

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
MESTRADO EM ECONOMIA

DANILO SOARES MONTE-MOR

APLICAÇÃO DA TEORIA DE VALORES EXTREMOS EM ESTRATÉGIAS
“LONG-SHORT”.

VITÓRIA
ESPÍRITO SANTO - BRASIL
2010

DANILO SOARES MONTE-MOR

**APLICAÇÃO DA TEORIA DE VALORES EXTREMOS EM ESTRATÉGIAS
“LONG-SHORT”.**

Dissertação apresentada ao Programa de Mestrado em Economia do Centro de Ciências Jurídicas e Econômicas da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Economia.
Orientador: Prof. Dr. Rogério Arthmar.
Co-Orientador: Prof. Dr. Marco Aurélio S. Sanfins.

**VITÓRIA
ESPÍRITO SANTO - BRASIL
2010**

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
MESTRADO EM ECONOMIA

APLICAÇÃO DA TEORIA DE VALORES EXTREMOS EM ESTRATÉGIAS
“LONG-SHORT”

Prof. Dr. Rogério Arthmar
ORIENTADOR

Prof.Dr. Marco Aurélio S. Sanfins
EXAMINADOR EXTERNO

Prof. Dr. Gutemberg Hespanha Brasil
EXAMINADOR INTERNO

ΕΠÍΓΡΑΦΕ

*“All Models Are Wrong But
Some Are Useful”*

George E.P. Box

DEDICATÓRIA

Aos meus pais Luzia, que amo muito, e Izailton, o melhor pai que um filho pode ter, e a JESUS CRISTO, que me emprestou seu colo nas horas de aflição.

AGRADECIMENTOS

Manifesto minha gratidão, em primeiro lugar, ao meu co-orientador Dr^o. Marco Aurélio dos Santos Sanfins que foi, sem dúvida, peça importante na elaboração deste trabalho. Agradeço, principalmente, pelo tempo a mim dispensado.

Agradeço ao professor e orientador Dr^o. Rogério Arthmar pelo seu enorme interesse, sempre indicando a direção a ser tomada nos momentos de maior dificuldade que, certamente, foram fundamentais para o desenvolvimento de determinados pontos.

Em especial, aos amigos Mateus, Dante, Rodrigo, César, Livia e Otávio, por toda a dedicação, amizade e atenção que me foram prestados. Ao meu irmão Fernando, que se mostrou sempre incansável e prestativo durante este trabalho e durante toda a minha vida acadêmica. À minha namorada Carol, pilar psicológico, emocional e afetivo dessa caminhada. E aos meus pais, sem os quais não conseguiria.

Gostaria de agradecer a CAPES pelo financiamento concedido, que tornou viável a realização deste trabalho, assim como a execução dos trabalhos de outros colegas do mestrado. Por fim, agradeço, também, a maneira cordial com que fui recebido nas cidades de Vitória e Niterói, principalmente no departamento de Estatística da Universidade Federal Fluminense, o que tornou a residência e a pesquisa nessas localidades uma agradável experiência.

RESUMO

Cada vez mais tem surgido no mercado de investimento fundos de retorno absoluto (*Hedge Funds*) que têm como objetivo principal melhorar seus desempenhos através de estratégias de arbitragem, como é o caso das estratégias *long-short*. É o comportamento desproporcional e até mesmo antagônico dos preços dos ativos que permite aos players estruturar estratégias para gerar retornos adicionais, superiores aos custos de oportunidade e independentes ao movimento do mercado. Neste trabalho foi utilizada a Teoria de Valores Extremos (TVE), um importante ramo da probabilidade, para que fossem modeladas as séries da relação direta entre preços de dois pares de ativos. Os quantis obtidos a partir de tal modelagem, juntamente com os quantis fornecidos pela normal, foram superpostos aos dados para períodos subsequentes ao período analisado. A partir da comparação desses dados foi criada uma nova estratégia quantitativa *long-short* de arbitragem, a qual denominamos *GEV Long-Short Strategy*.

Palavras-chave: Arbitragem, *Hedge Fund*, Estratégias *Long-Short*, Teoria dos Valores Extremos, *Pair Trading*.

ABSTRACT

Increasingly has appeared on the market of investment Absolute Return Funds (Hedge Funds), which have the main objective to improve their performance through arbitrage strategies, as long-short strategies. It is the disproportionate evolution and even antagonistic of active prices that allows the players to structure strategies to generate additional returns, higher than the opportunity costs and independent of the movement of the market. In this work we used Extreme Value Theory (EVT), an important segment of probability, to model the series of direct relationship between prices of two pairs of assets. The quantiles obtained from such modeling and the quantile provided by normal were superimposed on data for periods subsequent to the period analyzed. From the comparison of such data we created a new quantitative long-short arbitrage strategy, called GEV Long-Short Strategy.

Key-words: Arbitragy, Hedge Fund, Long-Short Strategy, Extreme Value Theory, Pair Trading.

Sumário

1	INTRODUÇÃO	1
2	FINANÇAS	4
2.1	Taxonomia	4
2.2	Finanças Tradicionais	5
2.3	Finanças Modernas	9
2.3.1	Teoria da Diversificação de Ativos em Carteiras	10
2.3.2	Modelo de Determinação dos Preços dos Ativos	14
2.3.3	Fama(1970) e a Eficiência dos Mercados	17
2.4	Finanças Comportamentais	19
2.5	Crítica às Hipóteses Simplificadoras	21
3	<i>HEDGE FUNDS E AS ESTRATÉGIAS LONG-SHORT</i>	24
3.1	<i>Hedge Fund</i>	24
3.1.1	<i>Hedge Funds</i> Norte-Americanos	26
3.1.2	<i>Hedge Funds</i> brasileiros	29
3.1.3	Diferenças entre Brasil e EUA	30
3.2	Estratégias <i>Long-Short</i>	31
3.2.1	Abordagens Long-Short	32
3.2.2	Fundos Neutros e Fundos Long-Short	34
3.2.3	Estratégia Pair Trading: Arbitragem Estatística	36
3.2.4	Estratégia Neutra ao Mercado <i>Long-Short</i> com Portfólios <i>Índex Tracking</i>	39
4	TEORIA DOS VALORES EXTREMOS (TVE)	42
4.1	Modelagem Univariada dos Máximos	42
	43	
4.1.2	Distribuições Limite para Máximos	45
4.1.3	Max-Estabilidade das Distribuições Extremas	48
4.1.4	Domínio de Atração da Distribuição do Máximo	50
4.1.5	Distribuição Generalizada de Valor Extremo	50
4.2	Estimação	53
4.2.1	Estimadores de Máxima Verossimilhança	53

4.2.2	Estimadores de L-Momentos	54
4.3	Estatísticas de Teste	55
5	<i>GEV Long-Short Strategy: uma nova modalidade quantitativa</i>	58
5.1	Exemplos de Pair Tradings	58
5.2	Proposta de Arbitragem	60
5.3	Modelagem	65
5.3.1	<i>Pair Trading</i> PETR4/VALE5	65
5.3.2	<i>Pair Trading</i> VIVO4/TCSL4	70
5.4	Impactos do quantil <i>GEV Long-Short</i> : expectativas e comportamento	77
6	Conclusão	80
7	APÊNDICE	83
A	Critérios de Convergência	83
B	Distribuições Max-estáveis	85
B6	Máximo Domínio de Atração da Weibull	87
E	Máximo Domínio de Atração da Gumbel	88
F	Convergência do Máximo de Sequências Estacionárias	90
G	Índice Extremal	93

Lista de Figuras

1	Figura adaptada de Anache (2008).	14
2	Figura gerada pelo programa estatístico R referente aos gráfico das distribuições extremas com parâmetro de forma iguais a 0, 1 e -1, respectivamente.	47
3	Figura gerada pelo programa estatístico R referente aos gráfico das fdp das distribuições extremas com parâmetro de forma iguais a 0, 1 e -1, respectivamente.	49
4	Figura gerada pelo programa estatístico R referente a comparação dos gráficos das fdp das distribuições $F(x) = (1 - \exp(-x))^{40}$ e Gumbel para $\mu = 3, 7$ e $\sigma = 1$	51
5	Figura gerada pelo programa estatístico R referente a comparação dos gráficos das fdp das distribuições $F(x) = \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \right]^{40}$ e Gumbel para $\mu = 1, 9$ e $\sigma = 0, 4$	51
6	Movimentação dos preços da PETR4 e VALE5 durante as 76 semanas do período de 28/12/2007 a 05/06/2009.	58
7	Série histórica da VIVO4.SA para o período compreendido entre 01/06/2006 e 15/12/2007.	59
8	Série histórica da TCSL4.SA para o período compreendido entre 01/06/2006 e 15/12/2007.	59
9	Razão histórica de preços $\frac{PETR4.SA}{VALE5.SA}$ para o período compreendido entre 01/01/2005 e 02/07/2007.	65
10	QQ-plot dos resíduos para a modelagem dos máximos.	66
11	QQ-plot dos resíduos para a modelagem dos mínimos.	66
12	Densidades empírica versus GEV ajustada para a modelagem dos máximos para n=15.	67
13	Densidades empírica versus GEV ajustada para a modelagem dos mínimos para n=22.	67
14	Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos máximos.	68
15	Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos mínimos.	68

16	Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos máximos.	68
17	Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos mínimos.	68
18	Performance da razão $k = \frac{PETRA.SA}{VALE5.SA}$ para os 131 dias subsequentes ao período compreendido entre os 01/01/2005 e 02/07/2007.	69
19	Rentabilidade acumulada das posições a partir do quantil da normal.	70
20	Rentabilidade acumulada a partir da inversão das posições.	70
21	Razão histórica de preços $\frac{VIVO4.SA}{TCSL4.SA}$ para o período compreendido entre 01/06/2006 e 15/12/2007.	70
22	QQ-plot dos resíduos para a modelagem dos máximos.	71
23	QQ-plot dos resíduos para a modelagem dos mínimos.	71
24	Densidades empírica versus GEV ajustada para a modelagem dos máximos para n=15.	72
25	Densidades empírica versus GEV ajustada para a modelagem dos mínimos para n=22.	72
26	Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos máximos.	73
27	Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos mínimos.	73
28	Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos máximos.	73
29	Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos mínimos.	73
30	Performance da razão $k = \frac{VIVO4.SA}{TCSL4.SA}$ para os 54 dias subsequentes ao período compreendido entre 01/06/2006 a 15/12/2007.	74
31	Rentabilidade acumulada das posições a partir do quantil da normal.	75
32	Rentabilidade acumulada a partir da inversão das posições.	75

Lista de Tabelas

- 1 Tabela de valores críticos para as estatísticas teste D^+ , D^- , D e V a níveis de significância de 1% e 5%. 56

1 INTRODUÇÃO

Cada vez mais tem surgido no mercado de investimento fundos exclusivos que têm como objetivo principal melhorar seus desempenhos através de estratégias de arbitragem. No Brasil esses fundos são classificados de acordo com a Associação Nacional de Bancos de Investimentos (Anbid), dentre os quais se destacam os *Hedge Funds*, também conhecidos como fundos multimercados ou fundos de retorno absoluto.

Carol J. Loomis foi quem usou publicamente pela primeira vez o termo *Hedge Fund* em seu artigo de 1966 na revista Fortune. Ao introduzir esse termo, Loomis referia-se às estratégias de investimento adotadas por Alfred W. Jones baseadas na compra de ações sub-avaliadas e na venda a descoberto de ações sobre-avaliadas com o intuito de proteger-se dos riscos de mercado. É exatamente essa busca por proteção que inspirou o nome *Hedge* (proteção) desses fundos.

Com a crise da década de 70, entretanto, alguns *Hedge Funds* foram extintos e apenas na metade da década de 90 é que esses fundos foram reintroduzidos nos fundos de interesse dos investidores. Essa retomada atraiu a atenção de pesquisadores pelo assunto, principalmente no que diz respeito à mensuração das reais possibilidades de retornos superiores. Schneeweis et al (2000) encontra evidências de que portfólios de fundos americanos com *Hedge Funds* em sua composição apresentam desempenho superior em termos de retorno esperado, volatilidade e índice de Sharpe, se comparado a portfólios com apenas fundos tradicionais de renda fixa e ações. Outros artigos, como o de Liang (1999), encontram evidências que *Hedge Funds* conseguem prover retornos positivos, com volatilidade inferior aos índices de ações e com uma correlação baixa com esses mesmos índices.

No Brasil, a partir de 2002, a indústria de *Hedge Funds* tem apresentado um crescimento contínuo expressivo, principalmente em função da não ruptura do modelo econômico iniciado na década de 90 e das constantes quedas nos juros reais a partir da estabilidade e do equilíbrio macroeconômico. Segundo dados da Anbid, o segmento dos Hedge Funds foi um dos que apresentou maior crescimento no período de Dezembro de 1997 a Março de 2008, com um aumento de quase doze vezes o valor da movimentação do setor para período considerado. No período de Setembro de 2008 a Dezembro de 2009, os *Hedge Funds* foram uma das opções dos investidores para proteção contra a crise financeira mundial precipitada, dentre outras razões, pela falência do tradicional banco de investimento estadunidense Lehman Brothers, fundado em 1850.

Diante de expressivo crescimento, o *Hedge Fund* tornou-se um termo que incorpora inclusive fundos não tradicionais como, por exemplo, fundos que exploram ineficiências de mercado através de estratégias sofisticadas de análises e com apostas não necessariamente direcionadas, como é o caso dos fundos que utilizam a estratégia *long-short*.

De uma forma geral, uma estratégia *long-short* se baseia na manutenção de um portfólio *long* financiado por um portfólio *short* que, além de servir como uma proteção contra riscos de declínios dos mercados, fornece ao gestor a possibilidade de obter retornos superiores aos retornos globais pelo uso de critérios de seleção de ações. Impulsionado por esse objetivo, pesquisas têm sido feitas com o intuito de ajustar modelos que propiciem a escolha dos ativos bem como o *timing* das operações.

Por muitos anos, entretanto, os gestores avaliaram os co-movimentos e interdependência entre os mercados e construíram seus portfólios a partir de análises como as que envolvem o conceito linear de correlação para modelar as interdependências entre ativos financeiros. Nos últimos anos progressos foram feitos a partir de outras técnicas como a de cointegração para modelagens de séries de tempo (Lucas(1997) e Alexander(1999)) e utilização da distribuição normal para modelagem probabilística de séries de pares de ativos.

Observe entretanto que mesmo pares de ativos não cointegrados podem oferecer no curto prazo oportunidades de arbitragem e que a distribuição normal subestima os quantis extremos de séries financeiras, uma vez que a modelagem dessas séries requer análises que envolvam distribuições com caudas pesadas. Neste contexto, torna-se necessário o desenvolvimento de uma estratégia quantitativa *long-short* focada nos valores extremos para estimarmos a probabilidade de ocorrência de eventos fora dos padrões normais da relação entre os preços de pares de determinados ativos, de modo a fornecer ao mercado novos indicadores que auxiliem os *players* em suas decisões de arbitragem. É nessa perspectiva que a Teoria dos Valores Extremos (TVE) desempenha um papel fundamental, dada a grande capacidade obtida através da modelagem de dados financeiros.

O principal objetivo desta dissertação é utilizar a Teoria dos Valores Extremos para se estabelecer uma nova modalidade quantitativa *long-short*, de modo que os fundos tenham a capacidade de gerar ganhos positivos com redução dos riscos de investimento.

Os objetivos específicos visam, por um lado, abordar a partir de um contexto com periodicidade sequencial alguns dos principais conceitos teóricos e empíricos que mais contribuíram para a evolução da teoria de finanças, na tentativa de apresentar conceitos que enfatizem a necessidade da construção de modelos que diminuam o risco e aumentem os

retornos dos fundos de investimento diante das anomalias do mercado. Para tanto serão apresentados, inclusive, modelos tradicionais como o Modelo de diversificação de Ativos em Carteiras, elaborada por Markowitz (1952), o CAPM (*Capital Asset Pricing Model*), proposto por Sharpe (1964) e abordagens como a de Fama (1970).

Por outro lado, os objetivos específicos visam apontar em um primeiro momento as características dos Hedge Funds e, em um segundo, apontar a necessidade da criação de um modelo probabilístico *ad hoc*¹ (baseado na análise de extremos) para as estratégias *long-short*, a partir da exposição de algumas das principais técnicas de *long-short* existentes e de suas ineficiências reconhecidas.

A partir da verificação de tal necessidade, será apresentado uma revisão sobre a Teoria dos Valores Extremos² na tentativa de fomentar teoricamente as simulações da nova modalidade quantitativa *long-short* que será proposta. Os pares de ativos $PETRA \times VALE5$ e $VIVO4 \times TCSL4$ serão utilizados para a aplicação da nova modalidade quantitativa proposta. Todas as ações utilizadas são preferenciais³.

Observe que o primeiro caso é um exemplo de arbitragem entre empresas de setores diferentes. Esse par foi escolhido dada a grande representatividade que ambas as ações possuem no total de negociações do mercado brasileiro. O segundo par se refere à arbitragem entre ativos de um mesmo setor. A escolha desse par foi feita devido ao importante crescimento apresentado pelo setor de comunicações no Brasil.

¹A expressão latina *ad hoc* significa literalmente para isto e em um sentido amplo pode ser traduzida como específico ou especificamente. Por exemplo, um instrumento *ad hoc* é uma ferramenta elaborada especificamente para uma determinada aplicação.

²Para continuação dos estudos da Teoria dos Valores Extremos recomenda-se a tese de doutorado intitulada “Cópulas para distribuições generalizadas de valores extremos multidimensionais (Sanfins (2009)).

³Uma ação preferencial, também denotada por PN, oferece preferência na distribuição de resultados ou no reembolso do capital em caso de liquidação da referida companhia. O detentor dessas ações não possui, entretanto, poder de voto nas assembleias da companhia

2 FINANÇAS

2.1 Taxonomia

Embora não exista um consenso na literatura clássica com relação à origem do termo finanças, alguns autores consideram que sua etimologia provém do latim medieval a partir da palavra *financia*, que na época indicava os diferentes meios necessários para a realização de despesas públicas pelo Estado, e outros a partir do verbo *finire*, do qual surgiram posteriormente o verbo *finare* e o termo *finatio*. Na concepção atual, Haugen (1999a) elaborou uma taxonomia referente ao contexto abordado, na qual as finanças ficam divididas em três fases a partir das diferentes óticas de análise do comportamento do mercado financeiro, a saber:

1. finanças antigas ou tradicionais, tratadas no período de 1930 a 1960, na qual os estudos baseavam-se na análise de demonstrações financeiras e na natureza dos títulos de crédito, tendo como base de referência a Contabilidade e o Direito;
2. finanças modernas, tratadas no período de 1960 a 1990, na qual era avaliado o comportamento dos agentes a partir do princípio de racionalidade, com destaque para os paradigmas de otimização e de irrelevância, para os modelos de precificação e para os pressupostos de eficiência dos mercados;
3. novas finanças, tratadas a partir do final da década de 90, na qual a abordagem é baseada a partir de modelos comportamentais interpretados com auxílio da Estatística, Econometria, Psicologia e com modelagem *ad hoc* sob o ponto de vista de ineficiência dos mercados.

Neste capítulo será feito uma abordagem a partir de um contexto com periodicidade sequencial de alguns dos principais conceitos teóricos e empíricos que mais contribuíram para a evolução da teoria de finanças. Baseado nessa abordagem, serão apresentados fatores que enfatizem a necessidade de construção de ferramentas mais realistas diante das anomalias do mercado. Tais fatores estarão fundamentados:

- na Teoria de Diversificação de Ativos em Carteiras, elaborada por Markowitz (1952), na qual os fatores risco e retorno devem ser conciliados para análises de investimentos em ativos, de modo que a composição das carteiras forneça o menor risco possível para um dado nível de retorno esperado. A análise desses fatores é que permitiu Markowitz

definir fronteira eficiente e verificar que um investidor deve preocupar-se não apenas com o risco individual de um determinado ativo, mas com a sua contribuição ao risco total da carteira ao qual o ativo pertence;

- na contestação de modelos simplificados e que até hoje são utilizados para análise de dados financeiros, como é o caso do CAPM (Capital Asset Pricing Model), proposto por Sharpe (1964), cujas premissas não se verificam, uma vez que o mercado não apresenta competição perfeita e existem fricções que impedem determinados investimentos;
- na visão de Fama (1970) que, de um ponto de vista que melhor refina a hipótese de eficiência de mercado, reconhece a existência conjunta de investidores e de árbitros plenamente racionais, de forma que os mercados por vezes podem se comportar com alguma ineficiência;

2.2 Finanças Tradicionais

A Teoria das Finanças Tradicionais está fundamentada basicamente na busca de rentabilidade superior aos ganhos médios de mercado ao considerar que o mercado pode muitas vezes se comportar de maneira irracional ou ineficiente, ou seja, quando a maior parte dos investidores pagam pelos ativos mais do que eles valem (ativos sobre-avaliados) ou, do contrário, não pagam o real valor dos ativos (ativos sub-avaliados) definidos por seu fluxo de caixa descontado (MACEDO JR., 2003).

A base da análise tradicional consiste na formação de carteiras de investimento capazes de auferir remuneração maior que a média dos índices de mercado. As estratégias de investimento desse tipo de análise visam a antecipação de eventos futuros para que sejam incorporados ou eliminados *ex-ante* ativos que por ventura sejam identificados como sub-avaliados ou sobre-avaliados, respectivamente.

Existem dois tipos de investidores tradicionais, também denotados por investidores ativos: os observadores de mercado (*market timers*) e os selecionadores de títulos (*stock picking*).

Os observadores do mercado são investidores que se baseiam na análise das séries históricas de preços para prever a direção dos mercados de ações e derivativos. Esse tipo de análise, também conhecida como análise técnica, pressupõe no geral a repetição consistente dos padrões passados de comportamento do mercado e de investidores. Esse pressuposto faz

com que os resultados gráficos permitam aos *market timers* projetar e predizer o comportamento dos preços e de suas variações futuras.

Os selecionadores de títulos, por sua vez, são analistas que se baseiam nas características das empresas e dos setores aos quais elas pertencem a partir da comparação de resultados, demonstrativos contábeis, linhas de produtos e serviços, planos de investimento, experiência e competência dos administradores, liquidez e capitalização de um ativo, dentre outros, para verificar quais ativos encontram-se fora dos padrões coerentes de precificação (DAMODARAN, 1999). Esse tipo de análise, também conhecida como análise fundamentalista, procura valer-se de toda a informação disponível no mercado sobre determinada empresa para obter o preço justo da companhia e assim predizer acerca do comportamento dos preços de suas ações.

Observe que a análise técnica, diferentemente da análise fundamentalista, não considera aspectos internos das empresas e dos setores estudados, mas sim o comportamento dos preços de ativos e a participação dos investidores no mercado.

Os primeiros modelos da análise técnica foram elaborados por volta de 1850 pelo jornalista Charles Henry Dow (1851-1902), que juntamente com Edward Jones e Charles Bergstresser fundou a *Dow Jones & Company*, editora do *The Wall Street Journal*.

Em 1896 Dow criou o índice *Dow Jones Industrial Average*, que representava o desempenho médio das 12 empresas que ele considerava serem as mais relevantes do mercado americano. Esse índice foi criado com o intuito de monitorar os movimentos das ações negociadas na Bolsa de Nova York (NYSE) e a partir de 1928 passou a ser composto pelo desempenho das 30 maiores empresas. Além desse índice, Dow também criou um índice médio para o setor de ferrovias, que mais tarde passou a ser conhecido por *Dow Jones Transports Average*.

Com a fundação do *The Wall Street Journal*, Dow passou a publicar periodicamente um informativo sobre o comportamento do mercado financeiro e sobre as formas de entender e analisar o comportamento dos preços dos ativos. Após sua morte, em 1902, Willian Peter Hamilton o sucedeu como editor e compilou essa série de informativos no livro *The Stock Market Barometer*. Dow afirmava que todas as informações relevantes sobre uma empresa já estavam incorporadas e eram refletidas pelo preço das ações e fazia suas análises a partir de alguns princípios básicos, a saber:

1. os índices representam a performance geral do mercado, uma vez que incorporam as decisões conjunta dos investidores, desde os mais bem informados e com melhor

capacidade de gestão das informações até os mais inexperientes. Dessa forma, todo fator que afeta o mercado está refletido (descontado) nos índices e os acontecimentos inesperados podem ser rapidamente avaliados e absorvidos pelos investidores;

2. o mercado possui tendências de movimento, a saber tendências primária, secundária e terciária. As tendências primárias são movimentações mais gerais do mercado e indicam fases de valorização ou desvalorização no médio e longo prazo. Durante esses períodos coexistem as tendências secundárias, nas quais correções contrárias ao movimento de valorização ou desvalorização afetam os preços, embora não alterem as trajetórias primárias do movimento. As tendências secundárias permanecem por períodos menores de tempo se comparadas às primárias. Um raciocínio análogo pode ser tomado para se caracterizar as tendências terciárias, ou seja, correções de curto prazo contrárias às trajetórias secundárias e que não alteram as tendências secundárias do movimento;
3. tanto os movimentos de alta como os de baixa dos preços apresentam aspectos psicológicos marcantes e inerentes a cada uma dessas fases. O início do processo de valorização é geralmente acompanhado por investidores mais qualificados e mais bem informados, uma vez que são os primeiros a perceberem o potencial de retorno desses ativos. Enquanto a euforia da valorização não se instala, previsões pessimistas ainda dominam os investidores, o que permite aos analistas de elite comprarem papéis subavaliados. Após esse período, ocorre uma aceleração mais acentuada do movimento de valorização com o aumento da compra dos ativos. A fase subsequente é marcada por grandes altas, dada a manutenção da segurança de lucratividade desses ativos. É nesse período que os investidores mais bem preparados começam a vender suas posições e que se instala a possibilidade de queda dos preços. O processo de desvalorização ocorre sob uma ótica inversa e mesmo nesses períodos há a possibilidade de rentabilidade superior;
4. a confirmação das tendências primárias deve ser feita mediante o acompanhamento dos principais índices financeiros, uma vez que um índice só efetivamente rompe o nível de resistência ou suporte ⁴ do mercado caso o mesmo ocorra para no mínimo um

⁴Suportes e resistências, de uma maneira mais simplificada, são limites de zonas de preços nos quais a trajetória atual do mercado tem grande probabilidade de se reverter. Os suportes indicam os limites nos quais o interesse de compra supera a pressão vendedora, ou seja, nos quais a trajetória de desvalorização tende a mudar. As resistências indicam a região na qual o interesse de vender supera a pressão compradora,

outro índice de composição distinta, de modo que a posição de um índice confirme a tendência do outro;

5. o volume de negociações se relaciona de maneira direta com as tendências de alta e de maneira indireta com as tendências de baixa dos ativos financeiros, ou seja, em períodos em que existe a tendência primária de alta dos preços dos ativos o volume de compra aumenta nos períodos de valorização e diminui durante a desvalorização e em períodos em que existe a tendência primária de queda dos preços dos ativos o volume de compra aumenta durante a desvalorização e diminui durante a valorização;

Na década de 30 os princípios introduzidos por Dow passaram a ser conhecidos por Teoria de Dow a partir da publicação do livro *The Dow Theory* escrito por Robert Rhea. Essa teoria basicamente procurava identificar mudanças nas principais tendências de médio e longo prazo do mercado, ou seja, preocupava-se com a trajetória de valorização ou desvalorização das cotações das ações.

Dow acreditava que a configuração gráfica da série dos preços dos ativos apresentava suas possíveis trajetórias e tendências futuras, como explicam Rudge e Cavalcante (1998):

A Teoria Dow é hoje apenas a primeira semente de um conjunto de estudos de gráficos de preços e volumes, que evoluiu e se sofisticou enormemente, e que medem as forças de oferta e procura nos mercados de ações, para determinar se o investidor deve: comprar, vender, não comprar ou não vender (Rudge e Cavalcante (1998), p. 267-268).

Após o grande crash da bolsa em 1929, entretanto, alguns autores como Graham e Dodd (1934) começaram a analisar em conjunto a situação financeira, econômica e mercadológica de uma empresa, com a finalidade de formular a partir desses indicadores expectativas e projeções mais precisas para o futuro.

De uma forma mais ampla, o objetivo principal dessa metodologia de análise, que hoje é conhecida como análise fundamentalista, é avaliar as características da empresa e do referido setor, bem como o desempenho macroeconômico geral da economia, com o objetivo de determinar o valor intrínseco de cada ação.

O valor intrínseco se refere à avaliação da cotação total do patrimônio de uma empresa e é calculado a partir do desempenho e posição assumida no respectivo setor de atuação. Para ou seja, na qual a trajetória de valorização tende a se reverter.

tanto normalmente leva-se em consideração variáveis como a intensidade da concorrência, os tipos de produtos e serviços alternativos, o grau de atualização tecnológica do empreendimento, o nível de intervenção estatal (incentivos fiscais, subsídios, etc), programas de investimento, políticas de distribuição de dividendos e cálculo a valor presente dos ativos e passivos futuros estimados.

Observe que a partir desse tipo de análise o valor justo para as ações de uma empresa passa a ser definido pela capacidade da empresa de auferir lucros no futuro, dada a posição e a perspectiva de crescimento que ela possui, de modo que as expectativas de retorno de suas ações ficam condicionadas à sub ou sobre avaliação dos seus ativos por parte do mercado financeiro, ou seja, à precificação do valor teórico da firma.

Seja sob uma análise técnica ou fundamentalista, as ineficiências no processo de valoração dos ativos é que justificam os esforços dos investidores que gerenciam uma carteira baseados nos preceitos da teoria tradicional, no sentido de tornar passível de previsão o comportamento do mercado e projetar resultados futuros.

2.3 Finanças Modernas

As finanças modernas surgiram entre o final da década de 1950 e o início da década de 1960 na medida em que foi alterada a forma de avaliação com relação à decisão de investimento. A escolha, anteriormente baseada em elementos puramente institucionais, passa a se relacionar ao comportamento do investidor, que começa a ser visto como um ser econômico e racional e que contempla a utilização da racionalidade otimizadora⁵ neoclássica.

