independent component analysis. **Oxford Journals**, 13, January 2003. *Bioinformatics*, v.19, n.8, p. 981-986.

ELTON, E. J.; GRUBER, M. J. **Modern portfolio theory and investment analysis**. 5<sup>th</sup> ed. USA: John Wiley, 1995. 715 p.

ERIKSSON, Jan; KARVANEN, Juha and KOIVUNEN, Visa. **Source distribution adaptive maximum likelihood estimation of ICA model**. Helsinki University of Technology – Academy of Finland. Disponível em: <[http://citeseer.ist.psu.edu/cache/papers/cs/21618/http://zSzzSzwoster.hut.fizSzstatsp.zSzpaperszSzICA2000\\_jme.pdf/eriksson00source.pdf](http://citeseer.ist.psu.edu/cache/papers/cs/21618/http://zSzzSzwoster.hut.fizSzstatsp.zSzpaperszSzICA2000_jme.pdf/eriksson00source.pdf)>. Acesso em: 05 out. 2007.

FAMA, Eugene F. The behavior os stock market prices. **The journal of business**, v. 38, n. 1, p. 34-105, jan., 1965.

\_\_\_\_\_. Efficient capital markets. **Journal of finance**, v. 25, p. 383-417, 1970.

\_\_\_\_\_. Efficient capital markets II. **Journal of finance**, v. 462, n. 5, 1991.

\_\_\_\_\_. Market efficiency, long-term returns, and behavioral finance, **Journal of financial economics**, v. 49, p. 283-306, 1998.

FIORI, Simone. Hybrid independent component analysis by adaptive LUT activation function neurons. **Elsevier – Neural Networks**, n. 15, p. 85-94, 2002.

FIRMINO, A. L. G.; SANTOS, A. G. Q.; MATSUMOTO, A. S. Dividendos interessam? Uma constatação empírica recente sobre a relevância da política de dividendos na bolsa de valores de São Paulo (1996 a 2002). In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII Enanpad, 2004, Curitiba (PR). **Anais...** Rio de Janeiro: ANPAD, 2004.

FU, Tak-chung et al. **Pattern discovery from stock time series using self-organizing maps**. Disponível em: <<http://www.sigkdd.org/kdd2001/Workshops/fcl.pdf>>. Acesso em : 17 out. 2007.

GABRIELI, M. F.; SAITO, R. Recompra de ações no Brasil: regulamentação e impacto sobre minoritários. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVII ENANPAD, 2003, Atibaia. **Anais...**, Rio de Janeiro: ANPAD, 2003.

GARCIA, F. G.; SATO, L. G.; CASELANI, C. N. O impacto da política de transparência sobre o valor das empresas brasileiras. In: ENCONTRO NACIONAL DE PRO-

GRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII ENANPAD, 2004, Curitiba (PR). **Anais...**, Rio de Janeiro: ANPAD, 2004.

GIANNALOPOULOS, Xavier. **Comparison of adaptive independent component analysis algorithms**. École Polytechnique Fédérale de Lausanne. Disponível em: <<http://citeseer.ist.psu.edu/giannakopoulos98comparison.html>>. Acesso em: 05 out. 2007.

HALFELD; SILVA. Análise dos estilos de investimentos na América Latina. EP/USP, trabalho em andamento, SP, 1999. In: HAUGEN, R. A. **Os segredos da bolsa**. Tradução Nota Assessoria. São Paulo: Pearson Prentice Hall, 2000. 175 p.

HANSEN, Lars Kai, LARSEN, Jan and KOLENDA, Thomas. **On independent component analysis for multimedia signals**. Technical University of Denmark – Department of Mathematical Modelling, Scandinavia, Disponível em: <[http://mole.imm.dtu.dk/thko\\_project/hansen.mmica.pdf](http://mole.imm.dtu.dk/thko_project/hansen.mmica.pdf)>. Acesso em: 05 out. 2007.

HARITOPOULOS, Michel; YIN, Hujun and ALLINSON, Nigel M. Image denoising using self-organizing map-based nonlinear independent component analysis. **Elsevier – Neural Networks**, n. 15, p. 1085-1098, 2002.

HATEREN, J. H. van and SCHAAF, A. van der. Independent component filters of natural images compared with simple cells in primary cortex. **The royal society**. The Netherlands, n. 265, p. 359-366, November 1997.

\_\_\_\_\_ and RUDERMAN, D. L. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. **The royal society**, n. 265, p. 2315-2320, 1998.

HAUGEN, R. A. **Os segredos da bolsa**. Tradução Nota Assessoria. São Paulo: Pearson Prentice Hall, 2000. 175 p.

HAVRAN, C. et al. **Independent component analysis for face authentication**. Université Catholique de Louvain – Electricity Dept. Belgium. Disponível em: <<http://www.tele.ucl.ac.be/publi/1162.pdf>>. Acesso em: 05 out. 2007.

HERRERA, P.; PROCIANOY, J. Avaliação dos retornos das ações das empresas concordatárias. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXII ENANPAD, 1998, Foz do Iguaçu (PR). **Anais...**, Rio de Janeiro: ANPAD, 1998.

HOLTHAUSEN, F. S.; GALLI, O. C. Lançamento de DRs por empresas brasileiras no mercado norte-americano: valorização de mercado, volatilidade e performance ajustada ao risco. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXV ENANPAD, 2001, Campinas (SP). **Anais...**, Rio de Janeiro: ANPAD, 2001.

HOPWOOD, Anthony G. **An accounting system and managerial behaviour**. Hampshire: Biddles Limited, 1973.

HUANG, De-Shuang and ZHENG, Chun-Hou. Independent component analysis-based penalized discriminant method for tumor classification using gene expression data. **Oxford Journals**, 2006. *Bioinformatics*, v.22, n.15, p. 1855-1862.

HYVÄRINEN, Aapo. Independent Component Analysis for Time-dependent Stochastic Processes. **Laboratory of Computer and Information Science**, Finland, 1999. Disponível em: <[http://citeseer.ist.psu.edu/cache/papers/cs/1194/http:zSzzSzwww.cis.hut.fizSz~aapo zSzpszSzICANN98\\_INNO.pdf/hyv98independent.pdf](http://citeseer.ist.psu.edu/cache/papers/cs/1194/http:zSzzSzwww.cis.hut.fizSz~aapo zSzpszSzICANN98_INNO.pdf/hyv98independent.pdf)>. Acesso em: 05 out. 2007.

