

CONTROLANDO O PÂNICO

Rubens Hossamu Morita[§]
Rodrigo De Losso da Silveira Bueno[□]
Ricardo Antônio Pires[†]

RESUMO

Este artigo aplica o algoritmo de Danielsson e De Vries (1997) e métodos de estimação paramétricos para calcular o “*value-at-risk*” baseado na distribuição dos extremos dos índices Ibovespa e MSCI Industrial. Mostra que as previsões fora da amostra, obtidas pelos métodos paramétrico e não-paramétrico, são consideravelmente melhores que o método convencional, em que se usa a Normal, para calcular o “*value-at-risk*” no caso de carteiras de ativos, cuja metodologia de composição está bastante consolidada. O artigo sugere a integração dos métodos de cálculo de “*value-at-risk*” em condições de normalidade e extremas.

Palavras-chave: teoria de valores extremos, VaR.

ABSTRACT

This article applies Danielsson and De Vries (1997) algorithm and parametric estimation methods to calculate the value-at-risk based upon the extreme distribution of Ibovespa and Industrial MSCI indexes. It shows that out of sample forecasts of both methods are better than using Normal distribution. The article suggest integrating both methods to calculate the value-at-risk in normal and extreme conditions.

Keywords: extreme value theory, VaR.

JEL classification: C14, C15, G12.

§ Mestre em Economia pela EESP/FGV. O autor agradece o apoio financeiro do Centro de Estudos Bancários da FGV. Email: rhmorita@gvmail.br.

□ Pesquisador GVCepe. O autor agradece o apoio do Centro de Estudos Bancários da FGV e GVPesquisa. Endereço para contato: CFC/EAESP – Fundação Getúlio Vargas – R. Itapeva, 474 – 8º andar – Bela Vista. CEP: 01332-000 – São Paulo – S.P. E-mail: rodrigo.bueno@fgv.br.

† Mestre em Economia pela FEA/USP – Universidade de São Paulo e consultor pela PRA. Email: ricardo@pra.com.br.

Recebido em agosto de 2007. Aceito para publicação em março de 2008.

1 INTRODUÇÃO

O mundo moderno é sujeito a flutuações extremas, em geral difíceis de antecipar. São as crises, as quais, com frequência, causam pânico nos agentes econômicos. Não é possível evitar as crises, mas é possível controlar o pânico que elas geram. Assim, este artigo aplica metodologias de mensuração de risco, condicionais a um estado extremo de crise. Em seguida, compara as metodologias em termos de erro de previsão fora da amostra, de forma a mensurar o grau de acurácia e eficiência de cada uma.

Um estudo sistemático de eventos críticos é de suma importância para o processo decisório dos investidores e de suas políticas de contingência. Mais importante é notar que o desempenho bom ou ruim de uma carteira de ativos é gerado justamente naqueles poucos dias em que houve retornos acentuados. Sem ferramentas de controle adequadas, decisões impróprias podem ser tomadas, justamente numa situação em que os erros custam mais caro. O estado de pânico, ou crise, decorre, essencialmente, da dificuldade de se estimar com precisão os riscos que o evento proporciona.

Os estudos empíricos e os modelos de mensuração de riscos são, em geral, associados ao primeiro e segundo momentos da distribuição, tais como a variância e retorno médio. Menos atenção tem sido dada ao movimento de uma variável aleatória, condicional ao extremo de sua distribuição, objeto da Teoria de Valores Extremos – TVE. Essa observação empírica contrasta com o fato de que, na verdade, os investidores importam-se mais com os riscos extremos do que com a média dos resultados, pois é nos extremos que o sucesso ou o fracasso da gestão de investimentos emerge.

Este artigo usa a Teoria de Valores Extremos – TVE – para examinar os movimentos dos índices Ibovespa e MSCI Industrial (índice de retorno de ações do setor industrial calculado pela Morgan Stanley Capital International) numa perspectiva semelhante à de Longin (1996), e compara o desempenho de metodologias de estimação concorrentes. A primeira baseia-se num método paramétrico, em que se maximiza a função de verossimilhança, usando a distribuição generalizada de Pareto. A outra se baseia num método não-paramétrico para determinar o início da cauda de uma distribuição. Determinar o desempenho desses métodos em relação à hipótese tradicional de normalidade dos retornos é o objetivo desse trabalho. Pretende-se, assim, sugerir um método para cálculo mais preciso dos riscos envolvidos em situações de pânico.

A TVE tem sido usada em diversas áreas do conhecimento científico, como Economia, Finanças, Meteorologia, Astronomia e Biologia, desde o século XVII.¹ Bortkiewicz (1922) iniciou a formalização da TVE, e Fisher e Tippett (1928) alcançaram resultados importantes para casos univariados em 1928. Os casos bivariados foram estudados por Oliveira (1958, 1962), Sibuya (1960) e Gumbel (1954, 1958). Em Economia, Rothschild e Stiglitz (1970) atribuíram pesos às caudas de duas variáveis a fim de propor uma melhor mensuração de um risco. McCulloch (1978) estudou as descontinuidades dos preços, associadas a grandes quedas e elevações. Jansen e De Vries (1991) usaram os extremos para o estudo das distribuições das caudas.

Pesquisas recentes têm envolvido a TVE em questões empíricas de Finanças e Ciências Atuariais em razão da necessidade de maior precisão a respeito dos riscos envolvidos em situações críticas, as quais não são consideradas no escopo de uma distribuição normal. Nesses casos, é preciso

1 Ver Gumbel (1958).

estudar a distribuição dos eventos quando estes se situam nas caudas das distribuições convencionais (e.g. Normal, *t*-student). O problema prático, sobre o qual este artigo se debruça, é encontrar o ponto da distribuição a partir do qual se situa o extremo, ou seja, uma situação econômica de crise. Em Finanças, particularmente, as pesquisas ainda estão florescendo, especialmente em mercados emergentes. Silva e Mendes (2003) analisam o mercado de ações asiático. Ferreira, Mendes e Duarte (2000) aplicam a metodologia para mercados latino-americanos. Nenhum desses trabalhos, no entanto, confronta diferentes métodos de estimação de valores extremos.

A análise tradicional usa estatística e probabilidade baseadas no teorema do limite central. Preocupa-se, portanto, com eventos associados ao centro da distribuição. Assim, revelam-se ineficazes sob condições extremas, justificando o uso de métodos mais sofisticados de avaliação. No contexto financeiro, é preciso modelar elevações e quedas abruptas, que podem causar grandes perdas, haja vista que a distribuição Normal subestima os riscos associados a eventos extremos.

Além desta introdução, o artigo divide-se da seguinte forma: dado que a TVE não é largamente conhecida, há uma breve exposição de seus fundamentos na Seção 2. Os métodos de estimação são discutidos na Seção 3. A parte empírica ficou reservada à Seção 4. A última seção dedica-se às conclusões.

2 FUNDAMENTOS DE TVE

Como a Teoria de Valores Extremos – TVE – não é largamente conhecida, esta seção discute os fundamentos básicos da TVE, seguindo Gumbel (1958) e Mendes (2004).

O termo **valor extremo** tem dois significados: ou é o máximo (mínimo), isto é, o maior (menor) valor de uma série; ou representa os excedentes, isto é, os maiores valores de um conjunto de dados a partir de um limiar suficientemente alto. Dessa forma, a TVE modela os extremos usando ou a distribuição do máximo (mínimo) ou dos excedentes. É preciso, assim, formalizar os conceitos de máximo e de excedente

No estudo do máximo, a amostra é dividida em subperíodos. De cada subperíodo, extrai-se o valor máximo, para formar o conjunto de dados extremos. O problema do procedimento é que os subperíodos são de mesmo tamanho, e o valor extremo de um subperíodo pode ser menor do que valores inferiores ao extremo de outro. Em conseqüência, a freqüência de extremos fica mal mensurada. Nesse sentido, a metodologia de excedentes é mais adequada, porque utiliza os maiores valores de **toda a amostra**.

Os valores dos excedentes são determinados de acordo com um limiar adotado, conforme a metodologia *POT* (*peaks over threshold*), cuja discussão é reservada a uma seção específica adiante. A escolha do limiar não é trivial, porque dele depende o número final de observações a serem consideradas. Se houver muitas observações, haverá uma enorme variabilidade dos dados, o que não é desejável pelo aumento da imprecisão das estimativas. Se o limiar gerar um reduzido número de observações, as estimativas poderão ser poucos confiáveis. Logo, uma forma endógena de definir o

limiar é recomendável e esta é a razão do desenvolvimento do método não-paramétrico a ser discutido mais tarde.