Conforme Mullainathan e Thaler (2000), todas as teorias que sustentam as Finanças Modernas têm como base a teoria microeconômica neoclássica na medida em que dão ênfase ao pressuposto de que os agentes são capazes de decidir a partir da racionalidade ilimitada⁶ (*unbounded rationality*) e na medida em que fazem uso da maximização da utilidade e do acúmulo de aprendizado. De uma forma mais ampla, é possível considerar que nas Finanças Modernas os investidores agem de forma racional no mercado e o fazem de forma avessa ao risco, no sentido de buscar a maximização da utilidade esperada (a partir da curva de utilidade). Note que essa maximização é fundamentada a partir do momento que se considera

⁵A racionalidade otimizadora ou mecanicista neoclássica recebeu diferentes adjetivos, dentre os quais podemos citar a racionalidade instrumental e a racionalidade substantiva ou ilimitada. Ver Fernandes (2000).

⁶Decisão racional, segundo a moderna teoria das finanças, é aquela em que o agente age com o objetivo de maximizar a sua função utilidade, dado um grupo de restrições inerentes à sua situação problema.

informação perfeita e expectativas racionais como hipóteses verificáveis. Nesse sentido, os mercados passam a ser entendidos como altamente eficientes.

A utilização dessas hipóteses fez com que surgisse dentro dos departamentos de economia diversos trabalhos que se propuseram a explicar os fenômenos econômicos. Segundo Haugen (2000), as teorias de Markowitz (1952), Modigliani e Miller (1958, 1961), Sharpe (1964), Lintner (1965), Mossin (1966), Fama (1970) e a Teoria de Precificação de Opções de Black (1973), Scholes (1973) e Merton (1973) foram os principais trabalhos que ostentaram esse avanço teórico e que conjuntamente formam o marco da ruptura entre as Finanças Antigas e as Finanças Modernas.

A Teoria da Diversificação de Ativos em Carteiras, elaborada por Markowitz (1952), o Modelo de Determinação dos Preços dos Ativos (Capital Asset Pricing Model-CAPM), proposto por Sharpe (1964) e algumas características da visão de Fama (1970) serão abordadas na sequência, dada a maior representatividade conferida a essas teorias no sentido de se compreender a importância dos pressupostos das Finanças Modernas.

2.3.1 Teoria da Diversificação de Ativos em Carteiras

O americano Harry Markowitz iniciou seus estudos em finanças na Universidade de Chicago ao se matricular na pós-graduação de Economia, onde incorporou as idéias de racionalidade dos agentes econômicos. Sua área de atuação, entretanto, era inicialmente voltada para a programação linear e até aquele momento Markowitz não demonstrava interesse em relação ao mercado de ações. A solicitação de um corretor para que Markowitz utilizasse programação linear em problemas de investimento acionário fez com suas pesquisas se voltassem para essa área.

Durante o doutorado, Markowitz conheceu a obra *The Theory of Investment Value*, de John Burr Williams, e despertou certa curiosidade na relação entre o risco e o retorno inerentes a certos investimentos. Com a finalidade de explicar os motivos pelos quais investidores por vezes abriam mão de rentabilidades superiores, Markowitz, em seu artigo de 1952, intitulado *Portfolio Selection* e apresentado no jornal *The Journal of Finance*, construiu um raciocínio no qual o investidor racional toma suas decisões considerando não apenas o retorno do investimento, mas também o risco inerente. Esse artigo, por abordar exatamente essa relação, é hoje considerado um marco na história da moderna teoria de finanças.

Nesse artigo, Markowitz sumariza a relação fundamental entre a variação do retorno esperado de um ativo financeiro e o respectivo impacto no risco de uma carteira e o faz

ao considerar que os fatores risco e retorno devem ser analisados conjuntamente e que as respectivas carteiras devem ser formadas no sentido de fornecer o menor nível de risco para um determinado nível de retorno esperado. Seu raciocínio foi baseado em um modelo de programação quadrática capaz de maximizar os retornos e minimizar a variância para um dado nível de retorno.

Em termos de definição, Markowitz considerou que o retorno esperado de um investimento pode ser obtido através da média dos retornos esperados dos ativos que compõe a respectiva carteira e ponderados pelas respectivas probabilidades de ocorrência. O risco, embora não tenha sido mencionado de forma explícita, pode ser obtido através da dispersão dos respectivos retornos em relação ao seus retornos esperados.

De forma geral, as idéias propostas por Markowitz incorporam a noção de que um investidor deve formar uma carteira de portfólios com base nos respectivos retornos esperados e na variância desses retornos, de modo a maximizar o retorno esperado e minimizar a variância. A otimização de tais objetivos seria feita, segundo Markowitz, por meio da diversificação da carteira a partir de ativos com reduzido nível de variância e covariância.

O início do artigo *Portifólio Selection* sumariza essas idéias:

O processo de seleção de um portfólio pode ser dividido em dois estágios. O primeiro estágio começa com observação e experiência e termina com crenças sobre o desempenho futuro de ações disponíveis. O segundo estágio começa com crenças relevantes sobre desempenho futuro e termina com a escolha do portfólio. Este artigo está preocupado com o segundo estágio. Nós primeiro consideramos a regra de que os investidores maximizam ou deveriam maximizar o valor esperado, ou retornos antecipados. E a regra de que o investidor considera ou deveria considerar os retornos esperados como uma coisa desejável e a variância dos retornos como algo indesejado. Esta regra possui muitos pontos de vista, ambos como uma máxima para uma hipótese sobre o comportamento do investidor. Nós ilustramos geometricamente a relação entre crenças e escolha de um portfólio segundo a regra dos retornos esperados e a variância dos retornos (Markowitz(1952)).

Segundo Bernstein (1997), os fundamentos das idéias de Markowitz formam uma síntese das principais idéias de autores bastante conhecidos na época, como Pascal, de Moivre, Bayes,

Laplace, Gauss, Galton, Daniel Bernoulli, Jevons e Von Neumann e Morgenstern. De acordo com Sharpe, Alexander e Bayley (1998), esses fundamentos podem ser sintetizados em 7 princípios pilares, que seguem:

1. Os investidores avaliam portfólios apenas com base no valor esperado e na variância (ou desvio padrão) das taxas de retorno sobre o horizonte de um período.
2. Os investidores nunca estão satisfeitos. Quando postos a escolher entre dois portfólios de mesmo risco, sempre escolherão o de maior retorno.
3. Os investidores são avessos ao risco. Quando postos a escolher entre dois portfólios de mesmo retorno, sempre escolherão o de menor risco.
4. Os ativos individuais são infinitamente divisíveis, significando que um investidor pode comprar a fração da ação, se assim desejar.
5. Existe uma taxa livre de risco, na qual um investidor pode tanto emprestar quanto tomar emprestado.
6. Os custos de transação e impostos são irrelevantes.
7. Os investidores estão de acordo quanto à distribuição de probabilidades das taxas de retorno dos ativos, o que assegura a existência de um único conjunto de carteiras eficientes.

Note que a teoria do portfólio de Markowitz postula que todos os investimentos podem ser analisados de acordo com a relação risco-retorno e que o risco pode ser compensado pela diversificação. Essas considerações eram contrárias ao pensamento até então vigente de concentração do investimento em um ativo específico e com maior retorno esperado. Ou seja, Markowitz, ao considerar tais premissas, contrariou a idéia de que o investidor estaria interessado apenas em maximizar o retorno esperado de sua carteira, uma vez que, se assim o fosse, todos os recursos seriam alocados em um ativo específico com maior retorno esperado.

Note entretanto que, de acordo com Markowitz (1952), embora seja possível reduzir o risco da carteira, o processo de diversificação não resulta na eliminação de todo o risco do investimento, ou seja, a diversificação não implica em variação nula para a carteira de portfólios.

Essa diminuição do risco da carteira, contudo, não seria possível a partir de diversificações tomadas de forma demasiada. Segundo Markowitz, se existir uma alta correlação positiva entre os retornos dos ativos, então as várias alocações múltiplas possíveis teriam resultado semelhante à uma alocação realizada em um único ativo. Ou seja, o processo de diversificação só acarreta redução da variância do retorno de uma carteira se os ativos escolhidos forem não-correlacionados, ou seja, se apresentarem baixos índices de correlação.

De qualquer forma, é importante ressaltar que Markowitz rompe com a teoria tradicional ao considerar a variância dos retornos como fator intrínseco à qualidade dos ativos e, conseqüentemente, a diversificação como fator de proteção e de otimização, uma vez que esse mecanismo de escolha conjunta de ativos incorpora a idéia de que a dispersão dos retornos de uma carteira pode ser reduzida através da escolha de ativos específicos não-correlacionados e com menores variações individuais.

Observe que, de acordo com esse raciocínio, é possível que sejam avaliadas carteiras ótimas no que tange a relação risco-retorno para cada nível de retorno esperado, no sentido de se determinar o que ficou conhecido como fronteira eficiente de Markowitz. A fronteira eficiente de Markowitz pode ser definida como o conjunto das carteiras com melhor desempenho para cada nível de retorno associados, ou seja, as carteiras que possuem o menor risco para um dado nível de retorno.

Note que dessa forma a fronteira eficiente passa a ser concebida por carteiras que maximizam a relação risco-retorno, ou seja, nas quais há a maximização dos retornos e minimização dos respectivos riscos. Uma vez determinada a fronteira eficiente, todas as demais carteiras seriam ignoradas por parte dos investidores, ainda que individualmente um ou outro ativo apresente retorno esperado superior. Isso porque, de acordo com Markowitz (1952), ao se analisar um determinado ativo, um investidor deve se preocupar não apenas com o risco individual do ativo, mas com a contribuição desse risco ao risco total da carteira.

O coeficiente angular da curva que representa a fronteira eficiente é positivo, ou seja, verifica-se que em escolhas ótimas existe uma relação positiva entre o risco e o retorno, isto é, a obtenção de maiores retornos está vinculada à uma maior exposição (maior risco).

Dependendo da atitude do investidor perante o risco, será feita opção por determinado ponto da fronteira eficiente, uma vez que a curva representa as melhores carteiras para diferentes perfis de risco (e níveis de retorno). É importante ressaltar, entretanto, que embora a obtenção da fronteira eficiente independa do tipo de investidor, a escolha da carteira mais eficiente depende do comportamento, das expectativas e do perfil do investidor. Ou seja, de

acordo com o risco em que um determinado investidor esta disposto a se expor e de acordo com o nível de retorno que ele espera obter, determina-se o ponto (carteira) ótimo da fronteira eficiente.

Um vez determinada a fronteira eficiente, portanto, obtém-se a carteira ótima para um dado investidor a partir de seu perfil e expectativas particulares. Graficamente, essa carteira ótima que maximiza a relação risco retorno e que leva em conta o perfil e as expectativas do investidor é dada pelo ponto de tangência entre as curvas de utilidade particulares desse investidor e a fronteira eficiente. O gráfico da figura 1 ilustra esse raciocínio.

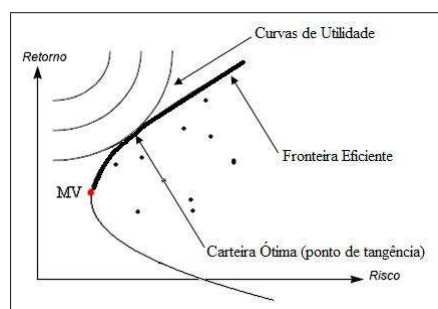


Figura 1: Figura adaptada de Anache (2008).

Note entretanto que, a partir das premissas apresentadas anteriormente, as hipóteses de informação perfeita e de expectativas racionais intrínsecas à fronteira eficiente fazem com que os mercados sejam entendidos como altamente eficientes. A contestação dessas hipóteses e a contestação estatística da premissa de número sete, entretanto, serão apresentadas no final deste capítulo, para que sejam justificadas as possibilidades de arbitragem dos mercados acionários frente às ineficiências reconhecidas.

2.3.2 Modelo de Determinação dos Preços dos Ativos

Como visto anteriormente, Markowitz (1952) introduziu a idéia de que a partir de algumas hipóteses é possível que sejam obtidas carteiras diversificadas nas quais verifica-se a otimização da relação risco-retorno. O conjunto dessas carteiras ficou conhecido como fronteira eficiente e graficamente representa a curva que contempla os pontos com melhor relação entre essas duas variáveis.

James Tobin (1958) introduziu uma análise complementar ao artigo *Portfólio selection* de Markowitz (1952) ao incorporar nas possibilidades de investimento aqueles investimentos nos quais o retorno positivo é garantido, ou seja, ao incorporar na análise de Markowitz a

taxa livre de risco⁷.

Basicamente as idéias de Tobin se relacionavam à decisão do investidor de aplicar seus recursos em ativos com taxa livre de risco e/ou numa carteira de ativos pertencentes à fronteira eficiente, cuja existência Markowitz havia demonstrado ser verificável. De uma maneira simplória, tais escolhas estavam relacionadas à magnitude dos respectivos retornos, ou seja, o investimento seria feito na carteira diversificada otimizada caso o retorno esperado fosse superior à taxa livre de risco, para um dado nível de risco que o investidor estivesse disposto a se expor.

O modelo de determinação dos preços dos ativos (*CAPM-Capital Asset Price Model*)⁸ teve início com William Sharpe e surgiu após a publicação de seu artigo de 1964, intitulado *Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*, no jornal *Journal of Finance*.

Nesse artigo Sharpe considera outras premissas além das premissas adotadas por Markowitz (1952). De acordo com Sharpe, Alexander e Bailey (1998), são 4 as hipóteses complementares, a saber:

1. todos os investidores possuem o mesmo horizonte de um período;
2. a taxa livre de risco é a mesma para todos os investidores;
3. a informação é livre e instantaneamente disponível para todos os investidores;
4. os investidores tem expectativas homogêneas, o que significa que eles têm as mesmas perspectivas.

De uma forma geral, o CAPM é um modelo que especifica as condições de equilíbrio no mercado de ativos de renda variável. Um dos princípios introduzidos por Sharpe em seu artigo de 1964 fundamenta-se na idéia de que deveria existir um equilíbrio entre os preços dos ativos no mercado de capitais. De acordo com Sharpe:

⁷A taxa livre de risco é a taxa de retorno cujo recebimento é garantido previamente, como, por exemplo, a taxa de juros de um título de renda fixa do Tesouro. Note que essa garantia não significa que o valor da taxa deva ser determinado *ex ante*.

⁸De acordo com Sharpe, Alexander e Bailey (1998), o modelo de determinação dos preços dos ativos originou-se dos trabalhos apresentados por Sharpe (1964), Lintner (1965) e Mossin (1966).

No equilíbrio, os preços dos ativos são ajustados de tal forma que o investidor que segue princípios racionais (primariamente o de diversificação) é capaz de alcançar qualquer ponto desejado através da linha do mercado de capitais (*capital market line*). Ele somente pode obter uma maior taxa esperada de retorno correndo riscos adicionais. De fato, o mercado apresenta para ele dois preços: o preço do tempo, ou a taxa pura de juros [...], e o preço do risco, o retorno esperado adicional por unidade de risco corrido. (Sharpe(1964))

A partir das suposições básicas citadas acima, deriva-se a seguinte equação, que sumaria os princípios do CAPM:

$$E(R_i) = RF + [E(RM) - RF]\beta_i \quad (1)$$

na qual $E(R_i)$ representa o retorno esperado de um ativo i com risco, RF o retorno de um ativo sem risco e $E(RM)$ o retorno esperado da carteira de mercado. β_i , por sua vez, é denominado coeficiente de risco sistemático e representa uma medida da sensibilidade do retorno do ativo i às variações nos retornos do mercado como um todo e matematicamente é dado pela covariância entre o retorno desse ativo e o retorno do mercado.

Note que essa equação equaliza a idéia de que o retorno esperado para um determinado ativo é dado pela taxa livre de risco acrescida do prêmio pela exposição do investimento. De acordo com Anache (2008), a simplicidade deste modelo reside na relação linear positiva existente entre risco (não diversificável) e rentabilidade. Além disso, é somente o risco de mercado que afeta o retorno. O investidor não recebe qualquer retorno adicional para assumir risco diversificável.

Segundo Haugen (1997), esta teoria é baseada na maneira como as ações são precificadas em relação ao seu risco, ou seja, basicamente está baseada no pressuposto de que todos os investidores, caso façam investimentos em ativos de renda variável, escolhem suas carteiras na fronteira eficiente e de acordo com a aversão ao risco, mas levando-se em consideração o retorno esperado do mercado e a taxa livre de risco.

Note que na equação (1) caso $\beta_i = 0$, o retorno esperado do título i é igual a taxa livre de risco e caso $\beta_i = 1$, o retorno esperado do título i é igual ao retorno esperado do mercado, já incluídas as taxas de risco diversificável e livre de risco.

As premissas simplificadoras assumidas por Sharpe (1964), entretanto, implicam que a combinação de ativos com risco para todos os investidores seria a mesma (o que na realidade

não se verifica), uma vez que expectativas homogêneas se traduzem em iguais expectativas por parte dos investidores com relação a retorno esperado, variância e covariâncias dos ativos e nível da taxa livre de risco. As hipóteses temporais 1 e 3, por sua vez, garantiriam que tais formulações fossem feitas ao mesmo tempo por parte de todos os investidores, ou seja, garantiriam a não defasagem temporal de informações e de expectativas com relação ao comportamento do mercado (o que também não se verifica, dados os diferentes níveis de gestão, formação e capacidade de análise dos investidores). A partir de tais premissas, seria possível considerar o mercado como um mercado eficiente e determinar a carteira otimizada desse mercado, formada por todos os títulos existentes e ponderados por seus respectivos valores de mercado. Isso porque a partir de tais hipóteses todos os investidores racionalmente escolheriam a mesma carteira de ativos com risco.

Note que essas considerações eliminam quaisquer possibilidades de arbitragem entre ativos financeiros (operação verificável no mercado). Tais premissas, portanto, não podem ser tomadas como hipóteses imutáveis, dado que mitigam importantes janelas de oportunidade, principalmente no que tange às operações de arbitragem entre ativos financeiros.

2.3.3 Fama(1970) e a Eficiência dos Mercados

Analisando o contexto de eficiência dos mercados, verifica-se que uma das grandes contribuições para as Finanças Modernas foi sistematizada por Fama em 1970 no artigo intitulado *Efficient capital markets: a review of theory and empirical work* e publicado no *Journal of Finance*, ao considerar que os mercados podem, eventualmente, se comportar com alguma ineficiência.

Esse artigo apresenta uma revisão da literatura acerca dos modelos de eficiência dos mercados. Em um primeiro momento, a teoria e o trabalho empírico são relacionados com as modificações dos preços dos ativos. Posteriormente, são feitas considerações acerca da operacionalização das formas de eficiência de mercado, baseadas em três conjuntos de informação: informações passadas (preços históricos), informações publicamente disponíveis e todo o conjunto de informação (informações passadas, públicas e privadas).

Nesse artigo, Fama considera um mercado eficiente como sendo um mercado no qual os preços dos ativos sempre refletem completamente todas as informações disponíveis. As

hipóteses de competição perfeita⁹, expectativas homogêneas¹⁰, racionalidade ilimitada¹¹, ausência de fricções (homogeneidade dos ativos quanto à divisibilidade e ausência de custos de transação) e aleatoriedade do comportamento dos ativos¹² formam a base das idéias desse artigo.

Note que essas hipóteses implicam que nenhuma informação (como preços passados, notícias, etc) poderia alterar a trajetória de um ativo e, portanto, trazer vantagens a determinado investidor. Isso porque quaisquer informações seriam instantaneamente incorporadas e refletidas nos preços dos ativos a que estivessem relacionadas. A partir de tal fato, todos os agentes econômicos rapidamente ajustariam suas ofertas para refletir novas informações e de forma a maximizar suas utilidades, mas sempre operando nos mercados independentemente uns dos outros.

Sob esse ponto de vista, o mercado seria um meio em que as empresas se comportariam como agentes tomadores das decisões de produção e investimento, e os investidores agentes apenas passíveis de escolher ativos representativos dessas empresas. Assim, os preços nos mercados financeiros incorporariam todas as informações disponíveis e representariam as melhores estimativas do valor fundamental dos ativos negociados. Os desvios de curto prazo, portanto, refletiriam apenas pequenas mudanças nos padrões de oferta e demanda, o que confere um caráter de imprevisibilidade ao comportamento dos ativos. Nesse sentido, a teoria de eficiência do mercado suprimiria quaisquer possibilidades de retornos adicionais.

Note que sob o pressuposto de eficiência de mercado, os preços dos ativos se modificariam aleatoriamente em resposta a novas informações, mas sempre em torno do seu valor intrínseco (valor fundamental). Esse comportamento requer, entretanto, a não ocorrência de quaisquer regularidades *ex post* nos retornos observados, isto é, a inexistência de padrões de comportamento no mercado que resultem em retornos anormais (LEMOS; COSTA JR., 1995).

Sob uma análise mais realista, entretanto, tais hipóteses não necessariamente se verificam para todos os investidores do mercado, uma vez que ao mesmo tempo em que investi-

⁹Em um mercado com competição perfeita nenhuma ação individual pode afetar de forma decisiva a trajetória dos preços dos ativos

¹⁰Expectativas homogêneas pressupõem que todos os investidores fazem a mesma previsão com relação a retornos, variâncias e covariâncias futuras.

¹¹Nesse tipo de racionalidade os agentes são totalmente informados e têm acesso igualitário aos mercados.

¹²Este raciocínio indica que o histórico passado de preços não é determinante dos preços futuros, uma vez que as mudanças de preços não necessariamente são seguidas da manutenção dos padrões de precificação, ou seja, não são necessariamente dependentes nem correlacionadas com as oscilações históricas dos ativos.

dores não plenamente racionais (*noise traders*) operam, coexistem *smart traders* (árbitros racionais) que rapidamente incorporam as informações de que dispõem e as usam de modo a antecipar a movimentação geral do mercado.

Fama (1970) apresentou uma nova visão para a questão da eficiência dos mercados. Ele afirmou que os mercados não são sempre totalmente eficientes, mas o mercado pode, por vezes, se comportar com alguma ineficiência.

A eficiência dos mercados, bem como a própria racionalidade dos investidores (importantes pilares da Moderna Teoria das Finanças), vêm sendo questionada pelas chamadas Finanças Comportamentais, como será exposto a seguir.

2.4 Finanças Comportamentais

Segundo Shefrin (2002), a finança comportamental se tornou evidente quando os avanços feitos pela psicologia chamaram a atenção dos economistas. Segundo esse autor, os trabalhos realizados por Paul Slovic (1969) e (1972) sobre corretores e investidores fizeram desse autor o primeiro a perceber a relevância dos conceitos comportamentais para as finanças. Entretanto, foram os artigos de Amos Tversky (1974) e Daniel Kahneman (1979) os que mais impressionaram os estudiosos da área de finanças.

No artigo *Judgement under Uncertainty: Heuristics and Biases* (1974), Kahneman e Tversky descrevem as três heurísticas empregadas para acessar probabilidades e para prever valores. No artigo *Prospect Theory: An Analysis of Decision under Risk*, Tversky e Kahneman apresentam uma crítica à teoria da utilidade esperada como um modelo descritivo de tomada de decisão e desenvolvem um modelo alternativo dentro de investimentos que envolvem riscos.

Após 1987, alguns estudiosos da área de finanças e psicologia como De Bondt, Thaler, Tversky, Kahneman, Shefrin, Statman, Shiller, entre outros, começaram a rever os pressupostos preconizados pelas finanças tradicionais e a estudar os fenômenos relacionados à fatores como volatilidade, anomalias, sobre-reação, aversão à perda, heurísticas e vieses cognitivos, fortalecendo o surgimento da área de finanças comportamentais.

Atualmente, a definição e os conceitos de finanças comportamentais ainda não estão consolidados no meio científico, dada a abrangência das áreas aos quais esses estudos estão vinculados. Segundo Macedo (2003), as finanças comportamentais fundem conceitos de economia, finanças e psicologia cognitiva na tentativa de construir um modelo mais detalhado do comportamento humano nos mercados financeiros. Basicamente o intuito desses estudos

se concentra na tentativa de explicar o que causa algumas anomalias que são observadas dentro dos modelos tradicionais de finanças, principalmente no que se refere ao modo como os investidores agem com a informação para tomar decisões de investimento¹³.

No período que compreende o final da década de oitenta e início da década de noventa, as hipóteses assumidas pelos modelos das finanças modernas começaram a ser questionadas, principalmente devido ao reconhecimento de que os indivíduos não são idênticos nem a informação perfeita. O processo de tomada de decisão passou a ser visto não mais apenas como um processo puramente técnico e dependente de análises históricas ou apenas como um processo baseado em mercados eficientes, no qual os preços atuais refletem toda a informação relevante, mas como um processo no qual agentes racionais e irracionais interagem entre si e com expectativas heterogêneas, de modo a influenciarem significativamente o preço dos ativos.

O processo cognitivo que envolve essa relação heterogênea de interpretação da realidade, em presença inclusive da irracionalidade entre os agentes, é que norteia os estudos das finanças comportamentais. A explicação dos vieses cognitivos decorrentes dos diversos comportamentos humanos, como excesso de confiança, arrependimento, ganância, medo, os quais fazem com que os investidores cometam erros de julgamento, pode ser observada e explorada por outros investidores, também atores do mesmo procedimento de mercado, para que previsões mais detalhadas possam ser produzidas.

Com relação à consideração de mercados eficientes, a premissa de que investidores são perfeitamente racionais, não havendo limites em relação ao processamento da informação com preferências estáveis e coerentes, onde os agentes maximizam a utilidade proveniente de tais preferências e constroem expectativas não-viesadas sobre acontecimentos futuros (MILANEZ, 2001), não se verifica quando analisamos a real influência do comportamento humano no mercado financeiro, principalmente no que se refere às modificações dos preços do ativos. Os preços dos ativos financeiros (ações, obrigações etc.) não refletem em cada momento toda a informação disponível, mas o conjunto de interpretações que cada investidor faz dessas informações.

De acordo com Fama (1998), as anomalias de mercado ocorreriam de forma aleatória, acabariam por se compensar e seriam consistentes com a eficiência de mercado, não sendo admitida então, a presença de tendências capazes de influenciar o preço. Entretanto, verifica-

¹³Um dos princípios básicos das finanças comportamentais postula que alguns fenômenos financeiros podem ser melhor compreendidos a partir de modelos nos quais os agentes não são completamente racionais.

se que tal equilíbrio hipotético não se materializa no mercado financeiro, tendo em vista a fragilidade das decisões humanas e o fato de que decisões erradas de um agente não necessariamente criam oportunidades para outros agentes. A racionalidade limitada e a informação imperfeita são, portanto, hipóteses mais pertinentes à adequação de um modelo que se proponha explicar o real funcionamento dos mercados financeiros.

As finanças comportamentais procuram, portanto, encontrar soluções viáveis para a compreensão das anomalias não explicadas pelos paradigmas tradicionais, baseando-se em modelos com agentes heterogêneos nos quais a irracionalidade se encontra presente e as emoções fazem parte do processo de tomada de decisão.

2.5 Crítica às Hipóteses Simplificadoras

A teoria de diversificação em ativos de Markowitz (1952) contraria corretamente a idéia de que o investidor estaria interessado apenas em maximizar o retorno esperado de sua carteira, uma vez que, se assim o fosse, todos os recursos estariam alocados em um ativo específico com maior retorno esperado.

Nesse universo de diversificação não se pode considerar, entretanto, que todos os investidores estejam de acordo quanto à distribuição de probabilidade das taxas de retorno dos ativos. Isso porque, na maior parte dos casos, a distribuição empírica dos preços dos ativos, bem como dos respectivos retornos, é desconhecida. O que os investidores fazem, na verdade, é aproximar em função das informações que têm disponíveis a distribuição empírica dos dados para distribuições conhecidas e que melhor se adequam ao tipo de análise requerida. Esse fato faz com que não sejamos capazes de assegurar, na prática, a existência de um conjunto de carteiras que seja eficiente a todos os investidores (o que não implica a não existência da fronteira eficiente). Por consequência são criadas janelas de arbitragem, dadas as diferentes expectativas que cada investidor faz acerca da carteira ótima para alocação de seus recursos e das tendências de evolução dos preços dos ativos financeiros.

A velocidade e a eficiência da interpretação das informações por parte dos investidores também não podem ser consideradas homogêneas, ainda que se considere que as informações estejam instantaneamente disponíveis a todos os investidores¹⁴.

A defasagem temporal causada pelo fator processamento e a qualidade de interpretação

¹⁴Na prática, dado o custo para obtenção de certas informações e dado que existe diferença entre disponibilidade de informação e decodificação de informação, não se pode considerar disponibilidade instantânea de qualquer informação para todos os investidores.

causada pelo fator eficiência fazem com que não se possa considerar, como o faz as finanças modernas, as hipóteses de racionalidade ilimitada e informação perfeita. Nesse sentido, os preços dos ativos financeiros não refletem em cada momento toda a informação disponível, mas o conjunto de interpretações que cada investidor faz dessas informações.

Note que a adoção das hipóteses de informação livre e perfeita e expectativas homogêneas (premissas simplificadoras adotadas também por Sharpe (1964)) implicariam em uma mesma combinação de ativos com risco para todos os investidores (o que na realidade não se verifica), uma vez que expectativas homogêneas se traduzem em iguais expectativas por parte dos investidores com relação a retorno esperado, variância e covariâncias dos ativos e nível da taxa livre de risco. Tais hipóteses, por sua vez, garantiriam que essas formulações fossem feitas ao mesmo tempo por parte de todos os investidores, ou seja, garantiriam a não defasagem temporal de informações e de expectativas com relação ao comportamento do mercado (o que também não se verifica, dados os diferentes níveis de gestão, formação e capacidade de análise dos investidores).

