

FUNDAÇÃO GETÚLIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

MARCIO CECILIO SANTOS

**REQUERIMENTOS DE CAPITAL PARA RISCO DE MERCADO NO BRASIL:
ABORDAGEM BASEADA NA TEORIA DE VALORES EXTREMOS.**

SÃO PAULO
2007

MARCIO CECILIO SANTOS

**REQUERIMENTOS DE CAPITAL PARA RISCO DE MERCADO NO BRASIL:
ABORDAGEM BASEADA NA TEORIA DE VALORES EXTREMOS.**

Dissertação de Mestrado apresentada à Escola de Economia de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção de título de Mestre em Economia Empresarial e Finanças.

Campo de conhecimento:
Gerenciamento de Risco

Orientador: Prof.Dr. Ricardo Matone

SÃO PAULO
2007

Santos, Marcio.

Requerimentos de Capital para Risco de Mercado no Brasil: Abordagem Baseada na Teoria de Valores Extremos. / Marcio Cecilio Santos. - 2007. 66 f.

Orientador: Ricardo Matone.

Dissertação (mestrado) - Escola de Economia de São Paulo.

1. Avaliação de riscos. 2. Valor (Economia). 3. Capital (Economia). 4. Investimentos - Análise. 5. I. Matone, Ricardo. II. Dissertação (mestrado) - Escola de Economia de São Paulo. III. Título.

CDU 330.14.011

MARCIO CECILIO SANTOS

**REQUERIMENTOS DE CAPITAL PARA RISCO DE MERCADO NO BRASIL:
ABORDAGEM BASEADA NA TEORIA DE VALORES EXTREMOS.**

Dissertação de Mestrado apresentada à Escola de Economia de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção de título de Mestre em Economia Empresarial e Finanças.

Data de aprovação:

__/__/__

Banca examinadora:

Prof. Dr. Ricardo Matone
FGV-EAESP

Prof. Dr. Richard Saito
FGV-EAESP

Prof. Dr. Franklin Gonçalves

SÃO PAULO
2007

À minha querida esposa Vivianne.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Ricardo Matone pela paciência, disposição irrestrita e por ajudar a encontrar os caminhos corretos nos momentos de maiores dificuldades durante a elaboração deste trabalho.

Aos amigos que conquistei durante o mestrado, em especial Daniel, Eliza, Flávio, Guido, Luciana e Rodrigo pelas inúmeras madrugadas que dedicamos estudando durante o curso.

Aos meus pais, grandes responsáveis pela base da minha formação acadêmica, que me incentivaram durante todo o curso.

Pelo apoio incondicional e paciência durante a aventura do mestrado, agradeço especialmente à minha esposa Vivianne.

Agradeço também ao Deutsche Bank S.A. por ter me proporcionado o apoio financeiro para a realização do mestrado e colegas de trabalho que direta ou indiretamente contribuíram para a conclusão deste curso.

RESUMO

Há forte evidência que os retornos das séries financeiras apresentam caudas mais pesadas que as da distribuição normal, principalmente em mercados emergentes. No entanto, muitos modelos de risco utilizados pelas instituições financeiras baseiam-se em normalidade condicional ou não condicional, reduzindo a acurácia das estimativas.

Os recentes avanços na Teoria de Valores Extremos permitem sua aplicação na modelagem de risco, como por exemplo, na estimação do Valor em Risco e do requerimento de capital.

Este trabalho verifica a adequação de um procedimento proposto por McNeil e Frey [1999] para estimação do Valor em Risco e conseqüente requerimento de capital às principais séries financeiras de retornos do Brasil. Tal procedimento semi-paramétrico combina um modelo GARCH ajustado por pseudo máxima verossimilhança para estimação da volatilidade corrente com a Teoria de Valores Extremos para estimação das caudas da distribuição das inovações do modelo GARCH.

O procedimento foi comparado através de *backtestings* com outros métodos mais comuns de estimação de VaR que desconsideram caudas pesadas das inovações ou a natureza estocástica da volatilidade.

Concluiu-se que o procedimento proposto por McNeil e Frey [1999] mostrou melhores resultados, principalmente para eventos relacionados a movimentos negativos nos mercados¹.

Futuros trabalhos consistirão no estudo de uma abordagem multivariada de grandes dimensões para estimação de VaR e requerimento de capital para carteiras de investimentos.

Palavras-chave: Requerimentos de Capital, Valor em Risco, Teoria de Valores Extremos, Modelo GARCH.

¹ Quedas de preços de ações, altas das taxas de juros e desvalorização da moeda nacional.

ABSTRACT

There is a strong evidence that financial return series are heavy-tailed, mostly in emerging markets. However, most of the risk models used by financial institutions are based in conditional or non-conditional normality, which reduces the accuracy of the estimates.

The recent advances in Extreme Value Theory permit its application to risk measuring, such as Value at Risk and capital adequacy estimates.

This work verifies the adequacy of a procedure proposed by McNeil and Frey [1999] to VaR and consequent capital requirement estimates for the main financial return series in Brazil. This semi parametric procedure combines a pseudo-maximum-likelihood fitting GARCH model to estimate the current volatility and the Extreme Value Theory (EVT) to estimate the tails of the innovations distribution of the GARCH model.

Using backtestings the procedure was compared to other common methods of VaR estimation that disregard heavy tails of the innovations or the stochastic nature of the volatility.

The procedure proposed by McNeil and Frey [1999] showed better results, mostly for negative events in the financial market².

Further works will consist of studying a high dimensional multivariate approach to estimate VaR and capital requirements for portfolios of investment instruments.

Keywords: Capital Requirements, Value at Risk, Extreme Value Theory, GARCH model.

² Stock prices fall, yields increase and local currency devaluation.

SUMÁRIO

<u>LISTA DE TABELAS E FIGURAS</u>	<u>10</u>
<u>1. INTRODUÇÃO</u>	<u>11</u>
1.1. INTRODUÇÃO	11
1.2. REVISÃO BIBLIOGRÁFICA.....	12
1.3. DESCRIÇÃO DO TRABALHO	17
<u>2. VALOR EM RISCO COM TEORIA DE VALORES EXTREMOS.....</u>	<u>19</u>
2.1. TEORIA DE VALORES EXTREMOS (TVE)	19
2.1.1 GEV	20
2.1.2 GPD	22
2.2. APLICAÇÃO EM VALOR EM RISCO	26
<u>3. EXIGÊNCIA DE CAPITAL</u>	<u>29</u>
3.1. INTRODUÇÃO	29
3.2. TVE PARA EXIGÊNCIA DE CAPITAL	34
<u>4. ESTUDO DE CASOS.....</u>	<u>37</u>
4.1. BASE DE DADOS DE MERCADO	37
4.2. DESCRIÇÃO DO PROCEDIMENTO	41
4.3. MODELAGEM	43
<u>5. CONCLUSÃO</u>	<u>49</u>
5.1. RESULTADOS DOS ESTUDOS DE CASOS	49
5.2. SUGESTÃO PARA TRABALHOS FUTUROS: CASO MULTIVARIADO	57
<u>6. ANEXOS</u>	<u>59</u>
<u>7. BIBLIOGRAFIA</u>	<u>64</u>

LISTA DE TABELAS E FIGURAS

<i>Figura 1 – Gráfico das Distribuições GEV: Weibull, Fréchet e Gumbel</i>	21
<i>Figura 2 – Representação Gráfica do procedimento Picks Over Threshold.</i>	23
<i>Figura 3 – Gráficos de preços e retornos da Taxa de Câmbio US\$/R\$ e Índice Bovespa.</i>	38
<i>Figura 4 – Gráficos de preços e retornos dos PUs de Juros “Pré” de 1 mês (1M), 6 meses (6M), 1 ano (1Y) e 2 anos (2Y).</i>	38
<i>Figura 5 – Gráficos de preços e retornos dos PUs de “Cupom Cambial Limpo” de 1 mês (1M), 6 meses (6M), 1 ano (1Y) e 2 anos (2Y).</i>	39
<i>Figura 6 - Histogramas das séries de Taxa de Câmbio US\$/R\$ (DOLP) e Índice Bovespa (IBP).</i>	40
<i>Figura 7 - Histogramas das séries de Juros “Pré” de 1M (P1MP), 6M (P6MP), 1Y (P1YP) e 2Y (P2YP).</i>	40
<i>Figura 8 - Histogramas das séries de Juros de Cupom Cambial de 1M (C1MP), 6M (C6MP), 1Y (C1YP) e 2Y (C2YP).</i>	41
<i>Figura 9 - QQ Plots dos Resíduos Padronizados do modelo AR(1) + GARCH(1,1) das séries de Taxas Pré-Fixadas de 1 mês (RESID_P1MP), 6 meses (RESID_P6MP), 1 ano (RESID_P1YP) e 2 anos (RESID_P2YP).</i>	49
<i>Figura 10 - QQ Plots dos Resíduos Padronizados do modelo AR(1) + GARCH(1,1) das séries de Cupom Cambial de 1 mês (RESID_C1MP), 6 meses (RESID_C6MP), 1 ano (RESID_C1YP) e 2 anos (RESID_C2YP).</i>	50
<i>Figura 11 - QQ Plots dos Resíduos Padronizados dos modelos AR(1) + GARCH(1,1) das séries de Dólar (RESID_DOLP) e Ibovespa (RESID_IBP).</i>	50
<i>Figura 12 - Hill Plot P1M – Max (esq.) e P1M – Min (dir.)</i>	51
<i>Figura 13 – Hill Plot P6M – Max (esq.) e P6M – Min (dir.)</i>	51
<i>Figura 14 – Hill Plot P1Y – Max (esq.) e P1Y – Min (dir.)</i>	52
<i>Figura 15 – Hill Plot P2Y – Max (esq.) e P2Y – Min (dir.)</i>	52
<i>Figura 16 – Hill Plot C1M – Max (esq.) e C1M – Min (dir.)</i>	52
<i>Figura 17 – Hill Plot C6M – Max (esq.) e C6M – Min (dir.)</i>	53
<i>Figura 18 – Hill Plot C1Y – Max (esq.) e C1Y – Min (dir.)</i>	53
<i>Figura 19 – Hill Plot C2Y – Max (esq.) e C2Y – Min (dir.)</i>	53

1. INTRODUÇÃO

1.1. INTRODUÇÃO

As metodologias tradicionais de gerenciamento de riscos de mercado costumam supor normalidade condicional ou não condicional dos retornos de séries financeiras. No entanto, há fortes evidências na literatura indicando que as distribuições desses retornos apresentam, em geral, caudas mais pesadas que a de uma distribuição normal. Os recentes avanços na Teoria de Valores Extremos e suas aplicações a finanças permitem uma sofisticação ao gerenciamento de risco pela possibilidade de introduzir modelos estatísticos paramétricos, especificamente desenvolvidos para as partes extremas das distribuições.

Este trabalho pretende examinar a adequação de um método univariado híbrido proposto por McNeil e Frey [1999] no contexto das séries de retornos dos principais instrumentos do mercado financeiro brasileiro. O método trata de um VaR condicional semi-paramétrico que consiste na modelagem da média e da volatilidade da série por um modelo auto regressivo e de variância condicional, cujas inovações são ajustadas por uma distribuição histórica para o centro da distribuição e uma distribuição generalizada de Pareto para as caudas.

O modelo acima foi aplicado às séries de taxa de câmbio comercial US\$ / R\$, Índice Bovespa e taxas de juros em Reais (chamadas Pré-fixadas) e em dólares (chamadas de cupom cambial) desde 1999 a 2006.

O modelo foi comparado com outros modelos de VaR que utilizam distribuição Gaussiana: um condicional e outro não condicional paramétrico. O modelo proposto apresentou resultados excepcionais, principalmente para as caudas inferiores das distribuições.

1.2. REVISÃO BIBLIOGRÁFICA

A teoria de valores extremos originou-se principalmente da necessidade de astrônomos utilizarem ou rejeitarem *outliers* em suas observações (ver Kotz e Nadarajah, 2000).

O conceito de distribuição de valores máximos foi introduzido por Bortkiewicz, em 1922, em artigo que tratava da distribuição do tamanho do intervalo entre máximos e mínimos a partir de uma amostra randômica da distribuição normal. Neste artigo concluiu-se que a distribuição de Poisson ajustava-se bem a eventos raros³.

Em seguida, a publicação mais relevante foi um artigo de Fréchet [1927] que introduziu a distribuição assintótica do máximo de uma amostra, identificando-a como sendo uma possível distribuição limite para estatísticas de ordem extrema. No ano seguinte, Fisher e Tippett [1928] mostraram que as distribuições do máximo poderiam somente pertencer a 3 tipos: Gumbel, Fréchet e Weibull. Von Mises [1936] apresentou as condições suficientes para convergência fraca das estatísticas de ordem extrema para cada uma das três distribuições limites propostas anteriormente por Fisher e Tippett. Em 1943, Gnedenko introduziu a distribuição generalizada de valores extremos e apresentou uma rigorosa fundamentação à Teoria de Valores Extremos, fornecendo as condições necessárias e suficientes para a convergência fraca em distribuição das estatísticas de ordem extrema.

Os resultados teóricos acima foram seguidos por vários estudos relacionados às aplicações práticas das estatísticas de valores extremos nas décadas de 1930 e 1940. Tais aplicações estendiam-se por uma variedade de áreas relacionadas a fenômenos naturais como poluição do ar, hidrologia, resistência dos materiais, emissões radioativas e análises pluviométricas. Dentre esses estudos destaca-se a contribuição de Gumbel [1941] ao chamar a atenção para a possível aplicação da teoria de valores extremos a certas distribuições que antes eram tratadas empiricamente. O primeiro caso envolvia fenômenos meteorológicos nos Estados Unidos da América (precipitações máximas, índices pluviométricos, inundações).

A aplicação de TVE em finanças também tem sido estudada por diversos autores mais recentemente, apesar da possibilidade de existência de caudas pesadas nos retornos de séries financeiras ter sido levantada já em 1963 por Mandelbrot.

Em 1996, Longin mostrou empiricamente que os retornos do índice que engloba as ações mais negociadas na Bolsa de Nova Iorque (*New York Stock Exchange*) para o período de 1885 a 1990 obedeciam a uma distribuição de Fréchet.

Danielsson e de Vries [1997] propuseram uma aplicação da TVE ao VaR. Os autores propuseram um método semi-paramétrico para cálculo de VaR, que modela parametricamente os eventos nas caudas da distribuição, enquanto que os eventos normais são modelados pela distribuição histórica. Em seus estudos, desenvolveram estimadores tanto para VaRs diários quanto para multi-períodos e demonstraram que para uma amostra de retornos de ações americanas, os métodos paramétricos condicionais subestimam o VaR e, conseqüentemente, o risco para os casos extremos. Eles concluíram que o modelo semi-paramétrico gerou melhores resultados que o método condicional de variância-covariância e que o método não-paramétrico de simulação histórica. Eles recomendaram o uso de VaR não condicional ao invés de modelos condicionais baseados em processos GARCH. Uma das razões para tal recomendação reside no fato da dificuldade de obtenção das estimações de VaR para multi-períodos⁴.

Além disso, propuseram duas abordagens simplificadas para a estimação de VaR para o caso multivariado (portfólio de ativos):

i) Post-fitting

Aplicam-se os pesos da carteira atual para os retornos históricos observados para os ativos que a compõe. Isso gera uma única série de retornos (como no caso da simulação histórica) que pode ser tratada como um caso univariado.

Depois, é necessário ajustar a distribuição das caudas da distribuição empírica obtida para a carteira.

³ Bortkiewicz utilizou, como base de dados, as mortes de oficiais da cavalaria Prussiana em acidentes com seus cavalos entre os anos de 1875 e 1894.

⁴ A hipótese de normalidade permite o uso do famoso e prático método da “raiz quadrada do tempo” implementado em RiskMetrics, ou seja, para obter-se a volatilidade para um período de n dias, multiplica-se a volatilidade de 1 dia pela raiz quadrada de n .

A grande desvantagem desse modelo é assumir que a matriz de correlação dos retornos dos ativos é constante ao longo de todo o período da amostra.

ii) Pre-sampling

Substituem-se os valores extremos por valores da distribuição das caudas ajustadas de forma independente para cada ativo da carteira. Depois, os retornos são reescalados realizando-se uma decomposição de Cholesky da matriz de retornos das séries:

Seja Σ_t a matriz de covariância dos retornos dos m ativos componentes da carteira para determinado período. Defini-se a matriz L_t como proveniente de uma fatoração de Cholesky tal que $L_t L_t' = \Sigma_t$.

Supõe-se também R_t^* uma matriz $m \times n$ de retornos simulados, obtidos do procedimento que une a distribuição empírica (retornos normais) e a cauda parametrizada por EVT (retornos extremos).