2.1 Distribuição do máximo

Para a distribuição do máximo,² suponha que $\{X_i\}_{i=1}^{sn}$ seja uma seqüência de retornos contínuos observados nos dias $1, 2, 3, \dots, s, s+1, \dots, sn$. O máximo dos primeiros s retornos contidos na amostra será chamado por $Y_{s,1}$; nas s observações seguintes, $X_{s+1}, X_{s+2}, X_{s+3}, \dots, X_{2s}$, o máximo será denotado por $Y_{s,2}$. Assim, para uma amostra sn observações, obtém-se a seqüência de máximos: $Y_{s,1}, Y_{s,2}, Y_{s,3}, \dots, Y_{s,n}$.

Se X_i é independente de X_j , $i \neq j$, e se têm a mesma distribuição,³ então a distribuição exata dos máximos, Y_s , pode ser escrita como uma função parente, F_X , descrita pela expressão: $F_{Y_s}(x) = [F_X(x)]^s$. De acordo com essa formulação, conclui-se que a distribuição exata de Y_s é nula para x menor que um limite superior u e igual a 1 se x é maior que u , o que torna a distribuição exata degenerada. Infelizmente, a distribuição das variáveis parentes não é precisamente conhecida, o que dificulta a modelagem da distribuição do máximo, embora possível mediante técnicas estatísticas avançadas, brevemente descritas a seguir.

Para encontrar a distribuição baseada nos extremos, portanto, é necessário que a variável que indica o máximo, Y_s , seja padronizada de acordo com os parâmetros $b_s \in \mathbb{R}$ e $a_s > 0$. Com isso, a distribuição de extremos padronizada, $\frac{Y_s - b_s}{a_s}$, não se degenera. De fato, Gnedenko, *apud* Embrechts, *et al.* (2004), prova o teorema de valores extremos, também conhecido como teorema de Fisher-Tippet, e explicita que há três possíveis formas de distribuição para $F_Y(x)$: A distribuição de Gumbel (tipo I),

$$F_Y(y) = \exp(-\exp(-y)), y \in \mathbb{R};$$

a distribuição de Fréchet (tipo II),

$$F_Y(y) = \begin{cases} 0, & y \leq 0 \\ \exp(-y^{-\alpha}), & y > 0 \text{ e } \alpha > 0 \end{cases}; \text{ e}$$

a distribuição de Weibull (tipo III),

$$F_Y(y) = \begin{cases} \exp(-(-y)^{-\alpha}), & y < 0 \text{ e } \alpha < 0 \\ 1, & y \geq 0 \end{cases}.$$

2 Do mesmo modo, pode-se trabalhar com a distribuição do mínimo, adotando o seguinte procedimento: $\text{Min}(\{X_i\}_{i=1}^n) = \text{Max}(\{-X_i\}_{i=1}^n)$.

3 Hipótese para passeio aleatório em preços e índices de ações.

Há como fundir esses casos num só. A *Distribuição Generalizada de valores Extremos*, doravante *GEV*, proposta por Jenkinson (1955) e conhecida também como distribuição de Von Mises-Jenkinson, usa a chamada γ -parametrização:

$$G_\gamma(y) = \exp\left[-(1 + \gamma y)^{-\frac{1}{\gamma}}\right], \quad 1 + \gamma y \geq 0.$$

O parâmetro γ é o índice de cauda e está relacionado com α pela equação $\gamma = \frac{1}{\alpha}$. Se $\gamma > 0$, então G_γ representa a distribuição de Fréchet; se $\gamma < 0$, trata-se da de Weibull. A distribuição de Gumbel ocorre quando $\gamma = 0$, que é o caso limite em que $\gamma \rightarrow 0$. A distribuição de Gumbel pode ser considerada como uma forma intermediária entre as de Fréchet e Weibull, porque, para valores muito pequenos de γ , as distribuições tipo I e II são muito próximas da de Gumbel.

2.2 Distribuição de excedentes

Para expor o conceito de excedente, considere a seqüência $\{X_i\}_{i=1}^n$, de variáveis aleatórias *i.i.d*. Os excedentes serão os valores X_i tais que $X_i > u$. O limiar u deve ser determinado de alguma forma; por hora, supõe-se que seja conhecido. Esses excedentes podem ser caracterizados em termos das estatísticas de ordem, as quais simplesmente ordenam os dados originais pela sua magnitude. Ou seja, a seqüência $\{X_i\}_{i=1}^n$, quando ordenada, transforma-se na seqüência $\{X_{i,n}\}_{i=1}^n$, tal que:

$$\begin{aligned} X_{1,n} = Y_n = \max\{X_1, \dots, X_n\} &\geq X_{2,n} \geq \dots \geq X_{n,n} = \\ &= \min\{X_1, \dots, X_n\}. \end{aligned} \quad (1)$$

A partir da série ordenada e com um limiar u predeterminado, denota-se a seqüência $\{Y_i\}_{i=1}^m$ como os excessos além de um limiar u ; ou ainda, quando $\{X_i\}_{i=1}^m - u \geq 0$, então $\{Y_i\}_{i=1}^m = \{X_i\}_{i=1}^m - u$. Portanto, a distribuição de probabilidade acumulada condicional de excessos $\{Y_i\}_{i=1}^m = \{X_i\}_{i=1}^m - u$, denotada por $F_u(y)$, é definida por:

$$\begin{aligned} F_u(y) &= P(Y \leq y | X > u) = P(X - u \leq y | X > u) = \\ &= \frac{F(u + y) - F(u)}{1 - F(u)}, \quad y > 0 \end{aligned} \quad (2)$$

As formas para a caracterização dos excedentes vão formar a base da metodologia conhecida como *POT*, que considera os valores acima de um limiar preestabelecido.

Na metodologia dos excedentes, também define-se uma distribuição generalizada. Se X for uma v.a. com função de probabilidade de valores extremos, então, sob um limiar u normalizado, as variáveis que representam os excessos, $X - u$, seguem uma *Distribuição Generalizada de Pareto*, doravante *GPD*. Usando o argumento de Embrechts *et al.* (2004), pelo qual, para um limiar u suficientemente alto, existe uma função $\beta(u) > 0$, tal que a distribuição de probabilidade acumulada de excessos (2) pode ser representada como uma boa aproximação da *GPD*, tem-se:

$$P_{\gamma, \beta(u)}(y) = \begin{cases} 1 - (1 + \gamma \frac{y}{\beta(u)})^{-\frac{1}{\gamma}}, & \text{para } \gamma \neq 0 \\ 1 - e^{-\frac{y}{\beta(u)}}, & \text{para } \gamma = 0 \end{cases} \quad (3)$$

em que

$$\beta(u) > 0;$$

$$y \geq 0 \text{ quando } \gamma \geq 0, \text{ e}$$

$$0 \leq y \leq -\frac{\beta(u)}{\gamma} \text{ quando } \gamma < 0.$$

Dessa forma, para um limiar u suficientemente alto, $F_u(y) \approx P_{\gamma, \beta(u)}(y)$,⁴ obtida a partir de um valor u específico, de forma que os demais parâmetros, γ e $\beta(u)$, serão calculados posteriormente pelos métodos de estimação.

Há uma estreita conexão entre as distribuições generalizadas *GEV*, que caracteriza o máximo, e *GPD*, que caracteriza os extremos, representada pela expressão:

$$P_\gamma(y) = 1 + \ln(G_\gamma(y)), \text{ com } \ln(G_\gamma(y)) > -1 \quad (4)$$

Essa característica explica por que as funções densidade de probabilidade da *GPD* e da *GEV* possuem caudas extremas assintoticamente equivalentes. Fundamental é que, para um dado limiar no valor de u , os parâmetros γ , b_s , e a_s da *GEV* determinam os valores γ e $\beta(u)$ da *GPD*. Em particular, o índice de cauda γ é o mesmo para a *GEV* e *GPD* e independem do limiar especificado u . Além disso, em função dessa associação entre a *GPD* e a *GEV*, pode-se inferir que a *GPD* possui três tipos de distribuições condicionadas ao parâmetro γ . Se $\gamma = 0$, então F tem uma distribuição $P_{\gamma, \beta(u)}$, é exponencial do tipo I e pertence à classe Gumbel; se $\gamma > 0$, então F tem uma distribuição $P_{\gamma, \beta(u)}$, é uma distribuição de Pareto do tipo II e pertence à classe Fréchet, e se $\gamma < 0$, então F tem uma distribuição $P_{\gamma, \beta(u)}$ (de Pareto) do tipo III e pertence à classe Weibull.