Note que essas considerações eliminam quaisquer possibilidades de arbitragem entre ativos financeiros (operação verificável no mercado). Tais premissas, portanto, não podem ser tomadas como hipóteses imutáveis, dado que mitigam importantes janelas de oportunidade, principalmente no que tange às operações de arbitragem entre ativos financeiros.

Note que sob o pressuposto de eficiência de mercado, os preços dos ativos se modificariam aleatoriamente em resposta a novas informações, mas sempre em torno do seu valor intrínseco (valor fundamental). Esse comportamento requer, entretanto, a não ocorrência de quaisquer regularidades *ex post* nos retornos observados, isto é, a inexistência de padrões de comportamento no mercado que resultem em retornos anormais (LEMOS; COSTA JR., 1995).

Sob uma análise mais realista, entretanto, tais hipóteses não necessariamente se verificam para todos os investidores do mercado, uma vez que ao mesmo tempo em que investidores não plenamente racionais (*noise traders*) operam, coexistem *smart traders* (árbitros racionais) que rapidamente incorporam as informações de que dispõem e as usam de modo a antecipar a movimentação geral do mercado.

É nesse contexto que o modelo de arbitragem de ativos que será proposto se fundamenta. Esse modelo não procurará justificar o modo como se processa a formulação dos preços dos ativos e o modo como se comportam os investidores frente a novas informações de mercado, mas será formulado com o intuito de determinar oportunidades de arbitragem

a partir de anomalias relacionadas a comportamentos fora dos padrões normais da relação de precificação entre ativos financeiros, de forma a fornecer ao mercado um novo indicador probabilístico que auxilie os investidores em suas decisões de arbitragem.

3 *HEDGE FUNDS E AS ESTRATÉGIAS LONG-SHORT*

Este capítulo consistirá em um primeiro momento em apontar as principais características dos *Hedge Funds*, que em termos práticos são fundos de investimento cujo objetivo é obter a maior rentabilidade possível e de forma independente à movimentação dos mercados financeiros a partir de estratégias sofisticadas de análises, como é o caso das estratégias *long-short*. Em um segundo momento, será apontada a necessidade da criação de um modelo probabilístico *ad hoc* (baseado na análise de extremos) para as estratégias *long-short*, a partir da exposição geral das principais características das estratégias *long-short* e de duas estratégias existentes: a estratégia *Pair Trading* e a estratégia Neutra ao Mercado *Long-Short* com Portfólios *Índex Tracking*, ambas construídas a partir do método de correção erros para séries cointegradas.

3.1 *Hedge Fund*

O termo *Hedge Fund* certamente é conhecido por investidores e acadêmicos de finanças, seja por causa dos lucros extraordinários que trouxeram a investidores internacionais, como George Soros, ou por iniciativas marcantes, como a do fundo *Long Term Capital Management*, que levou o banco central norte-americano *Federal Reserve* a formar em meio à crise russa de 1998 um *pool* de banqueiros para capitalizar o fundo e evitar pânico no mercado financeiro mundial.

O fato é que, independente de acontecimentos particulares, os *Hedge Funds* são fundos bastante expressivos e com estratégias bastante interessantes. Criados em 1949 por Alfred W. Jones, eram baseados na compra de ações sub-avaliadas e na venda a descoberto de ações sobre-avaliadas com o intuito de proteger¹⁵ a carteira contra riscos de mercado, ou seja, reduzir o risco das carteira de ações através de posições vendidas¹⁶ em outras ações. Assim, ao comprar ações que ele acreditava terem bom potencial de retornos e manter posição vendida em papéis nos quais estava pessimista, Jones acabou reduzindo a sua exposição ao mercado de ações.

¹⁵É exatamente essa busca por proteção que inspirou o nome *Hedge* (proteção) desses fundos.

¹⁶Uma posição vendida no mercado de ações indica que o investidor vendeu ações a descoberto, ou seja, as vendeu mesmo sem possuí-las, e o fez por acreditar em uma queda de suas cotações.

Trabalhando com tais posições Jones reduzia seu risco, uma vez que caso o mercado como um todo caísse, o fundo ganharia ao menos em parte de suas posições (no caso, as vendidas, ou *short*). Sua estratégia, entretanto, quando confirmadas as perspectivas traçadas, permitia que os fundos auferissem lucro nas duas operações, ou seja, com a alta das ações compradas e com a queda das vendidas a descoberto, uma vez que sua posição permitia adquirí-las por valor inferior.

Essa estratégia fez com que mesmo com a cobrança de uma taxa de remuneração em torno de 20% do retorno líquido, o que era considerada alta para os padrões da época, o fundo administrado por Jones apresentasse retorno maior do que qualquer fundo mútuo, o que despertou grande interesse e atraiu no final da década de 60 tanto investidores quanto profissionais de mercado que criaram centenas de novos fundos *Hedge*¹⁷.

Com a crise da década de 70, entretanto, alguns *Hedge Funds* foram extintos e apenas na metade da década de 90 é que esses fundos foram reintroduzidos nos fundos de interesse dos investidores. Mas se teoricamente as estratégias de investimentos adotadas por esse tipo de fundo reduziam os riscos, por que alguns investidores fazem associação dos *Hedge Funds* com investimentos de risco elevado e por que eles sofreram tanto com a crise da década de 70? A resposta reside no fato de que a estratégia mencionada acima pode não ser a única estratégia adotada por um *Hedge Fund*.

Por exemplo, muitos *Hedge Funds* adotam estratégias que envolvem alavancagem, ou seja, o uso de derivativos para multiplicar o desempenho do fundo. Foi exatamente a adoção de estratégias mais agressivas e de posições extremamente alavancadas é que levaram muitos *Hedge Funds* a fortes perdas no final da década de 60 e início da década de 70 nos EUA. Esse processo de acumulação de perdas durou até 1986, quando o desempenho do *Tiger Fund* trouxe os *Hedge Funds* de volta às alternativas de investimento de grandes investidores.

A participação dos *Hedge Funds* na indústria, entretanto, não é tão expressiva quando comparadas a de fundos tradicionais. Isso porque esses são fundos de acesso restrito, com investimento mínimo para pessoas físicas de US\$ 1 milhão. Como na prática não se recomenda que o investidor tenha mais de 10% do seu patrimônio nesses fundos, os *Hedge Funds* acabam por englobar pessoas físicas com fortunas mínimas de US\$ 10 milhões.

Embora requeiram investimentos elevados, os *Hedge Funds* são fundos bastante livres do ponto de vista da legislação. Os investidores podem fazer todo tipo de operação financeira,

¹⁷Embora os *Hedge Funds* tivessem sido criados em 1949 por Alfred W. Jones, Carol J. Loomis foi quem usou publicamente pela primeira vez o termo *Hedge Fund* em seu artigo de 1966 na revista Fortune.

a qualquer momento e em todos os mercados, sem nenhuma restrição quanto a volumes de negociação. As posições alavancadas permitem que esses investidores apostem em suas estratégias, literalmente, um montante que não possuem, ou seja, um patrimônio muito superior ao que de fato possuem e tanto superior quantas for desejado (alavancagem).

Note que se a operação for bem sucedida, o fundo auferirá lucro extraordinário. Mas caso contrário, poderá incorrer em prejuízos extraordinários. É por embutir esses riscos que esses fundos são reservados por legislação apenas a investidores de grandes capitais¹⁸.

Sob um ponto de vista mais teórico, conforme Pelosi (2003), poderíamos também entender como *Hedge Fund* o tipo de investimento o qual o gestor utiliza-se de estratégias para redução dos riscos, ou seja, um fundo de investimento de baixo risco. Não devemos considerar essa interpretação, entretanto, já que muitos *Hedge Funds* não necessariamente utilizam-se de estratégias totalmente hedgeadas e se expõe a elevados riscos.

Nos dias de hoje, o termo *Hedge Fund* tornou-se um termo genérico que abrange diversos fundos não tradicionais como, por exemplo, fundos que exploram ineficiências de mercado através de estratégias sofisticadas de abordagem ou com grandes apostas direcionais. Como na legislação de investimentos não existe uma definição clara sobre esses fundos, o portfólio de investimento de um *Hedge Fund* é bastante diversificado, englobando desde títulos públicos do governo, arbitragem, commodities, câmbio, futuros, opções ou mesmo outros *Hedge Funds*.

No Brasil, não existem na legislação *Hedge Funds*, mas fundos que por vezes atuam como *Hedge Fund*. De acordo com a nova classificação de fundos feita pelo Banco Central e baseada em risco (introduzida em março de 2005), tais carteiras estão entre os fundos genéricos. Fundos dessa categoria têm permissão de manterem posições alavancadas, ou seja, atuar em mercado de derivativos com a finalidade de proteger a carteira e incrementar ganhos adicionais. No Brasil, portanto, alguns fundos de derivativos podem comportar-se como os *hedge funds* norte-americanos.

3.1.1 *Hedge Funds* Norte-Americanos

O mercado norte-americano ocupa a posição de primeiro lugar no mundo quando analisamos o número de *Hedge Funds* e o volume financeiro dessa indústria. Embora tenha sido pioneiro na criação desses fundos, o que mais contribuiu para a expansão dessa indústria tem

¹⁸Os *Hedge Funds* oferecem aos seus gestores ganhos expressivos através da participação nos resultados. Esse fato contribuiu para que sejam atraídos para os *Hedge Funds* os melhores e mais criativos profissionais.

suporte na legislação de investimentos norte-americana, que impõe muito menos restrições ou fiscalizações para as ações desses fundos que nos demais países. Para que nos EUA um fundo possa ser qualificado como *Hedge Fund*, entretanto, o número de investidores não pode ser superior a 500, os gestores não podem fazer ofertas públicas ou qualquer outro tipo de propaganda e seus investidores devem entender de mercado e se qualificarem como *qualified investors buyers* (QIB).

De acordo com o Banco Central Norte-Americano (FED - *Federal Reserve*), o número de *Hedge Funds* e o volume de negociações têm aumentado de forma expressiva, principalmente na última década. Analisando o período que compreende os anos entre 1993 e 2003, verifica-se um salto de negociações de 50 para 600 bilhões de dólares, já ultrapassando hoje uma marca de 1,5 trilhão de dólares e dividida por mais de 8000 fundos. A indústria de *Hedge Funds*, embora administre menos recursos que a indústria de fundos mútuos (2,106 trilhões), deve ser adotada como alternativa de investimento, dado o crescimento expressivo que tem demonstrado.

A seguir serão apresentadas características comuns a todos os *Hedge Funds* norte-americanos:

1. esses fundos não são por definição restritos a uma classe de ativos específica, ou seja, possuem livre escolha com relação às classes dos ativos a formarem a carteira;
2. as estratégias e os estilos de investimento são livres e facultativas ao perfil do gestor;
3. os instrumentos de investimento podem ser diversificados, ou seja, o gestor pode optar tanto por moeda e ações quanto por derivativos;
4. embora exista a possibilidade de altas taxas de lucratividade, as taxas de administração e performance também são elevadas;
5. os gestores investem o próprio capital no seu fundo;
6. Muitos dos fundos apresentam baixa correlação com mercados tradicionais;
7. A transparência é um pouco restrita, uma vez que não há a obrigatoriedade da apresentação de relatórios ao tesouro americano.

Embora possuam essas características em comum, a classificação dos *Hedge Funds* norte-americanos está vinculada ao tipo de estratégia que seguem. Segundo a Revista Bloomberg Markets (Jan/2006), as principais estratégias são:

- *Long Short Equity*: constituir posições compradoras ou vendedoras sobre ações susceptíveis de se apreciarem ou de se depreciarem. Técnica que se alimenta de numerosas anomalias dos mercados, reforçadas em período de forte volatilidade. É necessária uma boa especialização do gestor no que tange à identificação de anomalias, pois o universo de investimento é bastante diversificado, o que possibilita diferentes estilos de gestão.
- *Event Driven Strategies*: neste tipo de estratégia de investimento o gestor tenta antecipar a ocorrência de determinados eventos de caráter corporativo que podem afetar o valor de determinados valores mobiliários. Os ganhos ocorrem caso os eventos se concretizem de acordo com as expectativas do gestor. São dois casos particulares desta estratégia:
 - *Merger Arbitrage*: que consiste em posições longas para ações de empresas que estejam a ser alvo de aquisição ou fusão e de posições curtas em ações das respectivas empresas adquirentes ou nas quais se vão incorporar as empresas a fusionar. Os ganhos esperados estão associados à expectativa de que as ações das empresas alvo de oferta de aquisição ou fusão se valorizem e de que as ações das empresas com posição dominante no evento corporativo se desvalorizem.
 - *Distressed Securities*: estratégia na qual o gestor investe em ativos de empresas que estejam em situação econômico-financeira ou operacional débeis. Os ganhos associados a esta estratégia sucedem caso as ações que foram adquiridas a desconto, em virtude da pressão vendedora que existe nestes casos, se venham a valorizar em função da reorganização ou da recuperação das empresas.
- *Fixed Income Arbitrage*: consiste na tomada de posições longas e curtas em ativos de rendimento fixo cujos preços se encontram historicamente correlacionados. Os ganhos de arbitragem são geralmente amplificados pela utilização de alavancagem, uma vez que as alterações de correlação são, em regra, muito reduzidas.
- *Convertible Arbitrage*: trata-se de uma estratégia em que o gestor combina posições longas em obrigações convertíveis em ações com posições curtas nessas mesmas ações. Os ganhos de arbitragem resultam do fato do preço do instrumento híbrido cair de forma menos intensa do que o da ação, uma vez que assume simultaneamente características de uma obrigação e ação.

- *Macro Investing*: trata-se da estratégia mais ampla de todos os *Hedge Funds*. Os gestores destes fundos atuam em uma escala global procurando identificar situações de ineficiência nos preços dos valores mobiliários e antecipar os efeitos que eventos macroeconômicos e políticos possam ter sobre esses preços.

3.1.2 *Hedge Funds* brasileiros

No Brasil os fundos que se comportam como *Hedge Funds* são classificados como fundos Multimercados, uma vez que na regulamentação brasileira esses são os fundos de maior abrangência operacional. Nem todos os fundos Multimercados, entretanto, podem ser considerados *Hedge Funds*. A seguir estão apresentadas as principais formas de operação dos fundos Multimercados:

1. Com Renda Variável e sem Alavancagem: Fazem parte desse segmento os fundos que buscam retorno no longo prazo através de investimento em diversas classes de ativos (renda fixa, câmbio, etc) incluindo renda variável (ações, etc.). Não admitem alavancagem.
2. Com Renda Variável e com Alavancagem: Possuem a mesma estratégia que a categoria anterior, podendo fazer uso, entretanto, de posições alavancadas.
3. Sem Renda Variável e sem Alavancagem: Classificam-se neste segmento os fundos que buscam retorno no longo prazo através de investimento em diversas classes de ativos (renda fixa, câmbio, etc), exceto renda variável (ações, etc.). Estes fundos não têm explicitado o mix de ativos com o qual devem ser comparados (*asset allocation benchmark*) e podem, inclusive, ser comparados a parâmetro de desempenho que reflita apenas uma classe de ativos (por exemplo: 100% CDI). Não admitem alavancagem.
4. Sem Renda Variável e com Alavancagem: Da mesma forma, classificam-se como fundos que buscam retorno no longo prazo através de investimento em diversas classes de ativos (renda fixa, câmbio, etc) exceto renda variável (ações, etc.), porém admitem alavancagem. Também não há um mix de ativos com o qual devem ser comparados (*asset allocation benchmark*).
5. Balanceados: Também buscam retorno no longo prazo através de investimento em diversas classes de ativos (renda fixa, ações, câmbio, etc). Estes fundos utilizam uma

estratégia de investimento diversificada e deslocamentos táticos entre as classes de ativos ou estratégia explícita de rebalanceamento de curto prazo. Estes fundos devem ter explicitado o mix de ativos (percentual de cada classe de ativo) com o qual devem ser comparados (*asset allocation benchmark*). Sendo assim, esses fundos não podem ser comparados a indicador de desempenho que reflita apenas uma classe de ativos (por exemplo: 100% CDI). Não admitem alavancagem.

6. Capital Protegido: Busca retornos em mercados de risco procurando proteger parcial ou totalmente o principal investido.

Normalmente os fundos da categoria 3 são os mais conservadores e os da categoria 2 os mais arriscados, já que esses são os verdadeiros Hedge Funds brasileiros, ou seja, a categoria onde as estratégias de investimento mais se aproximam do modelo americano. Porém, dentro de cada sub-categoria existem fundos com estratégias e graus de risco variáveis.

3.1.3 Diferenças entre Brasil e EUA

De acordo com Branco e Franco (2004), a principal diferença entre os *Hedge Funds* brasileiro e norte-americano se refere à estrutura administrativa das respectivas indústrias. Nos EUA, esses fundos são sociedades limitadas quanto ao número de cotistas, pouco reguladas e isentas dos controles da regulamentação *Company Act* de 1940 (restrições a alavancagem, venda a descoberto, concentração de risco, etc.), aplicada aos *mutual funds*. A falta de exigências quanto à transparência dos balanços e das operações dos *Hedge Funds* norte-americanos faz com que haja auditorias particulares. Essas auditorias, entretanto, segundo Liang (2003), são limitadas, uma vez que 40% de uma amostra ampla não é auditada adequadamente.

No Brasil, a rigidez na regulamentação das operações dos *Hedge Funds* é muito maior, uma vez que esses fundos recebem o mesmo tratamento legal que outros fundos, como os de renda fixa, por exemplo. Essa regulamentação, embora em alguns casos apresente restrições quanto às negociações, proporciona aos investidores maior segurança quanto à precificação das cotas dos respectivos fundos.

Com relação ao tempo de maturação, os *Hedge Funds* norte-americanos têm um tempo de vida médio (63 meses) bastante superior aos dos fundos brasileiros, cujas carteiras ainda são relativamente jovens.

Com relação às estratégias de investimento, os fundos norte-americanos são bem mais diversificados quanto ao grau de especialização e categorização (*opportunistic, event driven,*

futures and currencies arbitrage, market timing, market neutral, global, etc). No Brasil existe pouca divisão que, em geral, se dividem em macro (comprado ou vendido em Brasil) ou *long-short* em ações. As estratégias utilizadas costumam variar de acordo com alterações na equipe e o momento de mercado (Branco e Franco, 2004).

Embora existam tais diferenças e essa complexidade de estratégias, o importante como marco histórico é que a partir da estratégia de Jones estava criada o que hoje é conhecida como a estratégia *long-short* no mercado de ações.

3.2 Estratégias *Long-Short*

As estratégias *long-short*, objetivos do nosso estudo, compreendem a manutenção simultânea de compra e venda de ações e de derivativos, ambos susceptíveis a apreciação ou depreciação. Esse tipo de estratégia fundamenta-se nas inúmeras anomalias dos mercados, que são reforçadas em período de forte volatilidade. A utilização desse tipo de estratégia visa eliminar ou reduzir a exposição dos investimentos aos chamados riscos sistêmicos. Algumas estratégias *long-short* podem também ser classificadas como estratégias *market neutral*, dado o seu caráter de obtenção de retornos adicionais sem exposição a riscos direcionais de mercado.

De uma forma geral, uma estratégia *long-short* se baseia na manutenção de um portfólio *long* financiado por um portfólio *short*, que além de servir como uma proteção contra riscos de declínios dos mercados, fornece ao gestor a possibilidade de obter retornos superiores aos retornos globais pelo uso de algum critério de seleção de ações. Impulsionado por esse objetivo, pesquisas têm sido feitas com o intuito de ajustar modelos que propiciem a escolha dos ativos bem como o *timing* das operações.

Na maioria dos casos, os administradores de carteiras *long-short* compram ações e opções de algumas empresas e, simultaneamente, alugam ações de outras empresas para vendê-las em mercado e devolvê-las quando lhes interessar ou quando expirarem os contratos de aluguel. As posições vendidas também podem ser feitas via contratos futuros ou opções de ações.

As estratégias *long-short* podem se dividir em modalidades distintas com relação às especificidades setoriais das negociações, a saber:

1. Modalidade intra-setorial, na qual o administrador compra e vende ações de empresas de um setor específico. É uma modalidade que tende a apresentar oscilações menores

do que as demais, já que os impactos afetam o setor como um todo, o que minimiza os riscos gerais;

2. Modalidade intersetorial, na qual o administrador compra ações de um setor e vende ações de outro. Nesse tipo de modalidade, o conhecimento dos padrões setoriais é mais determinante até que o conhecimento particular de uma dada empresa;
3. Modalidade carteira contra carteira, na qual o investidor compra uma carteira formada por ações de diferentes setores e vende outro grupo de ações, sem que haja relações entre os setores. Nesse tipo de modalidade, assume-se a posição vendida em uma das carteiras, a parte *short*, e com os recursos obtidos na venda de uma das carteiras assume-se posição comprada na outra, a parte *long*;
4. Modalidade quantitativa, cujas decisões são influenciadas por análises estatísticas.

Note entretanto que nas estratégias long-short, sejam elas de qualquer modalidade setorial, mais importante que a valorização ou desvalorização das ações é o desempenho relativo entre as posições compradas e vendidas. Ou seja, seja com mercados em alta ou em períodos de baixa, o potencial de rendimentos diferenciados nas estratégias long-short está sempre presente.

3.2.1 Abordagens Long-Short

O gestor de um *Hedge Fund*, ao optar por adotar uma estratégia long-short, pode direcionar sua análise para uma abordagem técnica, com a finalidade de identificar empresas que estejam com relação de precificação fora dos padrões históricos, ou para uma abordagem fundamentalista, que tem o objetivo de avaliar situações financeiras particulares de determinadas empresas com o intuito de verificar o real valor de mercado e projeções futuras de crescimento.

A abordagem fundamentalista se baseia em uma análise acerca dos temas de investimento (*investment themes*) ou em uma análise fundamental.

A análise de temas de investimento procura identificar quais fatores macroeconômicos têm maior impacto na valoração futura dos ativos financeiros. Empresas com potencial de inovação e com planos de investimento direcionados são candidatas a assumirem posições *long*. Depois de verificar qual porção do mercado tem maiores possibilidades de crescimento,

os gestores *long-short* verificam quais companhias têm maiores tendências de acompanhar o desenvolvimento tecnológico e econômico.

Um exemplo da análise de temas de investimento envolve o debate acerca do grande crescimento da internet, principalmente no que se refere a serviços de imagens, vídeos e voz. Tais transmissões podem ser feitas via cabo-modem, fibra óptica ou satélite. O investidor que, sob análises particulares, verificar que a tecnologia cabo-modem ampliará as fronteiras do mercado, poderá assumir posições *long* em grandes empresas de telecomunicações, como a AT&T, que ampliou sua participação ao adquirir duas das maiores companhias de televisão a cabo do mundo. Esse é um exemplo de como gestores *long-short* podem incorporar fatores macroeconômicos para identificar bons candidatos a assumirem posições *long*.

Com relação à posição *short*, entretanto, os investidores devem ter o foco da análise em uma situação contrária. Na busca pelos candidatos à posições *short*, o gestor *long-short* provavelmente identificará companhias que terão sua competitividade comprometida em função de fatores como gestão ruim, queda no fluxo de caixa, perda de parcelas de mercado, débitos excessivos, etc. Esses fatores podem contribuir para que se aumente a pressão sobre essas companhias por parte do mercado, o que aumentariam as chances de um desempenho continuado negativo dessas empresas.

A análise fundamental ou fundamentalista, tal como explorada no capítulo 2, por sua vez, é um método de análise das finanças particulares de uma determinada companhia, com o intuito de avaliar seu real valor, suas projeções financeiras futuras e suas perspectivas de crescimento. Note que essa análise, embora esteja mais focada em particularidades internas de uma empresa, se baseia não só em aspectos qualitativos, mas também em aspectos quantitativos. Além disso, por ter o objetivo de verificar quais empresas têm tendência de crescimento (empresa *long*) e de desvalorização (empresa *short*), em partes se aproxima da análise de temas de investimentos, embora não tenha necessariamente atuação intra-setorial.

Uma análise quantitativa fundamentalista se concentra basicamente na verificação de dados estatísticos para que se possa averiguar a real situação financeira de uma determinada empresa e, assim, o seu real valor de mercado. Uma análise qualitativa fundamentalista, por sua vez, é mais subjetiva, uma vez que leva em consideração fatores de difícil mensuração numérica, como plano de negócios, nível de competitividade, aprovação pública, qualidade de gestão, etc.

Já a abordagem técnica pressupõe no geral a repetição consistente dos padrões passados de comportamento do mercado e de investidores, de modo que os resultados gráficos permi-

tam aos *market timers* projetar e prever o comportamento dos preços e de suas variações futuras. No caso da abordagem técnica *long-short*, não é o padrão de comportamento isolado de um ativo o foco de interesse, mas o padrão existente entre a relação de precificação entre ativos. Neste caso, os esforços não se concentram em determinar o real valor de mercado da empresa, mas a partir da relação histórica do desempenho entre empresas estimar modelos que mais se aproximem do comportamento relativo dos preços de seus ativos.

3.2.2 Fundos Neutros e Fundos Long-Short

É comum que se estabeleça uma certa confusão entre estratégias *long-short* e estratégias de mercados neutros. Estratégias de mercados neutros se referem à um grupo de estratégias de investimento que operam com posições *long* e posições *short* procurando neutralizar certos riscos de mercado. Dentre as principais estratégias de mercado neutro se destacam:

1. *Convertible Arbitrage*: A Arbitragem Conversível constrói portfólios *long* de valores mobiliários conversíveis e protege essa posição com a venda *short* de ações subjacentes de cada obrigação. Valores mobiliários conversíveis incluem inclusive *warrants*, que de uma forma mais geral são valores mobiliários que dão aos seus detentores o direito mas não a obrigação de comprar (*textitcall warrant*) ou de vender (*put warrant*) um determinado ativo (ativo subjacente), a um determinado preço (preço de exercício ou *strike*), numa data predeterminada ou até à data em que esse mesmo direito pode ser exercido (data de maturidade ou vencimento). Após esta data o direito até aí devido deixa de existir.
2. *Fixed-Income Arbitrage*: A Arbitragem de Renda Fixa é uma estratégia que procura explorar diferenças de avaliação e de preços entre vários instrumentos de renda fixa com base na ausência de padrão entre os modelos de precificação absoluta, previsibilidade de oferta/demanda de alguns ativos específicos e da natureza complexa dos ativos de renda fixa. Através da venda de uma quantidade pequena de valores de renda fixa com preços mais elevados e da compra de títulos de renda fixa mais baratos (operações com igual valor absoluto), arbitros de renda fixa se protegem de variações nas taxas de juros.
3. *Merger Arbitrage*: A Estratégia de Fusão normalmente envolve posições compradas em ativos de companhias que estão sendo adquiridas por outras companhias ou se fundindo

a outra companhia e posições vendidas em ações da companhia que está realizando a operação. Alguns *traders* fazem uso de opções quando essa operação permite que a arbitragem do par de ativos seja realizada a valores mais baixos.

4. *Equity Market Neutral*: Essa estratégia procura manter o valor de sua exposição próximo de zero, ou seja, toda a posição *short* é convertida em posição *long*. Em todos os portfólios *equity market neutral* os retornos são derivados do *spread long-short* obtido.
5. *Equity Hedge*: Esta não é propriamente uma estratégia de mercado neutro, mas utiliza uma filosofia de proteção parecida com a dos fundos neutros. Os gestores que utilizam a estratégia *equity hedge* constroem portfólios que combinam posições *long* com posições *short* em ações ou índices de opções de ações, mas com apostas não necessariamente direcionadas.

Com relação à diferença entre as estratégias neutras e as estratégias *long-short*, Ineichen (2001) mostra que apesar de não existir consenso, é possível separar claramente essas duas estratégias. A clareza acerca dessas definições é de fundamental importância no processo de escolha do investimento, uma vez que fundos *long-short* podem manter consideráveis posições direcionais no mercado de ações, o que faz com que o efeito diversificação proporcionado pelas partes *long* e *short* não tenha o resultado esperado, dada a alocação de recursos em estratégias que operam o mesmo mercado ou em ativos com forte relação de dependência.

Segundo Ineichen (2001), o principal objetivo de estratégias neutras ao mercado é obter, no contexto do modelo CAPM, um beta zero, ou seja, gerar retorno de *out performance* das posições *long* e *short* independente da direção que o mercado se move e com baixa exposição. Ainda, o retorno de tais estratégias, caracterizadas por serem neutras aos retornos do mercado, apresentam baixa correlação com o retorno do mercado. Ou seja, mesmo em tempos de alta volatilidade, as estratégias neutras ao mercado se comportam de maneira estável e os resultados geralmente são alcançados com baixa volatilidade.

Outra característica dos portfólios *long-short* neutros ao mercado é que os recursos obtidos com as posições *short* são convertidos quase que integralmente para que se possa assumir as posições compradas, ou seja, praticamente não há o envolvimento de quantias externas à operação. Esse fato faz com que tais portfólios sejam também conhecidos por *dollar neutral portfolios*.

As estratégias *long-short*, entretanto, geralmente são empregadas para explorar ineficiências de mercado e envolvem artifícios de exposição e alavancagem para aumentar a rentabilidade das operações. Note que este tipo de operação não implica necessariamente neutralidade a mercado, uma vez que ativos de quaisquer setores e empresas podem ser utilizados, ou seja, uma vez que não existe uma relação comprovada entre os dois grupos de ações que assegure um beta neutro ao risco.

A próxima seção se destina a apresentação de duas estratégias *long-short* utilizadas no mercado: a estratégia *Pair Trading* e a estratégia Neutra ao Mercado *Long-Short* com Portfólios *Index Tracking*, ambas fundamentadas em abordagens estatísticas.