\_\_\_\_\_; SÄRELÄ, Jaakko and VIGÁRIO, Ricardo. **Spikes and bumps: artifacts generated by independent component analysis with insufficient sample size**. Helsinki University of Technology – Academy of Finland, 1999b. Disponível em: <[http://citeseer.ist.psu.edu/cache/papers/cs/15045/http:zSzzSzwww.cis.hut.fizSz~aapo zSzpszSzgzzSzICA99\\_bumps.pdf/spikes-and-bumps-artefacts.pdf](http://citeseer.ist.psu.edu/cache/papers/cs/15045/http:zSzzSzwww.cis.hut.fizSz~aapo zSzpszSzgzzSzICA99_bumps.pdf/spikes-and-bumps-artefacts.pdf)>. Acesso em: 05 out. 2007.

\_\_\_\_\_. Survey on Independent component analysis. **Neural Computing Surveys**, 1999c. n.2, p. 94-128. Disponível em: <[http://www.icsi.berkeley.edu/ftp/pub/ai/jagota/vol2\\_4.pdf](http://www.icsi.berkeley.edu/ftp/pub/ai/jagota/vol2_4.pdf)>. Acesso em: 17 out. 2007.

\_\_\_\_\_ and OJA, Erkki. Independent component analysis: algorithms and applications. **Neural networks**, Finland, v.13, n.4-5, p. 411-430, 2000. Disponível em: <[http://www.cis.hut.fi/aapo/papers/IJCNN99\\_tutorialweb/](http://www.cis.hut.fi/aapo/papers/IJCNN99_tutorialweb/)>. Acesso em: 17 set. 2007.

\_\_\_\_\_. **One-unit learning rules for independent component analysis**. Helsinki University of Technology – Academy of Finland, 2002. Disponível em: <<http://citeseer.ist.psu.edu/cache/papers/cs/512/http:zSzzSzwww.cis.hut.fizSz~aapo zSzpszSzNIPS96.pdf/one-unit-learning-rules.pdf>>. Acesso em: 05 out. 2007.

JENSEN, M. C.; Some anomalous evidence regarding market efficiency. **Journal of financial economics**, v.6, n.1, p. 95-101, mar., 1978.

JONES, E. **Company Investment Announcements and the Market Value of the Firm**. (2000) SSRN: disponível em <<http://www.ssrn.com>>. Acessado em: 25 de maio de 2006.

JUNG, Tzyy-Ping et al. Extended ICA removes artifacts from electroencephalographic recordings. **Advances in Neural Information Processing Systems**, Cambridge, v. 10, p. 894-900, 1998.

\_\_\_\_\_. Independent component analysis of single-trial event-related potentials. In: INTERNATIONAL WORKSHOP ON INDEPENDENT COMPONENT ANALYSIS AND BLIND SIGNAL SEPARATION, France, 1999. **Anais...** International Workshop on Independent Component Analysis, 1999. p. 173-178.

\_\_\_\_\_. Independent component analysis of biomedical signals. In: INTERNATIONAL WORKSHOP ON INDEPENDENT COMPONENT ANALYSIS AND BLIND SIGNAL SEPARATION, n.2, 2000, Finland. **Anais...** Finland: International Workshop on ICA, 2000. p. 633-644.

\_\_\_\_\_. Independent component analysis of biomedical signals. In: INTERNATIONAL SYMPOSIUM ON INDEPENDENT COMPONENT ANALYSIS AND BLIND SIGNAL SEPARATION, n.4, 2003, Japan. **Anais...**Japan: International Workshop on ICA, april 2003. p. 1077-1082.

JUTTEN, C. and HÉRAULT, J. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. **Signal Processing**. n.º 24, p. 1-10, 1991.

KANO, Manabu; HASEBE, Shinji and HASHIMOTO, Iori. Evolution of multivariate statistical process control: application of independent component analysis and external analysis. In: PROCEEDINGS FOUNDATIONS OF COMPUTER-AIDED PROCESS OPERATIONS, 2003. **Anais...** FOCAPO, 2003. p. 385-388.

KELM, M. L. **A sensibilidade do preço das ações de empresas de capital aberto às informações contábeis divulgadas**. 1996. Dissertação (Mestrado em Administração), Programa de Pós-Graduação em Administração, Escola de Administração, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1996.

KENDALL, M. The analysis of economic time series, part I: prices. **Journal of the royal statistic society**, n.96, p. 11-25, 1953.

KHOTARI, S. P.; WARNER, J. B. Econometrics of event studies. (working paper) forthcoming in B. Espen Eckbo (ed.). Handbook of corporate finance: empirical corporate finance. v. A. **Handbooks in finance series**. Ch. 1. North-Holland/Elsevier, 2006.

KLOECKNER, Gilberto de Oliveira. Estudos de evento: a análise de um método. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XIX ENANPAD, set. 1995, João Pessoa (PB). **Anais...**, p. 261-70. Rio de Janeiro: ANPAD, 1995.

LAMB, R. A. **Decisão de Investimento nas Empresas e a Reação do Mercado de Capitais**: uma Abordagem Informacional. 1993. Dissertação (Mestrado em Administração) – Programa de Pós-Graduação em Administração, Escola de Administração, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1993.

LAPPALAINEN, Harri and HONKELA, Antti. Bayesian nonlinear independent component analysis by multi-layer perceptrons. **Advances in independent component analysis**. Ed. by Mark Girolami, p. 93-121, Springer, 2000.

LEAL, R. P. C. Retornos anormais e sinalização nas aberturas de capital. **Revista Brasileira do Mercado de Capitais**. Rio de Janeiro, v. 14, n. 40, p. 33-48, 1988/1989.

\_\_\_\_\_; AMARAL, A. S. Um momento para o insider trading: o período anterior ao anúncio de uma emissão pública de ações. **Revista Brasileira do Mercado de Capitais**. Rio de Janeiro, v. 15, n. 41, p. 21-6, 1990.



LEE, Jae Sung et al. Blind separation of cardiac components and extraction of input function  $H_2^{15}O$  dynamic myocardial PET using independent component analysis. **The journal of nuclear medicine**, USA, June 2001. v. 42. p. 938-943.

LEE, Su-In and BATZOGLOU, Serafim. Application of independent component analysis to microarrays. **Genome biology**, v.4, n.11, October, 2003.

LEE, Te-Won; GIROLAMI, M.; BELL, A. J. and SEJNOWSKI, T. J. A unifying information-theoretic framework for independent component analysis. **Computers and mathematics with applications**, n. 39, p. 1-21, 2000.

LEITE, Hélio de P.; SANVICENTE, Antônio Z. Valor patrimonial: usos, abusos e conteúdo informacional. **Revista de Administração de Empresas**. São Paulo, v. 30, n. 3, p. 17-31. jul/set, 1990.

LEITE, Luciana. **Análise de Componentes Independentes Aplicada à Identificação de Regiões Lesionadas em Mamogramas**. 2005. Dissertação (Mestrado em Engenharia Elétrica). Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia. Universidade Federal do RJ, Rio de Janeiro, 2005.