É preciso mostrar a que convergem cada uma das distribuições a partir dos cálculos amostrais. Isto é feito usando o conceito de Domínio Máximo de Atração. Como há uma relação entre a distribuição do máximo e de excedentes, descrita pela equação (4), pode-se estender esse conceito para a distribuição de excedentes. Em Gnedenko, *apud* Embrechts, *et al.* (2004), encontram-se as condições para que uma distribuição amostral particular convirja para um dos três tipos de distribuições de extremos explicitadas. As condições de Gnedenko são:

Condição 1 – Para que F_Y convirja para uma Gumbel, é preciso que

$$\lim_{s \rightarrow \infty} n[1 - F_X(xa_s + b_s)] = e^{-x} \quad (5)$$

Condição 2 – Para que F_Y convirja para uma Fréchet, é preciso que

$$\lim_{t \rightarrow \infty} \frac{1 - F_X(tx)}{1 - F_X(t)} = y^\alpha \quad (6)$$

4 Isto significa que a *GPD* para um limiar alto converge para uma distribuição de excessos.

em que $t > 0$ e $\alpha > 0$.

Condição 3 – Para que F_Y convirja para uma Weibull, é preciso que

$$\lim_{t \rightarrow 0} \frac{1 - F_X(tx + u)}{1 - F_X(x + u)} = x^\alpha \tag{7}$$

em que u é o limiar superior da distribuição acumulada de $X (F_X(u) = 1)$ definido anteriormente, e $t > 0$ e $\alpha < 0$.

O parâmetro de formato α reflete o peso da cauda da distribuição da variável X , e os parâmetros α , a_s e b_s podem ser diferentes quando se estuda o máximo. Fundamental notar dessas condições que apenas as distribuições não limitadas à direita podem convergir para uma distribuição de Fréchet. Observe que qualquer distribuição estatística recai em um desses critérios. Em particular, a Tabela (1) ilustra para que distribuição de extremos convergem os valores da cauda de algumas distribuições estatísticas bem conhecidas, dadas as condições (5), (6) e (7).

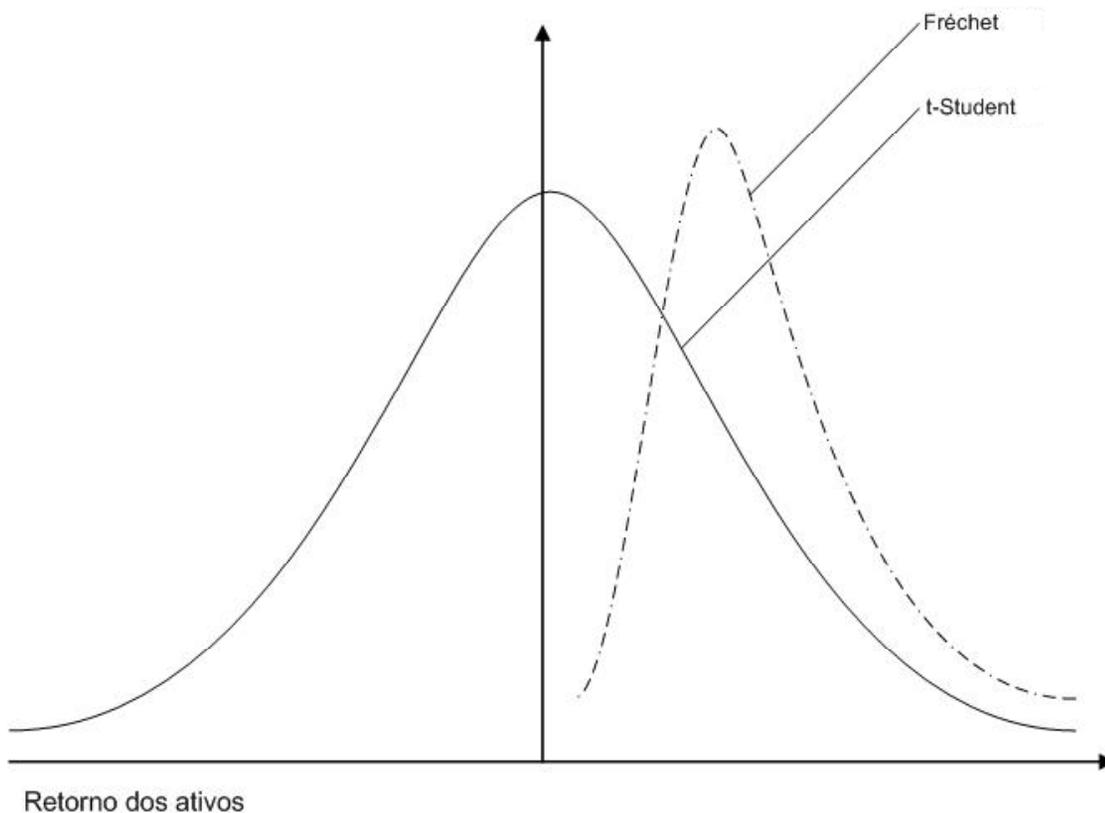
Tabela 1 – Atração e distribuição

Domínio Máximo de Atração	Distribuição
Fréchet	Processos ARCH
	t-Student
	Log-gama
	Pareto
	Cauchy
	Burr
Gumbel	Exponencial
	Benktander I e II
	Gama
	Normal
	LogNormal
Weibull	Uniforme
	Beta

Os três tipos de distribuições de valores extremos representados na Tabela (1) podem ser denominados casos particulares de uma específica *GPD*.

Para ter uma idéia das diferenças entre uma distribuição de caudas pesadas e uma distribuição de extremos, a Figura (1) mostra como se comporta a distribuição convencional perante uma situação de extremos, modelada pela sua contraparte.

Figura 1 – Distribuição de cauda pesada, *t*-student, e de extremos, Fréchet



3 MÉTODOS DE ESTIMAÇÃO

Existem diversos métodos disponíveis de estimação dos parâmetros $\beta(u)$ e γ da distribuição de valores extremos, que podem ser divididos em duas possibilidades. A primeira usa métodos paramétricos, pressupondo que os dados sejam realizações de uma distribuição particular, *e.g.*, a *GPD*. As estimativas dos parâmetros da distribuição de excessos serão baseadas na função de máxima-verossimilhança.

A segunda usa métodos não-paramétricos sem suposições a respeito da distribuição amostral. Um dos estimadores não-paramétricos mais conhecidos é o de Hill, aperfeiçoado por Danielsson e De Vries (1997). O estimador de Hill é fundamentado no conceito de erro quadrático médio e de

bootstrap, e parece ser bastante eficiente. Por isso será usado neste artigo e servirá para encontrar o parâmetro de cauda.

3.1 Estimação paramétrica

3.1.1 Máxima verossimilhança⁵

Os parâmetros estimados pelo método da máxima verossimilhança são consistentes, assintoticamente normais e eficientes.⁶ No caso de excedentes, as estimações são feitas dado um limiar u . Há duas possibilidades de função de verossimilhança de acordo com o valor de γ . Para o caso em que $\gamma = 0$, maximiza-se a função:

$$\mathcal{L}(\gamma, \beta(u); y_1, \dots, y_m) = -m \ln(\beta(u)) - \beta(u)^{-1} \ln\left(\sum_{i=1}^m y_i\right)$$

Quando $\gamma \neq 0$, estima-se:

$$\mathcal{L}(\gamma, \beta(u); y_1, \dots, y_m) = -m \ln(\beta(u)) - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^s \ln\left(1 + \gamma \frac{y_i}{\beta(u)}\right)$$

A determinação de qual função de extremos será usada é feita via gráfico quantílico, em que são confrontados os resultados das estimativas da *GPD* e dos dados originais dos excedentes ordenados.

No segundo caso, espera-se que, na situação em que $y_i \geq 0$, então se obtenha $\gamma > 0$; por outro lado, se $0 \leq y_i \leq -\frac{\beta(u)}{\gamma}$, então $\gamma < 0$.