3.2.3 Estratégia Pair Trading: Arbitragem Estatística

A partir dos anos 80 a estratégia que envolve arbitragem entre pares de ativos, conhecida como *pair trading*, começou a ser utilizada com grande frequência no mercado financeiro e principalmente nos fundos de investimentos. O precursor dessa estratégia, Nunzio Tartaglia, a utilizou pela primeira vez em *Wall Street*, mais especificamente no Morgan Stanley.

O objetivo de Tartaglia era desenvolver uma técnica quantitativa que fosse baseada em métodos estatísticos para selecionar pares de ativos que fossem temporalmente relacionados e que por vezes oferecessem janelas de arbitragem. Sua estratégia consistia basicamente na procura de pares de ativos que tinham tendência de longo prazo conjunta. Selecionados esses pares, o monitoramento diário do comportamento relativo dos preços dos respectivos ativos indicaria a existência no curto prazo de anomalias para o diferencial desses preços. Ao identificar tais discrepâncias no comportamento relativo de precificação, Tartaglia arbitrava com esses ativos até que a diferença relativa estivesse corrigida. Tal arbitragem consistia na venda do ativo relativamente sobrevalorizado e na compra do ativo relativamente subavaliado.

Note que a estratégia de Tartaglia contornava a dificuldade em se identificar ativos cujos preços não estavam de acordo com a real precificação, ou seja, com os valores justos de mercado. Isso porque o objetivo da estratégia não era estruturar metodologias de correção de precificação, mas utilizar o histórico relativo dos preços e as precificações presentes já estabelecidas para verificar discrepâncias no comportamento relativo de curto prazo¹⁹.

Os avanços na estruturação dessa estratégia foram impulsionados pelos trabalhos dos econometricistas Engle e Granger (1987). Tais avanços estão relacionados à forma com que os pares de ativos passaram a ser selecionados. Engle e Granger (1987) verificaram que apesar

¹⁹É exatamente nesse fundamento que se baseia a metodologia de arbitragem que será proposta.

de duas séries temporais serem não estacionárias²⁰, uma relação linear que seja estacionária pode existir para essas duas séries (processo que ficou conhecido como cointegração²¹). Esse processo passou a ser utilizado para que se verificasse se dois ativos têm relação histórica linear relativa de precificação.

A utilização de modelos de regressão que envolvem séries temporais não estacionárias pode se relacionar ao problema que ficou conhecido como regressão espúria. Uma regressão é dita espúria quando apresenta um coeficiente de ajuste bastante elevado, mas que na verdade não indica relação significativa entre as variáveis. Conforme Gujarati(2000), essa incoerência de informações ocorre devido ao fato de que a presença de uma tendência (crescente ou decrescente) em ambas as séries produz um alto valor do coeficiente de ajuste, mas não necessariamente a presença de uma relação de dependência linear verdadeira entre as séries. Como os fundamentos da estratégia *Pair Trading* estão baseados em tal relação de dependência, os testes de cointegração assumiram grande relevância no campo das finanças, uma vez que permitem verificar se existe equilíbrio (ou relação) de longo prazo entre os preços dos ativos.

Para melhor entender o papel da cointegração na escolha de pares de ativos com relação de longo prazo, é necessário que se compreenda os fundamentos desse processo. O primeiro passo para verificar a existência de cointegração entre duas series temporais é verificar a ordem de integração das variáveis, que pode ser obtida através de um teste de raiz unitária²².

Verificada a igualdade da ordem de integração para duas séries temporais p_t^x e p_t^y , estima-se a regressão de cointegração

$$u_t = p_t^x - \lambda p_t^y + \epsilon_t$$

Caso a série dos resíduos ϵ_t (também chamada de erro de equilíbrio) seja estacionária, então dizemos que as séries p_t^x e p_t^y são cointegradas, ou seja, que existe uma relação de equilíbrio

²⁰Dizemos que um conjunto de variáveis aleatórias $\{X_t\}_{t \in N}$ é uma série estacionária se $E(X_t)$ é constante, se $Var\{X_t\}$ é finita e se a covariância independe dos índices t_i e t_j referentes a cada variável, mas da diferença $k = t_i - t_j$.

²¹Basicamente, há três tipos de testes de cointegração. Um dos mais utilizados é o teste de Engle-Granger, desenvolvido por Engle e Granger (1987). Os outro dois testes são o de Phillips-Ouliaris, apresentado por Phillips e Ouliaris (1990) e o teste de Johansen, desenvolvido por Johansen e Juselius (1990).

²²A ordem de integração de uma série se refere ao número de diferenciações necessárias para que a série se torne estacionária. Entre os principais testes de raízes unitárias, os mais utilizados são os testes Dickey-Fuller Aumentado (ADF), conforme apresentado em Dickey e Fuller (1979 e 1981), e o teste não paramétrico Phillips-Perron (PP) desenvolvido por Phillips e Perron (1988).

no longo prazo para essas variáveis²³.

O mecanismo que relaciona o desequilíbrio de curto prazo com o equilíbrio no longo prazo é conhecido como correção de erro e é representado por:

$$\begin{aligned} p_t^y - p_{t-1}^y &= \alpha^y (p_{t-1}^y - \lambda p_{t-1}^x) + \epsilon_t^y \\ p_t^x - p_{t-1}^x &= \alpha^x (p_{t-1}^y - \lambda p_{t-1}^x) + \epsilon_t^x \end{aligned}$$

Note que os lados esquerdo das equações acima mostram o incremento para cada série em cada intervalo de tempo. α^x e α^y representam as taxas de correção do erro dos ativos X e Y, respectivamente, e indicam a velocidade com que as séries temporais voltam ao equilíbrio de longo prazo.

Considerando o resultado de um portfolio F com posição comprada no ativo Y e vendida em X, cada qual em quantidades iguais a λ , temos:

$$F = [\log p_{t+j}^y - \log p_t^y] - \lambda [\log p_{t+j}^x - \log p_t^x]$$

Rearranjando os termos do lado direito, obtemos:

$$[\log p_{t+j}^y - \lambda \log p_{t+j}^x] - [\log p_t^y - \lambda \log p_t^x] = \text{spread}_{t+j} - \text{spread}_t$$

onde u_t representa o spread da operação.

A partir da equação acima, temos que o retorno do portfolio F será o incremento ao valor do *spread* no período de tempo j . A análise desse *spread* é fator de decisão para o árbitro: o gestor de investimentos poderá fazer sua estratégia quando o *spread* entre as ações escolhidas chegar a um determinado patamar e se desfazer dela quando o *spread* for nulo.

Note entretanto que o *spread* pode aparecer para ambos os lados do ponto de equilíbrio, ou seja, o gestor pode dar início à estratégia *long-short Pair Trading* quando houver um desvio de ϕ para mais ou para menos em relação ao equilíbrio de longo prazo μ (considerando a relação de retornos de X sobre Y).

Note que a utilização de cointegração para a escolha do *pair trading* indica a fundamentação dessa estratégia nas hipóteses de reversão à média e estacionariedade. Embora esse tipo de análise seja bastante coerente, não parece ser viável a sua utilização para a análise de todos os possíveis pares do mercado acionário, uma vez que o procedimento utilizado, tal

²³Note que a verificação de uma relação de equilíbrio de longo prazo não exclui no curto prazo as possibilidades de desequilíbrio

como descrito acima, se mostra bastante trabalhoso. Além disso, o fato de séries de preços de ativos não serem cointegradas não exclui a possibilidade de janelas de arbitragem de curto prazo²⁴.

É de extrema importância, portanto, que ferramentas menos trabalhosas e com maior poder de abrangência sejam utilizadas para a detecção de janelas de arbitragem entre pares de ativos. Em trabalhos posteriores poderão ser comparadas as eficiências da metodologia de cointegração e da abordagem probabilística de valores extremos para pares de ativos cointegrados. Poderão ser identificados ainda a partir da nova metodologia proposta pares de ativos não cointegrados que ofereçam melhores oportunidades de arbitragem no curto prazo.

3.2.4 Estratégia Neutra ao Mercado *Long-Short* com Portfólios *Índex Tracking*

Uma estratégia bastante utilizada pelos gestores de fundos de investimento se baseia na construção de um portfólio que em termos de retorno e volatilidade se comporta como uma réplica de um determinado *benchmark*. Ao elaborar tal estratégia (denominada *Índex Tracking*) o gestor deve ficar atento à diferença entre o retorno do portfólio²⁵ e o retorno do *benchmark* a ser replicado.

A composição do portfólio não necessariamente precisa ser formada por todas as ações que compõem o índice. Se o portfólio é composto por um número de ações menor que o número que compõe o índice, então possivelmente existirá um erro de trajetória (*tracking error*). Esse erro de trajetória faz com que o retorno do portfólio não seja exatamente o mesmo retorno do índice. O retorno dessa estratégia está sustentado por esse erro, ou seja, consiste em formar uma carteira que, embora se comporte como uma réplica do índice, produza um *tracking error* positivo, ou seja, um retorno superior ao retorno do índice.

Em termos mais gerais, o *tracking error* pode ser interpretado como uma medida da discrepância ou uma margem de erro que se pode esperar no processo de construção da carteira. Estatisticamente podemos definir esse erro como o desvio padrão da diferença entre os retornos da carteira e do índice. A análise dessa medida é importante uma vez que ela indica a capacidade do portfólio em imitar os retornos do mercado.

²⁴Tal fato será explorado no capítulo 5 a partir da arbitragem do *pair trading* não cointegrado *VIVO4.SAxTCSL4.SA*

²⁵Uma carteira formada por ações que buscam seguir um determinado índice é denominada carteira *índex tracking* ou *tracking basket*.

Considerando o *tracking error* como uma variável aleatória, seria interessante a verificação de algumas características, a saber:

- em uma situação ideal, o comportamento dessa variável se daria segundo um ruído branco²⁶. Esse fato faria com que o portfólio *tracking* não apresentasse desvios inconsistentes com relação às variações do *benchmark*.
- Como o objetivo da estratégia é replicar o *benchmark* em termos de volatilidade e retorno, seria interessante que o portfólio fosse contruído de forma que seus retornos fossem altamente correlacionados com os retornos do *benchmark*. Note que como na maior parte das vezes o número de ativos que compõem o portfólio é inferior ao número de ativos que compõem o índice, deve ser frequente o rebalanceamento do peso dos ativos na carteira bem como da própria diversificação do portfólio.

Note que se as alocações em um portfólio são feitas para replicar um índice, então o portfólio deve ser cointegrado com o índice (no curto prazo o portfólio pode até se desviar do índice, mas eles devem estar linearmente relacionaodos no longo prazo). O uso de cointegração no lugar de correlação para a construção do portfólio permite o uso total das informações contidas nos preços das ações, o que faz com que a composição do portfólio esteja baseada no comportamento de longo prazo dos ativos. Além disso, o processo de cointegração reduz a quantidade de transações para reequilibrar o portfólio, uma vez que garante maior estabilidade dos pesos das ações no portfólio e convergência do *tracking error* (por construção) para a média (Alexander, 2002).

A construção de um portfólio *index tracking* está fundamentada em duas partes igualmente importantes:

- A primeira conste na seleção das ações que irão compor o portfólio. Note que a qualidade do *index tracking* é altamente dependente desta seleção de ações e que, na ausência de um critério que torne essa seleção endógena, várias composições alternativas devem ser testadas até a escolha final do portfólio.
- A segunda etapa visa determinar os pesos de cada ação no portfólio. Esses pesos podem ser determinados pelos coeficientes da equação de cointegração.

²⁶Um processo estacionário com média zero e variância constante é denominado ruído branco

Uma extensão natural do uso da cointegração na construção de portfólios *index tracking* é empregá-la para o desenvolvimento de estratégias *long-short*. Nesse caso, o objetivo é construir portfólios que sejam capazes de replicar um *benchmark* para a parte *long* e outro para replicar a parte *short* da estratégia, assumindo posição comprada no primeiro e posição vendida no segundo.

Entretanto, nem sempre é estabelecida uma relação entre os grupos de ativos na abertura da estratégia *long-short*. Normalmente espera-se que os ativos subvalorizados se valorizem mais ou se desvalorizem menos que os ativos sobrevalorizados. Porém, isto não implica em neutralidade a mercado, a não ser que seja estabelecida alguma relação entre os grupos de ações individualmente. Estratégias *long-short* baseadas em cointegração envolvem apenas ativos com comprovada interdependência.

A construção dos portfólios *long* e *short* surge naturalmente a partir da estratégia *index tracking*. Agora, entretanto, podem ser construídos dois portfólios cointegrados e relacionados ao *benchmark*, a saber um portfólio plus e um minus. Os portfólios plus e minus são construídos a partir de ativos que estão relacionados ao *benchmark* a partir de um excesso de retorno anual de $\delta\%$, pra mais e para menos, respectivamente, e distribuídos uniformemente para os retornos diários.

A partir daí são construídas as séries de preços dos portfólios plus e minus, que devem cointegrar com o *benchmark* estabelecido. Evidentemente quanto maior for δ mais difícil se torna a construção de portfólios cointegrados.

Mesmo sendo razoável esperar que portfólios compostos por muitas ações cointegrem com o índice de mercado de referência, não necessariamente verificamos esse fato para quaisquer composições. O procedimento de cointegração, mais uma vez, torna-se bastante trabalhoso do ponto de vista prático e mitiga janelas de oportunidade de possíveis portfólios que, embora não relacionados no longo prazo, oferecem oportunidades de retornos superiores através da arbitragem.

O próximo capítulo se destina à apresentação da Teoria dos Valores Extremos, que será utilizada para modelar as séries financeiras de ativos não necessariamente cointegrados, para que sejam verificados *pair tradings* e, em trabalhos futuros, portfólios que, embora não repliquem o índice de mercado, ofereçam no curto prazo oportunidades de arbitragem.

4 TEORIA DOS VALORES EXTREMOS (TVE)

4.1 Modelagem Univariada dos Máximos

Alguns autores se referem ao artigo de Bortkiewicz (1922) como o marco inicial do desenvolvimento da TVE. Nesse artigo, Bortkiewicz aborda a distribuição do tamanho do intervalo entre o máximo e o mínimo, mas numa amostra com distribuição normal (citado por Mendes(2004)). Em 1927, Fréchet introduz a distribuição assintótica para o máximo de uma amostra.

Entretanto, os fundamentos básicos da Teoria dos Valores Extremos foram inicialmente expostos por Fisher e Tippett (1928), que introduziram os três tipos possíveis de distribuição assintótica dos valores extremos, hoje conhecidas como Gumbel, Fréchet e Weibull. Avanços para o caso bivariado foram obtidos por Gumbel (1954), (1958), Tiago de Oliveira (1958), (1962) e Sibuya (1960). Entretanto, o primeiro a estudar e formalizar a aplicação estatística dessa teoria foi Gumbel, cuja metodologia tem sido frequentemente aplicada. Outras contribuições importantes para o estudo dos valores extremos foram dadas por Genedenko (1943), que mostrou as condições necessárias e suficientes para a existência das distribuições assintóticas dos valores extremos e introduziu a distribuição de valores extremos generalizada.

Vários outros artigos envolvendo aplicações dos resultados citados acima foram desenvolvidos nesse período, principalmente no que se refere à análises de eventos meteorológicos que envolvem inundações e precipitações máximas e outros como fenômenos naturais envolvendo poluição atmosférica, correntes oceânicas, resistência de materiais, etc. Embrechts, Kluppelberg e Mikosch (1997) relatam um acontecimento da década de 50 no qual o governo alemão estabelece um comitê Delta que fica encarregado de modelar estatisticamente inundações catastróficas futuras. Podemos citar ainda as contribuições de Leadbetter, Lindgren e Rootzén (1983), Embrechts, Kluppelberg e Mikosch (1997), e Reiss e Thomas (1997) dentre várias outras²⁷.

Na última década tem crescido a utilização dessa teoria nas áreas de finanças, atuária e engenharia principalmente devido à sua capacidade em melhor quantificar as probabilidades de ocorrência de eventos raros e fora dos padrões de normalidade, uma vez que ela estuda o comportamento estocástico de extremos associado a um conjunto de variáveis aleatórias com mesma função de distribuição F . A distribuição desses extremos é determinada pelas

²⁷Para estudos mais aprofundados da Teoria de Valores Extremos recomenda-se a tese de doutorado intitulada “Cópulas para distribuições generalizadas de valores extremos multidimensionais (Sanfins (2009)).

caudas extremas da distribuição F . Como a modelagem dessas séries requer análises que envolvam distribuições com caudas pesadas, não é pertinente a simples suposição conveniente de normalidade.

Atualmente no mercado financeiro a teoria dos valores extremos tem sido empregada em grande parte para gestão de riscos, embora esta seja uma entre as diversas aplicações em que essa teoria pode ser utilizada dentro de finanças.

Dentre os autores que têm estudado a aplicabilidade da TVE em finanças podemos citar Longin (1996), que estimou as caudas da distribuição de um índice americano, Danielson e Dvries (1997), que calcularam o VaR para retornos financeiros, McNeil (1998), que estudou a estimação do índice de cauda e calculou quantis através da abordagem Block Máxima, Smith (1999) que realizou uma revisão das técnicas utilizadas na TVE, Embrechts (2000) que abordou benefícios e limitações do cálculo do VaR pela TVE e Mendes (2000), que combinou técnicas robustas para calcular riscos financeiros.

Como estamos interessados na análise de quantis extremos para a relação de precificação entre pares de ativos, torna-se necessário o desenvolvimento de técnicas focadas nos valores extremos (máximos ou mínimos) das amostras ou nos valores que excedam determinado limiar para estimarmos a probabilidade de ocorrência de eventos extremos e fora dos padrões da normalidade. É nesse contexto que a Teoria dos Valores Extremos desempenha um papel fundamental, dada a grande capacidade obtida através da modelagem de eventos extremos.

4.1.1 Distribuição Exata do Máximo²⁸

A TVE é um ramo da probabilidade que estuda o comportamento estocástico de extremos associados a um conjunto de variáveis aleatórias com distribuição comum F . As características da distribuição desses extremos (distribuição do máximo/mínimo) são determinadas pelas caudas da distribuição F .

Se analisarmos uma amostra como um todo, podemos encontrar apenas um máximo (ou mínimo) absoluto associado a um número finito de observações. Pode-se resolver esse problema e tornar a disposição das observações passível de análise a partir de valores extremos dividindo a amostra em sub-períodos disjuntos e de mesmo tamanho.

²⁸As definições e principais resultados deste capítulo podem ser vistos em Mendes (2004). As demonstrações dos teoremas e proposições podem ser vistas em Embrechts Kluppelberg e Mikosch (1997).

Para tanto, considere X_1, X_2, \dots variáveis aleatórias i.i.d.²⁹ com função de distribuição F e um conjunto $X = \{X_1, X_2, \dots, X_n\}$ formado por n dessas variáveis. A partir desse conjunto podemos obter uma nova sucessão ao rearranjarmos seus termos em ordem crescente de magnitude, a saber

$$(X_{(n)}) = \{X_{(1)}, X_{(2)}, \dots, X_{(n)}\}$$

onde $X_{(1)} = \min X$ e $X_{(n)} = \max X$.

Definição 4.1 Chamamos de estatística de ordem correspondente á amostra aleatória X o sistema de inequações $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$.

Definição 4.2 A função $F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = [P(X \leq x)]^n = F^n(x)$, $x \in \mathbb{R}, n \in \mathbb{N}$ é dita função de distribuição do máximo.

Observe que na definição acima $P(X_{(n)} \leq x) = [P(X \leq x)]^n$, uma vez que dada a independência das variáveis $P(\bigcap_{i=1}^n [X_i \leq x]) = \prod_{i=1}^n [P(X_i \leq x)]$.

Observação 4.1 Embora na maior parte das vezes tratemos a TVE a partir de uma abordagem das observações máximas, os mesmos resultados podem ser utilizados na abordagem dos mínimos, dada a conversão imediata

$$\min(X_1, X_2, \dots, X_n) = -\max(-X_1, -X_2, \dots, -X_n)$$

A seguir serão apresentadas as definições de convergência em probabilidade e convergência em distribuição que serão utilizadas para que se verifique que o máximo converge em probabilidade para o limite superior x_F do suporte da distribuição F .

Definição 4.3 Uma sequência de variáveis aleatórias (X_n) converge para uma variável aleatória X em probabilidade se para todo $\epsilon > 0$ tivermos

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) = 0$$

e denota-se $X_n \xrightarrow{p} X$ ou $\text{plim} X_n = X$.

²⁹O apêndice F estende a aplicação da TVE para sequências de variáveis aleatórias (X_n) que apresentam estrutura de dependência. A extensão dos resultados obtidos a partir de sequências i.i.d. para sequências fracamente estacionárias é extremamente importante, dado que a maior parte das séries financeiras apresentam dependência temporal.

Corolário 4.1 Se $X_n \xrightarrow{p} X$, então para todo $\epsilon > 0$ tem-se $\lim_{n \rightarrow \infty} P(|X_n - X| < \epsilon) = 1$

Definição 4.4 Sejam (F_1, F_2, \dots) uma sequência de funções de distribuição relacionadas às variáveis aleatórias X_1, X_2, \dots , respectivamente, e F a função de distribuição de X . Dizemos que (X_n) converge em distribuição para X e denotamos $X_n \xrightarrow{d} X$ se

$$\lim_{n \rightarrow \infty} F_n(a) = F(a)$$

para todo $a \in C$, com $C = \{x \in \mathbb{R}/F \text{ é contínua em } x\}$.

Corolário 4.2 Se $X_n \xrightarrow{p} X$, então $X_n \xrightarrow{d} X$

Interessados na análise da cauda da distribuição F a partir da distribuição do máximo, notemos que o máximo converge em probabilidade para o limite superior x_F do suporte da distribuição F , ou seja

$$X_n \xrightarrow{d} x_F$$

quando $n \rightarrow \infty$ e para $x_F \leq \infty$, sendo $x_F = \sup\{x \in \mathbb{R}/F(x) < 1\}$, uma vez que para n suficientemente grande:

- se $x < x_F$, então $P(X_{(n)} < x) = F^n(x) \rightarrow 0$
- se $x \geq x_F$, com $x_F < \infty$, então $P(X_{(n)} \leq x) = F^n(x) \rightarrow 1$.

Note entretanto que para valores de n pequenos é necessário que se conheça previamente a distribuição F (que em geral é desconhecida) para que se estabeleça a distribuição do máximo e que, para valores de n suficientemente grandes, a função de distribuição do máximo torna-se degenerada, ou seja, $P(X_{(n)} \leq x) \rightarrow 0$.

Fisher e Tippett, contudo, estabeleceram em 1928 um resultado no qual a distribuição dos máximos padronizados por sequências de constantes (a_n) e (b_n) converge para certas distribuições limite, denominadas distribuições extremas ou distribuições EV (*Extreme Value*). Esse resultado é hoje conhecido como Teorema de Fisher-Tippett.

4.1.2 Distribuições Limite para Máximos

Teorema 4.1 (Teorema de Fisher-Tippett) Seja (X_n) uma sucessão de variáveis aleatórias *i.i.d.* com função de distribuição F . Se existir uma sequência (a_n) de termos positivos, uma sequência real (b_n) e uma função de distribuição H não degenerada tais que

$$P[X_{(n)} \leq a_n x + b_n] = F^n(a_n x + b_n) \rightarrow H(x)$$

então as únicas formas possíveis de H são as distribuições Gumbel, Fréchet ou Weibull, também denominadas distribuições do tipo I, II e III respectivamente.

Prova 4.2 Ver Embrechts Kluppelberg e Mikosch (1997).

Observe que o problema em se modelar séries que apresentem distribuições teóricas com caudas pesadas é solucionado pelo Teorema de Fisher-Tippett, uma vez que a convergência dos máximos torna possível o cálculo em probabilidade do valor do suporte máximo da distribuição, que passa a ser entendido como uma variável aleatória com função de distribuição conhecida. A seguir estão apresentadas as distribuições limite para máximos:

- Gumbel ($\lambda = 0, x \in \mathbb{R}$) : $H_I(x) = \exp[-\exp(-x)]$
- Fréchet ($\lambda > 0, x \in \mathbb{R}$) : $H_{II}(x) = \begin{cases} \exp(-x)^{-\lambda} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$
- Weibull ($\lambda < 0, x \in \mathbb{R}$) : $H_{III}(x) = \begin{cases} \exp[-(-x)^{-\lambda}] & , x \leq 0 \\ 1 & , x > 0 \end{cases}$

Proposição 4.1 Se $X > 0$, então as distribuições EV se relacionam da seguinte forma:

$$X \sim \text{Fréchet}(\lambda) \Leftrightarrow \ln X^\lambda \sim \text{Gumbel} \Leftrightarrow -X^{-1} \sim \text{Weibull}(-\lambda)$$

Note que o Teorema de Fisher-Tippett especifica as distribuições extremas limite para as quais há a possibilidade de convergência da distribuição do máximo padronizado $\frac{X_{(n)} - b_n}{a_n}$, embora não estabeleça as condições que F deve satisfazer para que ocorra tal convergência (ver apêndice A). A seguir seguem exemplos de casos particulares de distribuições que convergem para as respectivas distribuições limite:

- TIPO I: Exponencial, Gama, Normal, Log-normal e Gumbel;
- TIPO II: Pareto, Student, Cauchy, Log-gama, Fréchet;
- TIPO III: Beta, Uniforme, Reflected-power e Weibull;

Observe ainda que as distribuição EV podem ser estimadas sem que se faça necessário a utilização da distribuição F . Uma vez estimada a distribuição limite, entretanto, inferências

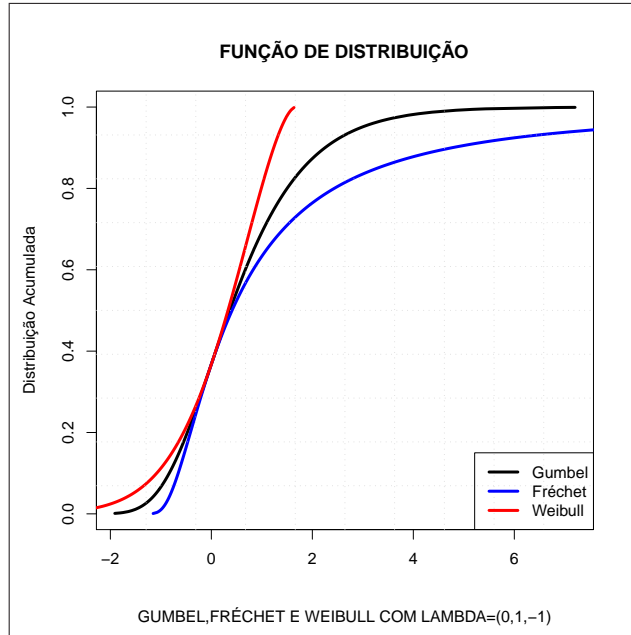


Figura 2: Figura gerada pelo programa estatístico R referente aos gráfico das distribuições extremas com parâmetro de forma iguais a 0, 1 e -1, respectivamente.

acerca da distribuição verdadeira F podem ser feitas, uma vez que de $H(x) = P(X_{(n)} \leq x) = F^n(x)$ obtém-se

$$F(x) = [H(x)]^{\frac{1}{n}}$$

Na figura 2 estão apresentados os gráficos das três distribuições EV e logo abaixo segue um exemplo da convergência da distribuição do máximo proveniente de observações com função de distribuição exponencial unitária para a distribuição Gumbel.

Exemplo 4.1 *Suponha que F seja a distribuição exponencial unitária, ou seja, $F(x) = 1 - \exp(-x)$. Então $[F(x)]^n = [1 - \exp(-x)]^n$. Como*

$$P \left[\frac{M_n - b_n}{a_n} \leq x \right] = P[M_n \leq a_n x + b_n] = [F(a_n x + b_n)]^n,$$

tomando $a_n = 1$ e $b_n = \ln n$, temos que

$$\begin{aligned} [F(x + \ln n)]^n &= [1 - \exp[-(x + \ln n)]]^n \\ &= [1 - \exp(-x - \ln n)]^n \\ &= \left[1 - \frac{\exp(-x)}{\exp(\ln n)} \right]^n \\ &= \left[1 - \frac{\exp(-x)}{n} \right]^n \end{aligned}$$

Passando ao limite e fazendo $\frac{-\exp(-x)}{n} = \frac{1}{z}$, temos

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[1 - \frac{\exp(-x)}{n} \right]^n &= \lim_{z \rightarrow \infty} \left[1 + \frac{1}{z} \right]^{z(-\exp(-x))} \\ &= \lim_{z \rightarrow \infty} \left[\left(1 + \frac{1}{z} \right)^z \right]^{-\exp(-x)} \\ &= \exp(-\exp(-x)) \end{aligned}$$

Ou seja, considerando $a_n = 1$ e $b_n = \ln(n)$ como seqüências normalizadoras, temos que quando n cresce indefinidamente a distribuição do máximo coletado em blocos de tamanho n de observações com função de distribuição unitária será a distribuição Gumbel.

□

A função densidade de probabilidade de cada uma das três distribuições extremas pode ser obtida por diferenciação e aplicação da regra da cadeia, como segue:

$$\begin{aligned} h_I(x) &= \frac{dH_I(x)}{dx} = \frac{d[\exp(-\exp(-x))]}{dx} \\ &= \exp(-\exp(-x))(-1)(-\exp(-x)) \\ &= H_I(x) \exp(-x), \quad x \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} h_{II}(x) &= \frac{dH_{II}(x)}{dx} = \frac{d[\exp(-x^{-\lambda})]}{dx} \\ &= \exp(-x^{-\lambda})(-\lambda)(-1)(x^{-\lambda-1}) \\ &= \lambda H_{II}(x) x^{-(\lambda+1)}, \quad x > 0 \end{aligned}$$

$$\begin{aligned} h_{III}(x) &= \frac{dH_{III}(x)}{dx} = \frac{d[\exp(-(-x)^{-\lambda})]}{dx} \\ &= \exp(-(-x)^{-\lambda})(-\lambda)(-1)(-x)^{-\lambda-1}(-1) \\ &= -\lambda H_{III}(x) (-x)^{-(\lambda+1)}, \quad x < 0 \end{aligned}$$

Na figura 3 estão apresentados os gráficos das fdp associadas a cada uma das três distribuições.