LEITE, P.; SANVICENTE, A. Z. Planos de estabilização e instabilidade nos mercados acionários. **Revista de Administração de Empresas**. Rio de Janeiro: FGV. V. 31, n. 2, p. 25-36, 1991.

LEONG, W. Y. and HOMER, J. Blind multiuser receiver for DS-CDMA wireless system. **IEE Proc.-Commun.**, vol. 153, n. 5, October 2006. Disponível em: <<http://web.ebscohost.com/ehost/pdf?vid=11&hid=15&sid=3446a6dd-89a3-497a-9c16-807bf93d640a%40SRCM1>>. Acesso em: 18 set. 2007.

LIAO et al. Network component analysis: reconstruction of regulatory signals in biological systems. **PNAS - Proceedings of the National Academy of Sciences**, EUA, v. 100, n. 26, p. 15522-15527, December 2003.

LIEBERMEISTER, Wolfram. Linear modes of gene expression determined by independent component analysis. **Oxford Journals**, August 2001. v.18, n.1, p. 51-60.

LIMA, João Batista Nast de; TERRA, Paulo Renato Soares. A reação do mercado de capitais brasileiro à divulgação das informações contábeis. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII Enanpad, 2004, Curitiba (PR). **Anais...** Rio de Janeiro: ANPAD, 2004.

\_\_\_\_\_. Governança corporativa e reação do mercado de capitais às informações financeiras. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII Enanpad, 2004, Curitiba (PR). **Anais...** Rio de Janeiro: ANPAD, 2004b.

\_\_\_\_\_. Governança Corporativa e a Reação do Mercado de Capitais às Informações Financeiras. **Revista Contabilidade & Finanças**. São Paulo, v. 4, p. 35-49, 2006.



LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The Review of Economics and Statistics**, n. 47, p. 13-37, 1965.

MACKINLAY, A. C.; Event studies in economics and finance. **Journal of economic literature**, v. 35, n. 1, p. 13-39, mar., 1997.

MADAN, Dilip B. and YEN, Ju-Yi J. **Asset allocation for CARA utility with multivariate levy returns.** Disponível em: <[http://www.fdic.gov/bank/analytical/cfr/apr\\_2005/CFRCP\\_2005-17\\_MadanYen.pdf](http://www.fdic.gov/bank/analytical/cfr/apr_2005/CFRCP_2005-17_MadanYen.pdf)>. Acesso em : 17 out. 2007.

MAKEIG, Scott et al. Independent component analysis of simulated ERP data. In: INTEGRATED HUMAN BRAIN SCIENCE, T. Nakada, ed., 2000. **Anais...** T. Nakada: Integrated Human Brain Science: Theory, Method, Applications (music), 2000, p. 1-24.

MALAROIU, Simona; KIVILUOLO, Kimmo and OJA, Erkki. **Time series prediction with independent component analysis.** Disponível em: <<http://www.cis.hut.fi/~kkluoto/publications/ait99ica.pdf>>. Acesso em : 05 out. 2007.

MARKOSKI, Ângela Silva; MORENO, Roberto. Mudanças na classificação de risco soberano brasileiro e efeitos no Ibovespa. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIX ENANPAD, 2005, Brasília (DF). **Anais...**, Rio de Janeiro: ANPAD, 2005.

MARKOWITZ, Harry M.; Portfolio selection. **The Journal of Finance**, v. 7, n. 1, p. 77-91, mar., 1952.

MARTOGLIO, Ann-Marie et al. A decomposition model to track gene expression signatures: preview on observer-independent classification of ovarian cancer. **Oxford Journals**, 2002. *Bioinformatics*, v.18, n.12, p. 1617-1624.

McGOVERN, Siobhain. Dealing with the Duhem-Quine thesis in financial economics: can causal holism help? **Cambridge Journal of Economics**, n. 30, p. 105-122, 2005.

McKEOWN, Martin J. Detection of consistently task-related activations in fMRI data with hybrid independent component analysis. **Ideal library – neuroimage**, n.11, p. 24-35, 2000.

\_\_\_\_\_; HANSEN, Lars Kai and SEJNOWSKI, Terrence J. Spatially independent activity patterns in functional MRI data during the stroop color-naming task. **PNAS - Proceedings of the National Academy of Sciences**, EUA, v. 95, p. 803-810, February 1998.

\_\_\_\_\_. Independent component analysis of functional MRI: what is signal and what is noise? **Elsevier – current opinion in neurobiology**, n. 13, p. 620-629, 2003.

MEDEIROS, Otávio Ribeiro de; MATSUMOTO, Alberto Shiguero. Emissões públicas de ações, volatilidade e insider information na Bovespa. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII Enanpad, 2004, Curitiba (PR). **Anais...** Rio de Janeiro: ANPAD, 2004.

MORITZ, Chad H. Whole-brain functional MR imaging activation from a finger-tapping task examined with independent component analysis. **AJNR Am J Neuroradiol**, Madison, n. 21, p. 1629-1635, October 2000.

MOSSIN, J. Equilibrium in a capital asset market. **Econometrica**, n. 34, 768-83, 1966.

MUTIHAC, Radu and HULLE, Marc M. Van. A comparative survey on adaptive neural network algorithms for independent component analysis. **Romanian reports in physics**, Romania, 2003. v.55, n.1, p. 43-67, 2003.

NAVARRO, Paulo Sérgio. **O Impacto de formas societárias sobre o valor das empresas de capital aberto: o caso da transformação de conglomerados financeiros em bancos múltiplos**. 1996. Dissertação (Mestrado em Administração), Programa de Pós-Graduação em Administração, Escola de Administração, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1996.

NAVARRO, Paulo Sérgio. O impacto de formas societárias sobre o valor das empresas de capital aberto: o caso da transformação de conglomerados financeiros em bancos múltiplos. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXI ENANPAD, 1997, Rio das Pedras. **Anais...**, Rio de Janeiro: ANPAD, 1997.

NEFTCI, S. N. **An introduction to the mathematics of financial derivatives**. San Diego, CA: Academic Press, 1996.

NOVIS NETO, J. A.; SAITO, R. Dividend yields e persistência de retornos anormais das ações: evidência do mercado brasileiro. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVI ENANPAD, 2002, Salvador (BA). **Anais...**, Rio de Janeiro: ANPAD, 2002.

OJA, E. **Independent component analysis: algorithms and applications**. Disponível em: <<http://www.stat.jyu.fi/icors2005/icorsabstracts/oja.pdf>>. Acesso em :17 set. 2007.