A maximização dessas funções resulta na obtenção dos parâmetros γ e $\beta(u)$. Particularmente, quando $\gamma > -\frac{1}{2}$, os estimadores de γ e $\beta(u)$ são bem eficientes e, nesse caso, as estimativas resultantes obedecem à seguinte convergência:

$$\sqrt{N_u} \left(\hat{\gamma} - \gamma, \frac{\hat{\beta}(u)}{\beta(u)} - 1 \right) \xrightarrow{d} N_2(0, M)$$

em que

N_u é o número de observações extremas, fixado u ;

\xrightarrow{d} representa convergência em distribuição;

$$M = (1 + \gamma) \begin{pmatrix} 1 + \gamma & 1 \\ 1 & 2 \end{pmatrix};$$

$N_2(a, A)$ é a distribuição normal bivariada com média $a \in \mathbb{R}^2$ e matriz de covariância A .

⁵ Maiores detalhe ver o Apêndice 1.

⁶ Ver Oliveira (1973).

3.2 Estimação não-paramétrica

Em vez de admitir uma distribuição predeterminada para os extremos, ou seja, uma *GPD* ou *GEV*, considere apenas que os dados pertençam a uma distribuição de valor extremo. Neste contexto, podem ser utilizados os estimadores não-paramétricos de Hill e de Picklands (omitido por ser inadequado para este estudo).

Como as séries financeiras são normalmente distribuídas como uma *t*-student, então o estimador de Hill é mais adequado nessas circunstâncias. De acordo com isso, a Tabela (1) permite inferir que as séries financeiras possuem uma distribuição de extremos do tipo Fréchet, cujo parâmetro de cauda é obtido pelo estimador mencionado.

O estimador de Hill para o índice de cauda é dado por:⁷

$$\hat{\gamma}_{n,k}^H = \frac{1}{k-1} \sum_{i=1}^{k-1} \ln \left(\frac{X_{i,n}}{X_{k,n}} \right) \quad (8)$$

em que

$X_{k,n}$ é uma série diária de retornos de forma crescente, pois é um estimador baseado nas estatísticas de ordem de uma variável X ;

k , também chamado de corte, é um inteiro que depende do número de observações dos retornos diários, n . O índice k denota o número de observações a serem consideradas na cauda, ou seja, o número de observações além de um limiar u .

O estimador $\hat{\gamma}_{n,k}^H$ é assintoticamente normal e sua variância assintótica, como mostrado por Dacorogna, *et al.* (1995), é dada por:

$$\text{Var}(\hat{\gamma}_{n,k}^H) \underset{n,k \rightarrow \infty}{\simeq} \gamma^2 \times \frac{1}{k}.$$

Pela expressão anterior, verifica-se que, quanto mais observações entrarem no estimador de Hill (maior k), isto é, quanto menor o limiar u adotado, menor será a variância. Por outro lado, o uso de muitas observações pode deslocar a cauda para o centro da distribuição, incorrendo num viés na estimativa de γ . Nota-se, portanto, que o valor de $\hat{\gamma}_{n,k}^H$ depende do número k de estatísticas de ordem usado e, assim, a escolha do valor de k é feita endogenamente por meio de um algoritmo de otimização. O princípio dessa otimização advém do fato de que, à medida que k varia, existe um “*trade-off*” entre viés e variância no estimador de Hill. Dessa forma, deve existir um ponto ótimo de equilíbrio entre esses dois elementos. Para a otimização, usar-se-á a versão proposta por Danielsson e De Vries (1997), pela qual se faz uso do método *Bootstrap* de Hall (1990).⁸

⁷ Danielson e De Vries (1997b).

⁸ Para maiores informações ver Apêndice 2.

4 RESULTADOS EMPÍRICOS

4.1 Dados

Seguindo a literatura, escolhemos índices de ações compostos para os testes. Particularmente, usam-se as séries Ibovespa e MSCI industrial⁹ para testar os modelos. O Ibovespa, discutivelmente, representa a carteira de mercado no Brasil e, portanto, tende a ser distribuído de forma mais simétrica em comparação com um índice setorial, tal como o MSCI Industrial. Por razões análogas, o MSCI Industrial deverá ter uma distribuição mais assimétrica e mais concentrada em uma categoria de ativos.

A frequência de dados é diária, em função da necessidade de um grande número de observações para que o estimador de Hill funcione bem. As amostras utilizadas foram coletadas entre os dias 02/01/1986 e 30/03/2006, para o caso do Ibovespa, e entre os dias 02/02/1999 e 19/09/2006, para o caso do MSCI industrial.

Dentro do período indicado, foram calculados os retornos diários da série utilizando o conceito de composição contínua:

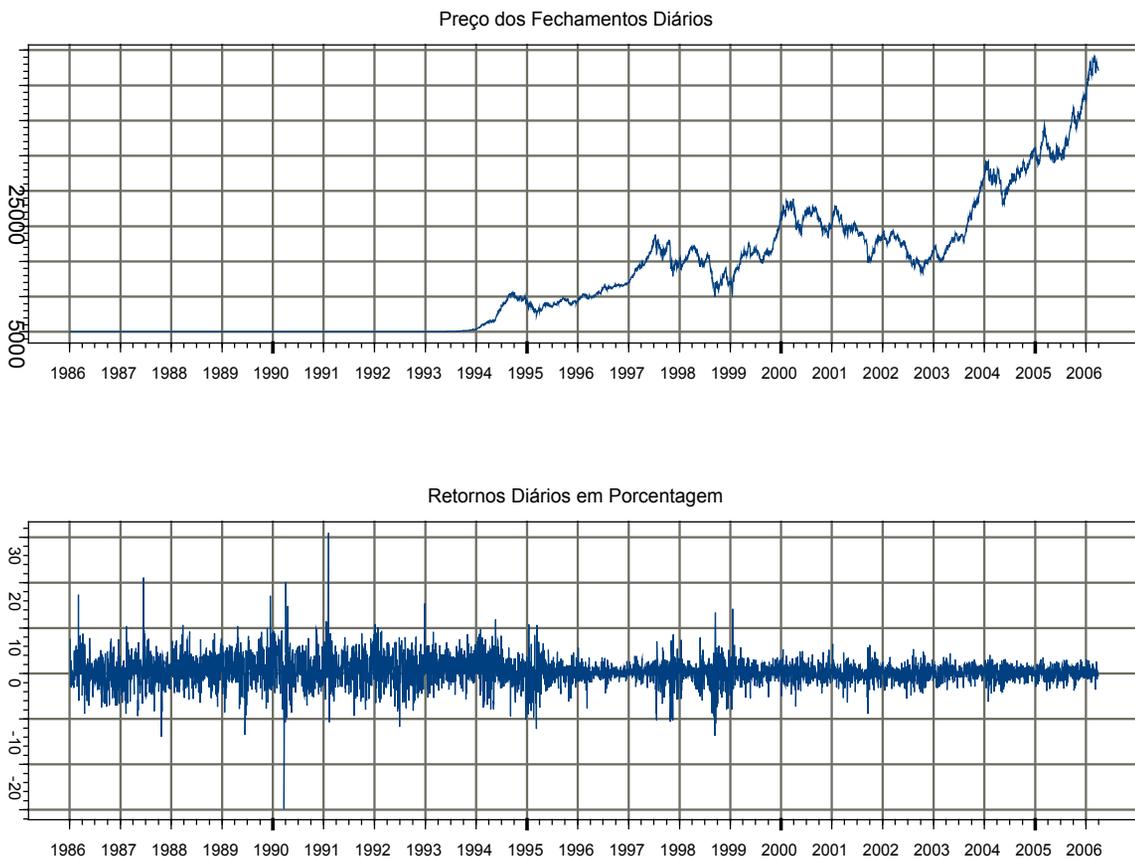
$$r_t = \ln \left(\frac{P_t}{P_{t-1}} \right)$$

em que P_t é o valor médio do índice na data em t .

Em razão das datas de coleta, é de se esperar variações acentuadas no Ibovespa, haja vista estar parcialmente inserido no período megainflacionário brasileiro. Por outro lado, o MSCI industrial está coletado somente após a flexibilização do dólar. Conseqüentemente, espera-se que os resultados da estimação de ambas as medidas sejam razoavelmente diferentes, porém permitindo avaliar o poder da estimação por extremos. De fato, isso pode ser visto na Figura (2).