4.1.3 Max-Estabilidade das Distribuições Extremas

Observe que as três distribuições limite foram apontadas como distribuições que convergem para elas mesmas. Isso porque a classe das distribuições extremas coincide com a classe das distribuições max-estáveis, ou seja, assim como no Teorema Central do Limite onde apenas as distribuições estáveis poderiam ser distribuições limites, temos que as distribuições EV, sendo limites, são estáveis (ver apêndice B). As definições de igualdade em distribuição e

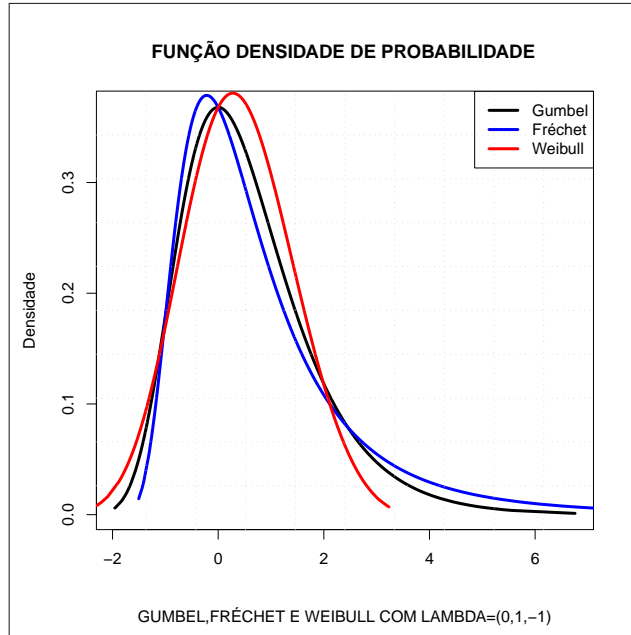


Figura 3: Figura gerada pelo programa estatístico R referente aos gráfico das fdp das distribuições extremas com parâmetro de forma iguais a 0, 1 e -1, respectivamente.

max-estabilidade estão apresentadas a seguir e serão utilizadas para que se possa verificar a max-estabilidade da função de distribuição Gumbel. A verificação da max-estabilidade das demais distribuições limite não será apresentada, embora seja um bom exercício.

Definição 4.5 Dizemos que duas variáveis aleatórias X_1 e X_2 são iguais em distribuição, e denotamos $X_1 \stackrel{d}{=} X_2$, se para todo x tivermos

$$P(X_1 \leq x) = P(X_2 \leq x)$$

Definição 4.6 Sejam X, X_1, X_2, \dots variáveis aleatórias i.i.d. com função de distribuição F e $b_n \in \mathbb{R}$ e $a_n > 0$ seqüências. Se $\max(X_1, X_2, \dots, X_n) \stackrel{d}{=} a_n X + b_n$, então dizemos que F é max-estável.

Note que para as distribuições EV temos $\lim_{n \rightarrow \infty} F_X^n(a_n x + b_n) = H(x)$, ou seja, se X assume uma distribuição extrema, então essa distribuição extrema sendo limite é estável. A seguir será verificada a max-estabilidade da Gumbel.

Exemplo 4.2 (Max-estabilidade da Gumbel) Se X tem uma Gumbel como função de distribuição, então para $a_n = 1$, $b_n = \ln n$ temos

$$\begin{aligned}
\lim_{n \rightarrow \infty} [F_X(a_n x + b_n)]^n &= \lim_{n \rightarrow \infty} [\exp(-\exp(-(a_n x + b_n)))]^n \\
&= \lim_{n \rightarrow \infty} \left[\exp \left[\frac{-\exp(-x)}{\exp(\ln n)} \right] \right]^n \\
&= \lim_{n \rightarrow \infty} \exp[-\exp(-x)n^{-1}]^n \\
&= \lim_{n \rightarrow \infty} \exp[-\exp(-x)] \\
&= \exp(-\exp(-x)) \\
&= H_I(x)
\end{aligned}$$

ou seja, a distribuição de Gumbel sendo limite é estável.

4.1.4 Domínio de Atração da Distribuição do Máximo

Uma implicação direta do Teorema de Fisher-Tippett é que se $F^n(a_n x + b_n)$ é não degenerada para n suficientemente grande e certas constantes $a_n > 0$ e $b_n \in \mathbb{R}$, então

$$\left| F^n(a_n x + b_n) - H \left(\frac{x - b_n}{a_n} \right) \right| \rightarrow 0$$

para alguma $H \in \{H_I, H_{II}, H_{III}\}$. Esse fato nos permite definir uma coleção de funções de distribuição que possuem a mesma distribuição limite.

Definição 4.7 *Chama-se domínio de atração a coleção das funções de distribuição F que possuem a mesma distribuição limite de extremos H e denota-se $F \in MDA(H)$, para alguma $H \in \{H_I, H_{II}, H_{III}\}$.*

Note que embora duas funções F_X e G_X pertençam ao mesmo domínio de atração $MDA(H)$, a velocidade e a qualidade com que essas funções se aproximam da distribuição limite H dependem das características particulares de cada função. A figura 4 compara a distribuição exata do máximo de uma exponencial unitária com sua distribuição limite Gumbel e a figura 5 faz o mesmo utilizando a distribuição normal. Os apêndices C, D e E caracterizam o domínio de atração de cada uma das três distribuições extremas, ou seja, estabelece condições que F deve satisfazer para que o máximo normalizado convirja fracamente para uma dada distribuição EV H .

4.1.5 Distribuição Generalizada de Valor Extremo

Observe que o problema em se modelar séries que apresentem distribuições teóricas com caudas pesadas é solucionado pelo Teorema de Fisher-Tippett, uma vez que a convergência

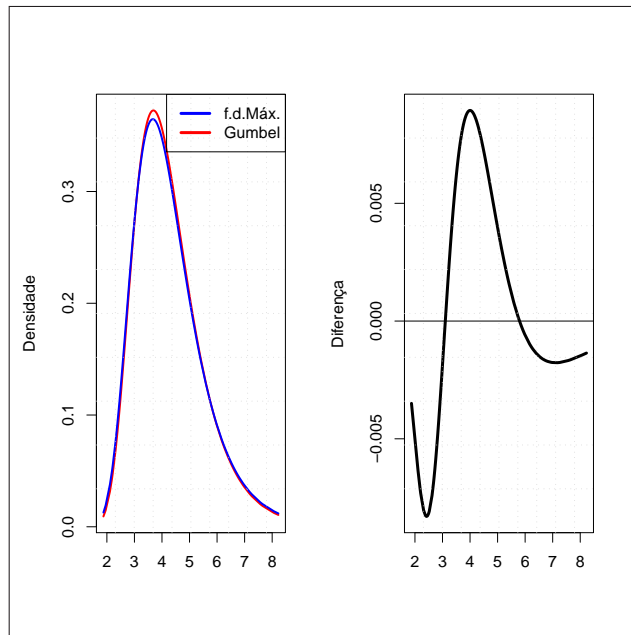


Figura 4: Figura gerada pelo programa estatístico R referente a comparação dos gráficos das fdp das distribuições $F(x) = (1 - \exp(-x))^{40}$ e Gumbel para $\mu = 3, 7$ e $\sigma = 1$.

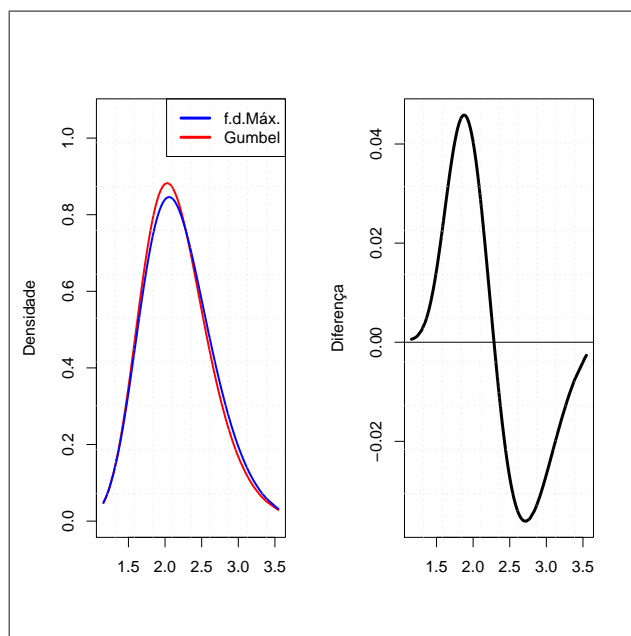


Figura 5: Figura gerada pelo programa estatístico R referente a comparação dos gráficos das fdp das distribuições $F(x) = \left[\frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \right]^{40}$ e Gumbel para $\mu = 1, 9$ e $\sigma = 0, 4$.

dos máximos torna possível o cálculo em probabilidade do valor do suporte máximo da distribuição, que passa a ser entendido como uma variável aleatória com função de distribuição conhecida.

Em trabalhos anteriores aos publicados por Jenkison e von Mises (1955), entretanto, a estimação do parâmetro λ que especifica a distribuição limite era feita supondo-se a convergência da distribuição teórica F para cada uma das distribuições limite, com posterior aplicação de testes para verificação da distribuição extrema com melhor qualidade de ajuste. O avanço teórico introduzido por esses autores foi bastante significativo, uma vez que sob a ξ -parametrização (ou representação de Jenkison-von Mises), as distribuições de Gumbel, Fréchet e Weibull podem ser generalizadas em uma forma denominada Distribuição Generalizada de Valor Extremo (GEV), que representa uma família de funções de distribuição de um único parâmetro ξ . Ou seja, a parametrização das funções extremas permitiu que fossem estimados os parâmetros da verdadeira distribuição limite. Abaixo segue a Distribuição Generalizada de Valor Extremo:

$$G_{\xi}(x) = \begin{cases} \exp\left(-\left(1 + \xi x\right)^{\frac{-1}{\xi}}\right), & \xi \neq 0, \quad 1 + \xi x > 0 \\ \exp[-\exp(-x)], & \xi = 0 \end{cases}$$

Note que para o caso limite ($\xi \rightarrow 0$), temos que G_{ξ} corresponde à distribuição Gumbel, uma vez que

$$\begin{aligned} \lim_{\xi \rightarrow 0} (1 + \xi x)^{\frac{-1}{\xi}} &= \lim_{z \rightarrow 0} (1 + z)^{-\frac{x}{z}} \\ &= \lim_{h \rightarrow \infty} \left(1 + \frac{1}{h}\right)^{-hx} \\ &= \lim_{h \rightarrow \infty} \left[\left(1 + \frac{1}{h}\right)^h\right]^{-x} = \exp(-x) \end{aligned}$$

Ainda, se $\xi < 0$, H_{ξ} corresponde à distribuição Weibull e se $\xi > 0$, H_{ξ} corresponde à distribuição Fréchet.

Na função de distribuição generalizada $G_{\xi}(x)$ é possível substituir x pela transformação de escala e locação $T(x) = \frac{x-\mu}{\sigma}$, $\mu \in \mathbb{R}$, $\sigma > 0$, de modo que a família de locação e escala $G_{\xi, \mu, \sigma}$ correspondente é:

$$G_{\xi}(T(x)) = G_{\xi, \mu, \sigma}(x) = \begin{cases} \exp\left(-\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{\frac{-1}{\xi}}\right), & \xi \neq 0, \quad 1 + \xi \frac{x-\mu}{\sigma} > 0 \\ \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], & \xi = 0 \end{cases}$$

A função densidade de probabilidade (fdp) da distribuição generalizada G_{ξ} pode ser obtida por diferenciação e aplicação da regra da cadeia e segue abaixo:

$$f_i(x) = \begin{cases} (1 + \xi \frac{x-\mu}{\sigma})^{\frac{-1}{\xi}-1} \exp \left[- (1 + \xi \frac{x-\mu}{\sigma})^{\frac{-1}{\xi}} \right] & , \xi \neq 0, \mu \in \mathbb{R}, \sigma > 0 \\ \exp \left[- \exp \left(- \frac{x-\mu}{\sigma} \right) \right] \exp \left(- \frac{x-\mu}{\sigma} \right) \frac{1}{\sigma} & , \xi = 0, x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0 \end{cases}$$

4.2 Estimação

Introduzidas as características e alguns dos principais conceitos acerca das distribuições limite, é necessário que sejam apresentados métodos para a estimação do vetor de parâmetros $\theta = (\xi, \mu, \sigma)$ da distribuição GEV. Os estimadores de L-momentos (ver Hoskings e Wallis(1997)) e Máxima Verossimilhança (ver Embrechts, Kluppelberg e Mikosch(1997)) serão utilizados para a estimação de θ , dadas as boas propriedades e os bons ajustes que esses métodos proporcionam.

4.2.1 Estimadores de Máxima Verossimilhança

Embora introduzido por Fisher(1922), o método de máxima verossimilhança passou a ser utilizado de maneira mais extensiva apenas a partir dos anos 80 em função do desenvolvimento dos computadores de grande potência. A importância desse método se baseia nas propriedades assintóticas dos estimadores, que são consistentes e assintoticamente eficientes. Seu princípio básico fundamenta-se na proposta de se encontrar parâmetros que maximizam a probabilidade de uma certa amostra melhor representar uma determinada população.

Para tanto, considere uma amostra aleatória i.i.d. $Y = \{y_1, y_2, \dots, y_n\}$ proveniente de uma população com função densidade de probabilidade $f(y; \vec{\theta})$, onde y é uma variável aleatória e $\vec{\theta}$ é o vetor de parâmetros dessa distribuição.

Dada a independência das variáveis em Y , temos que a função densidade de probabilidade conjunta dessa amostra é dada por $f(y_1, y_2, \dots, y_n; \vec{\theta}) = f(y_1; \vec{\theta})f(y_2; \vec{\theta}) \dots f(y_n; \vec{\theta})$. Note que essa função indica a probabilidade de obtermos a amostra Y .

Observe que antes da retirada da amostra cada observação é uma variável aleatória cuja fdp é igual a fdp da população. Nesse sentido, consideramos inicialmente $\vec{\theta}$ fixo e y_i variável. Entretanto, uma vez que tenha sido obtida a amostra específica Y , y_i torna-se fixo e $\vec{\theta}$ variável, de modo que a fdp conjunta passa a ser reinterpretada como função do vetor de parâmetros $\vec{\theta}$ e denominada função de verossimilhança ($L(\vec{\theta}; y)$).

A partir da função de verossimilhança $L(\vec{\theta}; y)$ podemos obter o vetor $\hat{\theta}$ que maximiza essa função, ou seja, que faz com que a probabilidade de que seja obtida a amostra Y seja

a maior possível. Na maior parte das vezes trabalha-se com o logarítmo natural da função de verossimilhança, dada a maior simplicidade obtida no processo de maximização. No caso da GEV, a função de log-verossimilhança é dada por

$$L(\xi, \mu, \sigma; m_1, m_2, \dots, m_m) = \ln \prod_{i=1}^m h_{\xi, \mu, \sigma}(m_i) = \sum_{i=1}^m \ln h_{\xi, \mu, \sigma}(m_i)$$

onde $\{m_1, m_2, \dots, m_m\}$ é um conjunto formado por máximos.³⁰

Os estimadores de máxima verossimilhança dos parâmetros ξ, μ e σ podem ser obtidos através da resolução de um sistema de equações não-lineares homogêneo obtido a partir da diferenciação da função de log-verossimilhança $L(\xi, \mu, \sigma; m_1, m_2, \dots, m_m)$ em relação a esses parâmetros.

Os estimadores $\hat{\xi}$, $\hat{\mu}$ e $\hat{\sigma}$ obtidos são os valores em $\mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$ que maximizam a função de log-verossimilhança, ou seja, que maximizam a probabilidade de a amostra formada pelos máximos m_1, m_2, \dots, m_m representar o verdadeiro conjunto de máximos.

4.2.2 Estimadores de L-Momentos

O método dos L-Momentos é uma modificação do *Probability Weighted Moments*(PWM) de Green Wood et al.(1979) e representa uma maneira alternativa de se escrever uma distribuição. Esse método consiste em equacionar um número fixo de L-Momentos amostrais às respectivas quantidades populacionais, de forma que os parâmetros de distribuição possam ser expressos em função dos L-Momentos.

Os L-Momentos $\lambda_{r+1}, r = 0, 1, \dots$ (ver Hosking(1990)) definidos para variáveis aleatórias com esperança finita podem ser escritos como

$$\lambda_{r+1} = \sum_{j=0}^r P_{r,j} B_j, \quad r = 0, 1, 2, \dots \quad (2)$$

onde $P_{r,j} = \frac{(-1)^{r-j}(r+j)}{(j)^2(r-j)}$. Os casos em que r assume os valores 0, 1 e 2 especificam os L-Momentos relacionados com a locação, escala e assimetria da distribuição, respectivamente.

Os L-Momentos amostrais podem ser definidos de maneira análoga e são estimadores não viciados de λ_r , como segue:

$$l_{r+1} = \sum_{j=0}^r p_{r,j} b_j, \quad r = 0, 1, 2, \dots, m-1$$

³⁰ m_i neste caso representa o máximo do bloco i .

No caso da GEV, as equações que relacionam um número fixo de L-Momentos amostrais às respectivas quantidades populacionais são

$$\lambda_1 = \mu - \frac{\sigma}{\xi}[1 - \Gamma(1 - \xi)]$$

$$\lambda_2 = -\frac{\sigma}{\xi}(1 - 2^\xi)\Gamma(1 - \xi)$$

$$\frac{\lambda_3}{\lambda_2} = 2\frac{1-3^\xi}{1-2^\xi} - 3$$

onde Γ representa a função Gama.

A estimação dos parâmetros μ , σ e ξ é feita a partir da substituição dos L-Momentos populacionais pelos respectivos L-momentos amostrais. A resolução da terceira equação nos fornece a estimativa de ξ , que segundo Hosking e Wallis(1997) pode ser aproximada por

$$\hat{\xi} = -7,8590 - 2,9554c^2$$

onde $c = \frac{2}{3 + \frac{l_3}{l_2}} - \frac{\ln 2}{\ln 3}$.

Os parâmetros $\hat{\mu}$ e $\hat{\sigma}$ podem ser obtidos de maneira explícita a partir das expressões

$$\hat{\sigma} = \frac{-l_2 \hat{\xi}}{(1-2^{\hat{\xi}})\Gamma(1-\hat{\xi})}$$

$$\hat{\mu} = l_1 + \frac{\hat{\sigma}}{\hat{\xi}}(1 - \Gamma(1 - \hat{\xi}))$$

4.3 Estatísticas de Teste

No caso da estimação do vetor de parâmetros θ da GEV pelo método de máxima verossimilhança, as estatísticas de Kolmogorov-Smirnov D^+ , D^- e D e a estatística de Kuiper V (Chandra et al(1981)) podem ser utilizadas para se testar a suposição de que a distribuição do máximo dos dados analisados converge para a distribuição GEV, sendo

$$D^+ = \max\left\{\frac{i}{m} - H_\xi(m_{(i)})\right\}$$

$$D^- = \max\left\{H_\xi(m_{(i)}) - \frac{i-1}{m}\right\}$$

$$D = \max\{D^+, D^-\}$$

$$V = D^+ + D^-$$

onde $m_{(i)}$ são os máximos ordenados e H_ξ a distribuição GEV com os parâmetros estimados.

A tabela 1 apresenta alguns valores críticos para as estatísticas teste D^+ , D^- , D e V a níveis de significância de 1% e 5% para amostras de tamanho $m = 50$ e $m = \infty$ (que neste caso representa amostras suficientemente grandes).

Nível de Significância	$\sqrt{m}D^+$	$\sqrt{m}D^-$	$\sqrt{m}D$	$\sqrt{m}V$
m=50				
1%	0,940	0,944	0,988	1,639
5%	0,796	0,796	0,856	1,428
m= ∞				
1%	0,957	0,957	1,007	1,671
5%	0,808	0,808	0,874	1,477

Tabela 1: Tabela de valores críticos para as estatísticas teste D^+ , D^- , D e V a níveis de significância de 1% e 5%.

No caso da estimação do vetor de parâmetros por L-Momentos, o teste de Sherman (Sherman(1957)) pode ser utilizado para se verificar a qualidade do ajuste da GEV, ou seja, para se testar a qualidade da substituição da distribuição exata pela distribuição assintótica.

Considerando a hipótese nula de que a distribuição dos extremos converge para uma GEV, temos a seguinte estatística teste

$$\frac{W_m - E_m}{D_m}$$

assintoticamente normal e cujo p-valor deve ser calculado e comparado com a cauda direita de uma normal padrão, onde

$$W_m = \frac{1}{2} \sum_{i=1}^{m+1} \left| H_\xi(m_{(i)}) - H_\xi(m_{(i-1)}) - \frac{1}{m+1} \right|$$

$$E_m = \left(\frac{m}{m+1} \right)^{m+1}$$

$$E_m^2 + D_m^2 = \frac{2m^{m+2} + m(m-1)^{m+2}}{(m+2)(m+1)^{m+2}}$$

com $H_\xi(m_{(0)}) = 0$ e $H_\xi(m_{(m+1)}) = 1$.

Além dos testes estatísticos formais, algumas análises gráficas podem ser feitas para que seja verificada a qualidade do ajuste da GEV, como:

1. comparação superposta dos gráficos das funções de distribuição empírica e GEV ajustada;
2. comparação das densidades não-paramétrica e estimada, ambas superpostas a um histograma;
3. análise dos resíduos através do QQ-Plot;
4. a comparação dos quantis empíricos com os quantis estimados pela GEV, a partir de um gráfico que tenha os eixos formados por tais quantis;

5 GEV *Long-Short Strategy*: uma nova modalidade quantitativa

5.1 Exemplos de Pair Tradings

Por muitos anos os gestores avaliaram os co-movimentos e interdependência entre os mercados e construíram seus portfólios a partir de análises como as que envolvem o conceito linear de correlação para modelar as interdependências entre ativos financeiros e as que envolvem a distribuição normal. Percebe-se, entretanto, que a modelagem probabilística das séries dos co-movimentos dos preços desses ativos requer distribuições com caudas pesadas e que mesmo pares de ativos não correlacionados podem por vezes oferecer no curto prazo oportunidades de arbitragem. Sendo assim, não é o mais pertinente a simples suposição conveniente de normalidade e a escolha de ações para arbitragem baseada apenas em ativos com relação de interdependência comprovada.

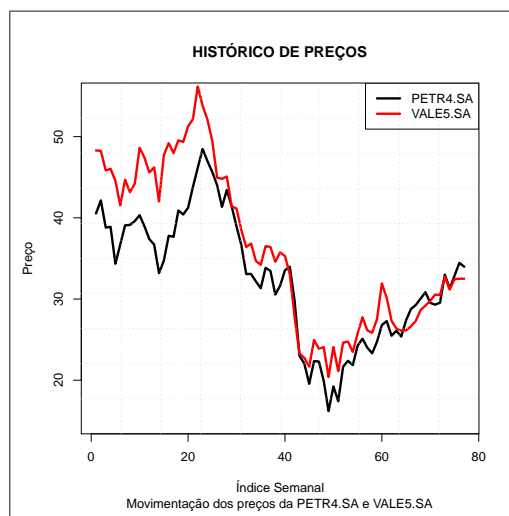


Figura 6: Movimentação dos preços da PETR4 e VALE5 durante as 76 semanas do período de 28/12/2007 a 05/06/2009.

A grande questão que envolve a nova modalidade quantitativa *long-short* que será proposta baseia-se na modelagem (através da TVE) da série obtida a partir do quociente dos preços de *pairs tradings* a serem arbitrados, de forma que seja detectado, com por exemplo 95% de confiança, o momento em que essa razão ocupe quantis extremos da GEV ajustada, ou seja, o momento em que os árbitros possam armar posições comprada (*long*) em

um ativo e vendida (*short*) em outro. Em trabalhos posteriores a TVE poderá ser utilizada para a construção de estratégias de arbitragem que envolvam mais de dois ativos.

Como exemplo das possibilidades de arbitragem *long-short* que existem entre o comportamento de alguns pares de ações, podemos observar a partir da figura 6 um sincronismo no movimento dos preços de duas *blue chips* do mercado brasileiro, a PETR4 e a VALE5, em algumas das 131 semanas do período de 01/01/2005 a 02/07/2007. Através do teste de Dickey e Fuller Aumentado (ADF) verificou-se que tais ativos não são cointegrados.

Note que em algumas dessas semanas, entretanto, existe uma falta de correlação entre os retornos das ações a partir das taxas de crescimento desproporcionais e às vezes até contrárias. É esse comportamento por vezes antagônico que permite aos players do mercado estruturar estratégias *long-short* para gerar retornos adicionais, superiores aos custos de oportunidade e independentes do movimento do mercado.

Embora no exemplo anterior exista um sincronismo entre os preços dos dois ativos considerados, esse sincronismo não é necessariamente requisito para que estratégias *long-short* possam ser construídas. Como mencionado anteriormente, mesmo pares de ativos que não possuam interdependência podem gerar no curto prazo janelas de arbitragem, o que será verificado a partir do exemplo posterior, que também não é cointegrado temporalmente.

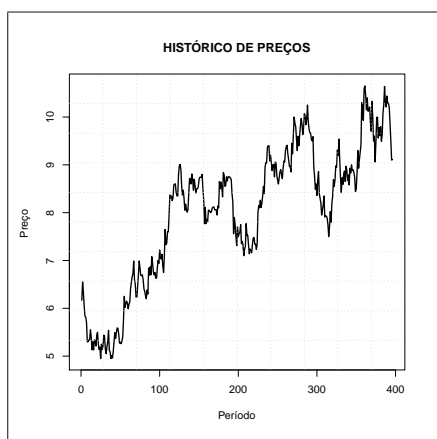


Figura 7: Série histórica da VIVO4.SA para o período compreendido entre 01/06/2006 e 15/12/2007.

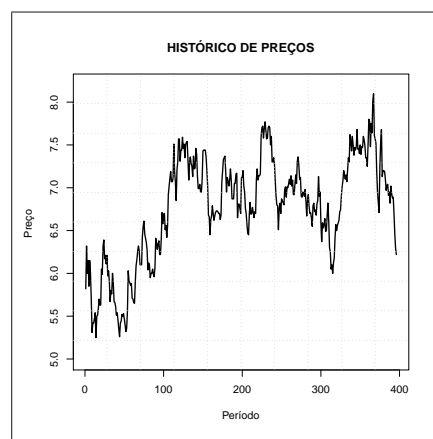


Figura 8: Série histórica da TCSL4.SA para o período compreendido entre 01/06/2006 e 15/12/2007.

A partir das figuras 7 e 8 podemos observar o movimento dos preços da VIVO4 e da TCSL4 em algumas das 81 semanas do período de 01/06/2006 a 15/12/2007. Note que esses ativos não possuem sincronismo para a evolução de seus preços.

Esses dois pares de ativos serão utilizados para a aplicação da nova modalidade qualitativa long-short que será proposta, a qual denominamos *GEV Long-Short Strategy*.

5.2 Proposta de Arbitragem

A estratégia *long-short* mais simples adotada pelos *players*, também chamada de estratégia *number one* ou estratégia *pair trading*, esta baseada no aluguel de uma determinada ação que está sobre-avaliada em relação ao preço de uma outra ação específica. Basicamente o objetivo dessa arbitragem é obter um retorno adicional e que não está relacionado à apreciação ou depreciação desses ativos, mas ao desempenho relativo entre as posições compradas (*long*) e vendidas (*short*).

Nessa operação as ações alugadas são imediatamente vendidas e o montante obtido é totalmente investido na compra da outra ação que compõe o *pair trading*³¹. Ao final do período de maturação do aluguel, espera-se que o *pair trading* esteja com razão de precificação (k) em relação à posição inicial menor o suficiente³² para cobrir os custos gerais da transação e gerar os retornos adicionais desejados, ou seja, espera-se que a venda das ações compradas seja superior ao valor necessário para compra das ações alugadas e que devem ser devolvidas acrescidas da taxa de negociação, que em geral é de 1%.

Para que esse raciocínio seja melhor compreendido, consideremos um exemplo fictício para esse tipo de arbitragem. Para tanto, sejam X e Y dois ativos quaisquer e p_X e p_Y seus respectivos preços. Se em um determinado momento o ativo X estiver sobre-precificado em relação a Y , então a razão $\frac{p_X}{p_Y}$ estará sobre-estimada. Note que em algum momento o mercado irá de forma natural corrigir essa supervalorização relativa, o que fará com que a diferença $\frac{p_X}{p_Y}$ diminua. As posições compradas e vendidas, portanto, seriam formadas em Y e X , respectivamente.

Observe que a sobre-valorização do ativo X em relação a Y não implica que X esteja com preço p_X acima do seu valor real e que Y esteja com preço p_Y abaixo do seu valor real, ou seja, não necessariamente X e Y devem se desvalorizar e valorizar no mercado,

³¹Em estratégias mais sofisticadas o montante obtido com a venda das ações *short* não necessariamente é totalmente convertido na compra das ações *long*. Nesse caso o gestor verifica a possibilidade de lucratividade superior a partir da diversificação.

³²Note que caso k seja obtido através da razão dos preços dos ativos sub-avaliado e sobre-avaliado, respectivamente, então espera-se que a razão k esteja maior o suficiente para que se possa alcançar os resultados desejados.

respectivamente. Isso porque mesmo movimentos de ambos os preços no mesmo sentido podem fazer com que essa razão diminua. Note que essa correção de precificação pode acontecer das seguintes formas:

1. caso p_X aumente proporcionalmente menos que p_Y , então a nova razão $\frac{p_X}{p_Y}$ será menor. Note que neste caso, embora ambos os ativos se valorizem no mercado, o ativo Y se valorizou mais que X .
2. caso p_X diminua proporcionalmente mais que p_Y , então a nova razão $\frac{p_X}{p_Y}$ será menor. Note que neste caso, embora ambos os ativos se desvalorizem no mercado, o ativo X se desvalorizou mais que Y .
3. caso p_X diminua e p_Y aumente, então a nova razão $\frac{p_X}{p_Y}$ será menor.