Se a matriz de covariância de R_t^* for dada por Ω_t , pode-se definir M_t como $M_t M_t' = \Omega_t$

Para obter os retornos simulados R_t com matriz de covariâncias Σ_t , basta reescalar R_t^* da forma : $R_t = L_t M_t^{-1} R_t^*$. Os retornos da carteira são finalmente obtidos ao aplicar-se o vetor de pesos w ($m \times 1$) na matriz de retornos gerados:

$$R_{\text{carteira}} = R_t^* w$$

Em 1998, McNeil estimou quantis nas caudas das distribuições de retornos de séries financeiras baseando-se no modelo de block-maxima. Concluiu que a abordagem de block-maxima requer uma quantidade de dados muito grande para se obter melhor precisão, já que grande parte das observações acaba sendo desprezada. No entanto, tal procedimento contorna o problema de aglomerados (*clusters*) de retornos.

Embrechts, Resnick e Samorodnitsky (1998a) defenderam o uso de uma estimação paramétrica baseada num resultado limite para a distribuição do excesso além de um limiar alto.

O modelo aplicado ao presente trabalho baseia-se, principalmente, no estudo de McNeil e Frey [1999]. Eles propõem um método de estimação de VaR modelando a cauda de uma distribuição condicional de uma série de retornos heterocedástica⁵. Tal abordagem combina a adequação de um modelo GARCH para estimar a estrutura de volatilidade e a teoria de valores extremos para estimar as caudas dos distúrbios desse modelo GARCH.

O procedimento é iniciado com o ajuste de um modelo GARCH à série de retornos através de uma estimação de pseudo-máxima verossimilhança. A partir dos retornos padronizados desse modelo, utilizaram a simulação histórica para o centro da distribuição e ajustaram uma distribuição de valores extremos às caudas.

A partir daí, estimaram VaR e perda esperada para algumas séries e, através de backtestings, concluíram que tal proposta gerava melhores resultados que os modelos GARCH com distúrbios gaussianos e VaR com EVT não condicional.

Ao final, adotaram um modelo híbrido de bootstrapping e simulação de Monte Carlo para a estimação de quantis de retornos para horizontes maiores que 1 dia, concluindo que tal método oferecia melhor performance que o método simplificado de “raiz quadrada do tempo”.

Brooks, Clare e Persaud (2000), com o intuito de calcular o requerimento mínimo de capital para risco de mercado, utilizaram uma abordagem semi-paramétrica, com as caudas sendo modeladas pela Distribuição Generalizada de Pareto e os retornos “normais”, através de distribuição empírica.

Eles compararam os resultados dessa abordagem com os resultados calculados por uma densidade não condicional e uma densidade condicional (GARCH(1,1)).

⁵ Com o intuito de tratar os casos de correlação serial e diferentes variâncias ao longo da série.

Nesse estudo, utilizaram as séries de preços de contratos futuros de taxa de juros em Libras esterlinas, contrato futuro de GILT (títulos públicos ingleses indexados à inflação) e contrato futuro de índice FTSE-100.

A distribuição de EVT por eles adotada nesse *paper* foi a Distribuição Generalizada de Pareto.

Analysaram as probabilidades de ocorrência dos maiores e menores eventos de cada uma das séries e concluíram que as probabilidades previstas pelo procedimento de valores extremos e os valores previstos pela distribuição empírica eram muito similares. Por outro lado, o modelo GARCH (1,1) tem desempenho ruim na modelagem dos eventos das caudas, se comparado com a abordagem de valores extremos.

Em seguida, utilizaram a metodologia de jackknife-after-bootstrap (EFRON & TIBSHIRANI, 1993) para calcular os intervalos de confiança para a estimativa de requerimento mínimo de capital para risco. A partir daí, compararam os retornos efetivos diários dos 3 contratos futuros com as previsões de requerimento de capital diárias.

A medida de desempenho dos modelos é, portanto, a contagem do número de vezes que o requerimento de capital prevê um valor menor que a perda efetiva dia para todo o período da amostra.

Além disso, eles realizaram também um teste fora da amostra, utilizando o mesmo critério descrito acima.

Por fim, concluem que:

- o modelo GARCH (1,1) estima sempre um requerimento de capital maior que o modelo de densidade não-condicional;
- os requerimentos de capital calculados pelos três modelos foram violados, em geral, menos do que o grau de confiança nominal de 5% adotado, portanto, os três mostraram-se adequados;

- no entanto, analisando mais profundamente, quanto menor o número de violações, mais rigoroso é o modelo, ou seja, maior a penalização imposta à instituição financeira, de forma desnecessária. Nesse aspecto, o modelo de EVT mostrou-se com número de violações um pouco maior que os demais e mais próximas do intervalo de 5%, impondo, portanto, menor penalização às instituições.

Hoje há diversas publicações sobre a TVE e suas aplicações, dentre elas, cabe mencionar Leadbetter, Lindgren e Rootzén [1983], Galambos [1987], Embrechts, Klüppelberg e Mikosch [1997], Kotz e Nadarajah [2000] e McNeil, Frey e Embrechts [2005]. No Brasil, destacam-se o livro de Mendes [2004], no qual grande parte da apresentação da teoria desse trabalho é baseada, e os artigos de Pinto e Valls Pereira [2006] e Souza e Silva [1999].

1.3. DESCRIÇÃO DO TRABALHO

Este trabalho pretende demonstrar a eficácia da aplicação da Teoria Univariada de Valores Extremos na mensuração do risco de séries financeiras brasileiras. A intenção é testar a proposta de uma modelagem híbrida de VaR condicional semi-paramétrico, cujo centro da distribuição é modelado através da distribuição empírica e as caudas são ajustadas para uma Distribuição Generalizada de Pareto. Os resultados do modelo serão comparados com os de outros modelos como: VaR condicional com distribuição Normal e VaR paramétrico com distribuição Normal. A partir do cálculo desse modelo de VaR, propõe-se a estimação do requerimento de capital, de acordo com as recomendações do BIS.

Para isso serão utilizadas séries de retornos da taxa de câmbio Reais por Dólar americano do mercado à vista, do Índice da Bolsa de Valores de São Paulo (Ibovespa) e das taxas de juros locais em Reais (“Pré”) e em Dólares (“Cupom Cambial”).

O presente trabalho está organizado da seguinte maneira: no Capítulo 2 é explicada a Teoria de Valores Extremos. Nas duas primeiras partes desse capítulo, após uma introdução genérica da TVE, aprofundam-se as explicações das duas diferentes classes de distribuições de extremos e suas aplicações: Distribuições Generalizadas de Extremo e Distribuições Generalizadas de Pareto, respectivamente. Na última parte, aborda-se a aplicação dessas teorias ao cálculo de VaR.

No Capítulo 3, discute-se o conceito de Exigência de Capital e relata-se sua história tanto internacional quanto no Brasil. A última parte desse capítulo trata da aplicação de TVE para Exigência de Capital.

No Capítulo 4 são realizados os estudos de casos das séries brasileiras. Na primeira parte há uma descrição das séries utilizadas, na segunda, explica-se o procedimento utilizado e na terceira, a modelagem estatística.

Por fim, no Capítulo 5 os resultados dos estudos de casos são apresentados.

2. VALOR EM RISCO COM TEORIA DE VALORES EXTREMOS

2.1. TEORIA DE VALORES EXTREMOS (TVE)

A teoria probabilística de valores extremos lida com o comportamento estocástico dos máximos e mínimos de variáveis randômicas independentemente e identicamente distribuídas.

O interesse deste trabalho é modelar as partes extremas das caudas da distribuição de uma variável X , que nesse caso, representa uma série de retornos diários de cada ativo. Alternativamente, X poderia ter uma variedade de interpretações tais como perdas de risco operacional, sinistros de seguros e perdas numa carteira de crédito durante um intervalo de tempo. Além disso, X pode representar tanto retornos de um único ativo financeiro quanto de uma carteira de ativos.

A Teoria de Valores Extremos engloba basicamente duas classes de distribuições:

- **Classes de Distribuições de Valores Extremos Generalizada (GEV):** que incluem as distribuições de Gumbel, Fréchet e Weibull.
- **Classes de Distribuições Generalizadas de Pareto (GPD):** que incluem distribuições como Beta, Exponencial e Pareto.

A GEV descreve a distribuição limite do máximo, enquanto a GPD descreve a distribuição limite dos excessos além de um limiar. Dessa forma, há duas abordagens principais na modelagem de distribuições de Valores Extremos:

- **Block Maxima (Máximos coletados em Blocos):** que consiste em analisar o conjunto de valores máximos (ou mínimos) coletados em blocos de observações. A esse método associam-se as distribuições GEV.
- **Picks Over Threshold (POT):** que consiste em ajustar uma distribuição de probabilidades aos valores que excedem determinado limiar. Em geral, associam-se as distribuições GPD.

O método de Block Maxima tem sido criticado (ver McNeil [1998]) por desperdiçar muitos dados, já que apenas as observações máximas de cada um dos blocos são retidas para análise. O método POT, por outro lado, utiliza todos os dados que excedem determinado limiar alto.

Por esta razão, neste trabalho serão apresentadas as duas abordagens, mas a modelagem proposta será baseada em POT.

2.1.1 GEV

Sejam X_1, X_2, X_3, \dots variáveis aleatórias independentes e identicamente distribuídas⁶ (iid) com função de distribuição $F_X : X \sim F_X$.

Define-se x_{F_X} o limite superior do suporte da distribuição F_X :

$$x_{F_X} = \sup\{x \in \mathbb{R} : F_X(x) < 1\}$$

Como a abordagem clássica da Teoria de Valores Extremos consiste em caracterizar as caudas da distribuição de F_X a partir da distribuição do máximo, define-se:

$$M_n = \max(X_1, X_2, \dots, X_n)$$

como o máximo de um conjunto de n variáveis aleatórias iid.

A função de distribuição do máximo⁷ é dada por:

$$F_{M_n}(x) = \Pr\{\max(X_1, \dots, X_n) \leq x\} = (\Pr\{X \leq x\})^n = F_X^n(x) \quad , \text{ para } x \in \mathbb{R} \text{ e } n \in \mathbb{N}.$$

⁶ Tal hipótese será relaxada no Capítulo 4 quando for considerada a modelagem da dependência temporal e heterocedasticidade através do VaR Condicional.

⁷ A distribuição do mínimo é obtida pela relação $\text{Min}(X_1, \dots, X_n) = -\text{Max}(-X_1, \dots, -X_n)$.

O Teorema de Fischer-Tippet [1928] fornece o resultado de convergência fraca para a distribuição do máximo padronizado.

Supondo X_n uma sequência de variáveis aleatórias iid. Se existem constantes de padronização $c_n > 0$ e $d_n \in \mathbb{R}$, e alguma distribuição não-degenerada H tal que:

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H$$

onde \xrightarrow{d} representa convergência em distribuição, então H pertence a uma das três funções de distribuição:

$$\text{Gumbel : } H_0(x) = \exp(-e^{-x}), \forall x \in \mathbb{R}$$

$$\text{Fréchet : } H_{1,\alpha}(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad \alpha > 0$$

$$\text{Weibull : } H_{2,\alpha}(x) = \begin{cases} \exp[-(-x)^{-\alpha}] & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha < 0$$

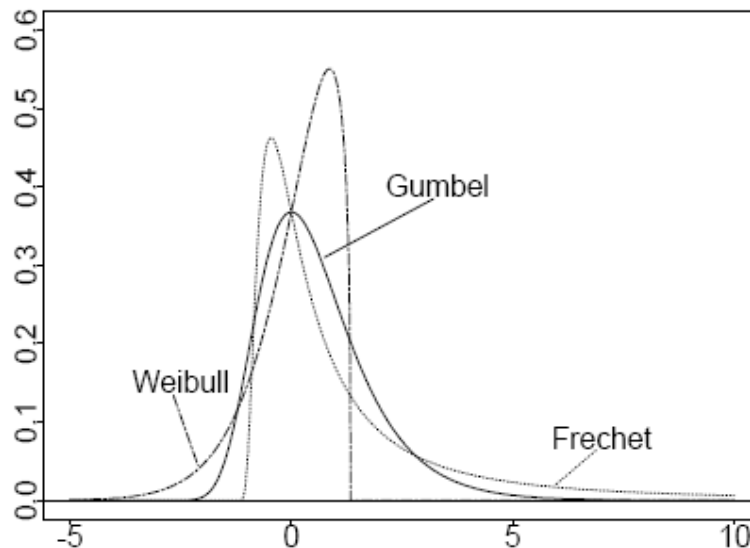


Figura 1 – Gráfico das Distribuições GEV: Weibull, Fréchet e Gumbel

Jenkinson [1955] definiu uma família paramétrica que representa as três distribuições acima descritas e dessa forma facilita a identificação da correta

distribuição do extremo de uma determinada amostra. Tal família é conhecida como **Distribuição Generalizada do Valor Extremo (GEV)** e definida como $H_{\xi,\sigma,\mu}$.

Primeiramente, define-se H_ξ como:

$$H_\xi(x) = \begin{cases} \exp\left[-(1+\xi x)^{-1/\xi}\right] & \text{se } \xi \neq 0 \\ \exp[-\exp(-x)] & \text{se } \xi = 0 \end{cases} \quad \text{onde } 1+\xi x > 0$$

É possível ainda substituir-se x na expressão acima pela transformação de escala $\frac{x-\mu}{\sigma}$ com $\mu \in \mathbb{R}$ e $\sigma > 0$, produzindo-se a distribuição:

$$\text{GEV} : \quad H_{\xi,\sigma,\mu}(x) = \exp\left\{-\left(1+\xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\}$$

Para o caso limite do índice de cauda $\xi \rightarrow 0$, $H_\xi(x)$ corresponde à distribuição Gumbel, quando $\xi < 0$, $H_\xi(x)$ corresponde à distribuição Weibull e quando $\xi > 0$, $H_\xi(x)$ corresponde à distribuição Fréchet.

$\xi = 0$	$\xi < 0$	$\xi > 0$
Gumbel	Weibull	Fréchet

2.1.2 GPD

Será apresentada neste capítulo a Distribuição Generalizada de Pareto e demonstrado que se X for uma variável aleatória com função distribuição $F_X \in \text{DMA}(H_\xi)$ então o excesso de X além de um limiar u ($X - u$) padronizado, segue a GPD.

Sejam X_1, X_2, \dots, X_N variáveis aleatórias iid com função de distribuição $F_X \in \text{DMA}(H_\xi)$ para $\xi \in \mathbb{R}$. Seja também x_{FX} o limite superior da distribuição F_X e x_0 um ponto (finito ou infinito) da extrema direita da cauda da distribuição F . Predefine-se um limiar alto u como sendo um suporte de X próximo de x_{FX} . Denominam-se excedentes os valores de X_i tais que $X_i > u$. Denominam-se também N_u como sendo o número de excedentes do limiar u e Y_1, Y_2, \dots, Y_{N_u} os valores $X_i - u \geq 0$.

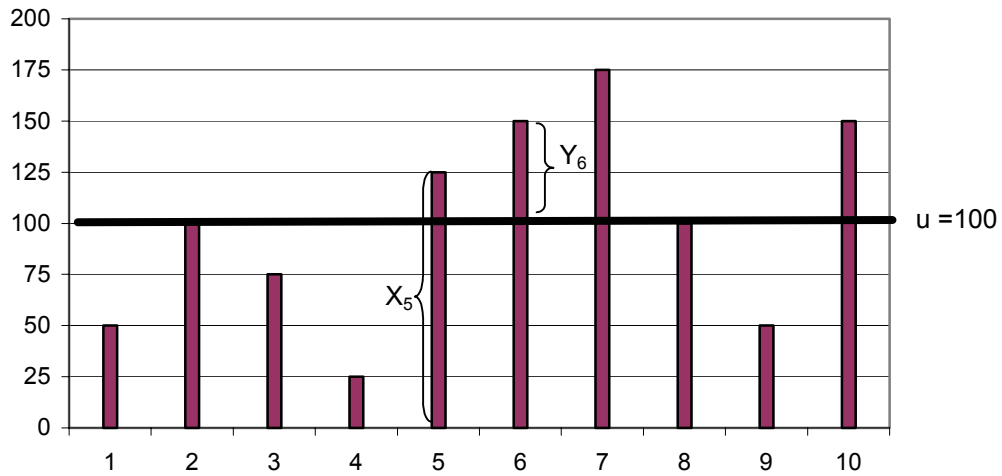


Figura 2 – Representação Gráfica do procedimento Pick's Over Threshold.

A distribuição condicional do excesso $Y = X - u$, representada por F_u é dada por:

$$F_u(y) = P\{Y \leq y \mid X > u\} = P\{X - u \leq y \mid X > u\} = \frac{F_X(u+y) - F_X(u)}{\bar{F}_X(u)}, \quad y \geq 0$$

Como $\bar{F}_u(y) = P\{X - u > y \mid X > u\} = \frac{\bar{F}_X(u+y)}{\bar{F}_X(u)}$, então $\bar{F}_X(u+y) = \bar{F}_u(y)\bar{F}_X(u)$.

Analogamente à distribuição GEV, que descreve a distribuição limite do máximo padronizado, a distribuição de Pareto Generalizada (GPD) descreve a distribuição limite dos excessos além de um limiar alto.