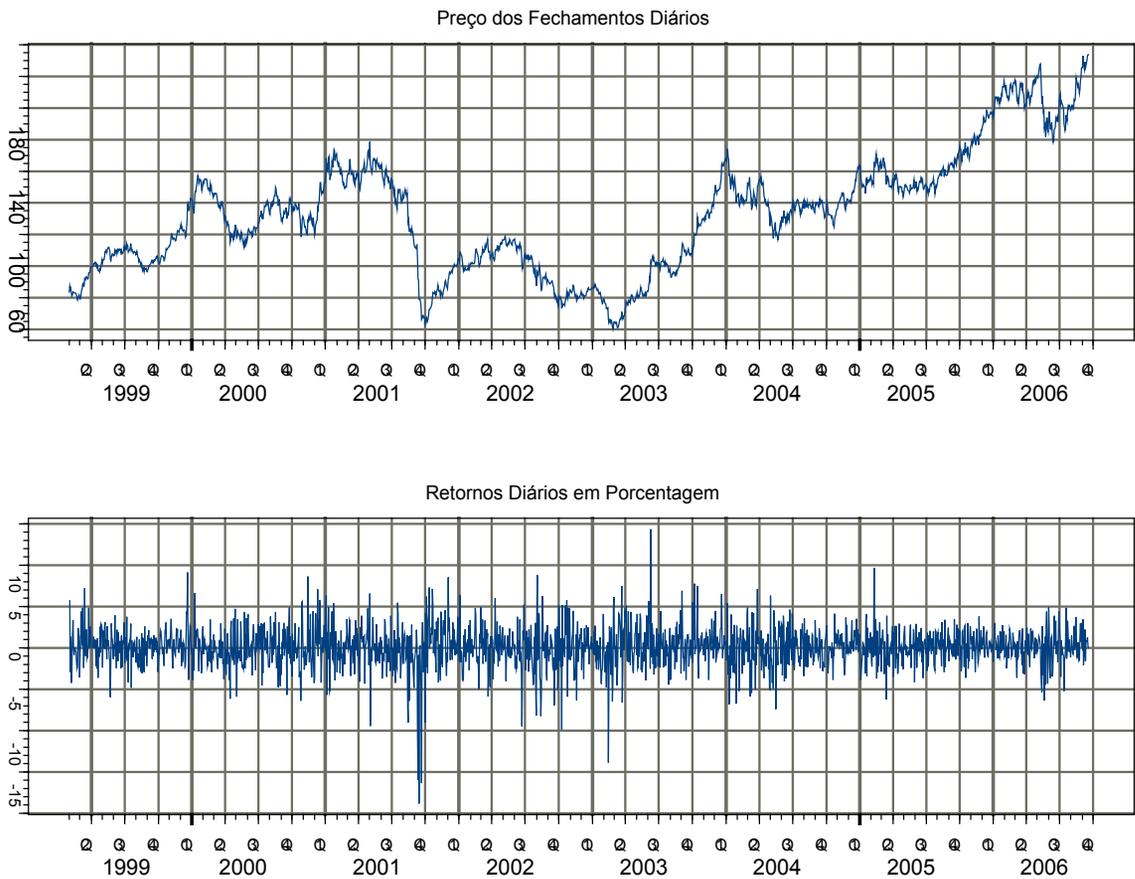
9 Da sigla de *Morgan Stanley Capital International*, é um índice criado pelo Banco Morgan Stanley para avaliar o desempenho das ações da bolsa de valores para o setor industrial brasileiro. Fonte: *Morgan Stanley Capital International Brazil Industrials Index*.

Figura 2 – Preço de fechamento médio e retorno médio diário do Ibovespa



Por outro lado, o MSCI Industrial tem um comportamento bastante diferente do Ibovespa, com variações bem menos acentuadas, embora tenha uma tendência igualmente crescente ao índice de ações no período comum a ambos.

Figura 3 – Índice de fechamento e variação diária do MSCI industrial



Claro que Ibovespa e MSCI industrial apresentam quedas acentuadas e grandes elevações. É justamente nessas quedas e elevações que as avaliações de risco baseadas nas distribuições convencionais *t*-student ou normal perdem poder preditivo.

Os dados indicam que os retornos diários do Ibovespa possuem média em torno de 0,42%, com amplitude entre 30,95% e -29,77%. O excesso de curtose é de 6,80, o que representa a rejeição da hipótese de normalidade. Os retornos diários do MSCI Industrial possuem média de 0,026%, amplitude entre 6,23% e -8,18%, curtose em 5,34, o que que leva à rejeição da hipótese de normalidade. As estatísticas básicas estão na Tabela (2).

Tabela 2 – Estatísticas descritivas: Ibovespa e MSCI para indústria

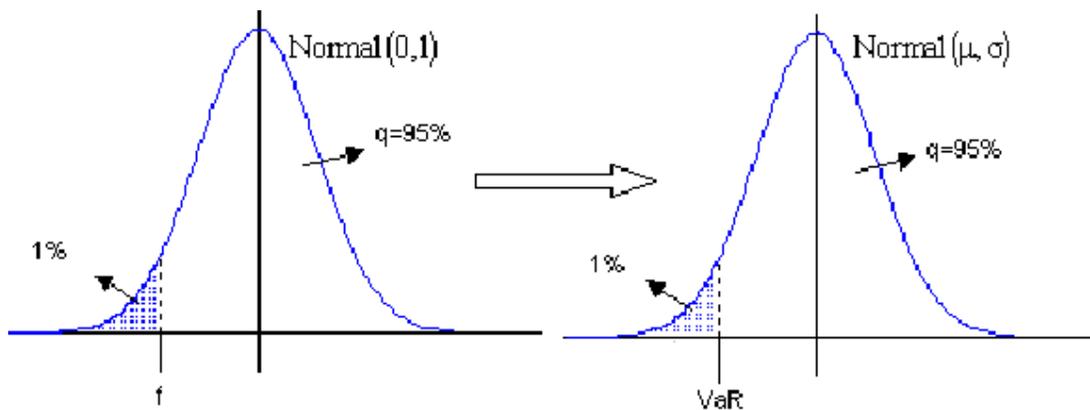
	Ibovespa	MSCI
Média	0,42%	0,026%
Mediana	0,21%	0,021%
Máximo	30,95%	6,23%
Mínimo	-29,77%	-8,18
Desvio Padrão	3%	1,07%
Assimetria	0,17	-0,53
Curtose	6,80	5,34

Outra ferramenta importante para a verificação da hipótese de normalidade é o gráfico quantílico,¹⁰ que confronta a distribuição normal convencional com os dados apresentados. Os gráficos quantílicos indicam que a hipótese de normalidade é rejeitada se houver um distanciamento da reta de 45° entre os dados e a distribuição normal ou *t*-student padrão. Para os dados dos retornos do Ibovespa e do MSCI Industrial, nota-se claramente um distanciamento da reta normal, caracterizando, pois, caudas mais pesadas quando comparadas à distribuição normal.

4.2 Resultados empíricos e previsões

O principal intuito deste estudo é encontrar uma forma mais precisa de mensuração de risco em ambientes de estresse. Neste contexto torna-se imprescindível a introdução de uma medida de risco que possibilite essa avaliação. A medida usual é o valor em risco, ou *VaR*. O *VaR* é o mais alto quantil da distribuição de perda para um determinado nível de significância, ou seja, dizemos que o VaR_q é o *q*-ésimo quantil da distribuição de *F*. Para ilustrar o que é o *VaR* considere a Figura (4).

Figura 4 – Relação entre a normal padronizada e o retorno crítico para o cálculo do *VaR* para o caso clássico da distribuição Normal



Portanto o *VaR* pode ser definido como:

$$VaR_q = f_X^{-1}(q),$$

em que

f_X é função densidade de probabilidade da distribuição;

f_X^{-1} é a inversa de f_X ;

q é a probabilidade complementar do nível de significância, ou seja, $1 - q$ é o nível de significância.

Para o caso clássico de uma distribuição normal, definimos o *VaR* como:

¹⁰ Estes gráficos podem ser visualizados no Apêndice 3.

$$\begin{aligned} VaR_q &= \mu + \sigma f_q; \\ X &\sim N(\mu, \sigma^2) \end{aligned}$$

em que

q é a probabilidade complementar do nível de significância;

f_q é o valor crítico para o nível de significância q , segundo uma distribuição normal;

μ é a média dos retornos dos ativos;

σ é o desvio padrão.