Oberve que o item 3 é o que traz maior rentabilidade para a operação, uma vez que a estratégia proporciona rentabilidade tanto na posição vendida como na posição comprada (maior spread para a operação). Neste caso o ativo X é comprado a um preço inferior e Y é vendido a um preço superior aos respectivos preços de mercado.

Observe ainda que, mesmo que o mercado opere em queda, é possível que se obtenha rentabilidade positiva. Basta que as posições sejam armadas de forma que o ativo short se desvalorize mais que o ativo long. Esse é, dentre outros, um dos importantes benefícios que as estratégias long-short podem proporcionar, uma vez que além de oportunizarem rentabilidades positivas, protegem a carteira contra riscos sistêmicos de decínio do mercado.

Muitos autores conceituam esse tipo de estratégia como uma estratégia neutra ao mercado (*market neutral*). Ineichen (2001) mostra que apesar de não existir consenso, é possível separar claramente as estratégias neutras e as estratégias long-short. A clareza acerca dessas definições é de fundamental importância no processo de escolha do investimento, uma vez que fundos long-short podem manter consideráveis posições direcionais no mercado de ações, o que faz com que o efeito diversificação proporcionado pelas partes long e short não tenha o resultado esperado (dentro do contexto do modelo CAPM), dada a alocação de recursos em estratégias que operam o mesmo setor no mercado.

O principal objetivo de estratégias neutras ao mercado é obter um beta zero, ou seja, gerar retorno de *out performance* para posições long e short independente da direção do mercado e com baixa exposição. O retorno de tais estratégias, caracterizadas por serem neutras aos retornos do mercado, apresentam baixa correlação com o retorno do mercado. Ou

seja, mesmo em tempos de alta volatilidade, as estratégias neutras ao mercado se comportam de maneira estável e os resultados geralmente são alcançados com baixa volatilidade.

As estratégias long-short, entretanto, geralmente são empregadas para explorar ineficiências de mercado e envolvem artifícios de exposição e alavancagem para aumentar a rentabilidade das operações. A rentabilidade desse tipo de operação vem do *spread* proporcionado entre as variações das cotações inicial e final tanto da parte long quanto da parte short. Note que este tipo de operação não implica necessariamente neutralidade a mercado, uma vez que ativos de quaisquer setores e empresas podem ser utilizados, ou seja, uma vez que não se requer uma relação comprovada entre os dois grupos de ações que assegure um beta neutro ao risco. A construção de um portfólio com beta zero só seria possível, portanto, caso fossem mantidas posições compradas (long) e vendidas (short) em ativos com relação comprovada de dependência.

Como mencionado na seção anterior, a nova modalidade quantitativa *long-short* proposta baseia-se na modelagem (através da TVE) da série obtida a partir do quociente dos preços de *pairs tradings* a serem arbitrados, para que seja detectado o momento em que os árbitros possam armar posições comprada (long) em um ativo e vendida (short) em outro.

Para tanto, será primeiramente estipulado um período para coleta dos dados diários de cada um dos ativos que compõem o *pair trading*. Esse período será escolhido de forma que contemple a existência de ambos os ativos, ou seja, de forma que tenhamos informações sobre os preços dos ativos para todas as datas do período escolhido³³.

O número de dados coletados deve ser suficientemente grande para que possamos dividir a série do quociente dos preços do *pair trading* a ser arbitrado em blocos (abordagem Block Máxima).

Para que possamos verificar o melhor ajuste da GEV e garantir uma melhor estabilidade da distribuição ajustada, será feita uma análise comparativa através de modelagens com tamanhos de bloco iguais a 5, 10, 15 e 22, de modo que sejam obtidas quatro séries com os valores das razões máximas e mínimas para períodos de 1, 2 e 3 semanas e 1 mês, respectivamente.

Tal análise comparativa será estruturada da seguinte forma:

1. Após obtidas as séries associadas aos blocos citados será feita uma verificação da exis-

³³As modelagens que serão feitas na próxima seção não incorporam em seu período de análise a crise financeira de 2008. Tal fato deve ser destacado, uma vez que tal crise teve impactos singulares no comportamento de muitos ativos. Esse fato não impede que tais impactos sejam estudados em trabalhos posteriores

tência de *outliers*, que são observações que se apresentam relativamente afastadas das demais ou que são inconsistentes com o processo gerador do processo estocástico analisado. Para tanto será utilizado o método exploratório gráfico Box Plot, que pode ser estruturado da seguinte forma:

- (a) Em um primeiro momento calcula-se a mediana, o quartil inferior (Q_1) e o quartil superior (Q_3);
 - (b) A seguir encontra-se o valor L de referência, dado por $L = Q_3 - Q_1$;
 - (c) Os valores que estiverem acima do quartil $Q_3 + 1,5L$ e abaixo do quartil $Q_1 - 1,5L$ serão considerados outliers. Tais valores serão retirados da respectiva série.
2. Após a retirada dos *outliers* serão obtidas a partir do método de máxima verossimilhança as estimativas para os parâmetros ξ , μ e σ da GEV ajustada, a partir da modelagem das séries dos máximos e mínimos. Um procedimento análogo será realizado para que esses parâmetros sejam estimados para as séries iniciais, caso seja verificada a presença de *outliers*³⁴ ;
 3. A estatística de Kolmogorov-Smirnov será utilizada para se testar a suposição de que a distribuição do máximo dos dados analisados converge para a distribuição GEV.
 4. Algumas análises gráficas comparativas serão feitas para que seja verificada ainda a qualidade do ajuste da GEV para os diferentes tamanhos de blocos considerados, como:
 - (a) a comparação superposta dos gráficos das funções de distribuição empírica e GEV ajustada;
 - (b) a comparação das densidades não-paramétrica e estimada, ambas superpostas a um histograma;
 - (c) a análise dos resíduos através do QQ-Plot.
 - (d) a comparação dos quantis empíricos com os quantis estimados pela GEV, a partir de um gráfico que tenha os eixos formados por tais quantis;

³⁴Caso seja verificada a presença de *outliers*, recomenda-se que sejam estimados intervalos com 95% de confiança para cada parâmetro estimado a partir da suposição de que cada parâmetro segue uma distribuição normal. Tais intervalos podem ser utilizados para que se possa verificar se o não tratamento dos dados representa fator prejudicial à estimação dos parâmetros e, assim, à modelagem da função GEV, ou seja, se a modelagem da GEV para o par de ativos considerado é ou não influenciada pela presença de *outliers*;

A análise dos dados bem como as estimações dos parâmetros e demais análises serão feitas com o uso do programa estatístico R-Project e a partir do pacote fExtremes, específico para análise dos valores extremos.

Ajustada a GEV para as séries dos máximos e mínimos e verificado para qual tamanho de bloco se tem maior max-estabilidade, serão calculados os quantis 95% e 5% para as séries dos máximos e mínimos referentes à esse bloco, respectivamente.

Definição 5.1 (Quantil *GEV Long-Short*) *Seja $\{X_1, X_2, \dots, X_n\}$ um conjunto de variáveis aleatórias que representam a razão de precificação entre ativos e com mesma função de distribuição F . Denominamos Quantil *GEV Long-Short* o quantil 95% (ou 5%) obtido através da distribuição extrema para o qual a distribuição dos máximos (ou mínimos) padronizados converge, respectivamente.*

Os quantis *GEV Long-Short*, juntamente com os quantis fornecidos pela normal, serão superpostos à série das razões para um período subsequente ao período analisado. Como inferências acerca da verdadeira distribuição F da série podem ser feitas a partir da GEV ajustada, serão superpostos também nesse gráfico os verdadeiros quantis da série, ambos obtidos a partir da GEV estimada. Esses quantis serão utilizados para que sejam verificados os momentos em que a razão de precificação ocupe quantis extremos.

Definição 5.2 (Saturação relativa do desequilíbrio) *Define-se por saturação relativa do desequilíbrio o momento em que ativos encontram-se com relação de precificação fora do intervalo definido pelos quantis *GEV Long-Short*.*

A partir da detecção de tais janelas (pontos de saturação relativa do desequilíbrio), serão estruturadas sub-estratégias de arbitragem que iniciem em posições compradas e vendidas e que liquidem ou aumentem as posições, a depender do comportamento dos retornos oferecidos pelo par de ativos arbitrado. Um nível de rentabilidade para a operação de 3%³⁵ será fixado como ponto de *stop* para a operação.

A partir das informações obtidas e da elaboração das sub-estratégias serão calculadas as rentabilidades fornecidas pela abordagem GEV e Normal, para que sejam verificados os reais potenciais de cada modalidade quantitativa. A comparação de tais estratégias será

³⁵O nível de rentabilidade para o desarme da operação não necessariamente deve ser de 3%, mas pode variar de acordo com o risco que o investidor está disposto a correr e com o percentual de rentabilidade desejado.

essencial para que se confirme a suposição de que o melhor ajuste dos quantis extremos através da GEV tem capacidade de fornecer sustentação probabilística mais refinada para que o retorno à média justifique a possibilidade de arbitragem *long-short*.

5.3 Modelagem

Os pares de ativos $PETRA \times VALE5$ e $VIVO4 \times TCSL4$ serão utilizados para a aplicação da nova modalidade quantitativa proposta. Observe que o primeiro caso é um exemplo de arbitragem entre empresas de setores diferentes. Esse par foi escolhido dada a grande representatividade que ambas as ações possuem no total de negociações do mercado brasileiro.

O segundo par se refere à arbitragem entre ativos de um mesmo setor. A escolha desse par foi feita devido ao importante crescimento apresentado pelo setor de comunicações no Brasil.

5.3.1 *Pair Trading* PETRA/VALE5

A seguir será feita uma modelagem comparativa da razão direta entre os preços da PETRA e VALE5 a partir do ajuste da distribuição GEV e Normal. Os dados coletados formam uma amostra de 651 observações dos valores diários dos preços desses ativos do período de 01/01/2005 a 02/07/2007. A série da razão direta $\frac{PETRA}{VALE5}$ está apresentada na figura 9.

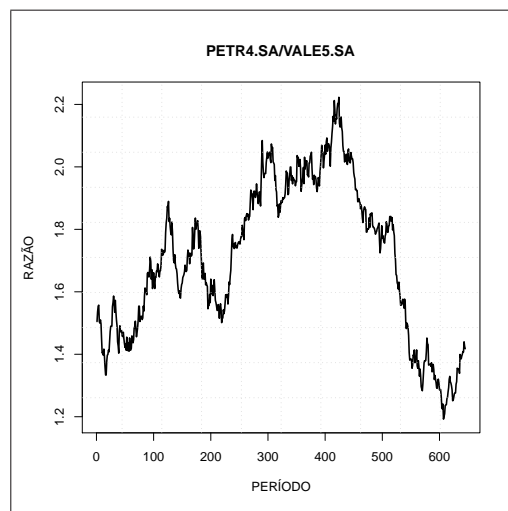


Figura 9: Razão histórica de preços $\frac{PETRA.SA}{VALE5.SA}$ para o período compreendido entre 01/01/2005 e 02/07/2007.

A partir do método de máxima verossimilhança foram obtidos as estimativas para os parâmetros ξ, μ e σ da GEV na modelagem dos máximos e mínimos, respectivamente, para blocos de tamanho 5, 10 15 e 22. Não houve a presença de *outliers* para o período considerado.

As figuras 10 e 11 se referem ao QQ-plot dos resíduos para as modelagens referentes ao máximo e mínimo para os diferentes tamanhos de blocos. Note que na modelagem do mínimo os blocos de tamanho 22 garantiram maior max-estabilidade para a GEV, uma vez que o plot dos resíduos se mostrou melhor ajustado. Já na modelagem dos máximos os blocos de tamanho 15 garantiram maior max-estabilidade.

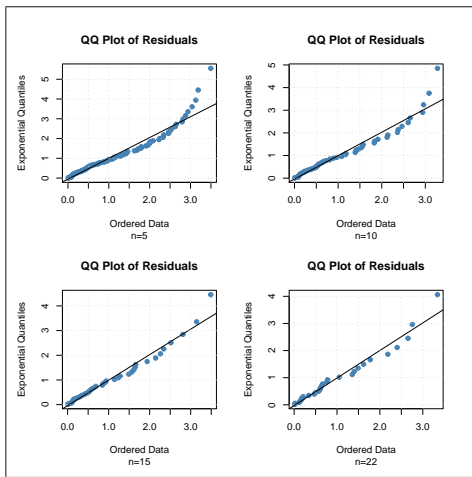


Figura 10: QQ-plot dos resíduos para a modelagem dos máximos.

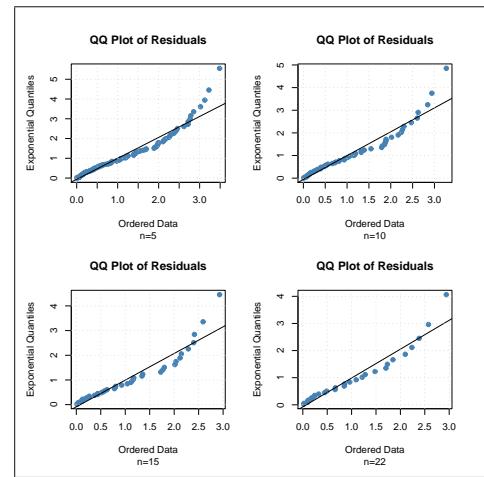


Figura 11: QQ-plot dos resíduos para a modelagem dos mínimos.

A seguir estão apresentados as estimativas dos parâmetros ξ, μ e σ para a modelagem dos máximos e mínimos com blocos de tamanho 15 e 22, respectivamente.

$$\hat{\xi} = -0.3945844 \quad \hat{\mu} = 1.6959661 \quad \hat{\sigma} = 0.2533173$$

$$\hat{\xi} = -0.5216825 \quad \hat{\mu} = 1.5910817 \quad \hat{\sigma} = 0.2604460$$

A estatística de Kolmogorov-Smirnov foi utilizada para se testar a suposição de que a distribuição do máximo e mínimo dos dados analisados converge para a distribuição GEV ajustada. Os p-valores obtidos na modelagem dos máximos e mínimos foram de 0.9429 e 0.9373, respectivamente. À esses valores estão associados as estatísticas D iguais a 0.0771 e 0.0943, respectivamente, todas elas menores que os valores críticos dados na tabela 1.

Note que os valores da estatística D são suficientes para garantirmos boas evidências da boa qualidade do ajuste da GEV.

A qualidade do ajuste da GEV pode ser verificada ainda através das figuras 12 e 13, relativas aos gráficos das densidades empírica versus a GEV ajustada, através das figuras 14 e 15, relativas à comparação superposta dos gráficos das funções de distribuição empírica, normal e GEV ajustada e através das figuras 16 e 17, relativas à comparação das densidades não-paramétrica e estimada, ambas superpostas a um histograma.

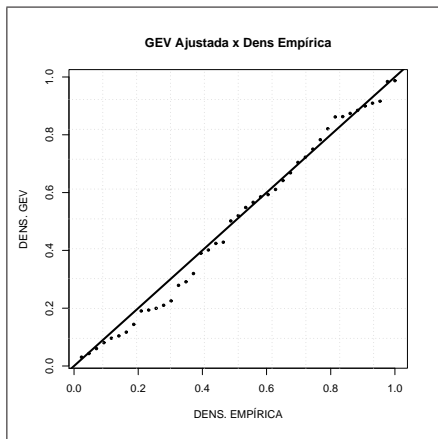


Figura 12: Densidades empírica versus GEV ajustada para a modelagem dos máximos para $n=15$.

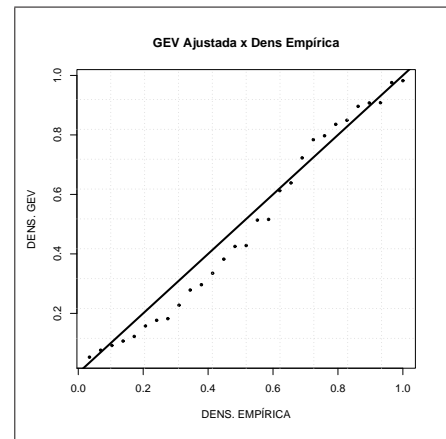


Figura 13: Densidades empírica versus GEV ajustada para a modelagem dos mínimos para $n=22$.

Observe a partir dos gráficos das figuras 12 e 13 que tanto para a modelagem dos máximos quanto dos mínimos a densidade ajustada pela GEV se aproximou da densidade empírica dos dados.

Os gráficos das figuras 14 e 15 apresentam uma comparação superposta das funções de distribuição empírica, normal e GEV ajustada para máximos e mínimos. Esses gráficos nos permitem comparar o ajuste das funções de distribuição GEV e normal. Observe que a GEV melhor estimou as probabilidades acumuladas em ambos os casos e que a normal subestimou as probabilidades acumuladas no caso da análise dos máximos e as sobre-estimou no caso dos mínimos.

Com relação à estimação dos quantis extremos, observe através dos gráficos das figuras 15 e 16 que as caudas da GEV (direita para máximos e esquerda par mínimos) são mais pesadas e melhor quantificam as probabilidades de ocorrência de eventos extremos.

A partir dos parâmetros obtidos, a GEV ajustada para mínimos e máximos nos fornece

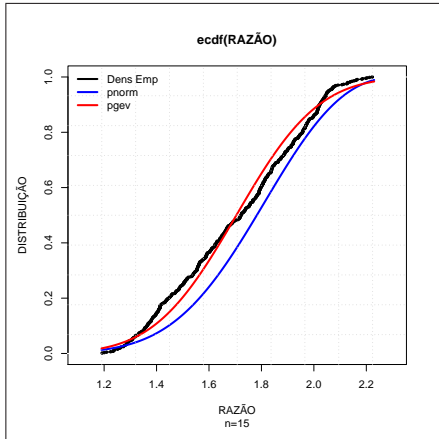


Figura 14: Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos máximos.

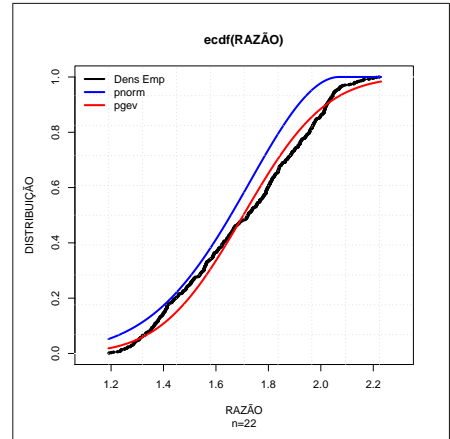


Figura 15: Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos mínimos.

1.183819 e 2.139097 como quantis 5% e 95% para a razão $k = \frac{PETR4.SA}{VALE5.SA}$, respectivamente. Utilizando-se uma distribuição normal, obtemos para o mesmo nível de significância os valores 1.299182 e 2.109515, respectivamente. A figura 18 compara os quantis de probabilidade obtidos com a performance da razão k para os 131 dias subsequentes ao período analisado. Nessa figura estão também inclusos as inferências acerca dos verdadeiros quantis de probabilidade (obtidos a partir dos quantis fornecidos pela GEV), a saber 1.4960 e 2.1948, respectivamente.

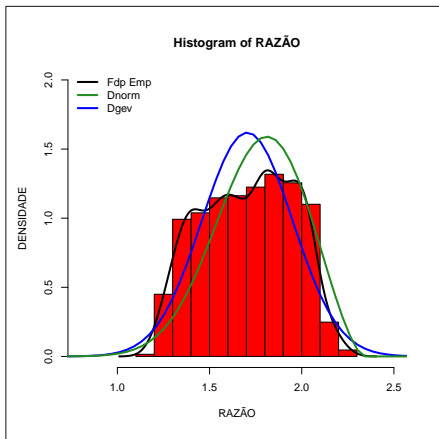


Figura 16: Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos máximos.

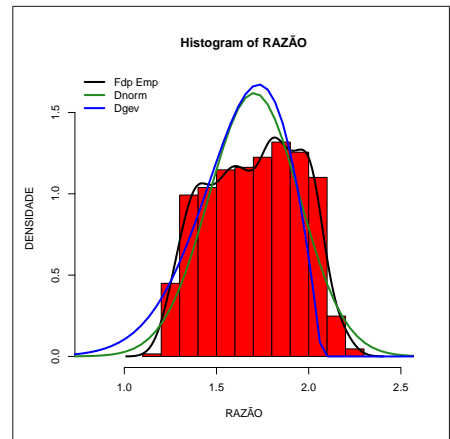


Figura 17: Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos mínimos.

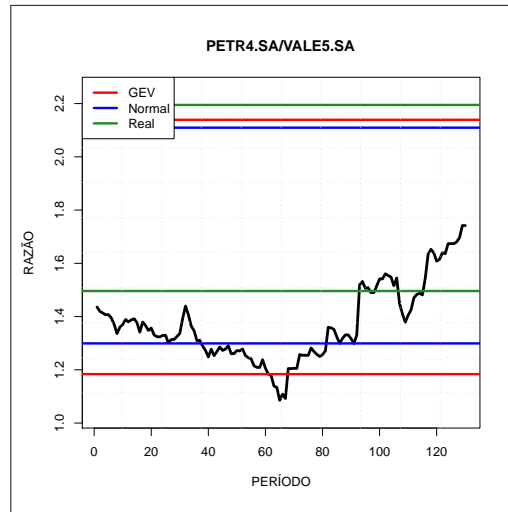


Figura 18: Performance da razão $k = \frac{PETRA.SA}{VALE5.SA}$ para os 131 dias subsequentes ao período compreendido entre os 01/01/2005 e 02/07/2007.

Note que caso os *players* tivessem conhecimento acerca dos quantis estimados pela normal e GEV, a partir do monitoramento diário da evolução da razão entre os ativos a seguinte estratégia de arbitragem poderia ser armada:

1. Acompanhando a evolução da razão $k = \frac{PETRA}{VALE5}$, verifica-se que no dia 08/08/2007 (período 27) poderia-se montar posições comprada em PETR4 e vendida em VALE5, uma vez que nesse dia a razão k atinge o quantil 5% estimado através da normal;
2. Fixando uma rentabilidade acumulada de 3% como ponto de *stop* da operação, verifica-se que o desarme de tais posições aconteceria no dia 15/08/2007 (período 32), cinco dias úteis depois do *start* da operação (ver figura 19), quando alcançamos uma rentabilidade acumulada de 3,98%;
3. Como a modelagem pela normal sobre-estima o quantil inferior em relação ao quantil da GEV, ao invés de simplesmente desarmar a posição, os *players* poderiam inverter as posições, ou seja, assumir posição comprada em VALE5 e vendida em PETR4, uma vez que a razão k ainda não atingiu o suporte inferior fornecido pela GEV;
4. A nova posição poderia se manter armada até que a razão atingisse o quantil inferior indicado pela GEV, o que aconteceu no dia 26/09/2007 (período 62), trinta dias úteis após a inversão, o que nos fornece 9,20% de rentabilidade. O gráfico da figura 20 mostra a rentabilidade acumulada a partir da inversão das posições armadas.

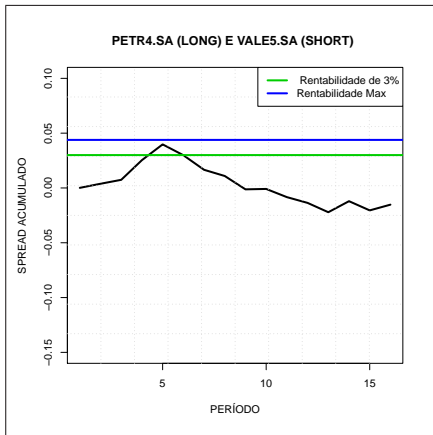


Figura 19: Rentabilidade acumulada das posições a partir do quantil da normal.

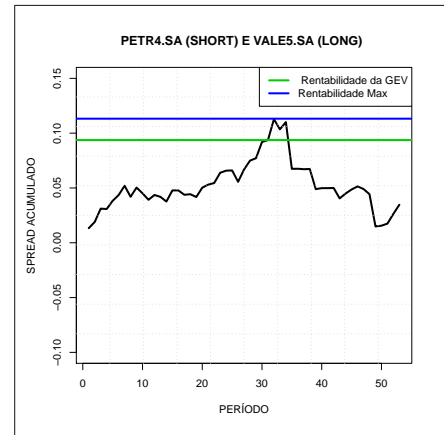


Figura 20: Rentabilidade acumulada a partir da inversão das posições.

Analisando a estratégia como um todo, obteríamos uma rentabilidade conjunta de 13,18% para um período de 35 dias úteis, o que equivale a uma taxa de 143,86%*a.a.*

5.3.2 *Pair Trading VIVO4/TCSL4*

A seguir será feita uma modelagem comparativa da razão direta entre os preços da VIVO4 e TCSL4 a partir do ajuste da distribuição GEV e Normal. Os dados coletados formam uma amostra de 396 observações dos valores diários dos preços desses ativos do período de 01/06/2006 a 15/12/2007. A série da razão direta $\frac{VIVO4}{TCSL4}$ está apresentada na figura 21.

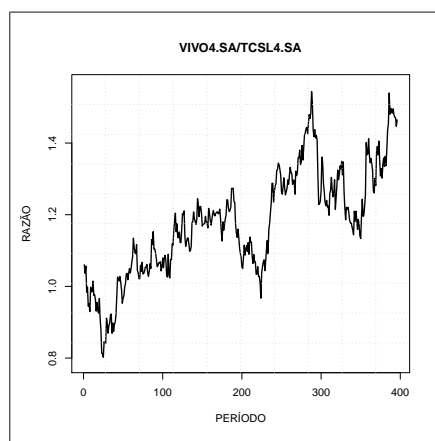


Figura 21: Razão histórica de preços $\frac{VIVO4.SA}{TCSL4.SA}$ para o período compreendido entre 01/06/2006 e 15/12/2007.

A partir do método de máxima verossimilhança foram obtidos as estimativas para os parâmetros ξ, μ e σ da GEV na modelagem dos máximos e mínimos, respectivamente, para blocos de tamanho 5, 10 15 e 22. Não houve a presença de *outliers* para o período considerado.

As figuras 22 e 23 se referem ao QQ-plot dos resíduos para as modelagens referentes ao máximo e mínimo para os diferentes tamanhos de blocos. Note que tanto na modelagem dos mínimos quanto na modelagem dos máximos os blocos de tamanho 5 garantiram maior max-estabilidade para a GEV, uma vez que o plot dos resíduos se mostrou melhor ajustado.

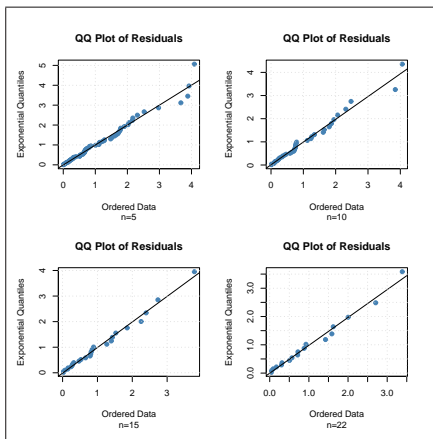


Figura 22: QQ-plot dos resíduos para a modelagem dos máximos.

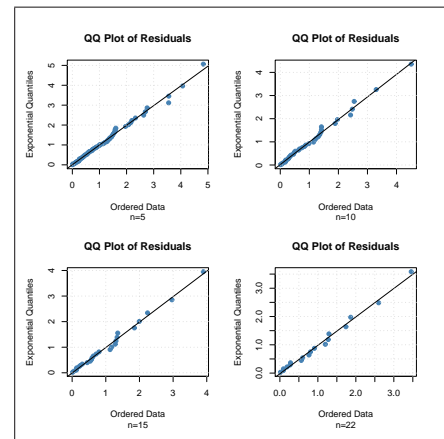


Figura 23: QQ-plot dos resíduos para a modelagem dos mínimos.

A seguir estão apresentados as estimativas dos parâmetros ξ, μ e σ para a modelagem dos máximos e mínimos com blocos de tamanho 5, respectivamente.

$$\hat{\xi} = -0.2814607 \quad \hat{\mu} = 1.1703674 \quad \hat{\sigma} = 0.1508817$$

$$\hat{\xi} = -0.2844813 \quad \hat{\mu} = 1.0923066 \quad \hat{\sigma} = 0.1460131$$

A estatística de Kolmogorov-Smirnov foi utilizada para se testar a suposição de que a distribuição do máximo e mínimo dos dados analisados converge para a distribuição GEV ajustada. Os p-valores obtidos na modelagem dos máximos e mínimos foram de 0.9398 e 0.9903, respectivamente. À esses valores estão associados as estatísticas D iguais a 0.0579 e 0.0495, respectivamente, todas elas menores que os valores críticos dados na tabela 1. Note que os valores da estatística D são suficientes para garantirmos evidências da boa qualidade do ajuste da GEV.

A qualidade do ajuste da GEV pode ser verificada ainda através das figuras 24 e 25, relativas aos gráficos das densidades empírica versus a GEV ajustada, através das figuras 26 e 27, relativas à comparação superposta dos gráficos das funções de distribuição empírica, normal e GEV ajustada e através das figuras 28 e 29, relativas à comparação das densidades não-paramétrica e estimada, ambas superpostas a um histograma.

Observe a partir dos gráficos das figuras 24 e 25 que tanto para a modelagem dos máximos quanto dos mínimos a densidade ajustada pela GEV se aproximou da densidade empírica dos dados.