PATROCINIO, Maurício Réa; KAYO, Eduardo Kazuo; KIMURA, Herbert. Intangibilidade e criação de valor nos eventos de fusão e aquisição: uma análise dos retornos anormais do período de 1994 a 2004. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIX ENANPAD, 2005, Brasília (DF). **Anais...**, Rio de Janeiro: ANPAD, 2005.

PAULOS, John Allen. **A lógica do mercado de ações: uma análise prática do funcionamento das bolsas de valores**. Tradução de Celso da Cunha Serra. 3<sup>th</sup> ed. Rio de Janeiro: Elsevier, 2004.

PAXSON, Deam; WOOD, Douglas. **Dicionário enciclopédico de finanças**. São Paulo: Atlas, 2001. 293p.

PENTEADO, Marco Antônio de Barros; FAMÁ, Rubens. Será que o beta que temos é o beta que queremos? **Caderno de Pesquisas em Administração**, São Paulo, v. 9, n. 3, julho/setembro, 2002.

PEROBELLI, F. F. C.; NESS JR., W. L.; Reações do mercado acionário a variações inesperadas nos lucros das empresas: um estudo sobre a eficiência informacional no mercado brasileiro. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXIV ENANPAD, 2000, Florianópolis (SC). **Anais...** Rio de Janeiro: ANPAD, set., 2000, 15 p.

PINDYCK, Robert S.; RUBINFELD, Daniel L. **Econometria: modelos e previsões**. 4ª edição. Rio de Janeiro: Elsevier, 2004.

PROCIANOY, J. L.; ANTUNES, M. Os Efeitos das Decisões de Investimento das Empresas Sobre os Preços de suas Ações no Mercado de Capitais. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXV ENANPAD, 2001, Campinas, SP. **Anais...** Rio de Janeiro: ANPAD, 2001.

\_\_\_\_\_; VERDI, R. S. Reação do Mercado à Alteração na Composição da Carteira de Índices da Bolsa de Valores Brasileiros. **Revista Brasileira de Finanças**, v. 4, p. 22-35, 2006.

PRUX JR., Jaime Luis. **Assimetria Informacional e Precificação das Ações das Empresas Negociadas na Bolsa de Valores de São Paulo: evidências a partir da faculdade de divulgar demonstrações contábeis em moeda constante a partir de 1996**. 1998. Dissertação. (Mestrado Interinstitucional PPGA – UCS), de Pós-Graduação em Administração, Escola de Administração, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1998.

QUIGLEY, Michelle A. Comparison of independent component analysis and conventional hypothesis-driven analysis for clinical functional MR image processing. **AJNR Am J Neuroradiol**, Madison, n. 23, p. 49-58, January 2002.

ROBERTS, H. Stock market patterns and financial analysis: methodological suggestions. **Journal of Finance**, n. 14, p. 11-25, 1959.

ROCHMAN, Ricardo Ratner; EID JR., William. Insiders conseguem retornos anormais? Estudos de eventos sobre as operações de insiders das empresas de governança corporativa diferenciada da Bovespa. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXX ENANPAD, 2006, Salvador (BA). **Anais...** Rio de Janeiro: ANPAD, 2006.

ROSS, S. A. The arbitrage theory of capital asset pricing. **Journal of Economic Theory**, v. 13, n. 3, p. 341-360, dez., 1976.

\_\_\_\_\_; WESTERFIELD, Randolph W.; JAFFE, Jeffrey, F. **Administração Financeira**. Tradução: Antônio Zoratto Sanvicente. 2<sup>th</sup> ed. São Paulo: Atlas, 2002. 776 p.

SÁ, Graciano. **O Valor das Empresas**. Rio de Janeiro: Expressão e Cultura, 2001.  
 SARLO NETO, Alfredo; LOSS, Lenita; NOSSA, Valcemiro. A capacidade informacio-  
 nal dos resultados contábeis no mercado brasileiro: a diferença entre as ações ordi-  
 nárias e as ações preferenciais. In: ENCONTRO NACIONAL DE PROGRAMAS DE  
 PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVIII Enanpad, 2004, Curitiba (PR).  
**Anais...** Rio de Janeiro: ANPAD, 2004.

SCHIEHLL, E. O efeito da divulgação das demonstrações contábeis no mercado de  
 capitais brasileiro: um estudo sobre a variação no preço das ações. In: ENCONTRO  
 NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XX  
 ENANPAD, set. 1996, Angra dos Reis (RJ). **Anais...**, p. 289-303. Rio de Janeiro:  
 ANPAD, 1996.

\_\_\_\_\_. **O Efeito da Divulgação das Demonstrações Financeiras no Mer-  
 cado de Capitais Brasileiro**: um estudo sobre a variação no preço das ações.  
 1996. Dissertação (Mestrado em Administração), Programa de Pós-Graduação em  
 Administração, Escola de Administração, Universidade Federal do Rio Grande do  
 Sul, Porto Alegre, 1996b.

SCHOLZ, M. et al. Metabolite fingerprinting: detecting biological features by inde-  
 pendent component analysis. **Oxford Journals**, abril 2004. *Bioinformatics*, v. 20, n.  
 15, p. 2447-2454.

SHANNON, C.; WEAVER, W. **A teoria matemática da comunicação**. Rio de Janei-  
 ro: Difel, 1975.

SHARPE, W. F. A Simplified Model for Portfolio Analysis. **Management Sci.**, n. 9, p.  
 277-293, 1963.

SHARPE, W. F. Capital asset prices: a theory of market equilibrium under conditions  
 of risk. **Journal of Finance**, n. 19, p. 425-42, 1964.

SOARES, R. O.; ROSTAGNO, L. M.; SOARES, K. T. C. Estudo de evento: o método  
 e as formas de cálculo do retorno anormal. In: ENCONTRO NACIONAL DE PRO-  
 GRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXVI ENANPAD, 2002,  
 Salvador (BA). **Anais...**, Rio de Janeiro: ANPAD, 2002, 14 p.

SPIEGEL, Murray R. **Probabilidade e Estatística**. Coleção Schaum. São Paulo:  
 Makron Books, 2001.

STONE, James V. Independent component analysis: an introduction. **Trends in  
 cognitive sciences**, U K, v. 6, n. 2, p. 59-64, February 2002.

TREYNOR, J. **Towards a Theory of the Market Value of Risky Assets**. Unpub-  
 lished manuscript, 1961.

VAN HORNE, James C. **Financial Management and Policy**. 11<sup>a</sup> ed. Upper Sadle  
 River: Prentice Hall, 1997.

VIEIRA, K. M.; PROCIANOY, J. L. Reação do mercado a stock splits e stock dividends: um estudo de evento e um teste para a hipótese de liquidez. In: ENCONTRO NACIONAL DE PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXII ENANPAD, 1998, Foz do Iguaçu (PR). **Anais...**, Rio de Janeiro: ANPAD, 1998, 15 p.