Para o caso de valores extremos, segundo o **método paramétrico**, define-se o VaR como:

$$VaR_q = u + \frac{\hat{\beta}(u)}{\hat{\gamma}} \left(\left(\frac{n}{k} (1-q) \right)^{-\hat{\gamma}} - 1 \right)$$

em que

$\hat{\gamma}$ é o estimador de Hill;

u é limiar;

k é o corte;

q é a probabilidade complementar do nível de significância;

$\hat{\beta}(u)$ é o parâmetro da *GPD*.

Para o caso de valores extremos, segundo o **método não-paramétrico**, define-se o VaR da mesma forma como na distribuição normal, porém o f_q , neste caso, será dado pelo algoritmo de Danielsson e De Vries (1997), conforme segue:

$$VaR_q = \mu + \sigma f_q$$

em que

f_q é o valor crítico segundo uma distribuição de valores extremos determinado endogenamente pelo algoritmo de Danielsson e De Vries (1997).

As estimativas do VaR são descritas nas tabelas (3) e (4). Os resultados para as estimações do VaR , no caso paramétrico, indicam que, com nível de significância de 5% de probabilidade, o retorno diário pode ser tão baixo quanto -4,20% para o Ibovespa e -3,54% para o MSCI Industrial; para o caso não-paramétrico, os resultados são -5,10% e -4,38%, respectivamente, para o Ibovespa e para o MSCI Industrial.

Tabela 3 – VaR com q igual a 0,99 e 0,95 para o Ibovespa

Ibovespa VaR_q			
q	TVE-PR	TVE-NPR	Normal
0,99 f_q	7,69% (2,70)	7,22% (2,55)	6,55% (2,32)
0,95 f_q	4,20% (1,54)	5,10% (1,84)	4,50% (1,64)

Similarmente, com nível de significância de 1% de probabilidade, as estimativas indicam, no caso paramétrico, que o retorno diário pode ser tão baixo quanto -7,69% para o Ibovespa e -6,65% para o MSCI Industrial, ou ainda na forma não-paramétrica, indicam -6,07% e -4,38%, respectivamente, para o Ibovespa e o MSCI Industrial. Usando a distribuição normal-padrão, os retornos diários, podem ser tão baixos quanto -4,50% e -6,55% para o Ibovespa e -3,99% e -5,67% para o MSCI Industrial, com 5% e 1% de probabilidade, respectivamente. É possível observar que, a 5% de significância, evidencia-se uma aproximação e até um valor mais elevado nos valores da distribuição normal em relação à distribuição de excedentes. Esse fato decorre do formato da densidade de probabilidade.

Tabela 4 – VaR com q igual a 0,99 e 0,95 para o MSCI Industrial

MSCI Industrial VaR_q			
q	TVE-PR	TVE-NPR	Normal
0,99 f_q	6,65% (2,72)	6,07% (2,48)	5,67% (2,32)
0,95 f_q	3,54% (1,46)	4,38% (1,80)	3,99% (1,64)

O passo seguinte foi comparar o poder preditivo dos modelos um passo à frente, no conceito fora da amostra. As tabelas (5) e (7) apresentam os resultados da estimação para o Ibovespa e para o MSCI Industrial com os piores valores, porém excluídos os valores antes de 1990, por se considerar que o período megainflacionário jamais existirá de novo. Mantivemos os anos megainflacionários antes de 1994 para mostrar que o método, ainda assim, é bastante poderoso.

Na primeira coluna, estão as datas em que ocorreram os piores retornos. A segunda coluna retorna os valores efetivos naquela data. Na terceira coluna, temos a previsão com base na distribuição normal. A quarta e a quinta colunas expressam os resultados da previsão de valores extremos na forma não-paramétrica – NP – e paramétrica – PR, respectivamente. As tabelas (6) e (8) mostram o erro quadrático médio das previsões fora da amostra. É por essa métrica que se comparam os diversos métodos de estimação.

Tabela 5 – Resultados para o Ibovespa

Data	Valor Verificado	Normal	Não-Paramétrico	Paramétrico
27/10/1997	-10,56%	-7,48%	-8,27%	-8,77%
20/03/1990	-10,41%	-7,94%	-8,77%	-9,31%
12/11/1997	-10,30%	-7,53%	-8,32%	-8,83%
18/07/1997	-10,27%	-7,49%	-8,28%	-8,79%
22/12/1994	-10,14%	-7,93%	-8,78%	-9,33%
30/10/1997	-10,02%	-7,50%	-8,29%	-8,80%
19/03/1990	-9,82%	-7,90%	-8,73%	-9,27%
09/04/1990	-9,74%	-8,54%	-9,43%	-10,00%
07/11/1997	-9,67%	-7,52%	-8,31%	-8,82%

Os resultados em termos de erro quadrático médio são apresentados na Tabela (6). Nota-se que o método paramétrico é mais eficaz do que o método não-paramétrico.

Tabela 6 – Tabela do erros quadráticos médios para o Ibovespa. NP = não-paramétrico, PR = paramétrico

	Normal	Não-Paramétrico	Paramétrico
Raiz do EQM	2,40%	1,63%	1,15%

A Tabela (7) mostra os resultados para o caso do MSCI Industrial.

Tabela 7 – Resultados MSCI Industrial

Data	Valor Verificado	Normal	Não-Paramétrico	Paramétrico
31/07/2002	-8,20%	-6,00%	-6,41%	-7,01%
17/05/2004	-7,40%	-6,22%	-6,64%	-7,26%
14/08/2001	-7,17%	-5,13%	-5,48%	-6,01%
18/09/2002	-6,95%	-6,11%	-6,53%	-7,14%
09/01/2004	-6,84%	-6,15%	-6,57%	-7,20%
29/01/2004	-6,73%	-6,18%	-6,60%	-7,22%
24/03/2003	-6,62%	-6,29%	-6,71%	-7,34%
25/02/2003	-6,49%	-6,25%	-6,67%	-7,29%
25/10/2000	-6,41%	-4,67%	-5,00%	-5,47%

O erro quadrático médio das previsões para o MSCI Industrial é apresentado a seguir.

Tabela 8 – Tabela do erros quadráticos médios MSCI Industrial

	Normal	Não-Paramétrico	Paramétrico
Raiz do EQM	1,29%	1,00%	0,76%

Para fins de uma análise mais profunda, define-se o seguinte Índice de Eficiência.

$$EI = 1 - \frac{EQM(\text{Modelo})}{EQM(\text{Referencial})}$$

Neste estudo, o $EQM(\text{Referencial})$ é o índice obtido a partir da Normal, e o $EQM(\text{Modelo})$ é o índice obtido segundo a Teoria de Valores Extremos. Quando a razão $\frac{EQM(\text{Modelo})}{EQM(\text{Referencial})}$ é igual a um, então o EI é igual a zero, significando que o modelo é tão eficiente quanto seu referencial. Se $\frac{EQM(\text{Modelo})}{EQM(\text{Referencial})}$ tende a zero, então o nível de eficiência do modelo tende ao máximo. Para os resultados anteriores, encontramos os seguintes níveis de eficiência:

Tabela 9 – Tabela dos níveis de eficiência dos modelos

	Não-Paramétrico	Paramétrico
Ibovespa	32,1%	52,1%
MSCI Industrial	22,5%	41,1%

As previsões do Ibovespa, segundo os métodos empregados, foram 52% e 32% melhores que os métodos baseados na distribuição normal, respectivamente aos métodos paramétrico e não-paramétrico. Para os dados do MSCI Industrial, os métodos são 22% e 41% melhores que o convencional. É possível ainda observar que as previsões usando a distribuição normal foram violadas em todas essas piores ocorrências.

5 CONCLUSÃO

Na primeira seção deste trabalho, foi exposto que o uso de uma f.d.p. de distribuição normal subestima as ocorrências extremas, de forma que isso se torna um fator limitante na previsão de eventos extremos. Sendo que os eventos extremos fazem parte de qualquer análise moderna de gestão de risco, argumentou-se que o intuito da Teoria de Valores Extremos – TVE – é a previsão de eventos situados nas caudas de uma distribuição. Por isso, a sua utilização para a previsão dos piores retornos parecia ser de grande utilidade na avaliação de grandes perdas potenciais.