O Teorema de Pickands–Balkema–deHaan⁸ (Balkema & de Haan 1974, Pickands 1975 apud McNeil 1997), assim como o Teorema de Fisher-Tippett em relação às distribuições dos Máximos, postula que a distribuição dos excessos das observações

⁸ BALKEMA, A. & de HAAN, L. (1974), "Residual life time at great age", Annals of Probability 2, 792-804 e PICKANDS, J. (1975), "Statistical inference using extreme order statistics", The Annals of Statistics 3, 119-131.

além de um determinado limiar alto, pode ser aproximada por uma GPD. Ou seja, quando u tende a x_0 , a distribuição tende a uma GPD, desde que a distribuição F satisfaça o teorema de Fisher-Tippett.

Formalmente:

$$\lim_{u \rightarrow x_0} \sup_{0 \leq x \leq x_0 - u} |F_u(x) - G_{\xi, \sigma(u)}(x)| = 0, \text{ se e somente se } F \in \text{DMA}(H_\xi).$$

Teorema Pickands–Balkema–deHaan: $F_X \in \text{DMA}(H_\xi)$ se e somente se, existir uma função $a(\cdot)$ positiva e mensurável tal que para $1 + \xi x > 0$ tem-se:

$$\lim_{u \rightarrow x_{F_X}} \frac{\bar{F}_X(u + \xi a(u))}{\bar{F}(u)} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{se } \xi \neq 0 \\ e^{-x} & \text{se } \xi = 0 \end{cases}.$$

Reescrevendo-se a expressão acima, têm-se que a distribuição limite para os excessos acima do limiar u (ajustados por um fator de escala $a(u)$) é:

$$\lim_{u \rightarrow x_{F_X}} P\left\{\frac{X - u}{a(u)} > x \mid X > u\right\} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{se } \xi \neq 0 \\ e^{-x} & \text{se } \xi = 0 \end{cases}.$$

A expressão acima dá uma aproximação para a distribuição dos excessos, apropriadamente reduzidos por um fator de escala, acima do limiar u .

Considere-se, agora, a distribuição padrão Generalizada de Pareto, cuja função distribuição G_ξ é dada por:

$$G_\xi(y) = \begin{cases} 1 - (1 + \xi y)^{-1/\xi}, & \text{se } \xi \neq 0 \\ 1 - e^{-y}, & \text{se } \xi = 0 \end{cases}, \text{ onde } y \geq 0 \text{ se } \xi \geq 0 \text{ e } 0 \geq y \geq -1/\xi \text{ se } \xi < 0.$$

A densidade de G_ξ é, por sua vez, é dada por :

$$G_\xi(y) = (1 + \xi y)^{-(1+1/\xi)}, \text{ para } 0 \leq y \text{ se } \xi \geq 0 \text{ e } 0 \leq y \leq 1/|\xi| \text{ se } \xi < 0.$$

Tem-se então que a GPD adequa-se à modelagem dos excessos além de um limiar alto pelo fato de a distribuição F_X de X pertencer ao domínio de atração da GEV.

É possível substituir-se y por uma transformação de localização e escala $(y - \delta) / \beta$, para $\delta \in \mathbb{R}$ e $\beta > 0$, obtendo-se a família $G_{\xi, \beta, \delta}$. Dentro da classe $G_{\xi, \beta, \delta}$, a distribuição de maior interesse nas aplicações práticas é a $G_{\xi, \beta, 0}$ para $\xi \geq 0$.

Assim, reescrevendo-se a distribuição GPD para o caso $G_{\xi, \beta, 0}$, tem-se:

$$G_{\xi, \beta}(y) = \begin{cases} 1 - (1 + \xi y / \beta)^{-1/\xi} & \text{se } \xi \neq 0 \\ 1 - e^{-y/\beta} & \text{se } \xi = 0 \end{cases}$$

Assim como no caso da GEV, a GPD contém três casos especiais ou classes de distribuições padrões ($\beta = 1$, $\delta = 0$, $\alpha = 1 / \xi$):

- i) Tipo I – Exponencial: distribuições cujas caudas decaem exponencialmente (como a Normal) quando $\xi = 0$.

$$G_0(y) = \begin{cases} 0 & \text{se } y < 0 \\ 1 - e^{-y} & \text{se } y \geq 0 \end{cases}$$

- ii) Tipo II – Pareto: distribuições cujas caudas decaem como um polinômio (como por exemplo, t-Student), quando $\xi > 0$.

$$G_{\alpha}(y) = \begin{cases} 0 & \text{se } y < 1 \\ 1 - y^{-\alpha} & \text{se } y \geq 1 \end{cases}$$

- iii) Tipo III – Beta: distribuições cujas caudas são finitas, quando $\xi < 0$.

$$G_{\alpha}(y) = \begin{cases} 0 & \text{se } y < -1 \\ 1 - (-y)^{-\alpha} & \text{se } -1 \leq y \leq 0 \\ 1 & y > 0 \end{cases}$$

A GPD pode ser ajustada aos dados de excessos de um limiar, através de uma variedade de procedimentos já propostos na literatura, sendo os mais comuns: Máxima Verossimilhança e Probability-Weighted Moments (PWM, ver Greenwood et

al. [1979]). Para uma comparação entre os méritos relativos dos dois métodos, sugere-se a consulta a Hosking e Wallis [1987].

O método mais utilizado e mais fácil de ser implementado é o de Máxima Verossimilhança, que gera estimadores com boas propriedades para GPD.

Escrevendo $G_{\xi,\beta}$ para a densidade da GPD, a log-verossimilhança pode ser calculada como:

$$\ln L(\xi, \beta, u; Y_1, \dots, Y_{N_u}) = \sum_{j=1}^{N_u} \ln g_{\xi, \beta}(Y_j)$$

$$\ln L(\xi, \beta, u; Y_1, \dots, Y_{N_u}) = -N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{Y_j}{\beta}\right),$$

e deve ser maximizada sujeita a $\beta > 0$ e $1 + \xi Y_j/\beta > 0$ para todo j . A resolução é obtida numericamente.

Hosking e Wallis [1987] demonstraram que, para $\xi > -0,5$ (que é o caso de todas as aplicações com caudas pesadas), os estimadores de máxima verossimilhança $(\hat{\xi}, \hat{\beta})$ são assintótica e conjuntamente normais:

$$N_u^{1/2} \left(\hat{\xi} - \xi, \frac{\hat{\beta}}{\beta} - 1 \right) \xrightarrow{d} N_2(0, M^{-1}), \quad N_u \rightarrow \infty,$$

onde

$$M^{-1} = (1 + \xi) \begin{pmatrix} 1 + \xi & 1 \\ 1 & 2 \end{pmatrix},$$

e onde $N_2(a, A)$ significa a distribuição normal bivariada com média $a \in \Re^2$.

2.2. APLICAÇÃO EM VALOR EM RISCO

De acordo com a definição de VaR encontrada em Jorion [2001]:

“O VaR sintetiza a maior (ou pior) perda esperada dentro de determinados período de tempo e intervalo de confiança”.

Em termos probabilísticos o VaR é definido como o percentil $(1 - \alpha)$ da distribuição de retornos dos preços dos ativos sobre um horizonte de tempo estipulado. Ou seja, dado um nível de confiança α , o VaR é o menor número v tal que a probabilidade de que uma perda V exceda v não seja maior que $(1 - \alpha)$:

$$VaR_{\alpha} = \inf\{v \in \mathcal{R} : P(V > v) \leq 1 - \alpha\} = \inf\{v \in \mathcal{R} : F_V(v) \geq \alpha\}, \quad \alpha \in (0,1)$$

De forma genérica, o VaR pode ser derivado da distribuição de probabilidade do valor futuro⁹ da carteira $f(v)$. A determinado nível de confiança α deseja-se descobrir a pior realização possível V^* , tal que a probabilidade de se exceder esse valor seja α :

$$\alpha = \int_{V^*}^{\infty} f(v) dv.$$

Ou, analogamente, a probabilidade $p = 1 - \alpha$ de um valor ser menor que V^* pode ser escrita como:

$$1 - \alpha = \int_{-\infty}^{V^*} f(v) dv = P(v \leq V^*) = p$$

Conforme Jorion [2001], o cálculo do VaR pode ser simplificado de forma considerável adotando-se uma abordagem¹⁰ paramétrica, na qual se supõe que a distribuição pertence a uma determinada família, como por exemplo a Normal. Nesse caso, o VaR pode ser facilmente derivado do desvio padrão da carteira, utilizando-se um fator multiplicativo que depende do nível de confiança adotado:

$$VaR_{\alpha} = \mu + \sigma \phi^{-1}(\alpha),$$

⁹ Para um horizonte determinado.

¹⁰ Outras abordagens comuns seriam as leituras de um quantil da distribuição empírica (chamada de simulação histórica) ou de uma distribuição simulada (chamada de simulação de Monte Carlo).

onde Φ denota a função distribuição de probabilidade Normal padronizada e $\Phi^{-1}(\alpha)$ é o quantil α de Φ . A prova desse resultado pode ser encontrada em McNei, Frey e Embrechts [2005].

Resultado similar ao acima pode ser obtido por qualquer família de distribuição com parâmetros de locação e escala definidos. No presente trabalho será utilizada a família de distribuições generalizadas de Pareto.

3. EXIGÊNCIA DE CAPITAL

3.1. INTRODUÇÃO

Entende-se por exigência de capital para risco de mercado a quantidade mínima de capital que uma instituição deve manter como um colchão para absorver perdas inesperadas em seu portfólio advindas de oscilações de preços dos instrumentos. A exigência de capital é imposta pelo órgão regulador do sistema financeiro que deve zelar pela solvência da instituição.

O grande marco na regulamentação financeira dos bancos foi a conclusão do Acordo da Basiléia em 15 de julho de 1988 pelos bancos centrais dos países membros do G-10¹¹, sob o patrocínio do BIS (Bank for International Settlements).

O principal objetivo do acordo de Basiléia foi fortalecer a estabilidade do sistema financeiro através da orientação sobre a manutenção de um capital mínimo por parte dos bancos, com o objetivo de reduzir os riscos de insolvência.

O presente trabalho visa aprofundar as discussões sobre a parcela desse capital associada ao risco de mercado dos ativos presentes nas carteiras das instituições financeiras no Brasil.

O Comitê de Supervisão Bancária de Basiléia publicou, em janeiro de 1996, uma alteração ao seu Acordo de Capital propondo duas abordagens para calcular a reserva de capital para cobertura de riscos de mercado:

- abordagem padronizada
- modelos internos

O Comitê propôs também uma série de critérios qualitativos e quantitativos que deveriam ser seguidos pelos bancos que optassem pelo uso de modelos proprietários.

¹¹ Bélgica, Canadá, França, Alemanha, Itália, Japão, Holanda, Suécia, Suíça, Reino Unido e Estados Unidos.

A ferramenta especificada nessa publicação para medida de risco de mercado foi o VaR. Os bancos poderiam, assim, utilizar seus modelos internos de estimação de VaR como base de cálculo do requerimento de capital.

A abordagem padronizada definia percentuais a serem aplicados sobre os diferentes tipos de exposições a taxas de juros, ações ou câmbio. A reserva de capital era então calculada pela soma das ponderações de cada exposição.

Com relação à adoção de modelos internos, o Comitê definiu padrões quantitativos e qualitativos que deveriam ser seguidos pelas instituições financeiras para administrar seus riscos de mercado. Assim, as seguintes diretrizes foram determinadas para cada tipo de risco:

a) Taxas de Juros

Os bancos devem definir fatores de risco correspondentes às taxas de juros em cada uma das moedas em que detêm exposições. Além disso, os sistemas de risco devem ser capazes de capturar as diferentes volatilidades para cada vencimento ao longo das curvas de taxas de juros. Ou seja, devem tratar cada um desses vencimentos como um fator de risco diferente.

O risco associado à diferença entre taxas de juros de swaps¹² e títulos públicos (chamado de *spread risk*) deve também ser incorporado.

b) Taxas de Câmbio

O sistema deve considerar cada uma das moedas a que a instituição está exposta como um fator de risco exclusivo.

c) Ações

Para capturar corretamente os riscos de posições em ações, as instituições devem considerar:

¹² De acordo com a definição dada por Hull [1995], *Swaps* são acordos privados entre duas partes para a troca futura de fluxos de caixa, respeitando uma fórmula preestabelecida.

i) ao menos um modelo simplificado que defina um fator de risco para cada mercado de ações e as posições em cada uma das ações individuais ou setores, podem ser expressos como beta-equivalentes¹³ em relação ao índice de mercado.

ii) uma abordagem intermediária, que consiste em definir fatores de risco diferentes para cada setor (por exemplo, telecomunicações, financeiro, energia, siderurgia, entre outros) e então considerar as posições em cada ação individual como beta-equivalentes em relação a cada índice setorial.

iii) uma abordagem mais sofisticada e extensa que atribua um fator de risco específico para cada ação individualmente.

A escolha entre os três modelos acima (mais simples, intermediário e mais sofisticado) deveria levar em consideração a natureza dos riscos de cada instituição.

De acordo com a abordagem baseada em modelos internos, o requerimento de capital deve ser computado multiplicando-se o VaR por um fator que varia de 3 a 4, dependendo da acurácia do modelo nos períodos anteriores.

O cálculo do VaR utilizado para requerimento de capital, segundo os reguladores, deve levar em conta um horizonte de 10 dias e nível de confiança de 99%, utilizando-se pelo menos um ano de base de dados históricos.

Assim, segundo o Comitê de Basileia, para bancos que optam pela utilização da abordagem de modelo interno para cálculo de capital regulatório para risco de mercado, a fórmula é:

$$RC^t(MR) = \max \left\{ VaR_{0.99}^{t,10}, \frac{k}{60} \sum_{i=1}^{60} VaR_{0.99}^{t-i+1,10} \right\} + C_{SR},$$

onde

$VaR_{0.99}^{t,10}$ corresponde ao VaR a 99% de confiança para um horizonte de 10 dias, calculado para data t ;

¹³ Uma posição beta-equivalente é calculada a partir de um modelo de mercado de retorno de ativos (como o modelo de CAPM), regredindo-se os retornos de uma ação individual (ou índice setorial) em relação ao índice de mercado e a taxa de juros livre de risco.

$3 \leq k \leq 4$ representa um fator multiplicador que é uma função da qualidade do modelo interno de risco da instituição;

C_{SR} é um componente de risco específico.

Seguindo as diretrizes do acordo de 1988, o Banco Central do Brasil publicou, em 1994, a Resolução 2.099 que estabeleceu, em seu Anexo IV, que as instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil, deveriam manter capital (PLA - Patrimônio Líquido Ajustado, depois alterado para PR – Patrimônio de Referência) compatível com o grau de risco da estrutura de seus ativos. Dessa forma, deveriam constituir Patrimônio Líquido Exigido (PLE) de pelo menos 8% de seus ativos ponderados pelo risco.

$$PLE = 0,08 \times (APr)$$

Onde APr – ativos ponderados pelo risco

Posteriormente, foram publicados diversos outros normativos que complementaram o cálculo do PLE, alterando as ponderações ou adicionando outros componentes de risco.

A Resolução 2139 foi publicada em dezembro de 1994 e incorpora um componente à expressão do PLE, relacionado ao risco de operações de swap. O novo PLE passou a ser:

$$PLE = 0,08 \times (APr) + 0,15 \times (\sum_{i=1}^n RCD_i)$$

Onde RCD – risco de crédito das operações de swap.

Até esse momento, os fatores APr e RCD cobriam apenas o componente de risco de crédito, sendo o primeiro associado ao risco corrente e o segundo (aplicado a derivativos, particularmente a swaps) associado ao risco futuro (representando o *add-on* de Basileia ou *potential future exposure*).

Em junho de 1997 a Resolução 2399 alterou dois fatores da expressão anterior: modifica o fator aplicável ao risco de crédito das operações de “swap” (Sw) de 0,15

para 0,16 e passa de 0,08 para 0,10 o fator aplicado sobre os ativos ponderados pelos fatores de risco (APr).

Esta mesma Resolução insere uma nova sistemática para o cálculo do RCDi (risco de crédito da i-ésima operação de “swap”). Novamente, em novembro de 1997 esses dois fatores são modificados pela Circular 2.784, a qual estabeleceu a seguinte expressão para o cálculo do PLE:

$$PLE = 0,11 \times (Apr) + 0,20 \times \left(\sum_{i=1}^n RCD_i \right)$$

Apenas em maio de 1999, com a Resolução 2.606 foi incorporado o primeiro componente relacionado ao risco de mercado: a exposição em ouro e operações referenciadas em câmbio.

$$PLE = 0,11 \times (Apr) + 0,20 \times \left(\sum_{i=1}^n RCD_i \right) + 0,50 \times \max \left[\left(\sum_{i=1}^n |Aprc_i| - 0,20 \times PLA \right), 0 \right]$$

Onde $\sum_{i=1}^n |Aprc_i|$ representa o somatório dos valores absolutos das posições líquidas em ouro e em cada uma das moedas.