\_\_\_\_\_. Reação dos investidores a bonificações: um estudo em países da América Latina. In: ENCONTRO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO, XXV ENANPAD, 2001, Campinas, SP. **Anais...** Rio de Janeiro: ANPAD, 2001.

YEN, Ju-Yi J. **Multivariate lévy processes for financial returns**. 2004. 158 f. Dissertation ( Doctor of Philosophy) – Faculty of the Graduate School of the University of Maryland. USA, 2004.

YU, Dan; SATTAR, Farook and MA, Kai-Kuang. Watermark detection and extraction using independent component analysis method. **EURASIP Journal on Applied Signal Processing – Hindawi Publishing Corporation**, n. 1, p. 92-104, 2002.

ZAMBONI, Nicola and SAUER, Uwe. Model-independent fluxome profiling from  $^2\text{H}$  and  $^{13}\text{C}$  experiments for metabolic variant discrimination. **Genome biology**, v.5, n.12, November, 2004.

## ANEXO A – Algoritmos elaborados no ambiente Matlab®

### 1- Algoritmo de simulação dos modelos geradores de retornos anormais

```
n=83; % numero de series a serem simuladas

t=245; % tamanho de cada serie a ser simulada

alpha1 = 0.05; % nivel de significancia
alpha2 = 0.01; % nivel de significancia

sim=1000; % numero de simulacoes a serem realizadas

clear rfr3_95;
clear rfo3_95;
clear rfc3_95;

clear rfr3_99;
clear rfo3_99;
clear rfc3_99;

for c=1:sim

    clear iid;

    for i=1:n
        iid_norm_serie(:,i)=random('norm',media(i),desviopadrao(i),t,1);
    end

    iid=iid_norm_serie;

    %iid=data;

    a=size(iid);

    media_dia=zeros(1,a(1));
    desvio_serie=zeros(1,a(2));

    for i=1:a(1)
        media_dia(i)=mean(iid(i,:)); % vetor para guardar a média de cada
dia
    end

    for i=1:a(2)
        desvio_serie(i)=std(iid(:,i)); % vetor para guardar o desvio padrao
de cada serie
    end

    %x = [0:0.0001:0.01]; % vetor com o incremento da anormalidade
x = [0:0.0001:0.02]; % vetor com o incremento da anormalidade
b = length(x); % guarda o tamanho do vetor de incrementos
ret0 = iid(a(1),:); % vetor com as observações do ultimo dia
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% MODELO AJUSTADO AO MERCADO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% matrix com retornos anormais
an0 = zeros(b,a(2));
% matrix com retornos anormais menos a media do mercado no ultimo dia
an1 = zeros(b,a(2));
% matrix padronizada
an2 = zeros(b,a(2));
% Vetores resposta
foo = zeros(1,b);
fo1 = zeros(1,b);
fo2 = zeros(1,b);

for i=1:b
    for j=1:a(2)
        an0(i,j) = (( 1+ret0(j))*( 1+x(i) )) - 1; % acrescenta o retor-
no anormal
        an1(i,j) = an0(i,j)-media_dia(a(1)); % desconta a media da ul-
tima observação
        an2(i,j) = an1(i,j)/desvio_serie(j); % padroniza a resposta
    end
    foo(i) = sum(an2(i,:)); % soma a ultima linha da tabela (n+1)
    fo1(i) = foo(i)/sqrt(a(2)); % calcula a estatistica do teste t
    if (fo1(i) > tinv(1-alpha1,a(2))) % testa se o resultado é signifi-
cativo
        fo2_95(i)= x(i);
    else
        fo2_95(i)=0;
    end

    if (fo1(i) > tinv(1-alpha2,a(2))) % testa se o resultado é signifi-
cativo
        fo2_99(i)= x(i);
    else
        fo2_99(i)=0;
    end

end

fo3_95 = fo2_95*100; % transforma os resultados para valor percentual
fo3_99 = fo2_99*100; % transforma os resultados para valor percentual

for i=1:length(x)
    if fo3_95(i)>0
        rfo3_95(c)=fo3_95(i); % imprime o primeiro valor de incremento
de anormalidade significativo na tela
        break
    end
end

for i=1:length(x)
    if fo3_99(i)>0
        rfo3_99(c)=fo3_99(i); % imprime o primeiro valor de incremento
de anormalidade significativo na tela
        break
    end
end
end

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% MODELO AJUSTADO AO RISCO E AO MERCADO
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% matrix com retornos anormais menos a regressão
ar1 = zeros(b,a(2));
% matrix padronizada
ar2 = zeros(b,a(2));
% matrix dos coeficientes da regressão
reg = zeros(a(2),2);
% Vetores resposta
fro = zeros(1,b);
fr1 = zeros(1,b);
fr2 = zeros(1,b);

for i=1:b
    for j=1:a(2)
        reg(j,:) = regress(iid(:,j),[media_dia; ones(1,t)]');
        ar1(i,j) = an0(i,j) - (reg(j,2) + reg(j,1)*media_dia(a(1)));
        ar2(i,j) = ar1(i,j)/desvio_serie(j);
    end
    fro(i) = sum(ar2(i,:));
    fr1(i) = fro(i)/sqrt(a(2));

    if (fr1(i) > tinv(1-alpha1,a(2)-1))
        fr2_95(i) = x(i);
    else
        fr2_95(i) = 0;
    end

    if (fr1(i) > tinv(1-alpha2,a(2)-1))
        fr2_99(i) = x(i);
    else
        fr2_99(i) = 0;
    end
end

fr3_95 = fr2_95*100;
fr3_99 = fr2_99*100;

for i=1:length(x)
    if fr3_95(i)>0
        rfr3_95(c)=fr3_95(i);
        break
    end
end

for i=1:length(x)
    if fr3_99(i)>0
        rfr3_99(c)=fr3_99(i);
        break
    end
end
end

```



```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% MODELO CAPM
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% matrix com retornos anormais menos a regressão
ac1 = zeros(b,a(2));
% matrix padronizada
ac2 = zeros(b,a(2));
% lê a matrix do CDI
CDI = 0.000566745;
% Vetores resposta
fco = zeros(1,b);
fc1 = zeros(1,b);
fc2 = zeros(1,b);

for i=1:b
    for j=1:a(2)
        ac1(i,j) = an0(i,j) - (CDI + reg(j,1)*(media_dia(a(1)) - CDI));
        ac2(i,j) = ac1(i,j)/desvio_serie(j);
    end
    fco(i) = sum(ac2(i,:));
    fc1(i) = fco(i)/sqrt(a(2));

    if (fc1(i) > tinv(1-alpha1,a(2)-1))
        fc2_95(i) = x(i);
    else
        fc2_95(i) = 0;
    end

    if (fc1(i) > tinv(1-alpha2,a(2)-1))
        fc2_99(i) = x(i);
    else
        fc2_99(i) = 0;
    end

end

fc3_95 = fc2_95*100;
fc3_99 = fc2_99*100;

for i=1:length(x)
    if fc3_95(i)>0
        rfc3_95(c)=fc3_95(i);
        break
    end
end

for i=1:length(x)
    if fc3_99(i)>0
        rfc3_99(c)=fc3_99(i);
        break
    end
end

end

resultado=zeros(length(x),7);