Os resultados empíricos apresentados na Seção 4 demonstraram que a metodologia que utiliza a distribuição generalizada de valores extremos é ineficiente em comparação com a metodologia que utiliza a distribuição generalizada de Pareto – *GPD*. Constatou-se, ainda, que as estimativas dos índices de cauda, quando parametrizados pela *GPD*, apresentam comportamento estável para um conjunto de limiares u . Para evidenciar a superioridade dos métodos aqui estudados, o critério usado foi o do erro quadrático médio das previsões fora da amostra, indicando que o desempenho das previsões de valor extremo foi superior em relação ao da normal.

Em relação aos métodos paramétricos e não-paramétricos, constatou-se que são totalmente compatíveis em resultados, desde que as séries escolhidas sigam os pressupostos do modelo. No caso do MSCI Industrial, nota-se uma total compatibilidade entre os métodos. Nos dados do Ibovespa isso não ocorre, pois não se tem certeza a qual tipo de distribuição de valores extremos pertencem os dados do Ibovespa, sendo que, para as estimações não-paramétricas, é imprescindível que os dados sejam distribuídos segundo uma distribuição de valores extremos do tipo Fréchet.

Estes motivos fazem com que a TVE possa ser considerada um instrumento muito útil na moderna gestão de risco dos investimentos, área na qual grande parte dos esforços é direcionada, justamente para o desenvolvimento de métodos que permitam avaliar e controlar grandes perdas, evitando pânico. Isso mostra a necessidade de se integrar aos métodos tradicionais as estimativas de risco por TVE, reconhecendo-se qual é o lugar apropriado dos métodos tradicionais e qual é o lugar da TVE. Ou seja, para a maior parte dos casos, os métodos tradicionais vão funcionar; porém, nos raros casos em que situações críticas ocorrem, deve-se dispor de um instrumento que permita aos gestores de risco ter idéia de quão críticos estes eventos podem ser.

A modelagem de eventos extremos multivariados é a extensão natural deste artigo, pois a inter-relação entre mercados e ativos far-se-ia explícita, com especial atenção aos eventos de cauda. Para isso, usa-se a teoria de cópulas, fugindo ao escopo deste artigo. Entretanto, o interessado pode consultar Kolev, Mendes e Anjos (2006).

REFERÊNCIAS

- BORTKIEWICZ, L. von. Variations und mittlerer Fehler. *Sitzungs-ber. Berli. Math. Ges.*, n. 21, p. 3-11, 1922.
- DACOROGNA, M., MULLER, U. PICTET, O; DE VRIES, C. *The distribution of Extremal foregein exchange rate returns in extremely large data sets*. Tinbergen Institute, 1995. (Working Paper, TI 95-70).
- DANIELSSON, Jon; DE VRIES, Casper. Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, v. 4, p. 241-257, 1997.
- _____. *Beyond the sample: extreme quantile and probability estimation*. LSE, 1998. (Manuscript).
- EMBRECHTS, P.; BASSI, F.; KAFETZAKI, M. *A survival kit on quantile estimation*. Dept. of Mathematics, Swiss Federal Institute of Technology, Zürich, 1996. (Working Paper).
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. *Modelling extremal events for insurance and finance*. Berlin: Springer-Verlag, 2004.
- FERREIRA, R. R.; MENDES, Beatriz V. M.; DUARTE JR., A. M. Extreme market events in Latin American stock markets. *Journal Of Emerging Markets*, v. 5, n. 1, p. 21-48, 2000.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largerst or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, n. 24, p. 180-190, 1928.
- GUMBEL, E. J. Statistical theory of extreme values and some practical applications. *Nat. Bureau of Standards Applications Mathematics Series*, n. 33, p. 1-151, 1954.
- _____. *Statistics of extremes*. New York: Columbia University Press, 1958.
- JANSEN, D. W.; DE VRIES, C. G. On the frequency of large stock returns: putting booms and busts into perspective. *Review of Economics and Statistics*, p. 18-24, 1991.
- JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, v. 87, p. 145-58.
- HALL, P. Using the bootstrap to estimate mean squared error and select smooting parameter in nonparametric problems. *Journal of Multivariate Analysis*, v. 32, p. 177-203, 1990.
- KOLEV, Nikolai; MENDES, Beatriz V. M.; ANJOS, Ulisses dos. Copulas: a review and recent developments. *Applied Stochastic Models in Business and Industry*, v. 4, n. 22, p. 25-55, 2006.
- LONGIN, François M. The asymptotic distribution of extreme stock market returns. *Journal of Business*, v. 69, n. 3, p. 383-408, 1996.
- MCNEIL, A.; SALADIN, T. The peaks over threshold method for estimating high quantiles of loss distributions. In: *Proceeding of XXVIIth International ASTIN*, Colloquium, Cairns, p. 23-43, 1997.
- MCCULLOCH, J. H. Continuous time processes with stable increments. *Journal of Business*, v. 51, n. 4, p. 601-619, 1978.
- MENDES, Beatriz V. M. *Introdução à análise de eventos extremos*. São Paulo: E-papers Editora, 2004.
- OLIVEIRA, J. T. de Extremal distributions. *Revista da Faculdade de Lisboa*, Lisboa, 2 ser. A, Mat., v. VII e Distribuições de Extremos, Curso Mira Fernandes, v. II, 1958.
- _____. Structure theory os bivariates extremes, extensions. *Est. Mat., Estat. e Econ.*, n. 7, p. 165-195, 1962.
- ROTHSCHILD, M.; STGLITZ, J. E. Increasing risk: I. A definition. *Journal of Economic Theory*, v. 2, p. 225-43, 1970.

SIBUYA, M. Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, n. 11, p. 195-210, 1960.

SILVA, André C.; MENDES, Beatriz V. M. Value-at-risk and extreme returns in Asian stock markets. *International Journal of Business*, v. 8, n. 1, p. 17-40, 2003.

APÊNDICE 1

Na subseção 2.2 concluiu-se que, para um dado limiar u , suficientemente grande, então os dados de excesso $\{Y_i\}_{i=1}^m$ podem ser entendidos como uma amostra para a estimação dos parâmetros γ e $\beta(u)$ de uma *GPD* com u determinado endogenamente.

A determinação de qual dos três tipos de distribuição está relacionado aos dados pode ser feita por uma simples análise gráfica via QQ-Plot, ou gráfico de análise quantílica, tomando como a referência de uma distribuição exponencial. Por essa metodologia, se os dados não apresentarem desvios muito grandes da linha padrão do gráfico quantílico, a *GPD* é uma exponencial do tipo Gumbel, com $\gamma = 0$; caso contrário, a *GPD* é do tipo Fréchet, se $\gamma > 0$, ou do tipo Weibull, se $\gamma < 0$.