A Resolução 2.692, publicada em fevereiro de 2000, adicionou outro componente relacionado ao risco de mercado ao requerimento de capital mínimo: a parcela de capital necessária para a cobertura de riscos decorrentes de operações pré-fixadas denominadas em reais. Para o cálculo dessa parcela de capital, utilizou-se um critério mais sofisticado, baseado no VaR das operações :

$$EC_{JurosPre} = \max \left[\left(\frac{M_t}{60} \times \sum_{i=1}^{60} VaR_{t-i}^{Padrão} \right), VaR_{t-1}^{Padrão} \right]$$

Onde:

Mt representa o multiplicador para o dia t, divulgado diariamente pelo Banco Central do Brasil, compreendido entre 1 e 3 e calculado através de uma função inversa da volatilidade.

VaR representa o Valor em Risco das operações pré-fixadas denominadas em reais.

Em fevereiro de 2000, a Resolução 2.692 altera definitivamente o cálculo do PLE como mostra a expressão abaixo, que prevalece até o presente:

$$PLE = 0,11 \times (Apr) + 0,20 \times \left(\sum_{i=1}^n RCD_i \right) + 0,50 \times \max \left[\left(\sum_{i=1}^n |Aprc_i| - 0,20 \times PLA \right) 0 \right] + \sum_{i=1}^n EC_i$$

O cálculo da parcela Aprc que corresponde ao risco de mercado relacionado às exposições em moedas estrangeiras e ouro é extremamente penalizante às instituições financeiras, uma vez que não leva em conta nenhum critério estatístico.

Por outro lado, o cálculo da EC, que corresponde ao risco de mercado das operações pré-fixadas denominadas em Reais, baseia-se em um modelo de VaR condicional, que leva em conta o comportamento estocástico da volatilidade. No entanto, assume que os resíduos do modelo de volatilidade condicional seguem uma distribuição normal, o que representa um pressuposto incorreto, conforme conclusão dos testes de normalidade dos resíduos feitos no Capítulo 5.

Outra deficiência do modelo atual de requerimento de capital para risco de mercado é que não considera demais fatores de riscos a que estão sujeitas a maior parte das instituições financeiras como, por exemplo, cupom cambial, cupons de inflação¹⁴, ações e mercadorias¹⁵.

Por fim, o modelo atual não considera nenhum tipo de correlação ou dependência entre as diferentes parcelas de capital, já que o PLE é calculado pela soma algébrica de todas as parcelas.

3.2. TVE PARA EXIGÊNCIA DE CAPITAL

O interesse então se volta ao cômputo do VaR com 99% de confiança para um período de 10 dias, conforme requerimento do BIS, a partir de um modelo ajustado a retornos diários.

¹⁴ Taxas de Juros reais relacionadas aos títulos e derivativos indexados ao IPCA ou IGP-M.

¹⁵ Exceto ouro que já é considerado na parcela Aprc.

Para conversão do VaR calculado com base em retornos diários independentes e identicamente distribuídos para o VaR representativo de um período maior, no caso 10 dias, é necessário multiplicá-lo pela raiz de 10.

Considerando X_1, X_2, \dots retornos diários independentes e identicamente distribuídos.

Sabe-se que $E(X_1 + X_2) = E(X_1) + E(X_2)$ e $V(X_1 + X_2) = V(X_1) + V(X_2) + 2\text{Cov}(X_1, X_2)$, onde E representa esperança e V representa Variância.

Se os retornos são independentes, tem-se que $\text{Cov}(X_1, X_2) = 0$, e se são distribuídos de forma idêntica, tem-se : $E(X_1) = E(X_2) = E(X)$ e $V(X_1) = V(X_2) = V(X)$.

Assim, $E(X_1 + X_2) = 2E(X)$ e $V(X_1 + X_2) = 2V(X)$.

E, portanto, $\sigma(X_1 + X_2) = \sqrt{2}\sigma(X)$.

A regra acima pressupõe uma distribuição Normal. Conforme será mostrado no Capítulo 4.3, a pressuposição de retornos auto-regressivos com distribuições diferentes da Normal, implica na proposição de uma regra alternativa para cômputo do VaR multi-períodos.

Outro aspecto relacionado à transformação do VaR em exigência de capital, segundo o Comitê de Basileia, diz respeito ao fator multiplicador $3 \leq k \leq 4$ determinado como função da qualidade do modelo interno adotado pela instituição.

A escolha de k é justificada na desigualdade de Chebyshev¹⁶ (Stahl[1997] apud Jorion [2001]). Para qualquer variável aleatória x de variância finita e conhecida σ^2 , a probabilidade de obtenção de uma observação fora de um intervalo específico é:

$$P(|x - \mu| > r\sigma) \leq \frac{1}{r^2}, \text{ para qualquer } r > 0, r \in \mathbb{R}.$$

Supondo-se que a distribuição seja simétrica e tomando-se os retornos negativos, tem-se:

$$P(x - \mu < -r\sigma) \leq \frac{1}{2r^2}.$$

¹⁶ STAHL, G. (1997). "Three Cheers". Risk, 10 October, p. 67-69.

Fixando-se a probabilidade em 1% (nível de significância correspondente ao grau de confiança de 99%), têm-se $r(99\%) = 7,071$. O VAR máximo é, portanto:

$$VAR_{m\acute{a}x} = r(99\%) \sigma$$

A desigualdade de Chebyshev vale para variáveis randômicas de qualquer tipo de distribuição, e pode-se calcular r apenas com conhecimento da média e variância.

Supondo-se um modelo de VaR que utiliza uma distribuição Normal, calcula-se o $VaR(99\%)$ multiplicando-se o quantil da distribuição Normal padrão pela volatilidade, obtendo-se:

$$VAR_N = 2,33 \sigma$$

Caso a verdadeira distribuição não seja especificada corretamente, o fator de correção será:

$$k = \frac{VaR_{m\acute{a}x}}{VaR_N} = \frac{7,071 \sigma}{2,33 \sigma} = 3,03$$

Justificando o fator multiplicador adotado pelo Comitê de Basiléia.

Para o caso de instituições cujos modelos de risco incorporem distribuições com caudas pesadas, o quantil será superior ao quantil da distribuição Normal, portanto, o fator multiplicador deveria ser recalibrado para baixo.

4. ESTUDO DE CASOS

4.1. BASE DE DADOS DE MERCADO

Foram analisadas as séries históricas de retornos logarítmicos diários dos seguintes ativos financeiros (fatores de riscos):

1) Cotação de fechamento US\$ / R\$ do mercado de câmbio à vista, obtida pelo terminal Bloomberg e através de corretoras.

2) PUs (preços unitários) referentes às taxas de Juros Locais em Reais (“Pré”), correspondentes aos vencimentos padronizados de 1 mês (P1MP), 6 meses (P6MP), 1 ano (P1YP), 2 anos (P2YP) obtidos através de interpolação (*flat forward*¹⁷) dos PUs dos contratos de DI futuros negociados na BM&F.

3) PUs (preços unitários) referentes às taxas de Juros Locais em Dólares (“Cupom Limpo”) correspondentes aos vencimentos padronizados de 1 mês (C1MP), 6 meses (C6MP), 1 ano (C1YP), 2 anos (C2YP) obtidos através de transformação das taxas dos contratos de FRC negociados na BM&F em taxas spot e interpolação (*flat forward*) dos respectivos PUs.

4) Índice Bovespa de fechamento.

Todas as séries utilizadas partem do dia 1º de março de 1999, após a mudança de regime cambial.

¹⁷ Tipo de interpolação que assume que a taxa de juros a termo, implícita entre dois prazos, é constante durante esse intervalo.

Os gráficos das Figuras 3, 4 e 5 mostram as evoluções dos preços e retornos de cada uma das séries descritas acima.

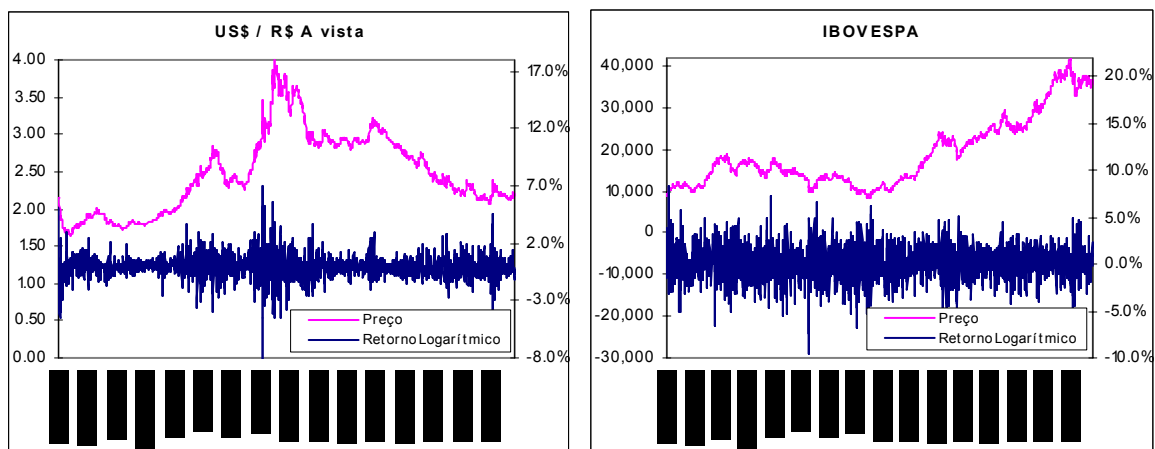


Figura 3 – Gráficos de preços e retornos da Taxa de Câmbio US\$/R\$ e Índice Bovespa.

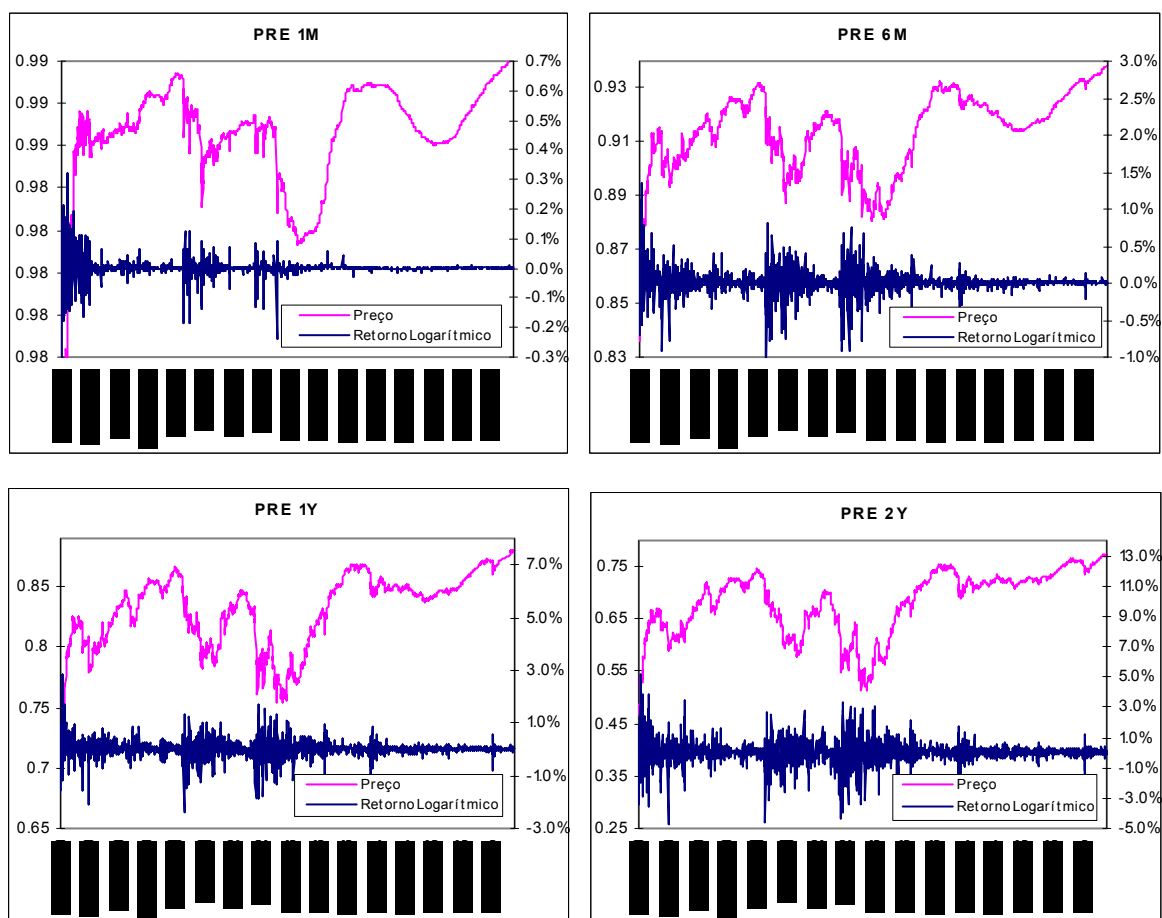


Figura 4 – Gráficos de preços e retornos dos PUs de Juros “Pré” de 1 mês (1M), 6 meses (6M), 1 ano (1Y) e 2 anos (2Y).

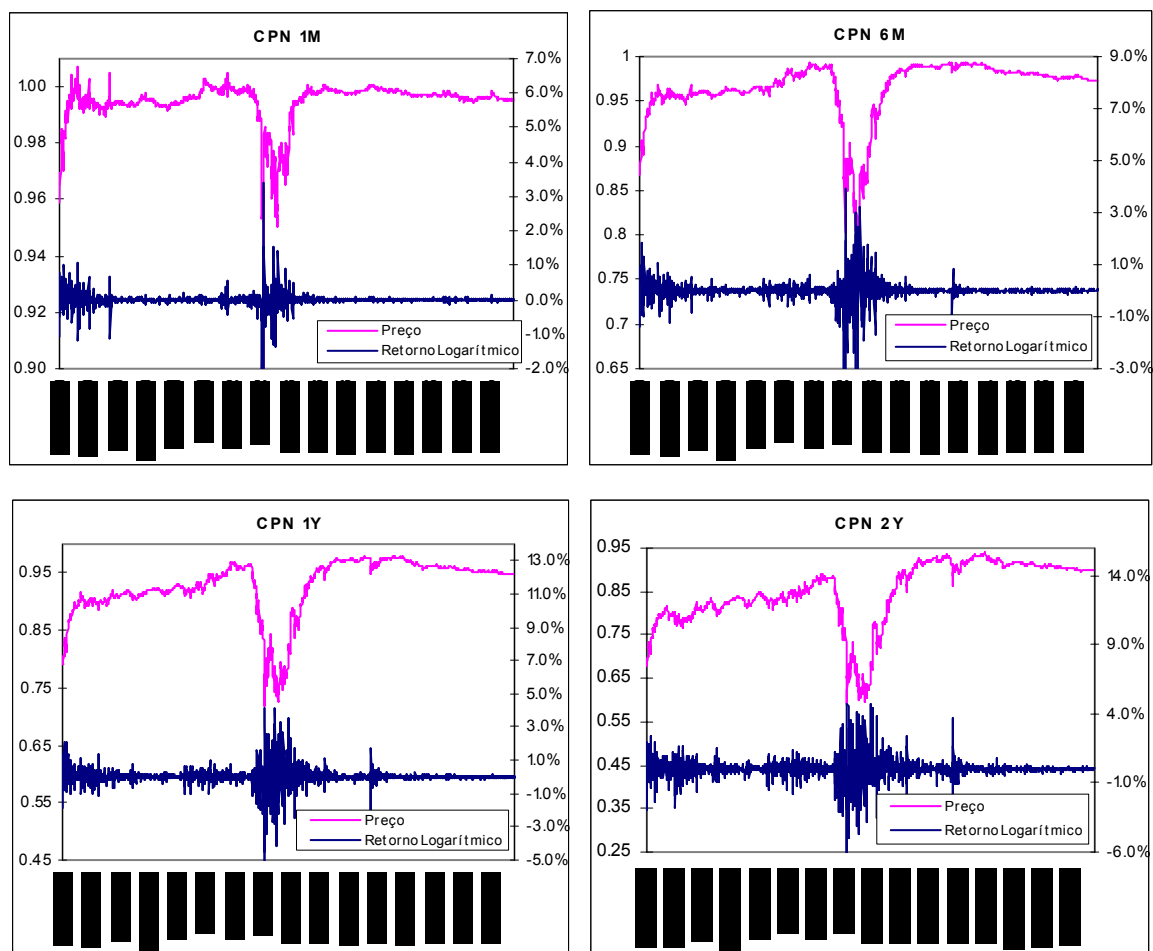


Figura 5 – Gráficos de preços e retornos dos PUs de “Cupom Cambial Limpo” de 1 mês (1M), 6 meses (6M), 1 ano (1Y) e 2 anos (2Y).

As figuras abaixo mostram as estatísticas descritivas e histogramas de cada uma das séries.

Os valores mostram que todas as séries apresentam níveis de curtose excessivamente altos se comparados com a distribuição normal, que tem curtose igual a 3.

O teste de Jarque Bera¹⁸ rejeitou a hipótese de normalidade para todas as séries analisadas.