```

```
resultado(:,1)=x*100;

for c=1:sim

    for j=1:length(x)

        if rfo3_95(c)==resultado(j,1)
            resultado(j,2)=resultado(j,2)+1;
        end

        if rfr3_95(c)==resultado(j,1)
            resultado(j,3)=resultado(j,3)+1;
        end

        if rfc3_95(c)==resultado(j,1)
            resultado(j,4)=resultado(j,4)+1;
        end

        if rfo3_99(c)==resultado(j,1)
            resultado(j,5)=resultado(j,5)+1;
        end

        if rfr3_99(c)==resultado(j,1)
            resultado(j,6)=resultado(j,6)+1;
        end

        if rfc3_99(c)==resultado(j,1)
            resultado(j,7)=resultado(j,7)+1;
        end

    end

end

resultado(:,2:7)=cumsum(resultado(:,2:7)/sim);

xlswrite('resultado_simu_fim_r.xls', resultado)
```

## 2- Algoritmo de simulação no FastICA de geração de séries aleatórias AR (1)

```

clear result;
clear anorm95;
clear anorm99;

n=83; % numero de series a serem simuladas

t=245; % tamanho de cada serie a ser simulada

sim=1000; % numero de vezes que a simulacao serah repetida

event=[0:0.01:0.1]; % variavel que dah o valor de cada anormalidade a ser
adicionada as series

for z=1:length(event)

    clear iid;
    clear iid_norm_serie;

    for i=1:n
        iid_norm_serie(:,i)=random('norm',media(i),desviopadrao(i),t,1);
    end

    iid=iid_norm_serie;

    for j=1:sim

        clear cheb95s;
        clear cheb95i;
        clear cheb99s;
        clear cheb99i;
        clear vvar;

        r_iidc = zeros(n,245); vvar=0;

        for i=1:n
            if mod(i,2)==1
                iid_norm_serie(:,i)=armasim(.7,0,t,media(i),desviopadrao(i));
            else
                iid_norm_serie(:,i)=armasim(-
                    .7,0,t,media(i),desviopadrao(i));
            end
        end

        iid=iid_norm_serie;

        iidc = [iid [zeros(t-1,1); event(z)]]; % adiciona o vetor de erros
ao conjunto de dados como uma outra serie

        iidc = (eye(n+1)+[zeros(1,n+1); 0.5*eye(n) zeros(n,1)])*iidc'; %
mistura as series como no paper de Baragona e Battaglia (2007)

```

```

    %r_iidc = fastica(iidc(1:83,:), 'maxNumIterations', 10000, 'verbo-
se', 'off'); % roda o fastica retirando a ultima serie

    %r_iidc = fastica(iidc, 'maxNumIterations', 10000, 'verbose',
'off'); % roda o fastica sem retirar a ultima serie

    % desigualdade de chebyshev

    kkk=0;

    for i=1:min(size(r_iidc));

        cheb95s=mean(r_iidc(i,1:(t)))+15*std(r_iidc(i,1:(t)));
        cheb95i=mean(r_iidc(i,1:(t)))-15*std(r_iidc(i,1:(t)));

        if (r_iidc(i,t)>cheb95s || r_iidc(i,t)<cheb95i)
            kkk(i)=1;
        else
            kkk(i)=0;
        end

    end

    % conta o numero de vezes que a fastICA detectou uma anormalidade
    % significativa a 15 desvios

    anorm95(j)=max(kkk);

    kkk=0;

    for i=1:min(size(r_iidc));

        cheb99s=mean(r_iidc(i,1:(t)))+20*std(r_iidc(i,1:(t)));
        cheb99i=mean(r_iidc(i,1:(t)))-20*std(r_iidc(i,1:(t)));

        if (r_iidc(i,t)>cheb99s || r_iidc(i,t)<cheb99i)
            kkk(i)=1;
        else
            kkk(i)=0;
        end

    end

    % conta o numero de vezes que a fastICA detectou uma anormalidade
    % significativa a 20 desvios

    anorm99(j)=max(kkk);

end

result(1,z)=event(z)*100;
result(2,z)=mean(anorm95)*100;
result(3,z)=mean(anorm99)*100;

end

xlswrite('result.xls', result);

```

### 3- Algoritmo de simulação de mistura de séries contínuas gaussianas

```

clear result;
clear anorm95;
clear anorm99;

n=83; % numero de series a serem simuladas

t=245; % tamanho de cada serie a ser simulada

sim=1000; % numero de vezes que a simulacao serah repetida

event=[0:0.001:0.009]; % variavel que dah o valor de cada anormalidade a
ser adicionada as series

for z=1:length(event)

    clear iid;
    clear iid_norm_serie;

    for i=1:n
        iid_norm_serie(:,i)=random('norm',media(i),desviopadiao(i),t,1);
    end

    iid=iid_norm_serie;

    for j=1:sim

        clear cheb95s;
        clear cheb95i;
        clear cheb99s;
        clear cheb99i;
        clear vvar;

        r_iidc = zeros(n,245); vvar=0;

        iidc = [iid [zeros(t-1,1); event(z)]]; % adiciona o vetor de erros
ao conjunto de dados como uma outra serie

        iidc=(eye(n+1)+[zeros(1,n+1); 0.5*eye(n) zeros(n,1)])*iidc'; % mis-
tura as series como no paper de Baragona e Battaglia (2007)

        %r_iidc = fastica(iidc(1:83,:), 'maxNumIterations', 10000, 'verbo-
se', 'off'); % roda o fastica retirando a ultima serie

        %r_iidc = fastica(iidc, 'maxNumIterations', 10000, 'verbose',
'off'); % roda o fastica sem retirar a ultima serie

        %r_iidc = fastica(iidc','numOfIC',1)

        % desigualdade de chebyshev

        kkk=0;