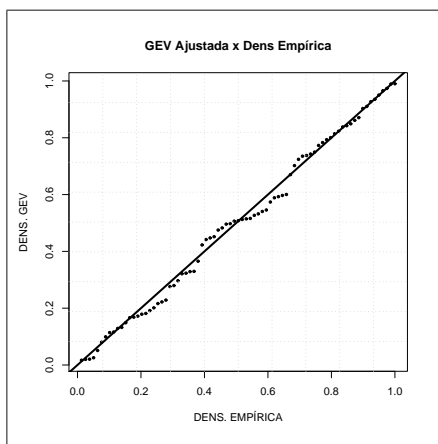


Figura 24: Densidades empírica versus GEV ajustada para a modelagem dos máximos para n=15.

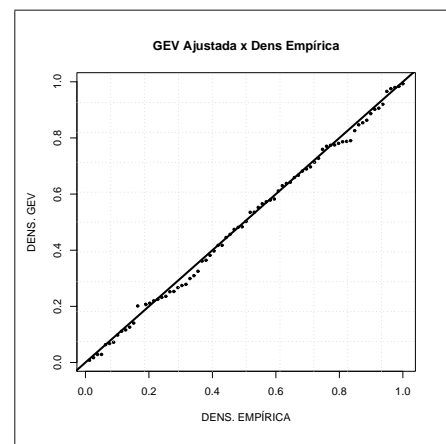


Figura 25: Densidades empírica versus GEV ajustada para a modelagem dos mínimos para n=22.

Os gráficos das figuras 26 e 27 apresentam uma comparação superposta das funções de distribuição empírica, normal e GEV ajustada para máximos e mínimos. Esses gráficos nos permitem comparar o ajuste das funções de distribuição GEV e normal. Observe que a GEV melhor estimou as probabilidades acumuladas em ambos os casos e que a normal subestimou as probabilidades acumuladas no caso da análise dos máximos e as sobre-estimou no caso dos mínimos.

Com relação à estimação dos quantis extremos, observe através dos gráficos das figuras 28 e 29 que as caudas da GEV (direita para máximos e esquerda par mínimos) são mais pesadas e melhor quantificam as probabilidades de ocorrência de eventos extremos.

A partir dos parâmetros obtidos, a GEV ajustada para mínimos e máximos nos fornece 0.9042828 e 1.455642 como quantis 5% e 95% para a razão $k = \frac{VIV04}{TCSL4}$, respectivamente. Utilizando-se uma distribuição normal, obtemos para o mesmo nível de significância os valo-

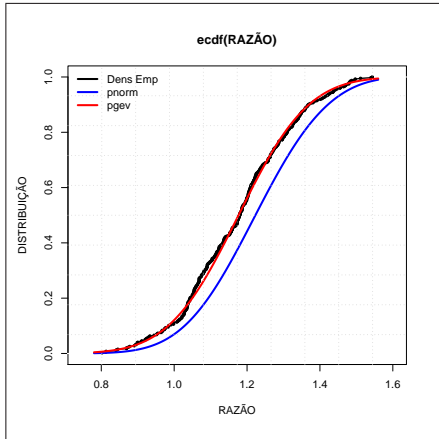


Figura 26: Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos máximos.

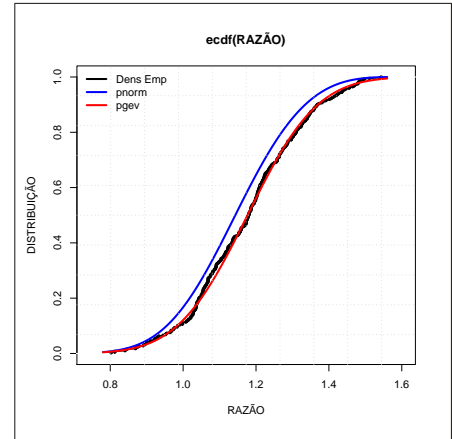


Figura 27: Comparação superposta da distribuição empírica, normal e Weibull ajustada para a razão dos mínimos.

res 0.9285019 e 1.424783, respectivamente. A figura 30 compara os quantis de probabilidade obtidos com a performance da razão k para os 61 dias subsequentes ao período analisado. Nessa figura estão também inclusos as inferências acerca dos verdadeiros quantis de probabilidade, a saber 1.052407 e 1.505114, respectivamente.

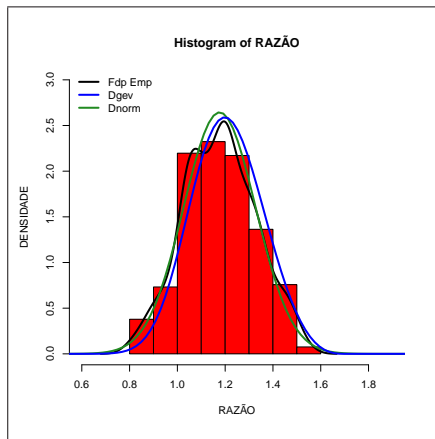


Figura 28: Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos máximos.

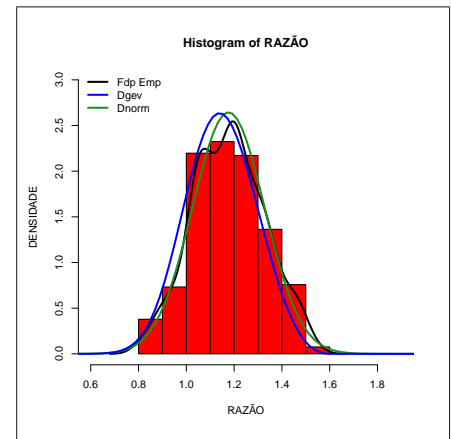


Figura 29: Comparação das densidades não-paramétrica e ajustadas (Gev e normal) para a razão dos mínimos.

Note que caso os *players* tivessem conhecimento acerca dos quantis estimados pela normal e GEV, a partir do monitoramento diário da evolução da razão entre os ativos a seguinte estratégia de arbitragem poderia ser armada:

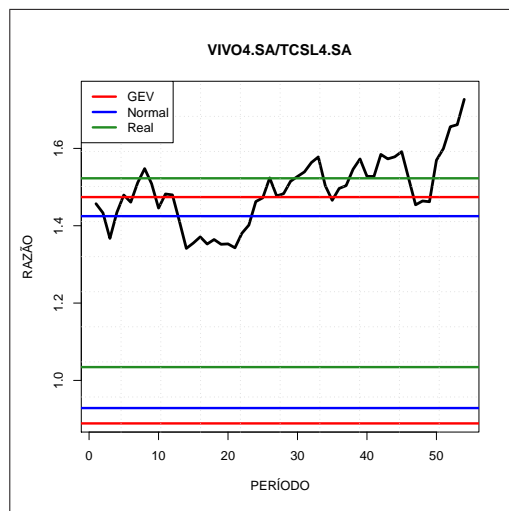


Figura 30: Performance da razão $k = \frac{VIVO4.SA}{TCSL4.SA}$ para os 54 dias subsequentes ao período compreendido entre 01/06/2006 a 15/12/2007.

1. Acompanhando a evolução da razão $k = \frac{VIVO4}{TCSL4}$, verifica-se que no período 2 a razão de precificação k mantém sua tendência de queda e assume valor bastante próximo do quantil 95% fornecido pela normal. Embora a razão esteja abaixo dos suportes superiores da GEV, o valor de k se encontra bastante próximo desses suportes, o que reforça a expectativa de queda continuada da razão de precificação. Diante disso poderia-se montar posições comprada em TCSL4 e vendida em VIVO4, fixando-se uma rentabilidade acumulada de 2% para inversão das posições. A partir desse limite, verifica-se que tal inversão aconteceria ao final do mesmo dia, ou seja, um dia útil após o *start* da operação (ver figura 31), a partir da obtenção de uma rentabilidade acumulada de 2,04%;
2. Alcançada a rentabilidade desejada assume-se posição comprada em VIVO4 e vendida em TCSL4. Note que a inversão das posições se justifica pela reação natural do mercado à tendência de queda contínua ao qual a razão k foi submetida.
3. Uma vez armadas as novas posições, pode-se determinar o ponto de *stop* da operação quando k atinja o real suporte superior da distribuição dos dados ou quando seja obtida uma rentabilidade acumulada de 2% contada a partir do ponto em que k rompe o limite superior fornecido pela GEV. Neste caso, esse momento se mostra presente 5 dias úteis após o início da inversão, quando a razão de precificação atinge o valor

real do suporte da distribuição. Para esse período a operação fornece rentabilidade acumulada de 5,38%³⁶ (ver figura 32).

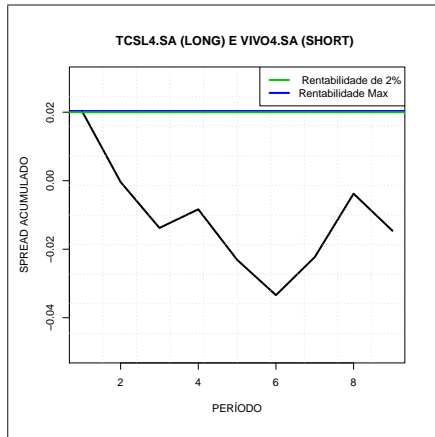


Figura 31: Rentabilidade acumulada das posições a partir do quantil da normal.

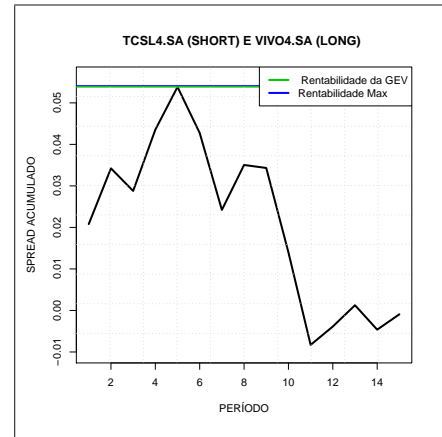


Figura 32: Rentabilidade acumulada a partir da inversão das posições.

Analisando a estratégia como um todo, obteríamos uma rentabilidade conjunta de 7.42% para um período de 6 dias úteis, o que equivale a uma taxa altamente significativa de 1921,05%*a.a* ou 30,01%*a.m*.

A estratégia utilizada poderia ainda ser ampliada de modo que fossem invertidas novamente as posições, uma vez que o suporte superior da real distribuição dos dados foi alcançado. Poderia-se determinar o ponto de *stop* da operação caso seja obtida uma rentabilidade acumulada de 2% contada a partir do ponto em que k rompe o limite superior fornecido pela GEV. Caso tal inversão fosse utilizada seriam incorporados à rentabilidade conjunta da operação 3,99% em 5 dias úteis.

Observe que esse par de ativos nos fornece ainda outro ponto de arbitragem para o qual poderia ser utilizada a estratégia descrita acima, como pode ser verificado na figura 29 (período 13).

Árbitros mais agressivos poderiam adotar como momento de *stop* da estratégia acima o ponto em que seja obtida uma rentabilidade acumulada de 2% contada a partir do momento em que k rompe o limite superior fornecido pela normal (e não pela GEV). Observe que, caso as expectativas desses árbitros se confirmassem, o mercado retornaria à situação inicial

³⁶Note que tal rentabilidade é igual à rentabilidade máxima que poderia ser obtida. Esse fato é mais um dos fatores que comprova o melhor ajuste dos quantis extremos a partir da GEV.

em que se encontrava no início de toda a estratégia, uma vez que estaria em tendência de queda e rompendo o quantil superior fornecido pela normal. Note que esse fato se confirma para o período analisado.

5.4 Impactos do quantil *GEV Long-Short*: expectativas e comportamento

Como verificado nas seções anteriores, o ajuste dos quantis extremos através da GEV foi capaz de fornecer informações mais refinadas acerca do *timing* de reversão da relação de precificação entre pares de ativos. Esse refinamento probabilístico permitiu à *GEV Long-Short Strategy* obter rentabilidades significativas através da arbitragem. Observe entretanto que essas janelas de oportunidade até o momento não têm sido exploradas pelos *players* de mercado. Essa seção procura apontar alguns dos aspectos comportamentais e possíveis consequências que a incorporação dessas novas informações pode trazer às decisões dos investidores, principalmente com relação à formulação de expectativas, e apresentar algumas propostas de trabalhos futuros.

Hong e Stein (1999) apresentaram uma linha de pensamento na qual os investidores de mercado são divididos entre observadores de notícias e *traders* de momento. Segundo esses autores, os observadores de notícia fazem previsões baseadas nos sinais que eles observam privadamente sobre os fundamentos futuros, mas ignoram preços passados e correntes. Já os *traders* de momento condicionam seus investimentos unicamente à alterações de preços no passado, ignorando quaisquer outros tipos de informações.

Observe que quando apenas os observadores de notícias estão ativos, as informações privadas são difundidas gradualmente entre o conjunto de agentes, de forma que os preços se ajustam de forma vagarosa às novas informações. Neste caso, existe apenas reação retardada e não reação em excesso³⁷. Isso porque ao rever suas crenças os indivíduos tendem naturalmente a dar maior peso à informações mais recentes e peso menor à informações mais antigas.

Quando os *traders* de momento ficam ativos, são feitas arbitragens sobre as reações retardadas causadas pelos observadores de notícias. Ao agirem dessa forma, os *traders* de momento acentuam as variações dos preços, que por sua vez ativam mais *traders* de momento, até o ponto em que a saturação do desequilíbrio inicial do mercado força os investidores a reavaliar as suas posições e, conseqüentemente, a própria precificação intrínseca dos ativos.

O que Hong e Stein justificaram é em termos gerais a essência comportamental teórica na qual se fundamenta a arbitragem *long-short*. Observe entretanto que a percepção das

³⁷O termo reação em excesso compreende os retornos anormais que não são consistentes com os padrões de retorno passados

ineficiências relativas dos preços dos ativos é, neste caso, percebida de forma gradativa pelos *traders* e sempre em um momento *ex-post* à ocorrência das reações retardadas. Um dos grandes ganhos que a utilização dos quantis teóricos *GEV Long-Short* fornecem está exatamente vinculado ao monitoramento *ex-ante* da difusão gradual das informações privadas e dos respectivos ajustes relativos dos preços, de forma que os *players* tenham capacidade de rever suas expectativas e perceber com defasagem temporal nula o *timing* oportuno em que devem operar como *traders* de momento.

No âmbito de uma abordagem heterodoxa, é possível considerar esse mecanismo a partir de um contexto de regulação, vinculada a uma convergência rumo a uma eventual posição de equilíbrio. De acordo com Herscovici (2004a), tal regulação se explica a partir da reação dos agentes em relação à uma posição inicial de desequilíbrio: é a reavaliação das expectativas em função de seu grau de realização, ou seja, da diferença entre o *ex-ante* e o *ex-post*, que constitui os mecanismos que permitem alcançar o equilíbrio.

Segundo Herscovici (2004a), na economia do desequilíbrio, o equilíbrio não é postulado *ex-ante*, mas é dado como o produto da reação dos agentes diante de uma situação inicial de desequilíbrio. Este conceito de equilíbrio torna-se diferente do conceito neoclássico pelo fato de estar dissociado do de maximização micro ou macroeconômica, uma vez que os processos de ajustamento dos agentes a um desequilíbrio inicial implicam um processo iterativo de modificação dessas decisões para assegurar o processo de gravitação rumo ao equilíbrio e de estabilidade das trajetórias. No caso da economia keynesiana, ainda segundo Herscovici, a decisão de investimento se baseia sobre as expectativas de longo prazo de lucro, de forma que o ajustamento em função de um desequilíbrio inicial se traduz por uma modificação das expectativas, que neste caso é provocada pelo conhecimento do quantil *GEV Long-Short*.

Diante das ineficiências de mercado e dos constantes momentos de desequilíbrio de curto prazo, a introdução do quantil *GEV Long-Short* se comporta como informação privada relevante, de forma a permitir a detecção de pontos de saturação do desequilíbrio e consequentemente a arbitragem pelos *traders* de momento. Observe que nesse contexto o universo tem que ser ergódico³⁸ para que as expectativas dos agentes possam se realizar. Não obstante, neste caso, o desequilíbrio provocado pelos efeitos imprevisíveis de comportamento a partir da incerção gradual dos demais *traders* implica uma forma de racionalidade limitada

³⁸Em um universo ergódico não há a presença de incerteza. Neste caso, a probabilidade de ocorrência dos eventos se aproxima da respectiva frequência empírica de ocorrência.

e não uma racionalidade substantiva. De acordo com Herscovici (2004a), isso ocorre porque longe do equilíbrio os processos se tornam irreversíveis e a natureza da informação se modifica igualmente, de forma que a informação não permite mais desconsiderar a incerteza. Na *GEV Long-Short Strategy* tal raciocínio é refletido a partir da inversão ou *stop* das posições, fixados as porcentagens de rentabilidade acumulada.

A utilização do quantil *GEV Long-Short*, ao possibilitar a exploração de janelas de oportunidade que não têm sido percebidas pelos investidores, poderá conferir aos *players* altas rentabilidades e a possibilidade de revisão de expectativas de longo prazo. Tais revisões de expectativas de longo prazo poderão exercer impactos singulares no mercado de renda variável e em sua própria regulamentação, uma vez que, como observou Keynes (1964), as flutuações diárias nos lucros dos investimentos existentes, que obviamente são fenômenos efêmeros e insignificantes, tendem a ter conjuntamente uma influência excessiva e mesmo absurda no mercado. Nesse sentido, pesquisas empíricas posteriores poderão ser realizados com o intuito de se avaliar os impactos e as influências que a utilização da nova metodologia de análise e do quantil *GEV Long-Short* possam trazer ao mercado e às decisões de investidores.

Outros enfoques de aplicabilidade poderão ser introduzidos a partir da utilização dessa metodologia de análise como, por exemplo, metodologias que avaliam desvios macroeconômicos extremos. A utilização do quantil *GEV Long-Short* também poderá ser utilizada como, por exemplo, medida indicadora da necessidade de intervenção no mercado com finalidade de regulação, uma vez que empresas que encontram com seus ativos com tendência à saturação do desequilíbrio podem intervir a partir da emissão de papéis e segurar o valor de seus ativos.

Em trabalhos futuros um novo sistema de análise poderá também ser estudado a partir das cópulas extremas para que portfólios que, ainda que repliquem um determinado índice de mercado, ofereçam no curto prazo oportunidades de arbitragem e ganhos superiores aos dos próprios índices. Tal análise poderá ser ainda utilizada para que se defina quantitativamente pontos de saturação do desequilíbrio do próprio mercado, análise indicativa inclusive de possibilidades de crises financeiras de curto prazo ³⁹.

³⁹Para início desse estudo recomenda-se a tese de doutorado intitulada “Cópulas para distribuições generalizadas de valores extremos multidimensionais (Sanfins (2009)).

6 Conclusão

No universo quantitativo da detecção de janelas de arbitragem não se pode considerar que todos os investidores estejam de acordo quanto à distribuição de probabilidade dos preços dos ativos. Isso porque, na maior parte dos casos, a verdadeira distribuição dos preços, bem como dos respectivos retornos, é desconhecida. O que os investidores fazem, na verdade, é aproximar em função das informações que têm disponíveis a distribuição empírica dos dados para distribuições conhecidas e que melhor se adequem ao tipo de análise requerida. Esse fato faz com que não sejamos capazes de assegurar, na prática, a existência de um conjunto de carteiras e de estratégias que sejam eficientes a todos os investidores (o que não implica a não existência teórica da fronteira eficiente e de estratégias maximizadoras de rentabilidade). Por consequência são criadas janelas de arbitragem, dadas as diferentes expectativas que cada investidor faz acerca da carteira ótima para alocação de seus recursos, das tendências de evolução dos preços dos ativos financeiros e das estratégias a serem utilizadas.

A velocidade e a eficiência da interpretação das informações por parte dos investidores também não podem ser consideradas homogêneas, ainda que se considere que as informações estejam instantaneamente disponíveis a todos os investidores⁴⁰.

A defasagem temporal causada pelo fator processamento e a qualidade de interpretação causada pelo fator eficiência fazem com que não se possa considerar ainda, como o faz as finanças modernas, as hipóteses de racionalidade ilimitada e informação perfeita. Nesse sentido, os preços dos ativos financeiros não refletem em cada momento toda a informação disponível, mas o conjunto de interpretações que cada investidor faz dessas informações.

Nesse sentido, a busca por metodologias de análise que melhor incorporem as reações dos investidores frente à alterações do comportamento do mercado se tornam cada vez mais necessárias. Como verificado através dos exemplos anteriores, a modelagem (através da TVE) das séries obtidas a partir dos quocientes dos preços dos *pairs trading* arbitrados mostrou que a GEV melhor ajustou os quantis de probabilidades extremos, detectando, com 95% de confiança, um *timing* mais oportuno para a reversão da tendência de queda e de aumento da razão de precificação entre os pares de ativos escolhidos.

O par de ativos PETR4 e VALE5 confere grande representatividade no total de negociações do mercado brasileiro. A rentabilidade anual de 143,86%*a.a* obtida, por ser extrema-

⁴⁰Na prática, dado o custo para obtenção de certas informações e dado que existe diferença entre disponibilidade de informação e decodificação de informação, não se pode considerar disponibilidade instantânea para qualquer informação para todos os investidores.

mente atraente, credita à utilização da nova modalidade quantitativa proposta importância significativa, ainda que os ativos analisados pertençam a diferentes setores de mercado.

Embora nesse exemplo exista um sincronismo entre os preços dos dois ativos considerados, verificou-se a partir do par VIVO4 e TCSL4 que tal sincronismo não é necessariamente um requisito para que estratégias *long-short* possam ser construídas, como é o caso da *GEV Long-Short Strategy* proposta. Tal fato se sustenta uma vez que a modelagem probabilística da razão de precificação através da GEV não procura estabelecer relação de dependência entre os processos geradores de cada série de preços dos ativos, mas verificar probabilisticamente, a partir da existência de cada processo gerador, momentos em que a razão de precificação ocupa quantis extremos e se mostra fora dos padrões de normalidade. Essa interpretação confere à metodologia de análise um avanço em termos de oportunidades de arbitragem com relação à metodologia que utiliza cointegração, uma vez que o fato de séries de preços de ativos não serem cointegradas não exclui a possibilidade da existência de janelas de arbitragem no curto prazo.

Além das altíssimas rentabilidades obtidas para os pares de ativos estudados, nas quais se destaca a de 30,01%*a.m* referente ao par VIVO4 e TCSL4, a detecção mais precisa do *timing* de armação, inversão e desarme das operações mostra que a modelagem a partir da GEV é capaz de oferecer sustentação para que pares de ativos que não possuam interdependência de longo prazo possam gerar no curto prazo janelas de arbitragem.

Esses fatores tornam-se essenciais para que possamos confirmar a suposição de que o melhor ajuste dos quantis extremos através da GEV tem capacidade de fornecer sustentação probabilística mais refinada para que o retorno à média justifique a possibilidade de arbitragem *long-short*.

É importante observar ainda que, embora a modelagem dos dados através da distribuição normal tenha superestimado os quantis inferiores e subestimado os quantis superiores, tais quantis podem ser consideradas como informação para que sejam verificadas inversões secundárias de movimentos extremos.

A arbitragem, como verificado pelos exemplos citados, não deve levar, entretanto, apenas em consideração os quantis fornecidos pela normal. Isso porque movimentações precipitadas por parte dos investidores poderiam ser tomadas e importantes janelas poderiam não ser exploradas, o que acarretaria em resultados de rentabilidade pouco atraentes.

É importante ressaltar que as estratégias *long-short* não necessariamente devem estar vinculadas à arbitragem entre pares de ativos. Existem outras formas de se arbitrar

com posições compradas e vendidas, inclusive fazendo o uso de alavancagem, que podem ser adotadas de maneira isolada ou conjunta pelos gestores de *hedge funds* para se auferir rentabilidades superiores, como é o caso de arbitragem entre índices, venda a descoberto, investimentos em ações de empresas próximas ao colapso (*distressed securities*), etc.

Com tantas opções, fica claro que não existe uma linha única de arbitragem que deva ser adotada entre os *hedge funds*, tanto em termos de risco como expectativas de rentabilidade. Cabe ao investidor identificar junto aos gestores destes fundos qual a estratégia adotada e buscar classificá-la, principalmente quanto ao do ponto de vista de risco (direcionais, geralmente mais agressivos, ou neutros em relação ao mercado, mais conservadores).

A utilização dos quantis teóricos da GEV pode fornecer informações para que seja feito um monitoramento *ex-ante* da difusão gradual das informações privadas e dos respectivos ajustes relativos dos preços. Tal monitoramento torna-se muito importante na medida em que possibilita que os *players* revejam suas expectativas e percebam com defasagem temporal nula o *timing* oportuno em que devem operar como *traders* de momento. Além disso, devido às altas rentabilidades obtidas pelos exemplos considerados e à grande capacidade de modelagem de dados extremos pela TVE, entretanto, torna-se interessante o monitoramento diário da razão de precificação k de outros pares de ativos e dos quantis fornecidos pela GEV para que os players possam arbitrar com esses ativos de maneira *long-short*. As altas performances obtidas através do monitoramento das duas ações de maior peso do mercado brasileiro e de duas ações importantes no crescente setor de comunicações confirmam a importância da nova estratégia quantitativa *long-short* proposta como metodologia de análise dentro do mercado de arbitragem.

7 APÊNDICE

A Critérios de Convergência

O Teorema de Fisher-Tippett especifica as distribuições extremas limite para as quais há a possibilidade de convergência da distribuição do máximo padronizado $\frac{X_{(n)}-b_n}{a_n}$, embora não estabeleça as condições que F deve satisfazer para que ocorra tal convergência. Esse apêndice tem o objetivo de elucidar essa questão, ou seja, de estabelecer quais condições F deve verificar para $P(X_{(n)} \leq u_n)$ convirja para algum valor do intervalo $(0, 1)$, quando $n \rightarrow \infty$.

O resultado abaixo, denotado por aproximação de Poisson, é o ponto de partida para essa convergência fraca do máximo⁴¹.

Proposição A.1 (Aproximação de Poisson) *Para um dado $\tau \in [0, \infty]$ e uma sequência (u_n) de números reais, as seguintes implicações são equivalentes:*

$$n\bar{F}(u_n) \rightarrow \tau \quad (3)$$

$$P(X_{(n)} \leq u_n) \rightarrow \exp(-\tau) \quad (4)$$

Prova A.1 *Para $\tau \in [0, \infty]$ e (u_n) sequência de números reais, temos que se $n\bar{F}(u_n) \rightarrow \tau$, então $P(X_{(n)} \leq u_n) = F^n(u_n) = (1 - \bar{F}(u_n))^n = \left(1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right)^n$, donde obtemos*

$$P(X_{(n)} \leq u_n) \rightarrow \exp(-\tau)$$

Por outro lado, se a implicação (5) se verifica, então $\bar{F}(u_n) \rightarrow 0$ (pois do contrário existiria uma subsequência (u_{n_k}) tal que $\bar{F}(u_{n_k})$ seria limitada e assumiria a partir de um certo n_{k_0} valores estritamente positivos. Desse modo, $P(X_{(n_k)} \leq u_{n_k}) = (1 - \bar{F}(u_{n_k}))^{n_k}$, de forma que teríamos $P(X_{(n_k)} \leq u_{n_k}) \rightarrow 0$). Aplicando-se o logaritmo na implicação (5), temos

$$-n \ln(1 - \bar{F}(u_n)) \rightarrow \tau$$

Como $-\ln(1 - x) \sim x$ quando $x \rightarrow 0$, então $n\bar{F}(u_n) \rightarrow \tau + o(1)$, o que nos dá $n\bar{F}(u_n) \rightarrow \tau$.

□

⁴¹As proposições, os lemas, corolários, definições e teoremas apresentados nesse e nos demais apêndices podem ser todos vistos em Embrechts, Kluppelber e Mikosch (1997).

Como verificado na sessão Distribuições Extremas do capítulo 2, temos que o máximo converge em probabilidade para o limite superior x_F do suporte da distribuição F , ou seja

$$X_x \xrightarrow{d} x_F$$

quando $n \rightarrow \infty$ e para $x_F \leq \infty$. Segue que

$$P(X_{(n)} \leq x) = \begin{cases} 0, & x < x_F \\ 1, & x > x_F \end{cases}$$

Esse fato nos leva a considerar o seguinte resultado:

Corolário A.1 *Se $x_F < \infty$ e $\bar{F}(x_{F-}) = F(x_F) - F(x_{F-}) > 0$, então para toda sequência (u_n) tal que $P(X_{(n)} \leq u_n) \rightarrow \rho$ temos que $\rho = 0$ ou $\rho = 1$.*

Prova A.2 *Se $\rho \in [0, 1]$, então existe um $\tau \in [0, \infty)$ tal que $\rho = \exp(-\tau)$. Pela proposição A.1, temos que $n\bar{F}(u_n) \rightarrow \tau$ quando $n \rightarrow \infty$. Se u_n é limitada superiormente por x_F ($u_n < x_F$ para todo n maior que um dado n_0), então para esses valores de n temos que $\bar{F}(u_n) \geq \bar{F}(x_{F-}) > 0$, donde obtemos $\tau = \infty$. Se caso contrário ($u_n \geq x_F$) para todo n maior que um dado n_0 , então $n\bar{F}(u_n) = 0$, donde obtemos $\tau = 0$. Desses dois casos ($\tau = \infty$ e $\tau = 0$), obtemos $\rho = 0$ e $\rho = 1$, respectivamente.*

□

Este resultado mostra que nenhum limite não degenerado existe para funções de distribuição com saltos no seu suporte finito, quaisquer que sejam as sequências normalizadoras escolhidas. O teorema que segue é similar a esse resultado e se estende também para funções de distribuição com suporte infinito.