¹⁸ Ver Anexos.

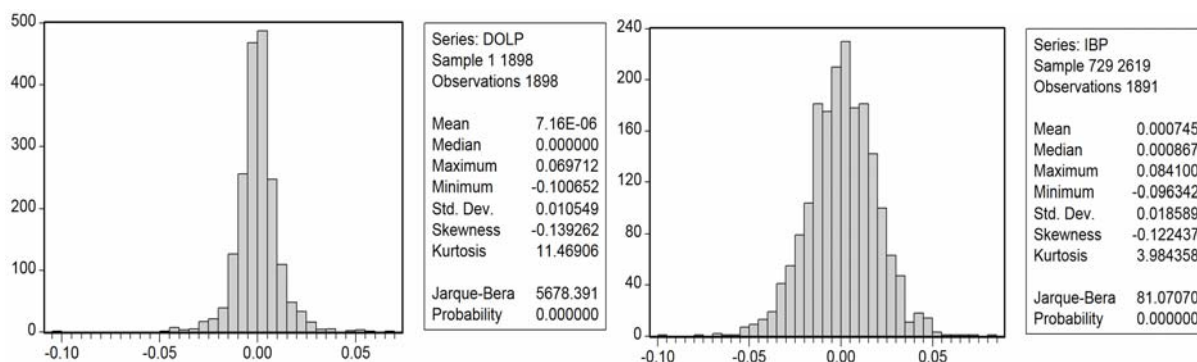


Figura 6 - Histogramas das séries de Taxa de Câmbio US\$/R\$ (DOLP) e Índice Bovespa (IBP).

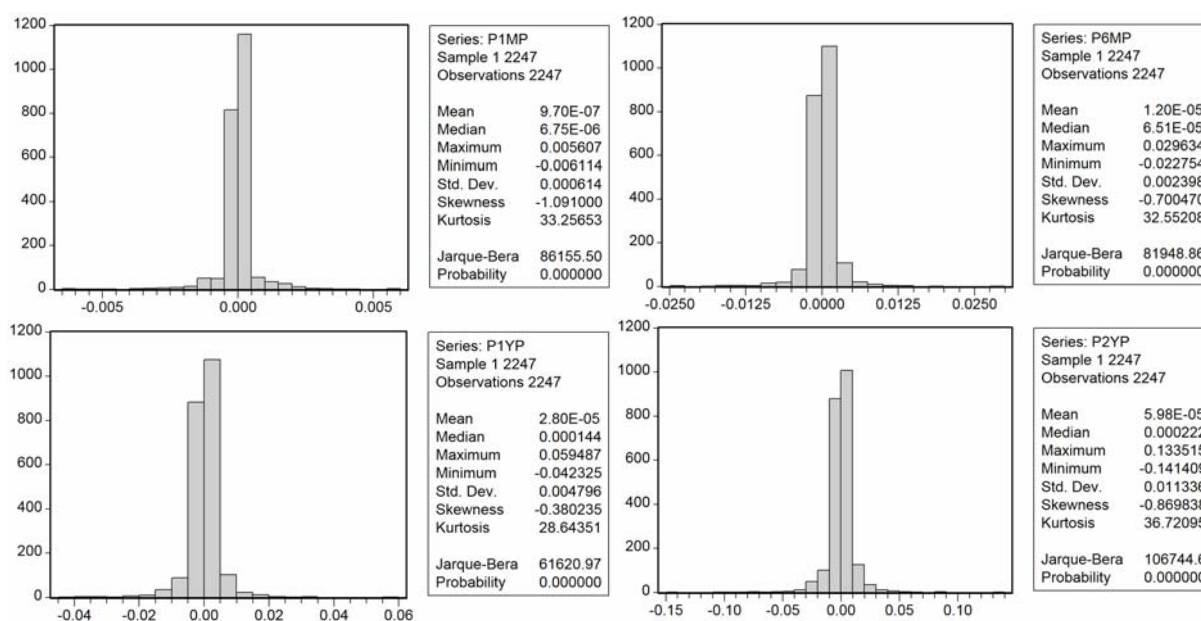
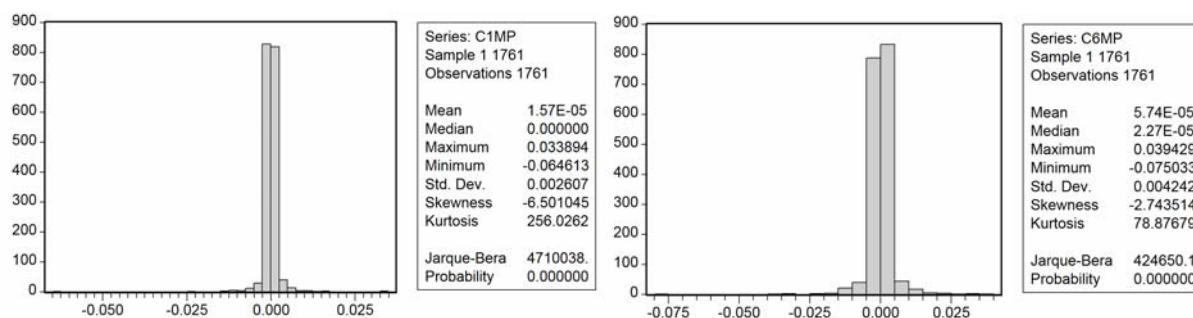


Figura 7 - Histogramas das séries de Juros "Pré" de 1M (P1MP), 6M (P6MP), 1Y (P1YP) e 2Y (P2YP).



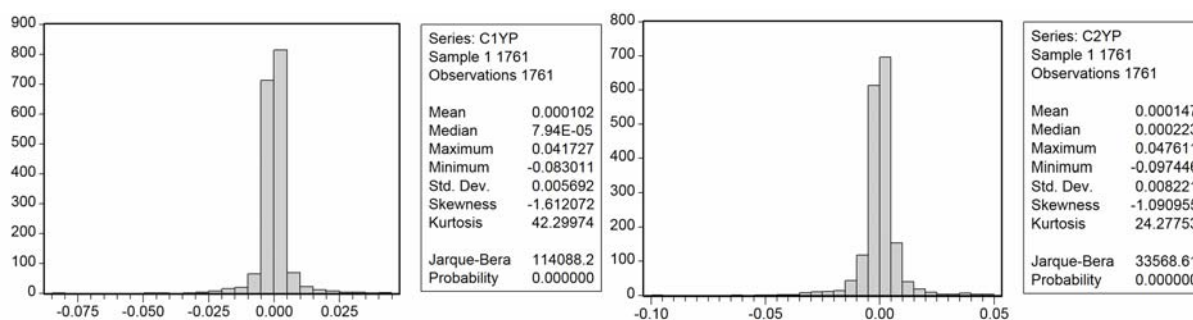


Figura 8 - Histogramas das séries de Juros de Cupom Cambial de 1M (C1MP), 6M (C6MP), 1Y (C1YP) e 2Y (C2YP).

4.2. DESCRIÇÃO DO PROCEDIMENTO

Utilizou-se uma janela móvel de dados históricos de 750 dias úteis, já que o BIS recomenda um mínimo de 250 dias para estimação do VaR e a modelagem de extremos requer uma quantidade maior de dados¹⁹. Por uma questão de processamento computacional, os parâmetros dos modelos foram re-estimados a cada 50 dias, mantendo-os constantes durante esse intervalo²⁰.

Primeiramente, realiza-se a análise estatística descritiva das diversas séries utilizadas no estudo, assim como o teste de Jarque-Bera para rejeitar (ou não) a normalidade dos retornos das séries. Isso é necessário, já que o modelo a ser proposto pressupõe a existência de caudas pesadas nos retornos dos ativos financeiros. O modelo será mais adequado quanto maior for a curtose da distribuição.

O próximo passo é verificar as dependências temporais das séries, através dos gráficos das auto-correlações estimadas para as observações. Os correlogramas, em geral, mostram dependência temporal para todas as séries.

Em seguida, ajusta-se um modelo auto-regressivo para capturar a dependência na média e um modelo GARCH (Engle [1982], Bollerslev [1986]) para capturar a dependência na variância através de estimações de quasi-máxima verossimilhança. Nesse estudo, optou-se por seguir à risca a proposição mais parcimoniosa de McNeil e Frey [1999], ou seja, um modelo AR(1) para a média e GARCH(1,1) para a dependência no segundo momento. No entanto, outros modelos mais complexos

¹⁹ A variância da estimação cai com o aumento do tamanho da amostra.

²⁰ Tal procedimento simplificado já foi também adotado por Souza [1999] e Pinto e Valls.

poderiam ser igualmente utilizados. Testou-se a adição de outros *lags*, mas os testes de Akaike e Schwarz não evidenciaram melhora substancial a ponto de se optar por um modelo mais complexo.

Calcula-se os resíduos padronizados do ajuste $AR(1) + GARCH(1,1)$ aos quais será ajustada uma GPD.

O próximo passo é gerar gráficos QQ-Plots desses resíduos versus a distribuição normal para confirmar a suposição de que a normalidade condicional não é realística e que a distribuição das inovações do processo apresenta caudas pesadas (leptocúrticas).

A partir daí, a questão que surge é a partir de que ponto u da distribuição dos resíduos, a região já pode ser considerada como cauda, ou seja, pode-se utilizar as hipóteses assintóticas da teoria de valores extremos.

A escolha de u alto reduz o viés, enquanto que k (número de eventos acima do limiar) alto reduz a variância das estimações dos parâmetros. À medida que se aumenta o valor de k , avança-se ao centro da distribuição, saindo do verdadeiro foco de atenção representado pelas caudas e aumenta-se o viés. Por outro lado, quando se diminui o número k , poucas observações são incorporadas à estimação e corre-se o risco de produzir um estimador com elevada variância.

Muitos estudos utilizam o 95% percentil da distribuição como ponto de partida, mas não existem, até o momento, estudos conclusivos sobre a melhor opção de escolha.

É importante mencionar que Danielsson e de Vries [1997] operacionalizaram um método de definição do ponto de corte baseado num procedimento de *Bootstrap* desenvolvido por Hall [1990]. Tal método já foi implementado no Brasil por Pinto e Valls [2006].

No presente trabalho optou-se por outro método sugerido por Embrechts et alii [1997] que consiste em traçar o gráfico conhecido como Hill-plot e procurar a região mais próxima das caudas onde o valor do índice de caudas estimado estabilize-se. Segundo Drees, de Haan e Resnick, tal método além de ser um dos mais populares e de fácil aplicação prática, deve ser adotado quando os dados aproximarem-se de uma distribuição de Pareto. Mesmo método também foi escolhido por Souza [1999].

Utilizou-se um número fixo de dados para representar a cauda igual a $N = k$ sendo $k > n(1 - q)$, de forma que o quantil adotado seja sempre maior que o limiar. O que na prática representa um limiar variável para a estatística de $(k+1)$ ordem. Analisando-se os gráficos de Hill-Plot, observa-se que para todas as séries, a região de estabilização do índice de caudas inicia-se por volta de 100, mas como $k = 100$ representa 13% das observações, optou-se pela parcimônia e adotou-se um limiar correspondente a $k = 50$ ($n = 749$).

Os resíduos são então ajustados a uma Distribuição Generalizada de Pareto para as caudas (valores acima do limiar) e uma distribuição empírica para a parte central dos dados. Duas técnicas foram utilizadas para estimação: Máxima Verossimilhança e *Probability Weighted Moments*. Apesar de nas últimas décadas ter sido amplamente utilizada, principalmente por sua simplicidade, a técnica de máxima verossimilhança é limitada para a estimação de parâmetros de distribuições de valores extremos por gerar estimadores viesados na média, particularmente para amostras pequenas. No entanto, Koch [1991] mostrou que tal viés não influencia significativamente a magnitude dos parâmetros e que a técnica de máxima verossimilhança é um método eficiente de estimação dos parâmetros das distribuições de extremos. No presente trabalho não se verificou diferença substancial nas estimações produzidas pelas duas técnicas, não se alterando o resultado final.

Após a estimação, calcula-se o Valor em Risco a partir da volatilidade estimada pelo modelo GARCH e o quantil associado estimado.

A maneira mais usual para se verificar a eficácia das estimativas de VaR é através de procedimento de *back-testing*, que consiste em comparar o numero de vezes que o retorno previsto pelo VaR é excedido pelo retorno efetivo (*outliers*) com o número esperado para determinado α .

4.3. MODELAGEM

Assume-se que os retornos diários (X_1, X_2, \dots, X_N) representam realizações de um processo estocástico estritamente estacionário, onde a média e a variância variem no tempo conforme:

$$X_t = \mu_t + \sigma_t Z_t ,$$

E que as inovações Z_t sejam variáveis aleatórias iid com distribuição F , média zero e variância unitária. Seja FX a distribuição de X_t e FX_{t+1} a distribuição preditiva um passo a frente.

Conforme a proposição de McNeil e Frey [1999], modela-se a média e a volatilidade condicional com os modelos AR(1) e GARCH(1,1), respectivamente.

Formalmente:

$$X_t = \phi X_{t-1} + \varepsilon_t$$

Onde as inovações ε_t seguem um modelo GARCH(1,1). A volatilidade ao quadrado é dada por:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Os estimadores convencionais de máxima verossimilhança não conseguem lidar com observações atípicas. No entanto, existem na literatura algumas propostas para estimação robusta como os M-estimadores (Huber[1981]) e CM-estimadores (Mendes e Duarte [2004] e Kent e Tyler [1991]). A intuição por trás desses estimadores é que não assumem nenhuma distribuição específica para as inovações.

O estimador utilizado nesse trabalho foi o de pseudo (ou *quasi*) máxima verossimilhança, considerado robusto já que produz estimações consistentes dos parâmetros mesmo que a distribuição seja incorretamente especificada. Mais detalhes sobre esse estimador podem ser obtidos em Gourioux, Monfort, e Trognon [1984].

Após a estimação do modelo, calculam-se os resíduos padronizados:

$$(z_{t-n+1}, \dots, z_t) = \left(\frac{X_{t-n+1} - \hat{\mu}_{t-n+1}}{\hat{\sigma}_{t-n+1}}, \dots, \frac{X_t - \hat{\mu}_t}{\hat{\sigma}_t} \right),$$

A seguir, assume-se que os resíduos excedentes seguem uma distribuição de Pareto (GPD) com função igual a:

$$G_{\xi, \beta}(y) = \begin{cases} 1 - (1 + \xi y / \beta)^{-1/\xi} & \text{se } \xi \neq 0 \\ 1 - e^{-y/\beta} & \text{se } \xi = 0 \end{cases},$$

onde $\beta > 0$, e $y \geq 0$ quando $\xi \geq 0$ e $0 \geq y \geq -\beta/\xi$ quando $\xi < 0$.

A escolha dessa distribuição é motivada por um resultado limite em EVT. Considerando uma fdp F e a correspondente distribuição de excessos acima de um limiar u dada por:

$$F_u(y) = P\{X - u \leq y \mid X > u\} = \frac{F(y+u) - F(u)}{1 - F(u)}$$

Para $0 \leq y < x_0 - u$, onde x_0 é o ponto final à direita de F .

A partir dos resíduos ajusta-se a GPD por máxima verossimilhança utilizando o estimador:

$$F_z(u+y) = \frac{k}{n} \left(1 + \xi \frac{y}{\beta} \right)^{-1/\xi}$$

Usando-se a inversa F_z^{-1} obtém-se um estimador do quantil:

$$z_\alpha = u + \frac{\beta}{\xi} \left(\left(\frac{\frac{\alpha}{k}}{\frac{n}{n}} \right)^{-\xi} - 1 \right)$$

Onde z_α é o $(1-\alpha)$ quantil da distribuição Z_t .

Com base no quantil z_α calcula-se a medida de risco VaR:

$$VaR_\alpha = \mu_{t+1} + \sigma_{t+1} z_\alpha$$

No entanto, o VaR acima calculado prevê a perda esperada para 1 dia, mas a exigência de capital deve ser calculada para um horizonte de 10 dias, segundo recomendação do BIS.

A seguir demonstra-se a modelagem para obtenção da estimação do VaR para multi-períodos, a partir de uma série de retornos diários.

Ajusta-se a série de retornos conforme o procedimento descrito anteriormente aos modelos AR(1) e GARCH(1,1) e obtém-se os resíduos. A partir dos resíduos, ajustam-se as duas caudas (inferior e superior) à função GPD.

A partir daí, simula-se o distúrbio dessa distribuição por uma combinação de *bootstrap* e simulação da GPD de acordo com o algoritmo abaixo proposto por Danielsson e de Vries [1997].

- i) Seleciona-se randomicamente um resíduo da série de resíduos.
- ii) Se o resíduo exceder o limiar máximo $z_{(k+1)}$, retorna-se $z_{(k+1)} + y1$, onde $y1$ é obtido a partir da distribuição do excesso superior $GPD(\xi_k^{(1)}, \beta_k^{(1)})$.
- iii) Se o resíduo for menor que o limiar mínimo $z_{(n-k)}$, retorna-se $z_{(n-k)} - y2$, onde $y2$ é obtido a partir da distribuição do excesso inferior $GPD(\xi_k^{(2)}, \beta_k^{(2)})$.
- iv) Caso contrário retorna-se o próprio resíduo.
- v) Repõe-se o resíduo na série e repete-se o procedimento a partir de i.