```

```

for i=1:min(size(r_iidc));

    cheb95s=mean(r_iidc(i,1:(t)))+15*std(r_iidc(i,1:(t)));
    cheb95i=mean(r_iidc(i,1:(t)))-15*std(r_iidc(i,1:(t)));

    if (r_iidc(i,t)>cheb95s || r_iidc(i,t)<cheb95i)
        kkk(i)=1;
    else
        kkk(i)=0;
    end

end

% conta o numero de vezes que a fastICA detectou uma anormalidade
% significativa a 15 desvios

anorm95(j)=max(kkk);

% 99%

kkk=0;

for i=1:min(size(r_iidc));

    cheb99s=mean(r_iidc(i,1:(t)))+20*std(r_iidc(i,1:(t)));
    cheb99i=mean(r_iidc(i,1:(t)))-20*std(r_iidc(i,1:(t)));

    if (r_iidc(i,t)>cheb99s || r_iidc(i,t)<cheb99i)
        kkk(i)=1;
    else
        kkk(i)=0;
    end

end

% conta o numero de vezes que a fastICA detectou uma anormalidade
% significativa a 20 desvios

anorm99(j)=max(kkk);

end

result(1,z)=event(z)*100;
result(2,z)=mean(anorm95)*100;
result(3,z)=mean(anorm99)*100;

end

xlswrite('result.xls', result);

```

#### 4- Algoritmo de simulação de séries contínuas gaussianas contaminadas com retornos anormais na última observação

```

clear result;
clear anorm95;
clear anorm99;

n=83; % numero de series a serem simuladas

t=245; % tamanho de cada serie a ser simulada

sim=1000; % numero de vezes que a simulacao serah repetida

event=[0:0.005:0.05]; % variavel que dah o valor de cada anormalidade a ser
adicionada as series

%event=.2;

for z=1:length(event)

    clear iid;
    clear iid_norm_serie;

    for i=1:n
        iid_norm_serie(:,i)=random('norm',media(i),desviopadrao(i),t,1);
    end

    iid=iid_norm_serie;

    for j=1:sim

        clear cheb95s;
        clear cheb95i;
        clear cheb99s;
        clear cheb99i;
        clear vvar;

        r_iidc = zeros(n,245); vvar=0;

        %iid=lognrnd(0,1,n,245);

        iidc = iid + [zeros((t-1),n); ((1+iid(t,1:n))*(1+event(z)))-1)-
iid(t,1:n)]; % contamina as series com a anormalidade

        % roda o fastica

        r_iidc = fastica(iidc', 'maxNumIterations', 10000, 'verbose',
'off');

        %r_iidc = fastica(iidc', 'numOfIC', 1)

        % desigualdade de chebyshev

```

```

kkk=0;

for i=1:min(size(r_iidc));

    cheb95s=mean(r_iidc(i,1:(t)))+15*std(r_iidc(i,1:(t)));
    cheb95i=mean(r_iidc(i,1:(t)))-15*std(r_iidc(i,1:(t)));

    if (r_iidc(i,t)>cheb95s || r_iidc(i,t)<cheb95i)
        kkk(i)=1;
    else
        kkk(i)=0;
    end

end

% conta o numero de vezes que a fastICA detectou uma anormalidade
% significativa a 15 desvios

anorm95(j)=max(kkk);

kkk=0;

for i=1:min(size(r_iidc));

    cheb99s=mean(r_iidc(i,1:(t)))+20*std(r_iidc(i,1:(t)));
    cheb99i=mean(r_iidc(i,1:(t)))-20*std(r_iidc(i,1:(t)));

    if (r_iidc(i,t)>cheb99s || r_iidc(i,t)<cheb99i)
        kkk(i)=1;
    else
        kkk(i)=0;
    end

end

% conta o numero de vezes que a fastICA detectou uma anormalidade
% significativa a 20 desvios

anorm99(j)=max(kkk);

end

result(1,z)=event(z)*100;
result(2,z)=mean(anorm95)*100;
result(3,z)=mean(anorm99)*100;

end

xlswrite('result.xls', result);

```



## **ANEXO B – CV Lattes resumido**

### **Alexandre Lerch Franco**

Doutor em Finanças pela UFRGS, mestre em administração de empresas pela UFRGS e engenheiro civil pela PUC/RS. Atualmente é professor de graduação na ESPM/RS e no Centro Universitário Metodista IPA, professor de pós-graduação convidado na Universidade de Passo Fundo e FSG e palestrante credenciado na corretora Solidus S/A CCVM. Como atividade profissional, é diretor executivo da S&F Engenharia de Avaliações e Consultoria Financeira S/S Ltda. Tem experiência na área de Finanças, atuando principalmente nos seguintes temas: finanças corporativas, mercado de capitais, avaliação de ativos, estatística inferencial, mercado imobiliário e métodos de pesquisa em finanças.

**Última atualização em 16/12/2008**

Endereço para acessar este CV:

<http://lattes.cnpq.br/2926973131384222>

### **Formação Acadêmica/Titulação**

**2004 - 2008** Doutorado em Administração.

Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil

Título: Aplicação da Análise de Componentes Independentes em Estudo de Eventos em Finanças, Ano de obtenção: 2008

Orientador: Gilberto de Oliveira Kloeckner, Ph.D.

**1998 - 2000** Mestrado Profissionalizante em Administração.

Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil

Título: Estratégia de Operações como Arma Competitiva em Serviços de Engenharia de Avaliações: Um Estudo de Caso, Ano de obtenção: 2001

Orientador: Prof. Jaime Evaldo Fensterseifer, Ph. D.

**1991 - 1997** Graduação em Engenharia Civil.

Pontifícia Universidade Católica do Rio Grande do Sul, PUC RS, Porto Alegre, Brasil

Título: Estudo de Viabilidade de Dois Empreendimentos Imobiliários: Uma Avaliação Comparativa de Rentabilidade

Orientador: Prof. Renato Solano

### **Formação complementar**

**1996 - 1996** Extensão universitária em Engenharia de Avaliações: Metodologia Científica.

Pontifícia Universidade Católica do Rio Grande do Sul, PUC RS, Porto Alegre, Brasil

**1996 - 1996** Curso de curta duração em Perícias Judiciais c/ Estatística Inferencial e Av.

Instituto Gaúcho de Engenharia Legal, IGEL, Brasil

**1997 - 1997** Curso de curta duração em Viabilidade Econômico-Financeira de Empreendimento.

Instituto Brasileiro de Avaliações e Perícias de Engenharia, IBAPE, Brasil

**1997 - 1997** Curso de curta duração em Análise da Viabilidade Econômica de Empreendimento.

Associação Brasileira de Engenheiros Mecânicos do RS, ABEMEC, Brasil

**1998 - 1998** Curso de curta duração em Engenharia Econômica.

Fundação de Pesquisa e Assessoramento à Indústria, FUPAI, Itajuba, Brasil

**1998 - 1998** Curso de curta duração em Introdução ao Mercado de Ações - Básico.

Associação Brasileira dos Analistas do Mercado de Capitais - Extremo Sul, ABAMEC-SUL, Brasil