Teorema A.3 *Sejam F uma função de distribuição com suporte $x_F \leq \infty$ e $\tau \in (0, \infty)$. Existe uma sequência (u_n) satisfazendo $n\bar{F}(u_n) \rightarrow \tau$ se, e somente se,*

$$\lim_{x \rightarrow x_F} \frac{\bar{F}(x)}{\bar{F}(x^-)} = 1 \quad (5)$$

e $F(x_{F-}) = 1$

B Distribuições Max-estáveis

Para verificarmos que a classe das distribuições max-estáveis coincide com a classe das possíveis distribuições limite não degeneradas é necessário que seja introduzido o teorema que segue:

Teorema B.1 *Sejam A, B, A_1, \dots variáveis aleatórias e $b_n > 0, \beta_n > 0, a_n \in \mathbb{R}$ e $\alpha_n \in \mathbb{R}$ constantes. Suponha que*

$$b_n^{-1}(A_n - a_n) \xrightarrow{d} A$$

Então

$$\beta_n^{-1}(A_n - \alpha_n) \xrightarrow{d} B$$

se, e somente se,

$$\lim_{n \rightarrow \infty} \frac{b_n}{\beta_n} = b \in [0, \infty) \quad e \quad \lim_{n \rightarrow \infty} \frac{(a_n - \alpha_n)}{\beta_n} = a \in \mathbb{R}$$

Teorema B.2 *A classe das distribuições max-estáveis coincide com a classe das possíveis distribuições limite não degeneradas para as quais há a convergência do máximo de variáveis aleatórias iid devidamente padronizados.*

Prova B.3 *Assumindo que para sequências normalizadoras apropriadas*

$$\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x), \quad x \in \mathbb{R}$$

para alguma H não degenerada e contínua, então para todo $k \in \mathbb{N}$

$$\lim_{n \rightarrow \infty} F^{nk}(c_n x + d_n) = \lim_{n \rightarrow \infty} (F^n(c_n x + d_n))^k = H^k(x), \quad x \in \mathbb{R}$$

Além disso, sabemos que $\lim_{n \rightarrow \infty} F^{nk}(c_{nk} x + d_{nk}) = H(x), x \in \mathbb{R}$. Pelo teorema anterior existem constantes $\tilde{c}_k > 0$ e $\tilde{d} \in \mathbb{R}$ tais que

$$\lim_{n \rightarrow \infty} \frac{c_{nk}}{c_n} = \tilde{c}_k \quad e \quad \lim_{n \rightarrow \infty} \frac{d_{nk} - d_n}{c_n} = \tilde{d}_k$$

Assim, para variáveis aleatórias Y, Y_1, \dots, Y_k com função de distribuição H , temos que

$$\max(Y_1, \dots, Y_k) \stackrel{d}{=} \tilde{c}_k Y + \tilde{d}_k$$

□

C Máximo Domínio de Atração da Fréchet⁴²

Definição C.1 *Uma função F possui variação regular com índice $-\alpha$, $\alpha > 0$, e denota-se $\overline{F} \in \mathfrak{R}_{-\alpha}$, se*

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(xt)}{\overline{F}(x)} = t^{-\alpha}, \quad t > 0.$$

Teorema C.1 (Máximo domínio de atração da Fréchet) *Uma função de distribuição F pertence ao domínio de atração da Fréchet se, e somente se, $\overline{F}(x) = x^{-\alpha}L(x)$ para alguma função de variação lenta L .*

Prova C.2 *Ver Embrechts, Kluppelber e Mikosch (1997).*

Note que esse resultado implica em particular que toda $F \in MDA(H_{II})$ tem suporte x_F infinito. Ainda, a sequência normalizadora (c_n) é uma sequência de variação regular, com $c_n = n^{\frac{1}{\alpha}}L_1(x)$ para alguma função L_1 de variação lenta, e a sequência (d_n) pode ser escolhida como nula (não há a necessidade de normalização).

Logo abaixo segue uma condição encontrada por Von Mises de mais fácil verificação acerca da convergência para a Fréchet que envolve a densidade da função de distribuição.

Corolário C.1 (Condição de Von Mises) *Seja F uma função de distribuição absolutamente contínua com densidade f . Se*

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{\overline{F}(x)} = \alpha > 0 \tag{6}$$

então $F \in MDA(H_{II})$.

⁴²As condições de Von-Mises apresentadas nesse apêndice e nos dois apêndices que seguem podem ser exploradas na tese de doutorado intitulada “Cópulas para distribuições generalizadas de valores extremos multidimensionais (Sanfins (2009)).

D Máximo Domínio de Atração da Weibull

Como observado na proposição 4.1, as distribuições Weibull e Fréchet possuem conexões bastante próximas, donde espera-se que seus domínios de atração também mantenham tal proximidade de caracterização. O teorema que segue confirma esse fato:

Teorema D.1 (Máximo domínio de atração da Weibull) *Uma função de distribuição F pertence ao domínio de atração da Weibull se, e somente se, $x_F < \infty$ e $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$ para alguma função de variação lenta L .*

Prova D.2 *Ver Embrechts, Kluppelber e Mikosch (1997).*

Note entretanto que, diferentemente das funções de distribuição que convergem para a Fréchet, o domínio de atração da Weibull é formado por distribuições com suporte superior x_F finito. A partir da transformação $T(F(x)) = F(x_F - x^{-1})$, o corolário C.1 pode ser reescrito como:

Corolário D.1 (Condição de Von Mises) *Seja F uma função de distribuição absolutamente contínua com densidade f positiva para algum intervalo finito (z, x_F) . Se*

$$\lim_{x \rightarrow x_F} \frac{(x_F - x)f(x)}{\bar{F}(x)} = \alpha > 0 \quad (7)$$

então $F \in MDA(H_{III})$.

E Máximo Domínio de Atração da Gumbel

O máximo domínio de atração de uma Gumbel engloba uma gama de distribuições bastante diversificada, que envolve desde distribuições com caudas moderadas (como é o caso da lognormal) à distribuições com caudas mais leves (como é o caso da normal). Algumas dessas distribuições possuem uma representação comum e são denominadas Função Von Mises.

Definição E.1 (Função Von Mises) *Seja F uma função de distribuição com suporte superior $x_F \leq \infty$. Se existir $z < x_F$ tal que*

$$\bar{F}(x) = c \cdot \exp \left(- \int_z^x \frac{1}{a(t)} dt \right), \quad z < x < x_F \quad (8)$$

para alguma constante positiva c e para alguma função positiva absolutamente contínua $a(\cdot)$ com densidade $a'(x)$ tal que $\lim_{x \rightarrow x_F} a'(x) = 0$, então dizemos que F é uma função de Von Mises e que $a(\cdot)$ é função auxiliar de F .

Exemplo: A distribuição Exponencial $F(x) = 1 - \exp(-\lambda x)$ é uma função Von Mises com função auxiliar $a(x) = \lambda^{-1}$, uma vez que ao considerarmos $z = 0$ e $c = 1$ temos

$$\bar{F}(x) = \exp \left(- \int_0^x \frac{1}{\lambda^{-1}} dt \right) = \exp(-\lambda x), \quad x \geq 0, \quad \lambda > 0$$

Definição E.2 (Função Quantil) *A generalizada inversa F^{\leftarrow} de uma função de distribuição F é dada por*

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} / F(x) \geq t\}, \quad 0 < t < 1$$

e é chamada de função quantil da distribuição F . O quantil $x_t = F^{\leftarrow}(t)$ define o t -quantil de F .

Proposição E.1 *Se F é uma função de Von Mises, então $F \in MDA(H_I)$. Uma escolha possível para as seqüências padronizadoras é*

$$d_n = F^{\leftarrow}(1 - n^{-1}) \quad c_n = a(d_n)$$

Prova E.1 *Ver Embrechts, Kluppelberg e Mikosch (1997)*

A caracterização do máximo domínio de atração da Gumbel $MDA(H_I)$ que segue é uma extensão modificada da expressão que define a função de Von Mises:

Teorema E.2 (Caracterização I do $MDA(H_I)$) *Uma função de distribuição F com suporte superior $x_F \leq \infty$ pertence ao domínio de atração da Gumbel se, e somente se, existe $z < x_F$ de forma que*

$$\bar{F}(x) = c(x) \exp \left(- \int_z^x \frac{g(t)}{a(t)} dt \right), \quad z < x < x_F \quad (9)$$

onde c e g são funções mensuráveis satisfazendo $c(x) \rightarrow c > 0$, $g(x) \rightarrow 1$ quando $x \rightarrow x_F$, e $a(x)$ uma função positiva absolutamente contínua com densidade $a'(x)$ tal que $\lim_{x \rightarrow x_F} a'(x) = 0$. Para F com a representação (9), as sequências (c_n) e (d_n) padronizadoras do máximo e a função $a(x)$ podem ser escolhidas de forma que

$$d_n = F^{\leftarrow}(1 - n^{-1})$$

$$c_n = a(d_n)$$

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, \quad x < x_F$$

Prova E.3 *Ver Resnick (1987), Corolário 1.7 e Proposição 1.9.*

Teorema E.4 (Caracterização II do $MDA(H_I)$) *Uma função de distribuição F pertence ao máximo domínio de atração de uma Gumbel se, e somente se, existe alguma função positiva \tilde{a} tal que*

$$\lim_{x \rightarrow x_F} \frac{\bar{F}(x + t\tilde{a}(x))}{\bar{F}(x)} = \exp(-t), \quad t \in \mathbb{R} \quad (10)$$

Uma escolha possível para \tilde{a} é tal que $\tilde{a}(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt$

Prova E.5 *Ver Hann (1970), teorema 2.5.1.*

F Convergência do Máximo de Sequências Estacionárias

O Teorema de Fisher-Tippett especifica as distribuições extremas limite para as quais há a possibilidade de convergência da distribuição do máximo padronizado de variáveis aleatórias i.i.d.. Neste apêndice serão apontadas condições que assegurem que o máximo $X_{(n)}$ de uma amostra proveniente de uma sequência de variáveis aleatórias fracamente estacionária (X_n) possua no limite um comportamento similar ao do máximo $\tilde{X}_{(n)}$ de uma amostra proveniente de uma sequência de variáveis aleatórias i.i.d. (\tilde{X}_n) correspondente com mesma função de distribuição F .

Note que se (\tilde{X}_n) é uma sequência i.i.d. associada a sequência estacionária (X_n) , então

$$\begin{aligned} P(\tilde{X}_{(n)} \leq u_n) &= P^n(\tilde{X} \leq u_n) \\ &= \exp(n \ln(1 - P(\tilde{X} > u_n))) \\ &\approx \exp(-n\bar{F}(u_n)) \end{aligned}$$

uma vez que $-\ln(1 - x) \sim x$.

Embora não seja possível a aplicação direta da definição 4.2 para o máximo de uma amostra proveniente de uma sequência de variáveis aleatórias fracamente estacionária, o próximo teorema mostra que a classe das possíveis distribuições limite para esses máximos padronizados coincide com a classe das distribuições limite EV, desde que uma condição específica de independência assintótica seja satisfeita.

Definição F.1 (Condição de Independência Assintótica) *Dizemos que uma sequência de variáveis aleatórias fracamente estacionária (X_n) satisfaz a condição $D_1(u_n)$ de independência asintótica se para quaisquer inteiros positivos p, q e n tais que*

$$1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$$

com $j_1 - i_p \geq l$, tivermos

$$\left| P\left(\max_{i \in A_1 \cup A_2} X_i \leq u_n\right) - P\left(\max_{i \in A_1} X_i \leq u_n\right) P\left(\max_{i \in A_2} X_i \leq u_n\right) \right| \leq \alpha_{n,l}$$

onde $A_1 = \{i_1, \dots, i_p\}$, $A_2 = \{j_1, \dots, j_q\}$ e $\alpha_{n,l} \rightarrow 0$ quando $n \rightarrow \infty$ para alguma sequência $l = l_n = o(n)$.

Intuitivamente, dizemos que uma sequência fracamente estacionária (X_n) satisfaz a condição $D_1(u_n)$ se no limite a probabilidade $P(X_{(n)} \leq u_n)$ é igual ao produto das probabilidades de que os máximos de quaisquer dois sub-blocos disjuntos sejam inferiores a u_n , para alguma (u_n) real. Ou ainda, se no limite tivermos para qualquer p inteiro que

$$P(X_{(n)} \leq u_n) = P(\max A_p \leq u_n) \cdot P(\max B_{n_p} \leq u_n)$$

onde (B_{n_p}) é uma subsequência de (X_n) e igual a (X_n) a menos dos p primeiros elementos de (X_n) e $A_p = \{X_1, \dots, X_p\}$.

Note ainda que a condição $D_1(u_n)$ possui implicação direta

$$P(X_{(n)} \leq u_n) = P^k(X_{(\frac{n}{k})} \leq u_n) + o(1) \quad (11)$$

onde k é uma constante de incremento lenta. Uma discussão mais aprofundada acerca da condição de independência assintótica pode ser vista em Leadbetter et al (1983).

Teorema F.1 (Convergência do Máximo de uma Sequência Estacionária) *Seja (X_n) uma sequência fracamente estacionária de variáveis aleatórias com sequência de distribuição associada (F_n) . Para uma amostra aleatória $X = \{X_1, X_2, \dots, X_n\}$, se existirem sequências padronizadoras (c_n) positiva e (d_n) real e uma variável aleatória G com função de distribuição F_G , tais que $c_n^{-1}(X_{(n)} - d_n) \xrightarrow{d} G$ e que a condição $D_1(c_n x + d_n)$ de independência assintótica seja satisfeita para todo $x \in \mathbb{R}$, então F_G pertence à classe das distribuições limite de valor extremo.*

Prova F.2 *Pelos teoremas B.1 e 4.1 temos que G é uma distribuição de valor extremo se, e somente se, G é max-estável. De (11) temos que*

$$P(X_{(nk)} \leq c_n x + d_n) = P^k(X_{(n)} \leq c_n x + d_n) + o(1) \rightarrow G^k(x)$$

para todo inteiro $k \geq 1$ e todo x em que G é contínua. Por outro lado,

$$P(X_{(nk)} \leq c_n x + d_n) \rightarrow G(x)$$

Agora basta procedermos como na prova do teorema B.1 para concluirmos que G é max-estável.

□

Observação F.1 *Note que do teorema F.1 não podemos afirmar a partir das relações*

$$c_n^{-1}(\tilde{X}_{(n)} - d_n) \xrightarrow{d} H$$

$$c_n^{-1}(X_{(n)} - d_n) \xrightarrow{d} G$$

que $G = H$, para alguma $H \in \{H_I, H_{II}, H_{III}\}$. Na verdade G é igual a H a menos de um índice extremal $\theta \in [0, 1]$, a saber $G = H^\theta$. O próximo apêndice trará mais informações acerca do índice extremal.

Note entretanto que assim como o Teorema de Fisher-Tippett, o teorema F.1 especifica a possibilidade de convergência da distribuição do máximo padronizado $\frac{X_{(n)} - b_n}{a_n}$ de uma amostra proveniente de sequência fracamente estacionária de variáveis aleatórias, embora não estabeleça as condições que a sequência de funções de distribuição associadas ($F_{(n)}$) deve satisfazer para que $P(X_{(n)} \leq u_n)$ convirja para um valor do intervalo $(0, 1)$, quando $n \rightarrow \infty$.

A seguir serão especificadas condições suficientes para que $n\bar{F}(u_n) \rightarrow \tau$, para algum $\tau \in [0, \infty)$ e para uma dada sequência de limiares (u_n), no caso de (X_n) possuir distribuição comum F . Note que a equivalência das implicações (3) e (4) da proposição A.1 não se verifica diretamente para o caso de (X_n) ser fracamente estacionária. Isso porque no máximo podemos obter

$$\liminf_{n \rightarrow \infty} P(X_{(n)} \leq u_n) \geq \exp(-\tau) \quad (12)$$

(ver Embrechts, Kluppelberg e Mikosch (1997)). A condição que segue é uma das condições que deve ser satisfeita para que tal equivalência seja verificada.

Definição F.2 (Condição anti-Conglomerado) *Dizemos que uma sequência fracamente estacionária (X_n) satisfaz a relação $D_2(u_n)$ para alguma (u_n) real se*

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=2}^{\frac{n}{k}} P(X_1 > u_n, X_j > u_n) = 0 \quad (13)$$

Proposição F.1 (Prob. Limite para o Máximo em Sequências Estacionárias) *Se (X_n) é uma sequência estacionária e u_n uma sequência de limiares tais que as condições $D_1(u_n)$ e $D_2(u_n)$ são satisfeitas e $\tau \in [0, \infty)$, então as seguintes implicações são equivalentes:*

$$n\bar{F}(u_n) \rightarrow \tau \quad (14)$$

$$P(X_{(n)} \leq u_n) \rightarrow \exp(-\tau) \quad (15)$$

G Índice Extremal

Em muitos casos verifica-se que sequências de variáveis aleatórias (i.i.d. ou fracamente estacionárias) possuem conglomerados de extremos. Na verdade, a maior parte dos eventos extremos tendem a ocorrer em conglomerados em virtude de uma certa dependência local dos dados. Intuitivamente, o índice extremal relaciona a estrutura de dependência desses dados com suas características extremas.

Este apêndice tem o intuito de apresentar o índice extremal θ , cuja aplicação incorpora informações que resultam em melhores estimativas para as caudas extremas e, conseqüentemente, para os quantis associados a probabilidades muito pequenas (Mendes (2004)).

Definição G.1 (Índice Extremal) *Seja (X_n) uma sequência fracamente estacionária e θ uma constante não negativa. Se para todo $\tau > 0$ existir uma sequência (u_n) tal que*

$$\lim_{n \rightarrow \infty} n\overline{F}(u_n) = \tau \quad (16)$$

e

$$\lim_{n \rightarrow \infty} P(X_{(n)} \leq u_n) = \exp(-\theta\tau) \quad (17)$$

então dizemos que θ é o índice extremal da sequência (X_n) .

Segundo Leadbetter, Lindgren e Rootzén (1983), o número de observações em um conglomerado depende diretamente das características (peso) das caudas da distribuição subjacente F . Para séries que apresentam conglomerados, a estimação dos quantis extremos deve ser feita utilizando-se o índice extremal para que haja a correção do impacto causado nas estimativas por tais conglomerados. Desse modo, temos que o quantil teórico de probabilidade $(1 - \alpha)$ da distribuição subjacente passa a ser dado por

$$x_\alpha = H_{\xi, \mu, \sigma}^{-1}((1 - \alpha)^{n\theta}) \quad (18)$$

onde $\theta \in [0, 1]$.

Sob este ponto de vista, $n\theta$ pode ser interpretado como o número de conglomerados independentes em n observações. Neste caso, quando não existem conglomerados de extremos, θ assume o valor 1, e quanto mais próximo de zero estiver θ , mais pesada será a cauda da distribuição subjacente.

Verifica-se ainda que o tamanho e a frequência dos conglomerados de valores extremos são características intrínsecas de cada sequência de variáveis aleatórias e que o índice extremal assimila as informações de tais característica.

Um estimador simples para o índice extremal θ pode ser obtido através da razão entre o número K_u de blocos nos quais o máximo excedeu um determinado limiar⁴³ u e o número total N_u de observações que excederam u durante todo o período analisado, ou seja:

$$\hat{\theta} = \frac{K_u}{N_u} \quad (19)$$

Vários outros estimadores estão propostos na literatura para a estimação do índice extremal θ . Em Embrechts, Kluppelber e Mikosch (1997) podem ser vistas propriedades de um estimador que apresenta melhores propriedades estatísticas que o estimador simples apresentado em (19), a saber:

$$\hat{\theta} = \frac{1}{m} \frac{\ln\left(1 - \frac{K_u}{m}\right)}{\ln\left(1 - \frac{N_u}{N}\right)} \quad (20)$$

onde N representa o total de observações e N_u , K_u e u são como definidos acima.

⁴³A determinação de u pode ser vinculada a um certo percentual de observações nas caudas, digamos 3%. Uma estimativa mais refinada de θ pode ser obtida como a média dos estimadores $\hat{\theta}$ vinculados a limiares provenientes de diferentes percentuais de observações nas caudas.

Referências

- [1] ALEXANDER, CAROL. *Optimal Hedging Using Cointegration*. Philosophical Transactions of the Royal Society, London, Series A, 357, p. 2039-2058, 1999.
- [2] BALKEMA, A.A.; HAAN, L. *Residual Life Time at Great Age*. Ann Probability, n.2, p.792-804, 1974.
- [3] BORTKIEWICZ, L.VON. *Variationsbreite und mittlerer Fehler*. Sitzungsber, Berli. Math. Gess., n.21, p.3-11, 1922.
- [4] COLES, S. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer: Berlin, 2001.
- [5] CHANDRA, M.; SINGPURWALLA, N.D.; STEPHENS, M.A. *Kolmogorov Statistics for Tests of Fit of the Extreme Value and Weibull Distribution*. Journal of the American Statistical Association, n.76, p. 729-731, 1981.
- [6] DAMODARAN, A. *Avaliação de investimentos: ferramentas e técnicas para a determinação do valor de qualquer ativo*. Rio de Janeiro: Qualitymark, 1999.
- [7] DANIELSON, J.; De VRIES, C.G. *Value-at-Risk and Extrem Returns*. Working Paper, Department of Economics, University of Iceland, 1997.
- [8] EMBRECHTS, P., KLUPPELBERG, C., MIKOSCH, T. *Modelling extremal events for insurance and finance*. Springer-Verlag: Berlin, 1997.
- [9] EMBRECHTS, P. *Extreme Value Theory: Potential and Limitation as an Integrated Risk Management Tool*. Department of Mathematik, ETH, Zentrum, CH 8092, Zurich, 2000.
- [10] FAMA, E.F. *Efficient Capital Markets: a review of theory and empirical work*. The Journal of Finance, Chicago: American Finance Association, v. 25, n. 2, p. 383-417, maio 1970.
- [11] FISHER, R.A. *The goodness of fit of regression formulae and the distribution of regression coefficients*. J. Roy Statist. Soc., n.85, p.597-612, 1922.
- [12] FISHER, R.A.; TIPPETT, L.H.C. *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*. Proceedings of the Cambridge Philosophical Society, n.24, p.180-190, 1928.

- [13] FUNG, WILLIAM; HSIEH, DAVID A. *Extracting Portable Alphas from Equity Long/Short Hedge Funds*. Journal of Investment Management, v.2, n.4, p.1-19, 2004.
- [14] GNEDENKO, B.V. *Sur la distribution limite du terme maximum d'une serie aléatoire*. Annals of Mathematics, n.44, p.423-453, 1943.
- [15] GRAHAM, B.; DODD D.L. *Security Analysis: The Classic 1934 Edition*. McGraw-Hill Trade, Nova York, 1996.
- [16] GREENWOOD, J.A.; LANDWEHR, J.M.; MATALAS, N.C.; WALLIS, J.R. *Probability Weighed Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form*. Water Resources Research, n.15, p. 1049-1054, 1979.
- [17] GUMBEL, E.J. *Statistics Theory of Extreme Values and Some Practical Applications*. Nat. Bureau of Standards Applications Mathematics Series, n.33, p.1-51, 1954.
- [18] GUMBEL, E.J. *Statistics of Extremes*. Columbia University Press, New York Journal of Business, n.63, p.383-408, 1958.
- [19] HAAN, L. *On Regular Variation and Its Application to Weak Convergence of Sample Extremes*. CWI Tract 32, Amsterdam, 1970.
- [20] HAUGEN, R.A. *The new finance: the case against efficient markets*. New Jersey: Prentice Hall, 1999. (a)
- [21] HAUGEN, R. *A Best on Wall Street: how stock volatility devours our wealth*. Upper Saddle River, NJ: Prentice Hall, 1999 (b).
- [22] HENRIKSSON, ROY D.; MERTON, ROBERT C. *On Market Timing and Investment Performance II. Statistical Procedures for Evaluating Forecasting Skills*. Journal of Business, v.54, n.4, p.513-533, 1981.
- [23] HENRIKSSON, ROY D. *Market Timing and Mutual Fund Performance: an empirical investigation*. Journal of Business, v.57, n.1, p.73-96, 1984.
- [24] HERSCOVICI, A. *Uma crítica do conceito neoclássico de capital: as contribuições de Keynes e de Sraffa*. In: ENCONTRO NACIONAL DE ECONOMIA POLÍTICA, 9, Uberlândia, 2004a

- [25] HOSKING, J.R.M. *L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics*. Journal of the Royal Statistical Society, n.52, p.105-124, 1990.
- [26] HOSKING, J.R.M.; WALLIS, J.R. *Regional Frequency Analysis*. [sl]: Cambridge University Press, 1997.
- [27] JENKINSON, A.F. *The frequency distribution of the annual maximum (or minimum) values of meteorological elements*. Quart. J. Roy. Meteo. Soc., v.81, p.158-171, 1955.
- [28] LEADBETTER, M.; LINDGREEN, G.; ROOTZÉN, H. *Extremes and Related Properties of Random Sequences and Processes*. Berlim: Springer-Verlag, 1983.
- [29] LEUSIN, LILIANA M.C.; BRITO, RICARDO D. *Market Timing e Avaliação de Desempenho dos Fundos Brasileiros*. Revista de Administração de Empresas, v.48, n.2, p.22-36, 2008.
- [30] LIANG, B. *On the performance of Hedge Funds*. 1999
- [31] LONGIN, F. *The Asymptotic Distribution of Extrem Stock Market Returns*. 1996.
- [32] LOOMIS, CAROL J. *The Jones Nobody Keeps Up With*. Fortune Magazine, p. 237-248, 1966.
- [33] LUCAS, ANDRE. *Strategic and Tactical Asset Allocation and the Effect of Long-Run Equilibrium Relations*. Research Memorandum, Vrije Universiteit Amsterdam, 1997.
- [34] MACEDO JR, JURANDIR S. *Teoria do Prospecto: Uma investigação utilizando simulação de investimentos*. 2003. 218f. Tese (Doutorado em Engenharia de Produção) Universidade Federal de Santa Catarina, Florianópolis.
- [35] MARKOWITZ, HARRY. *Portfolio Selection*. The Journal of Finance, v.7, n.1, p.77-91, 1952.
- [36] McNEIL, A.J. *Calculating Quantile Risk Measures for Financial Return Series Using Extreme Value Theory*. Zurich, Department Mathematik, ETH Zentrum, 1998.
- [37] MENDES, B.V.M. *Assessing the Bias of Maximum likelihood Estimates of Garch Models*. Journal of Statistical Computation and Simulation, v.67, p.359-376, 2000.

- [38] MENDES, B.V.M. *Introdução à Análise de Eventos Extremos*. E-papaers Serviços Editoriais Ltda, Rio de Janeiro, 2004.
- [39] MERTON, ROBERT C. *On Market Timing and Investment Performance I. An equilibrium theory of value for market forecasts*. Journal of Business, v.54, n.3, p.363-406, 1981.
- [40] PICKANDS, J. III. *Statistical Inference Using Extreme Order statistics*. Annals of Estatistics, n.3, p.119.131, 1975.
- [41] PICKANDS, J. III. *Bayes Quantile Estimation and Threshold Selection for the Generalized Pareto Family. Extreme Value Theory and Aplications*. [sl]: Kluver Academic Publishers, p.123-138, 1994.
- [42] REISS, R.D.; THOMAS, M. *Statistical Analysis of Extreme Values*. Birkhauser, Springer-Verlag, 1997.
- [43] RESNICK, S.I. *Extreme Values, Regular Variation, and Point Processes*. Springer, New York, 1987.
- [44] ROLL,R. *The International Crash of October 1987*. Financial Analysts Journal, p.19-35, 1988.
- [45] RUBEM, A.P.S. *Modelagem de extremos baseada nas r-maiores estatísticas de ordem*. Dissertação de Mestrado, IM-UFRJ, 2006.
- [46] RUDGE, L.F.; CAVALCANTE, F. *CNBV Mercado de Capitais*. Comissão Nacional de Bolsa de Valores, 1998.
- [47] SANFINS, M.A. *Cópulas para distribuições generalizadas de valores extremos multidimensionais*. Rio de Janeiro, UFRJ/IM, 2009.
- [48] SANFINS, M.A.; VALLE, G. *On the copula for the limiting distribution of the K largest order statistics of IID samples*. Rio de Janeiro, UFRJ/IM, 2009.
- [49] SCHNEEWEIS, THOMAS; MARTIN, GEORGE. *The Benefits of Hedge Funds: Asset Allocation for the Institutional Investor*. Lehman Brothers, 2000.

- [50] SCHNEEWEIS, THOMAS; MARTIN, GEORGE, KAZEMI, HOSSEIN B. *Understanding Hedge Fund Performance: Research Issues Revisited-Part II*. Journal of Alternative Investments, v.5, n.4, p.8-30, 2003.
- [51] SHARPE, WILLIAM F. *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*. The Journal of Finance, v.19, n.3, p. 425-442, 1964.
- [52] SHERMAN, L.K. *Percentiles of the Ω_π Statistic*. Annals of Mathematical Statistics, v.28, n.1, p.259-268, 1957.
- [53] SIBUYA, M. *Bivariate Extreme Statistics*. Annals of the Institute of Statistical Mathematics, n.11, p.195-210, 1960.
- [54] SMITH, R.L. *Measuring Risk with Extreme Value Theory*. Department of Statistics, University of North Carolina, Chapel Hill, 1999.
- [55] THOMAS, M. *Statistical Analysis of Extreme Values*. Birkhauser, Springer-Verlag, 1997.
- [56] TIAGO DE OLIVEIRA, J. *Extremal Distributions*. Revista da Faculdade Ciências, Lisboa, 2 ser. A, Mat., v.7, 1958.
- [57] TIAGO DE OLIVEIRA, J. *Structure Theory of Bivariate Extremes*. Extensions, Est. Mat. Estat. e Econ., n.7, p.165-195, 1962.
- [58] TREYNOR, J.L. *How to rate management investment funds*. Harvard Business Review, v.43, p.63-75, 1966.
- [59] VON MISES, R. *La distribution de la plus grande de n valeurs*. Reprinted in Selected Papers Volumen II, American Mathematical Society, Providence, R.I., 1954, pp.271-294, 1936.