O algoritmo acima gera pontos da distribuição:

$$\hat{F}_z(z) = \begin{cases} \frac{k}{n} \left(1 + \xi_k^{(2)} \frac{|z - z_{(n-k)}|}{\beta_k^{(2)}} \right)^{-1/\xi_k^{(2)}} & \text{se } z < z_{(n-k)} \\ \frac{1}{n} \sum_{i=1}^n 1_{\{z_i \leq z\}} & \text{se } z_{(n-k)} \leq z \leq z_{(k+1)} \\ 1 - \frac{k}{n} \left(1 + \xi_k^{(1)} \frac{z - z_{(k+1)}}{\beta_k^{(1)}} \right)^{-1/\xi_k^{(1)}} & \text{se } z > z_{(k+1)} \end{cases},$$

Que aproxima $F_z(z)$ para n grande.

Usando a distribuição acima que representa o distúrbio e o modelo GARCH estimado, simula-se 10,000 trajetórias de retornos para o período desejado $(x_{t+1}, x_{t+2}, \dots, x_{t+k})$ e a soma cumulativa correspondente dessas realizações.

O objetivo final é definir por qual fator deve-se multiplicar uma estimação condicional do retorno de um dia para se obter a estimação condicional para o retorno de k dias. Ou seja, necessita-se calcular a razão $x_q^t(k)/x_q^t$ para $k > 1$.

Para o caso de X_i serem iid, já estão disponíveis resultados teóricos sobre o fator de escala que transforma estimação de retorno de um dia em estimação de retorno de k dias.

Para o caso de distribuições estritamente estáveis, onde $X_1 + X_2 + \dots + X_k \stackrel{d}{=} k^{\frac{1}{\alpha}} X_1$ para $\alpha \in (0, 2]$, tem-se que $x_q(k)/x_q = k^{\frac{1}{\alpha}}$. O caso especial em que $\alpha = 2$, corresponde à famosa regra da “raiz quadrada do tempo” implementada em RiskMetrics.

Para os casos de distribuições com caudas pesadas, Feller [1971] demonstrou que os riscos são, para uma aproximação de primeira ordem, linearmente aditivos. Ou em termos probabilísticos, a aproximação para o risco do retorno de um único período é dada por $P[|X| > x] \approx ax^{-\xi}$, quando x é grande enquanto que a aproximação para o risco do retorno de k períodos é definido como $P[X_1 + X_2 + \dots + X_k > x] \approx Kax^{-\xi}$.

Isso implica no seguinte fator aproximado para grandes percentis: $x_q(k)/x_q \approx k^{\xi}$.

Partindo desses resultados teóricos para séries não condicionais, McNeil e Frey [1999] sugeriram que, para o modelo proposto, o fator $x_q^t(k)/x_q^t$ também poderia ser aproximado por uma função potência: $x_q(k)/x_q \approx k^{\lambda_t}$, para um k pequeno. No entanto, espera-se que λ_t dependa do nível da volatilidade inicial σ_t em relação à

média de longo prazo da volatilidade. Ou seja, para modelos estacionários com estrutura de volatilidade estocástica (no caso GARCH com $\alpha_1 + \beta < 1$), espera-se que quando σ_t for alta (baixa) a volatilidade futura será, em média mais baixa (alta) que σ_t .

5. CONCLUSÃO

5.1. RESULTADOS DOS ESTUDOS DE CASOS

A análise gráfica dos *QQPlots*²¹ dos resíduos padronizados demonstra claramente que as distribuições são leptocúrticas, uma vez que as curvas inclinam para cima do lado esquerdo do gráfico e para baixo do lado direito. A série de resíduos que mais se aproximou de uma Normal foi a de Ibovespa (Ver Figura 5 – RESID_IBP).

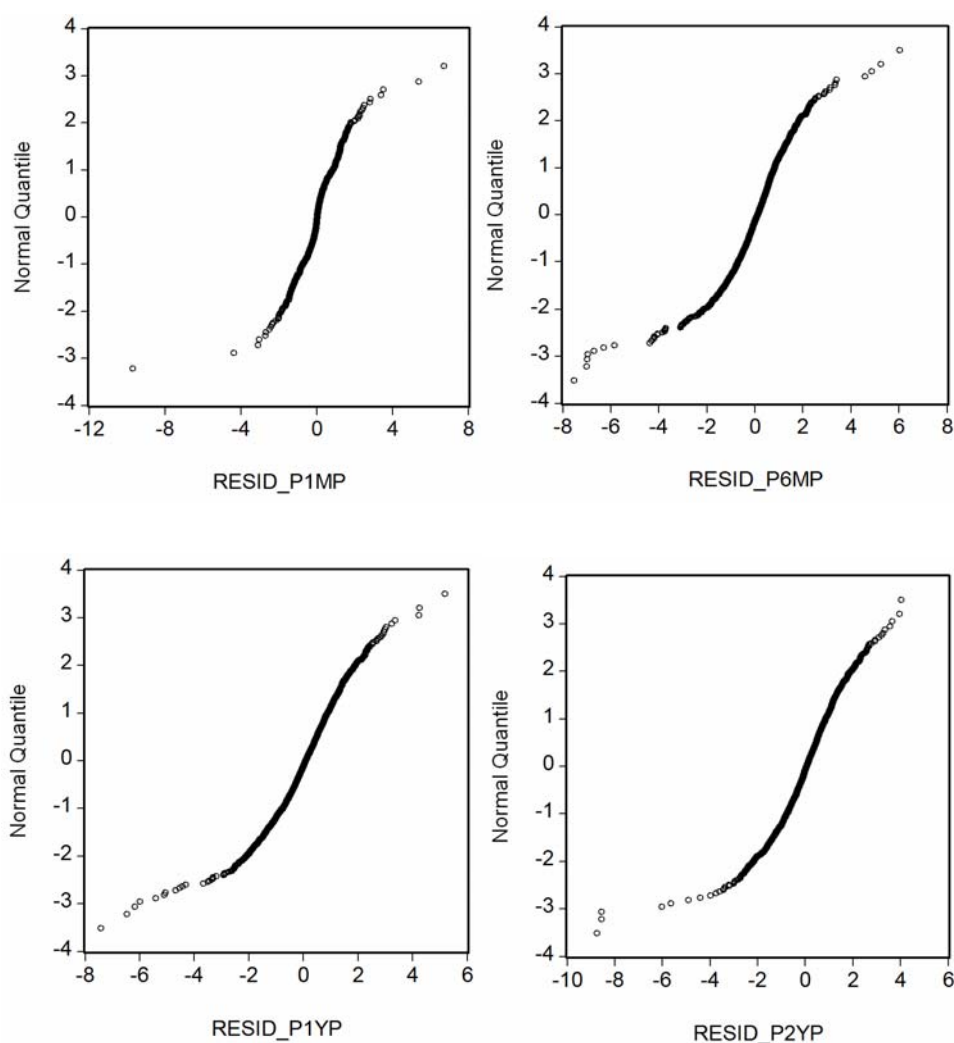


Figura 9 - *QQ Plots* dos Resíduos Padronizados do modelo AR(1) + GARCH(1,1) das séries de Taxas Pré-Fixadas de 1 mês (RESID_P1MP), 6 meses (RESID_P6MP), 1 ano (RESID_P1YP) e 2 anos (RESID_P2YP).

²¹ Ver explicação sobre Gráfico Quantil-Quantil (*QQPlot*) nos Anexos.

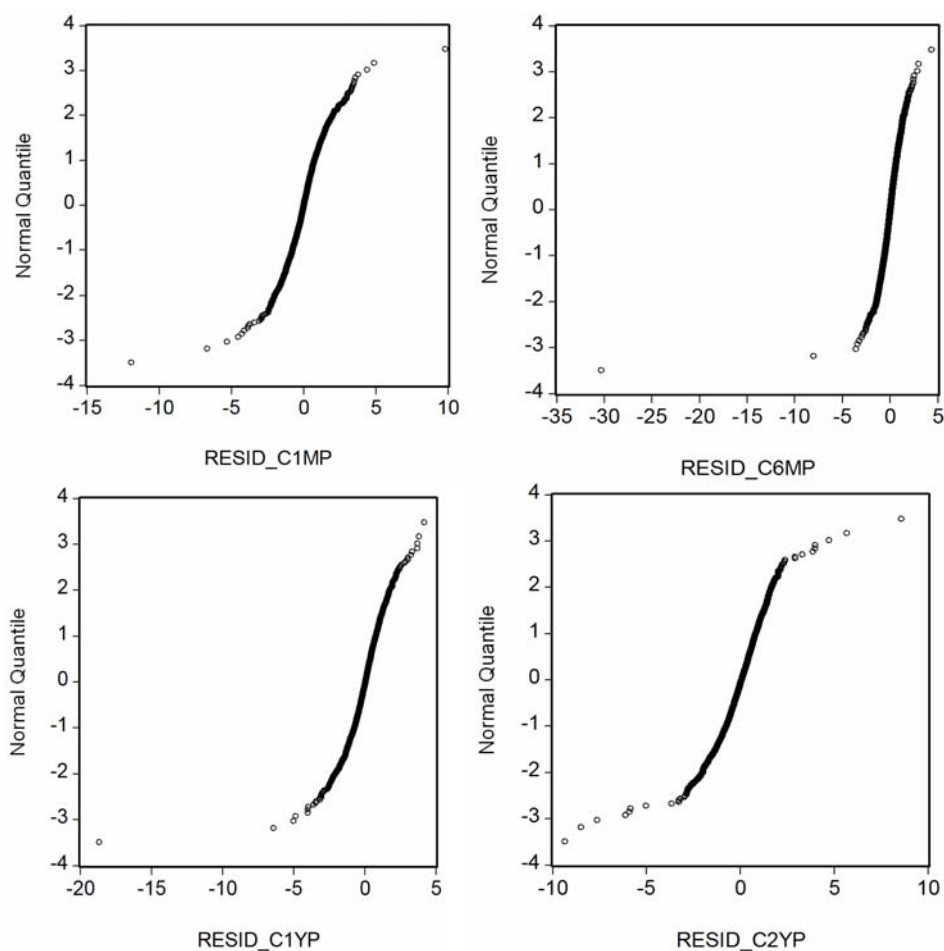


Figura 10 - QQ Plots dos Resíduos Padronizados do modelo AR(1) + GARCH(1,1) das séries de Cupom Cambial de 1 mês (RESID_C1MP), 6 meses (RESID_C6MP), 1 ano (RESID_C1YP) e 2 anos (RESID_C2YP).

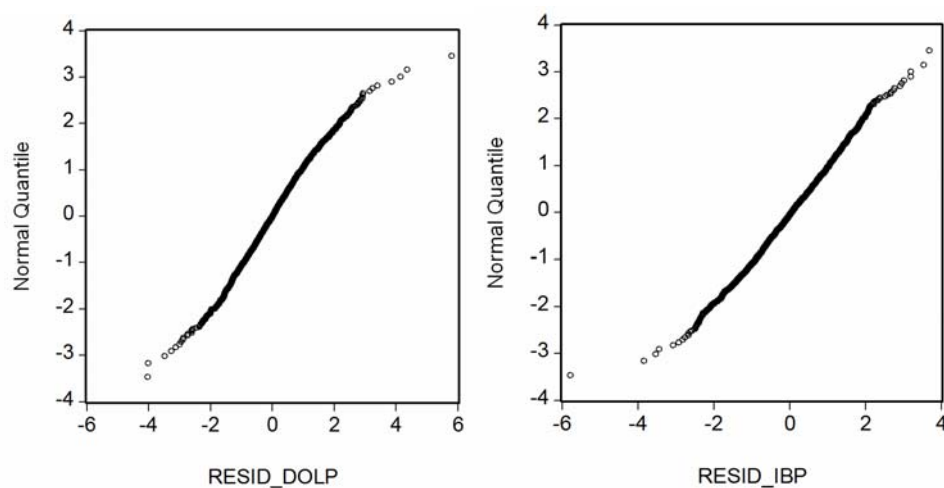


Figura 11 - QQ Plots dos Resíduos Padronizados dos modelos AR(1) + GARCH(1,1) das séries de Dólar (RESID_DOLP) e Ibovespa (RESID_IBP).

De acordo com a verificação gráfica, em todas as séries o estimador de Hill parece estabilizar-se para valores de k inferiores a 200. O limiar foi escolhido para todos os casos de forma que seja excedido por 50 observações ($k = 50$).

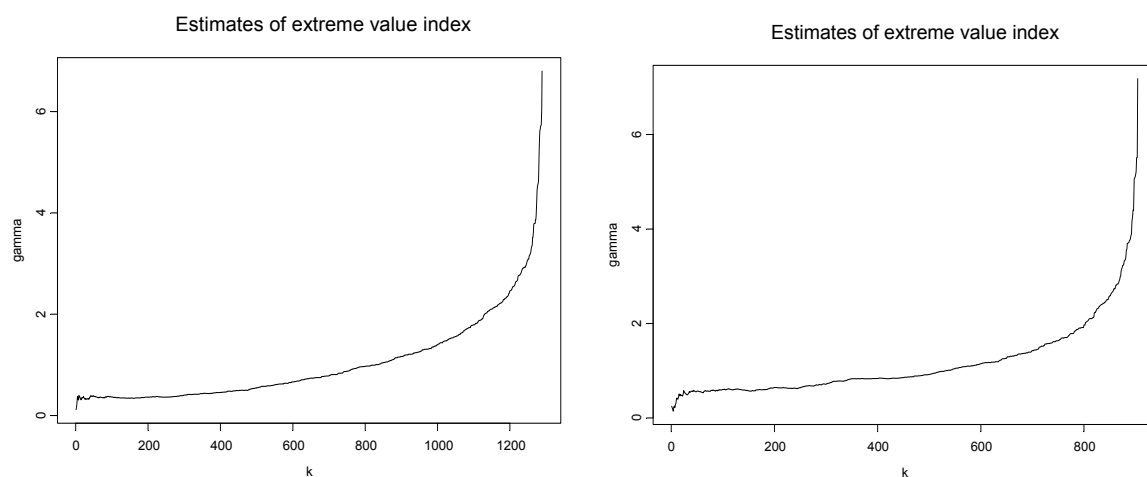


Figura 12 - Hill Plot P1M – Max (esq.) e P1M – Min (dir.)

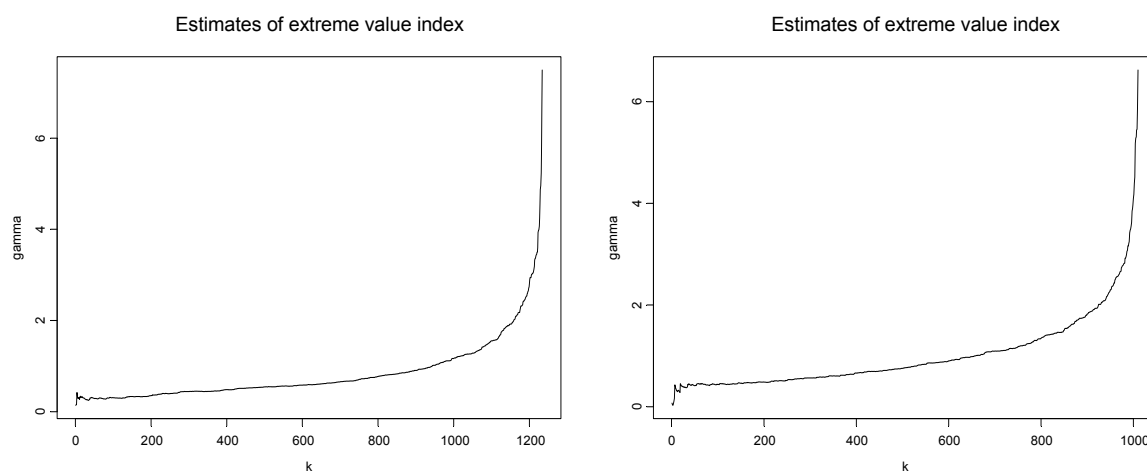


Figura 13 – Hill Plot P6M – Max (esq.) e P6M – Min (dir.)