**1999 - 1999** Curso de curta duração em Managing a Global Business.

HEC School of Management, HEC, França

**2001 - 2001** Curso de curta duração em Fundo de Comércio - Avaliação Empresarial.

Federação Brasileira das Associações de Peritos, Árbitros, Mediadores, FEBRAPAM, Brasil

**2003 - 2003** Curso de curta duração em Bolsa de Valores: da introdução ao avançado..

Treida, TREIDA, Brasil

### **Atuação profissional**

#### **1. Escola Superior de Propaganda e Marketing - ESPM**

##### **Vínculo institucional**

**2006 - Atual** Vínculo: Celetista , Enquadramento funcional: Professor Universitário , Carga horária: 12, Regime: Parcial

##### **Atividades**

**03/2006 - Atual** Graduação, Administração – Marketing e Relações Internacionais

*Disciplinas Ministradas:*

*Administração Financeira e Orçamentária III e IV , Avaliação de Operações Financeiras e Finanças de Mercado*

#### **2. Corretora Solidus S/A CCVM - SOLIDUS**

##### **Vínculo institucional**

**2005 - Atual** Vínculo: Colaborador , Enquadramento funcional: Palestrante , Carga horária: 2, Regime: Parcial

**Atividades****10/2005 - Atual** Treinamento*Especificação:**Módulo de Análise Técnica do Curso de Introdução ao Mercado de Capitais***3. S&F Engenharia de Avaliações e Consultoria Financeira S/S Ltda. - S&F****Vínculo institucional****1998 - Atual** Vínculo: Sócio-Gerente , Enquadramento funcional: Diretor Executivo , Carga horária: 20, Regime: Parcial**Atividades****11/1998 - Atual** Treinamento*Especificação:**Cursos e Treinamentos "in company" sobre Avaliação de Ativos***11/1998 - Atual** Direção e Administração*Cargos Ocupados:**Diretor Executivo***4. Faculdades Rio-Grandenses - FARGS****Vínculo institucional****2001 - 2006** Vínculo: Celetista , Enquadramento funcional: Coordenador de Curso de Graduação e Professor , Carga horária: 40,

Regime: Integral

**Atividades****10/2003 - 03/2006** Direção e Administração*Cargos Ocupados:**Coordenador de Curso de Graduação***08/2001 - 03/2006** Graduação, Administração de Empresas*Disciplinas Ministradas:**Estudo de Casos em Estratégia Empresarial e Análise e Avaliação de Investimentos***Revisor de periódico****1. Revista Economia e Gestão -****Vínculo****2007 - Atual** Regime: Parcial**2. REAd. Revista Eletrônica de Administração -****Vínculo****2004 - 2005** Regime: Parcial**Produção em C, T& A****Produção bibliográfica****Artigos completos publicados em periódicos****1. FRANCO, A. L.**

Armadilhas da Regressão Múltipla: erros e falhas que não estão nos livros. Think (Porto Alegre). , v.4, p.21 - 26, 2006.

**Trabalhos publicados em anais de eventos (completo)****1. FRANCO, A. L.**

Applying neural networks on intraday operations within the Brazilian Stock Options Market In: 2007 Annual BALAS Conference, 2007, San Jose

- Costa Rica.

**Anais de 2007 Annual BALAS CONFERENCE. , 2007.****2. FRANCO, A. L., GALLI, G.**

Método para Análise de Investimentos: alternativa para classificação de projetos com prazo e volume de recursos diferentes In: XXVII -

Encontro Nacional de Engenharia de Produção, 2007, Foz do Iguaçu.

**Anais do XXVII ENEGEP - Encontro Nacional de Engenharia de Produção. , 2007.****3. FRANCO, A. L.**

A contribuição da estatística nas avaliações imobiliárias em ambientes desconhecidos In: 1 Simpósio da Sociedade Brasileira de Engenharia

de Avaliações, 2006, Porto Alegre.

**Anais do 1 Simpósio da Sociedade Brasileira de Engenharia de Avaliações. , 2006.****4. FRANCO, A. L.**

A Estrutura de Capital e o Valor de Mercado das Empresas Brasileiras In: III CONVIBRA - Congresso Virtual Brasileiro de Administração, 2006,

Internet.

**Anais do III CONVIBRA. , 2006.****5. FRANCO, A. L.**

Inferência Estatística: uma arma perigosa In: 1 Simpósio da Sociedade Brasileira de Engenharia de Avaliações, 2006, Porto Alegre.

**Anais do 1 Simpósio da Sociedade Brasileira de Engenharia de Avaliações.** , 2006.

6. FRANCO, A. L.

Valor de um ativo: reflexos de um posicionamento In: 1 Simpósio da Sociedade Brasileira de Engenharia de Avaliações, 2006, Porto Alegre.

**Anais do 1 Simpósio da Sociedade Brasileira de Engenharia de Avaliações.** , 2006.

### **Artigos em jornal de notícias**

1. FRANCO, A. L.

Não Seja um Náufrago em um Mar de Incertezas: Artigo Sobre Metodologia de Avaliação de Marcas. Jornal do Comércio: Caderno Marcas de Quem Decide. Porto Alegre, 2003.

### **Artigos em revistas (Magazine)**

1. FRANCO, A. L.

A Palavra de Quem Já Fez: Aposta Certa. Revista Você S.A. e Exame. Brasil, 2000.

### **Orientações e Supervisões**

#### **Orientações e Supervisões concluídas**

#### **Trabalhos de conclusão de curso de graduação**

1. Diego da Rosa. **Análise do Comportamento do Governo Brasileiro na Aquisição de Medicamentos Excepcionais: Uma Contribuição**

**para Delinear uma Estratégia de Vendas de um Laboratório Multinacional.** 2005. Curso (Administração com Habilitação em Comércio Exterior) - Faculdades Rio-Grandenses

2. Thais Azolini. **Determinação dos Atributos Importantes em Uma Negociação no Varejo: Um Estudo de Caso.** 2005. Curso (Administração de Empresas) - Faculdades Rio-Grandenses

3. Diego Paiva. **Estudo de Viabilidade Econômico-Financeira para Implantação de uma Filial de uma Distribuidora de Materiais Elétricos**

**no Mercado Uruguaio.** 2005. Curso (Administração com Habilitação em Comércio Exterior) - Faculdades Rio-Grandenses

4. Vinícius Dorneles Valenti. **Proposta de Uma Metodologia para Análise de Concessão de Créditos pela Caixa Econômica Federal**

**Utilizando Inteligência Artificial.** 2005. Curso (Administração de Empresas) - Faculdades Rio-Grandenses

**Página gerada pelo Sistema Currículo Lattes em 16/12/2008 às 12:14:00.**