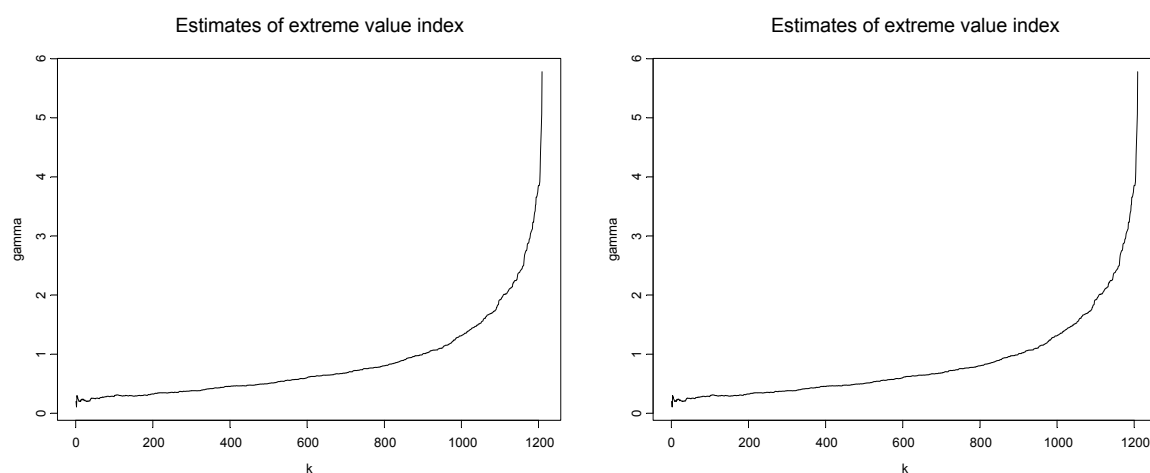


Figura 14 – Hill Plot P1Y – Max (esq.) e P1Y – Min (dir.)

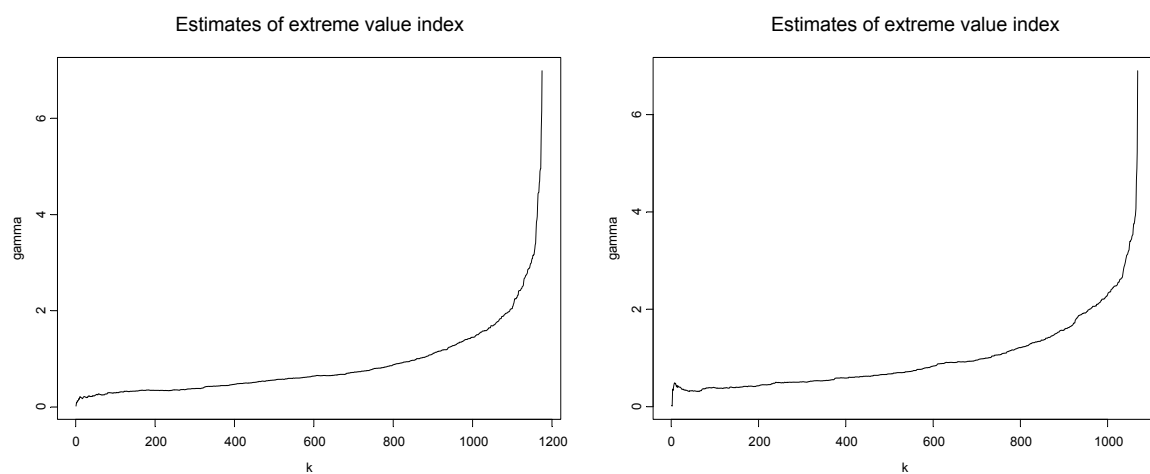


Figura 15 – Hill Plot P2Y – Max (esq.) e P2Y – Min (dir.)

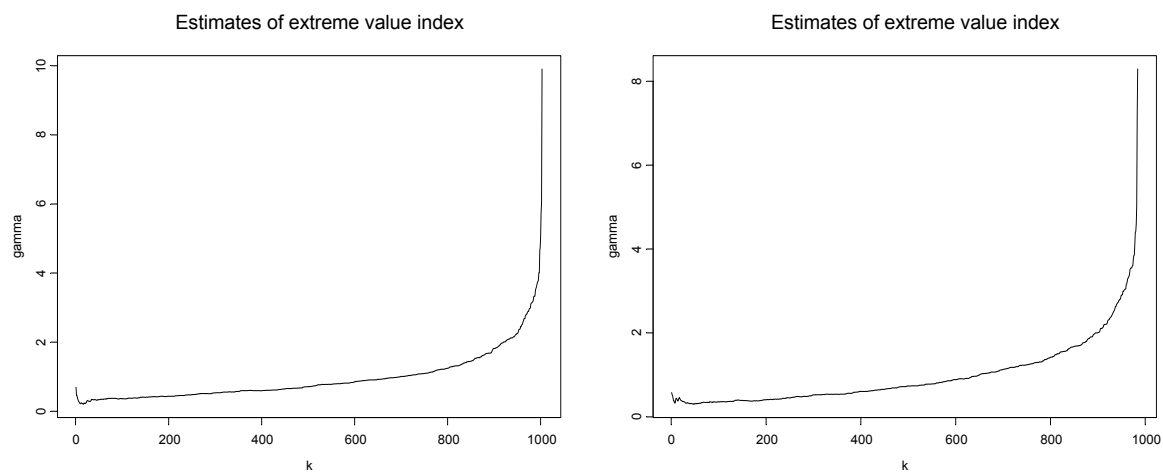


Figura 16 – Hill Plot C1M – Max (esq.) e C1M – Min (dir.)

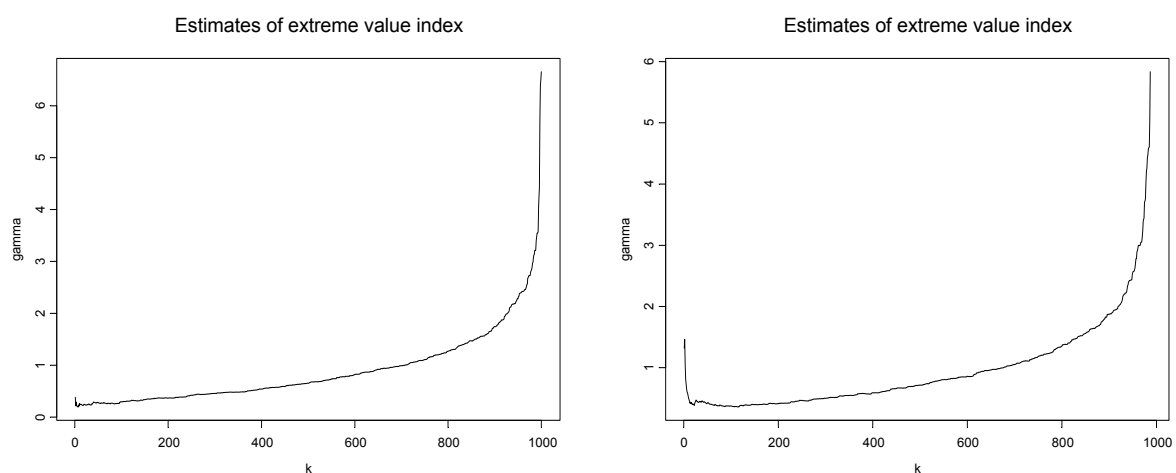


Figura 17 – Hill Plot C6M – Max (esq.) e C6M – Min (dir.)

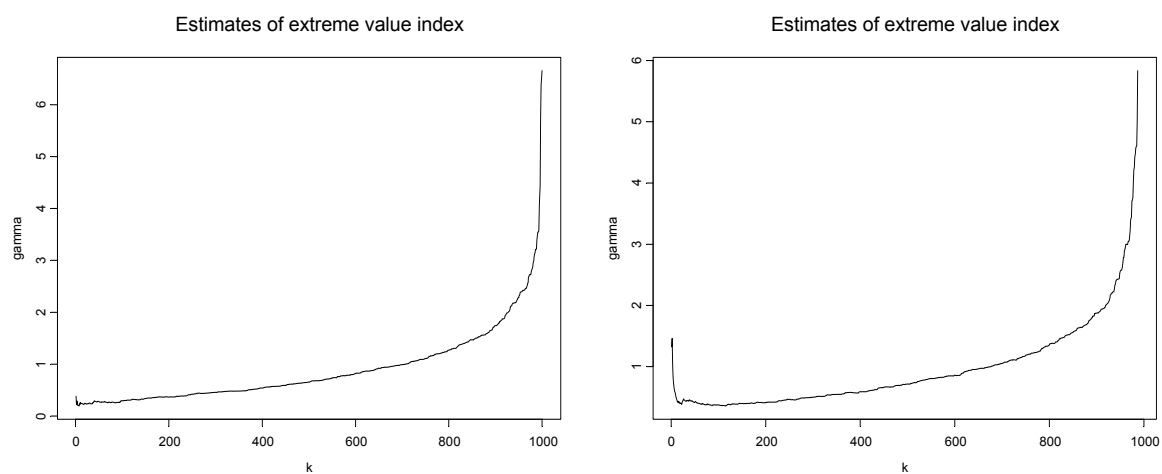


Figura 18 – Hill Plot C1Y – Max (esq.) e C1Y – Min (dir.)

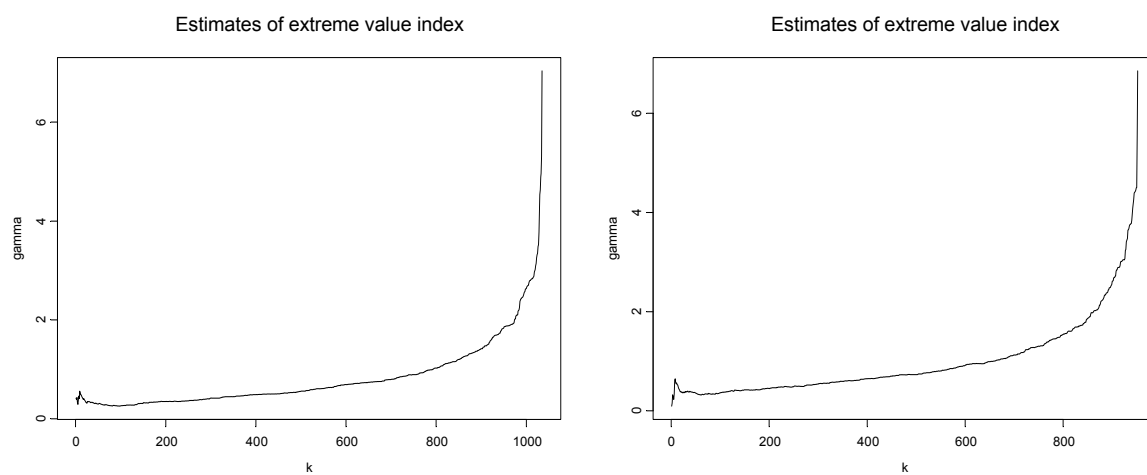


Figura 19 – Hill Plot C2Y – Max (esq.) e C2Y – Min (dir.)

O modelo proposto, de forma geral, mostrou resultados superiores ao modelo GARCH+Normal e Normal Não-Condiciona para as séries analisadas.

Para as séries de taxas de juros, os resultados de *backtestings* evidenciaram uma clara assimetria das caudas. Para a cauda inferior (mínimos), o modelo proposto GARCH+GPD produz menor número de *outliers* e mais próximos ao esperado que o modelo GARCH+Normal para todas as séries.

No entanto, para a cauda superior, os resultados não foram tão satisfatórios. Para as séries de taxas pré-fixadas os resultados foram muito similares para ambos os modelos, apresentando *outliers* abaixo do esperado (superestimando o risco). No entanto, os resultados para as séries de cupom cambial foram distintos para curto e médio prazo. Os números de *outliers* para as séries de 1M e 6M foram menores para o modelo GARCH+GPD, mas para as séries de 1Y e 2Y foram menores para o modelo GARCH+Normal.

Constatou-se, portanto, que movimentos adversos extremos com maior probabilidade de ocorrência que a prevista pela distribuição gaussiana acontecem principalmente na cauda inferior, ou seja, representando retornos negativos. Como a modelagem de taxas de juros baseia-se em PUs (preços unitários), conclui-se que tais movimentos adversos estão relacionados a momentos de crises ou reações negativas no mercado financeiro que levaram à alta das taxas de juros.

Analogamente, os resultados para a série de Dólar mostram que o modelo GARCH+GPD mostrou-se muito superior aos demais para a cauda direita, que representa movimentos de desvalorização cambial do Real.

O modelo GARCH+GPD apresentou piores resultados para a série de Ibovespa, possivelmente pelo fato de seus resíduos aproximarem-se da Normal (ver gráfico QQ-Plot), além de a própria série original apresentar o menor índice de curtose entre todas analisadas.

Backtesting

Máximos, $\alpha = 99\%$

	Nº Dias	Outliers Esperados		Outliers Verificados					
				GARCH + GPD		GARCH + Normal		Normal Não Condicional	
		Nº	%	Nº	%	Nº	%	Nº	%
Dol Comercial	1901	19.0	1.00%	16	0.84%	27	1.42%	49	2.58%
Ibovespa	1901	19.0	1.00%	11	0.58%	14	0.74%	14	0.74%
PRE 1MÊS	1901	19.0	1.00%	14	0.74%	15	0.79%	5	0.26%
PRE 6MESES	1901	19.0	1.00%	14	0.74%	16	0.84%	10	0.53%
PRE 1ANO	1901	19.0	1.00%	15	0.79%	14	0.74%	15	0.79%
PRE 2ANOS	1901	19.0	1.00%	14	0.74%	13	0.68%	14	0.74%
CPN 1MÊS	1761	17.6	1.00%	14	0.80%	24	1.36%	9	0.00%
CPN 6MESES	1761	17.6	1.00%	9	0.51%	11	0.62%	22	0.00%
CPN 1ANO	1761	17.6	1.00%	18	1.02%	12	0.68%	27	0.00%
CPN 2ANOS	1761	17.6	1.00%	14	0.80%	8	0.45%	25	0.00%

Teste de Razão de Verossimilhança (LR) - Máximos

	GARCH + GPD		GARCH + Normal		Normal Não Condicional	
	Teste LR	P-value	Teste LR	P-value	Teste LR	P-value
Dol Comercial	0.51	47.6%	3.00	8.3%	33.29	0.0%
Ibovespa	4.02	4.5%	1.47	22.6%	1.47	22.6%
PRE 1MÊS	1.47	22.6%	0.92	33.7%	14.77	0.0%
PRE 6MESES	1.47	22.6%	0.51	47.6%	5.22	2.2%
PRE 1ANO	0.92	33.7%	1.47	22.6%	0.92	33.7%
PRE 2ANOS	1.47	22.6%	2.16	14.2%	1.47	22.6%
CPN 1MÊS	0.80	37.0%	2.10	14.7%	5.18	2.3%
CPN 6MESES	5.18	2.3%	2.89	8.9%	1.02	31.1%
CPN 1ANO	0.01	92.6%	2.03	15.4%	4.35	3.7%
CPN 2ANOS	0.80	37.0%	6.65	1.0%	2.77	9.6%

Backtesting

Mínimos, alpha = 99%

	Nº Dias	Outliers Esperados		Outliers Verificados					
				GARCH + GPD		GARCH + Normal		Normal Não Condicional	
		Nº	%	Nº	%	Nº	%	Nº	%
Dol Comercial	1901	19.0	1.00%	18	0.95%	20	1.05%	34	1.79%
Ibovespa	1901	19.0	1.00%	13	0.68%	19	1.00%	14	0.74%
PRE 1MÊS	1901	19.0	1.00%	18	0.95%	31	1.63%	5	0.26%
PRE 6MESES	1901	19.0	1.00%	18	0.95%	27	1.42%	15	0.79%
PRE 1ANO	1901	19.0	1.00%	16	0.84%	23	1.21%	21	1.10%
PRE 2ANOS	1901	19.0	1.00%	15	0.79%	28	1.47%	23	1.21%
CPN 1MÊS	1761	17.6	1.00%	15	0.85%	23	1.31%	10	0.00%
CPN 6MESES	1761	17.6	1.00%	12	0.68%	23	1.31%	30	0.00%
CPN 1ANO	1761	17.6	1.00%	17	0.97%	23	1.31%	40	0.00%
CPN 2ANOS	1761	17.6	1.00%	12	0.68%	25	1.42%	38	0.00%

Teste de Razão de Verossimilhança (LR) - Mínimos

	GARCH + GPD		GARCH + Normal		Normal Não Condicional	
	Teste LR	P-value	Teste LR	P-value	Teste LR	P-value
Dol Comercial	0.06	81.4%	0.05	82.1%	9.67	0.2%
Ibovespa	2.16	14.2%	0.00	99.8%	1.47	22.6%
PRE 1MÊS	0.06	81.4%	6.42	1.1%	14.77	0.0%
PRE 6MESES	0.06	81.4%	3.00	8.3%	0.92	33.7%
PRE 1ANO	0.51	47.6%	0.79	37.3%	0.20	65.2%
PRE 2ANOS	0.92	33.7%	3.75	5.3%	0.79	37.3%
CPN 1MÊS	0.41	52.1%	1.52	21.8%	3.94	4.7%
CPN 6MESES	2.03	15.4%	1.52	21.8%	7.27	0.7%
CPN 1ANO	0.02	88.3%	1.52	21.8%	21.14	0.0%
CPN 2ANOS	2.03	15.4%	2.77	9.6%	17.91	0.0%

Para cálculo da exigência de capital, necessita-se obter a estimação do VaR para 10 dias. Utilizando-se o procedimento combinado de bootstrap e simulação de Monte Carlo descrito no capítulo anterior, chega-se às seguintes estimações para o parâmetro λ^{10} para cada uma das séries:

	Baixa		Media		Alta	
	Max	Min	Max	Min	Max	Min
p1m	0.56	0.33	0.52	0.28	0.51	0.30
p6m	0.65	0.47	0.60	0.43	0.56	0.40
p1y	0.82	0.62	0.74	0.52	0.74	0.48
p2y	0.63	0.59	0.58	0.54	0.58	0.50
c1m	0.66	0.64	0.56	0.58	0.51	0.51
c6m	0.52	0.35	0.47	0.31	0.39	0.26
c1y	0.81	0.79	0.70	0.67	0.72	0.64
c2y	0.78	0.65	0.68	0.58	0.69	0.60
dol	0.55	0.72	0.51	0.69	0.49	0.66
ibov	0.54	0.62	0.53	0.60	0.52	0.59

Os resultados encontrados (assim como em McNeil e Frey [1999]), não confirmam a conclusão de Danielsson, Hartman e de Vries [1998] de que o fator de escala para retornos que exibem caudas pesadas deveria ser menor que o fator de “raiz quadrada do tempo” utilizado para distribuições normais, provavelmente pelo fato deles não considerarem a dependência temporal em seu modelo.

Os resultados encontrados, mesmo para períodos de alta volatilidade são maiores que os utilizados para distribuições normais. Por exemplo, para o caso Pré 2 anos (Volatilidade Alta), cujo fator λ^{10} de mínimos é igual a 0,60 :

$$\text{VaR (Normal não condicional)}_{10 \text{ dias}} = \text{VaR}_{1 \text{ dia}} * 10^{0,5} = 3,1623 * \text{VaR}_{1 \text{ dia}}$$

$$\text{VaR (GARCH + EVT)}_{10 \text{ dias}} = \text{VaR}_{1 \text{ dia}} * 10^{0,6} = 3,9811 * \text{VaR}_{1 \text{ dia}}$$

5.2. SUGESTÃO PARA TRABALHOS FUTUROS: CASO MULTIVARIADO

O presente trabalho contribui para a aplicação da Teoria de Valores Extremos à mensuração de riscos no Brasil e não esgota totalmente os seus possíveis desdobramentos em estudos complementares ou mais aprofundados.

Uma possível sugestão seria a utilização de outros modelos da família GARCH que mais se adequem ao comportamento das séries brasileiras.

Outra sugestão seria a extensão para o caso multivariado. Em finanças, necessita-se de modelos de risco multi-dimensionais, uma vez que os portfólios de investimentos possuem centenas ou milhares de instrumentos, que por sua vez se traduzem em centenas ou milhares de fatores de riscos.

O modelo univariado proposto pode ser utilizado para uma carteira de ativos, assim como sugerido por Danielsson e de Vries [1997], bastando apenas aplicar o peso da carteira atual aos retornos históricos e gerando uma série única de retornos. Em seguida modela-se tal série como se fosse um fator de risco único. O grande problema dessa abordagem é que não considera a estrutura de dependência entre os fatores de risco, assumindo que a matriz de correlação é constante ao longo de todo o período.

Apesar de a maior parte dos estudos relacionados à Teoria de Valores Extremos basear-se em modelos univariados, já existem diversas publicações a respeito da teoria multivariada de extremos. No entanto, até o presente momento, não está disponível uma completa teoria multivariada aplicada a altas dimensões. As aplicações das teorias multivariadas de extremos concentram-se em geral em duas ou no máximo três dimensões, como por exemplo: modelos trivariados aplicados a estudos dos níveis da água do mar por Tawn [1990], modelos de máximos bivariados aplicados a retornos financeiros por Mendes [2004].

A modelagem da estrutura de dependência de dados multivariados nos extremos pode ser feita com bastante eficiência utilizando-se o conceito de cópulas.

6. ANEXOS

Teste Jarque-Bera de normalidade

O teste estatístico é feito através da seguinte expressão:

$$JB = \frac{n}{6} \hat{\gamma}^2 + \frac{n}{24} (\hat{\kappa} - 3)^2$$

Onde:

$$\hat{\gamma} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^3, \text{ representa skewness}$$

$$\hat{\kappa} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\hat{\sigma}} \right)^4, \text{ representa kurtosis}$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \text{ representa a variância da amostra.}$$

Foi provado que na hipótese de x_i representar observações independentes de uma distribuição normal, para n grande, a estatística do teste Jarque-Bera tende assintoticamente a uma distribuição de Chi quadrado. Isso possibilita testar a normalidade da amostra. Para uma dada amostra $x = (x_1, \dots, x_n)$ calcula-se a estatística JB e compara-se com o 95% quantil da distribuição Chi-quadrado que é igual a 5.99. Portanto, na hipótese nula de que os dados foram gerados por uma distribuição normal, sabe-se que em 95% dos casos valor da estatística JB será menor que 5.99. Consequentemente rejeita-se a hipótese de normalidade se o teste JB exceder 5.99.

Gráfico Quantil-Quantil (QQ-plot)

O QQ-plot serve para ilustrar graficamente a distância entre as funções de distribuição empíricas e teóricas. Mas pode também ser usado para comparar duas distribuições empíricas.

O gráfico é construído, comparando-se os quantis das duas séries.

Se as duas séries forem oriundas de populações com uma mesma distribuição, os pontos do gráfico estarão posicionados ao longo de uma linha de referência de 45°.

Quanto mais afastados dessa linha de referência estiverem os pontos, maior a chance das duas populações terem distribuições diferentes.

Distribuições Estáveis e Max-estáveis

Diz-se que uma distribuição é estável se uma combinação linear de suas variáveis (que devem ser randômicas iid) gerar uma mesma distribuição, exceto por possíveis diferenças nos parâmetros de escala e locação. São consideradas distribuições estáveis, as distribuições de Levy, Cauchy e Normal.

Similarmente às distribuições estáveis, as distribuições de valores extremos também se caracterizam por serem max-estáveis.

De acordo com Mendes (pg. 40), define-se Distribuições Max-estáveis como:

Sejam X, X_1, X_2, \dots variáveis aleatórias iid com função distribuição F , e sejam $d_n \in \mathbb{R}$ e $c_n > 0$ constantes apropriadas. Dizemos que F é max-estável se satisfaz a igualdade em distribuição

$$\max(X_1, \dots, X_n) \stackrel{d}{=} c_n X + d_n.$$

A definição mostra que toda distribuição max-estável é distribuição-limite para o máximo de variáveis aleatórias iid.

O teorema de Fisher-Tippett diz que se existirem constantes $c_n > 0$, $d_n \in \mathbb{R}$ tais que o máximo padronizado convirja para uma distribuição não degenerada, então esta deve ser de um dos três tipos (H_1 , H_2 , H_3). De fato, a classe das distribuições max-estáveis coincide com a classe das distribuições limite para o máximo padronizado. Isto é, temos também que:

$$\lim_{n \rightarrow \infty} F_X^n(c_n x + d_n) = H(x)$$

Por exemplo, considere X com distribuição Gumbel. Seja $d_n = \ln(n)$ e $c_n = 1$. Então:

$$\begin{aligned} \lim_{n \rightarrow \infty} F_X^n(c_n x + d_n) &= \lim_{n \rightarrow \infty} \left[\exp\left(-e^{-(x + \ln(n))}\right) \right]^n = \\ &= \lim_{n \rightarrow \infty} \left[\exp\left(-e^{-(x)} \frac{1}{n}\right) \right]^n = \lim_{n \rightarrow \infty} \left[\exp\left(-e^{-(x)}\right) \right] = H_1(x) , \end{aligned}$$

Mostrando assim que a distribuição Gumbel, sendo limite, é estável.

Gráfico de Hill

Sejam $X_1 > \dots > X_n$ estatísticas de ordem de variáveis randômicas independentemente e identicamente distribuídas. O estimador de Hill correspondente ao índice de cauda ξ usando a estatística de de ordem $k + 1$ é definido como:

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln\left(\frac{H_i}{H_{k+1}}\right) = \hat{\xi}$$

O gráfico de Hill é, por sua vez, definido pela seqüência de pontos:

$$\{(k, H_{k,n}^{-1}), 1 \leq k \leq n-1\}$$

O limiar u é selecionado a partir desse gráfico, buscando-se a área de estabilidade do índice de cauda. No entanto, essa escolha nem sempre é clara. Esse método aplica-se bem para distribuições GPD ou de família próxima à GPD. O estimador de Hill é uma aproximação do estimador de máxima verossimilhança para a GPD e como a distribuição de excessos além de um limiar converge a uma GPD, seu uso é justificado (ver resultados de Drees, de Haan e Resnick [2000]).

Teste de Razão de Verossimilhança (Likelihood Ratio – LR)

No presente trabalho o Teste LR²² (Christoffersen [1998] apud Souza [1999]) é utilizado como uma medida da frequência relativa das violações (outliers) de VaR. A frequência de violações deve ficar próxima da frequência esperada dado o intervalo de confiança adotado.

Como se espera que as observações tenham probabilidade α de exceder o VaR, assume-se uma distribuição de Bernoulli para a variável aleatória X_t , que por sua vez assume valor 1 (“sucesso”) caso o VaR seja violado e 0 (“fracasso”) caso contrário:

$$P(X_t) = \begin{cases} \alpha & \text{se } X_t > VaR \\ 1 - \alpha & \text{se } X_t \leq VaR \end{cases}$$

Assume-se que as realizações de X_t sejam independentes. Como $E(X_t) = \alpha$ e observa-se T realizações de X_t , o valor esperado para X no horizonte T é

$$E(X_T) = \sum_{t=1}^T E(X_t) = T\alpha.$$

Como $X_T = \sum_{t=1}^T X_t$ possui distribuição binomial de média $T\alpha$ e variância $T\alpha(1-\alpha)$,

pode-se construir um teste de hipótese para a proporção $\hat{p} = \frac{X_T}{T}$ das realizações

que superam o VaR. O teste de Razão de Verossimilhança testa o modelo binomial com a proporção de violações dada pela hipótese nula de o nível de significância ser o esperado do VaR ($p = \alpha$) contra o modelo onde α é dado pela proporção efetivamente observada ($p = \hat{p}$). A estatística do teste LR possui distribuição Qui-quadrado com 1 grau de liberdade:

$$LR = -2\ln[(1-\alpha)^{(1-\hat{p})T} \alpha^{\hat{p}T}] + 2\ln[(1-\hat{p})^{(1-\hat{p})T} \hat{p}^{\hat{p}T}] \sim \chi^2(1)$$

²² CHRISTOFFERSEN, P. F. (1998). “Evaluating Interval Forecasts”. *International Economic Review*. Vol. 39. No. 4, November 1998.

Domínio de atração

Obtêm-se as distribuições H de valores extremos como distribuições limites ($n \rightarrow \infty$) do máximo de um conjunto de variáveis aleatórias iid e são unicamente determinadas.

O teorema de Fisher-Tippett implica que se $F_X^n(c_n x + d_n)$ é não degenerada quando $n \rightarrow \infty$, para certas constantes $c_n > 0$, $d_n \in \mathbb{R}$, então:

$$\left| F_X^n(x) - H\left(\frac{x - d_n}{c_n}\right) \right| \rightarrow 0, \quad n \rightarrow \infty$$

para alguma H .

Gnedenko [1943] estabeleceu a relação entre a distribuição subjacente F_X e o tipo de distribuição limite. É possível, por exemplo, que máximos e mínimos coletados de uma mesma F_X tenham distribuições limites diferentes. O conjunto das F_X s tais que os seus respectivos máximos possuem a mesma distribuição limite não degenerada é chamada de domínio de atração.

7. BIBLIOGRAFIA

BROOKS, C.; CLARE, A.D. e PERSAND, G (2000). "An EVT Approach to Calculating Risk Capital Requirements". ISMA Centre, University of Reading, UK.

BROOKS, C.; PERSAND, G. (2000). "Value-at-Risk and Market Crashes". Journal of Risk, vol.2, no.4.

DANIELSSON, J.; de VRIES, C. G (1997b). "Beyond the Sample : Extreme Quantile and Probability with Applications to Financial Data". Tinbergen Institute, Rotterdam.

DANIELSSON, J.; de VRIES, C. G (2000). "Value-at-Risk and Extreme Returns". London School of Economics, Financial Markets Group Discussion Paper.

DREES, H.; HAAN, L.; RESNICK, S. (2000). "How to make a Hill Plot". Annals of Statistics, Institute of Mathematical Statistics, Hayward, CA, USA, vol. 28, n°1, pp. 254-274.

EMBRECHTS, P.; DE HANN, L.; HUANG, X. (2000). "Modelling Multivariate Extremes". Department of Mathematik, ETH Zentrum, CH 8092, Zürich.

EMBRECHTS, P.; RESNICK, S.; SAMORODNITSKY, G. (1999). "Extreme Value Theory as a risk management tool". North American Actuarial Journal

FELLER, W. (1971). "An Introduction to Probability Theory and Its Applications". Vol. 2, Wiley, NY.

GALAMBOS, J. (1978). "The Asymptotic Theory of Extreme Order Statistics". 2 ed., Melbourne, Florida: Kreier.

GOURIEROUX, C; MONFORT, A.; TROGNON, A. "Pseudo Maximum Likelihood Methods: Theory". *Econometrica*, Vol. 52, No. 3 (May, 1984), pp. 681-700 doi:10.2307/1913471

GREENWOOD, J.A.; LANDWEHR, J.M.; MATALAS, N.C.; WALLIS, J.R. (1979). "Probability Weighted Moments: Definition and Relation to parameters of Several Distributions Expressable in Inverse Form". Water Resources Research, n.15, p.1049-1054.

HALL, P. (1990). "Using the Bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems". *Journal of Multivariate Analysis*, 32:177-203.

HOSKING, J.; WALLIS, J. (1987). "Parameter and Quantile Estimation for the Generalized Pareto Distribution". *Technometrics* 29, 339-349.

HUBER, P.J (1981). "Robust Statistics". Wiley, New York.

JORION, P. (2001). "Value at Risk: The New Benchmark for Managing Financial Risk". 2/E, The McGraw-Hill Companies, Inc. ISBN 0-07-135502-2.

KENT, J.T.; TYLER, D.E. (1991). "Redescending M-Estimates of Multivariate location and Scatter". *The Annals of Statistics*, Vol.19, n.4, p. 2102-2119.

KOCH, S. P. (1991). "Bias Error in maximum Likelihood Estimation". *Journal of Hydrology JHYDA7*, Vol. 122, p 289-300.

KOTZ, S.; NADARAJAH, S. (2000). "Extreme Value Distributions – Theory and Applications". World Scientific Publishing. ISBN 1-86094-224-5.

LEADBETTER, M.; LINDGREN, G.; ROOTZÉN, H. (1983); "Extremes and Related Properties of Random Sequences and Processes". Berlin, Springer-Verlag.

LONGIN, F.M.. (1996). "The asymptotic Distribution of Extreme Stock Market Returns". *Journal of Business*, Vol. 69, N° 3 (Jul., 1996), pp. 383-408.

MANDELBROT, B. (1963b). "The variation of certain speculative prices". *The Journal of Business of the University of Chicago*: **36**, 394-419.

MCNEIL, A.J. (1997). "Estimating the Tails of Loss Severity Distributions using Extreme Value Theory". *ASTIN Bulletin*, 27, 117-137.

MCNEIL, A.J. (1998). "Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory". Preprint, ETH Zurich.

MCNEIL, A.J.; FREY, R. (1999). "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach". *Journal of Empirical Finance*, n.7, p.271-300.

MCNEIL, A.J.; FREY, R.; EMBRECHTS, P. (2005). "Quantitative Risk Management – Concepts, Techniques and Tools". Princeton University Press.

MENDES, B.V.M. (2000). "Computing Robust Risk Measures in Emerging Equity Markets Using Extreme Value Theory". *Emerging Markets Quarterly*, v.4, n.2, p.25-41.

MENDES, B.V.M. (2004). "Introdução à Análise de Eventos Extremos". E-Papers Serviços Editoriais Ltda.

MENDES, B.V.M.; DUARTE JR., A.M. (2004). "Robust Estimation for ARCH Models (with discussion)". *Brazilian Review of Econometrics*, v.19, n.1, p.139-180.

MORETTI, A. R.; MENDES, B. V. de M. (2003). "Sobre a Precisão das Estimativas de Máxima Verossimilhança nas Distribuições Bivariadas de Valores Extremos". *Pesquisa Operacional*, vol.23, no.2, p.301-324. ISSN 0101-7438.

PINTO, F. C.; VALLS PEREIRA, P. L. (2006). "Teoria de Valores Extremos: Aplicações em Valor em Risco". *Resenha BM&F* n. 162, p.35-50.

SOUZA, L.A.R. (1999). "Valor em Risco em Épocas de Crise". Dissertação de Mestrado, Faculdade de Economia e Administração, Universidade de São Paulo, São Paulo.

STARICA, C. (2000). "Estimation of the Extreme Value distribution for Constant Conditional Correlation Models". *Journal of Empirical Finance*.