Figura 5 – QQ-Plot Ibovespa

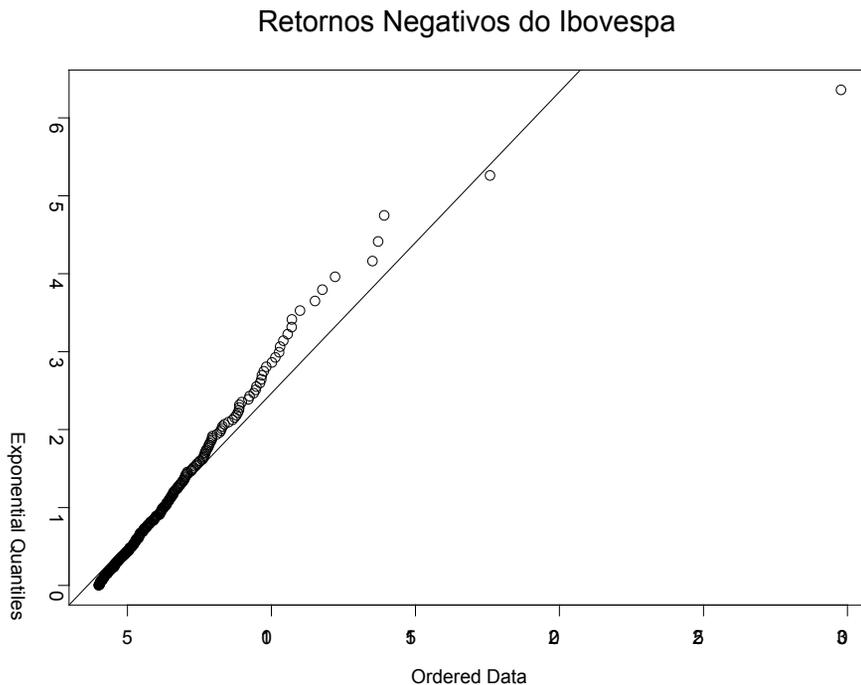
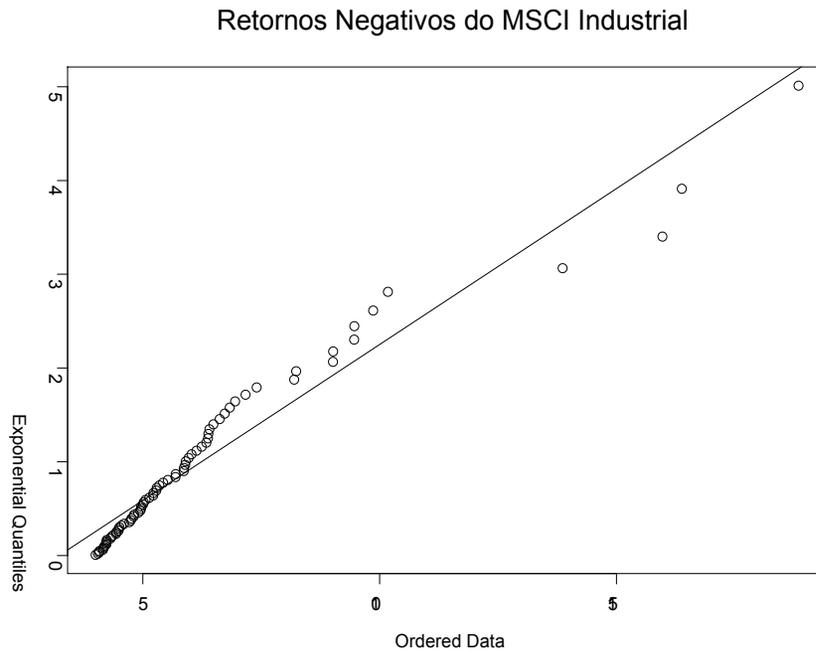


Figura 6 – QQ-Plot MSCI Industrial



Em ambos os casos ocorrem desvios, porém os mais acentuados ocorrem nos dados do MSCI Industrial, evidenciando assim que as caudas do MSCI Industrial são assintoticamente do tipo Fréchet ou Weibul, com $\gamma \neq 0$, ao passo que o Ibovespa pode ser confundido com uma distribuição do tipo Gumbel, com $\gamma = 0$.

Este apêndice descreve o algoritmo de Danielsson e De Vries, que consiste em ordenar a amostra original, obtendo assim as estatísticas de ordem. Posteriormente far-se-á a escolha de um ponto de corte inicial em u_0 , com k_0 dados para representar a cauda da distribuição. Danielsson e de Vries (1997) escolhem para k_0 o inteiro mais próximo de $0,01 \times n$, em que n é o número de observações. Com esse corte escolhido, calculou-se com o uso de (8):

$$\hat{\gamma}_{n,k_0}^H = \frac{1}{k-1} \sum_{i=1}^{k_0-1} \ln \left(\frac{X_{i,n}}{X_{k_0,n}} \right) \quad (10)$$

O valor $\hat{\gamma}_{n,k_0}^H$ é usado como uma primeira estimativa do parâmetro γ . Depois, por iteração, o valor dessa estimativa pode ser melhorado.

Da amostra original, sorteiam-se, com reposição, B subamostras de “*bootstrap*”. O tamanho de cada subamostra é n_1 , sendo $n_1 < n$. Cada amostra de *bootstrap* selecionada deve ser ordenadas de forma que:

$$\begin{aligned} & X_{1,n_1}^1, X_{2,n_1}^1, \dots, X_{n_1,n_1}^1 \\ & X_{1,n_1}^2, X_{2,n_1}^2, \dots, X_{n_1,n_1}^2 \\ & \vdots \\ & X_{1,n_1}^B, X_{2,n_1}^B, \dots, X_{n_1,n_1}^B \end{aligned}$$

Para cada uma das B subamostras, calcula-se o valor de $\widehat{\gamma}_{n_1, k_1}^H$, em que k_1 (com $k_1 < k$) é um ponto de corte de subamostra, a ser otimizado. Calculando B vezes, têm-se os valores:

$$\widehat{\gamma}_{n_1, k_1}^{H,1}, \widehat{\gamma}_{n_1, k_1}^{H,2}, \dots, \widehat{\gamma}_{n_1, k_1}^{H,B}.$$

De posse das estimativas $\widehat{\gamma}_{n_1, k_1}^{H,b}$ e com $\widehat{\gamma}_{n, k_0}^H$, pode-se calcular o erro quadrático médio definido por:

$$EQM(n_1, k_1) = \frac{1}{B} \sum_{b=1}^B \left(\widehat{\gamma}_{n_1, k_1}^{H,b} - \widehat{\gamma}_{n, k_0}^H \right)^2 \quad (11)$$

De posse desse erro quadrático, calcula-se o valor de k_1 , que minimiza o $EQM(n_1, k_1)$, denotado por \bar{k}_1 . Em seguida, calculam-se B vezes o valor $\tilde{\gamma}$, definido como:

$$\tilde{\gamma}_b = \frac{1}{v-1} \sum_{i=1}^{v-1} \ln \left(\frac{X_{i, n_1}^b}{X_{v, n_1}^b} \right), v = \bar{k}_1 - 1, b = 1, 2, \dots, B \quad (12)$$

Assim, obtemos: $\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_B$, e podemos calcular B vezes o valor \tilde{m}^j , definido como:

$$\tilde{m}_b^j = \frac{1}{w-1} \sum_{i=1}^{w-1} \ln \left(\frac{X_{i, n_1}^b}{X_{w, n_1}^b} \right)^j, w = \bar{k}_1 + 1, j = 1, 2, 3, 4, b = 1, 2, \dots, B. \quad (13)$$

Conseqüentemente, obtêm-se $\tilde{m}_1, \tilde{m}_2^j, \dots, \tilde{m}_B^j$ com $j = 1, 2, 3, 4$.

Usando os valores encontrados no passo anterior para cada subamostra, calcula-se B vezes o valor Δ , definido como:

$$\Delta_b = \frac{\tilde{m}_b^1 - \frac{\tilde{m}_b^2}{2\tilde{m}_b^1}}{\frac{\tilde{m}_b^3}{3\tilde{m}_b^2} - \frac{\tilde{m}_b^4}{4\tilde{m}_b^3}}, b = 1, 2, \dots, B \quad (14)$$

Em seguida, calcula-se uma estimativa β para cada subamostra:

$$\beta_b = \frac{1}{\tilde{\gamma}_b} \left(\sqrt{\Delta_b} - 1 \right), b = 1, 2, \dots, B \quad (15)$$

Finalmente, calculam-se as médias de $\tilde{\gamma}$ e $\hat{\beta}$:

$$\bar{\gamma} = \frac{1}{B} \sum_{b=1}^B \tilde{\gamma}_b \quad (16)$$

$$\bar{\beta} = \frac{1}{B} \sum_{b=1}^B \hat{\beta}_b \quad (17)$$

Com os valores de $\bar{\gamma}$ e $\bar{\beta}$, pode-se encontrar o ponto de corte ótimo \bar{k} de amostra inteira, usando

$$\bar{k} = \bar{k}_1 \left(\frac{n}{n_1} \right)^{\frac{2\bar{\beta}}{1+2\bar{\beta}}} \quad (18)$$

Com o ponto de corte ótimo \bar{k} de amostra inteira, calcula-se o estimador de Hill como:

$$\hat{\gamma}_{n,\bar{k}}^H = \frac{1}{\bar{k}-1} \sum_{i=1}^{\bar{k}-1} \ln \left(\frac{X_{i,n}}{X_{\bar{k},n}} \right) \quad (19)$$

Para introduzir um procedimento iterativo, usa-se o valor $\hat{\gamma}_{n,\bar{k}}^H$ encontrado no lugar do $\hat{\gamma}_{n,k_0}^H$. Depois, repete-se o algoritmo, encontrando um novo valor para o ponto de corte ótimo \bar{k} e, consequentemente, para o estimador de Hill. As iterações podem ser repetidas quantas vezes quisermos, a fim de melhorar o valor estimado do parâmetro γ .

APÊNDICE 3 – GRÁFICOS QUANTÍLICOS

Figura 7 – Gráfico quantílico em relação à distribuição normal dos dados dos retornos do Ibovespa

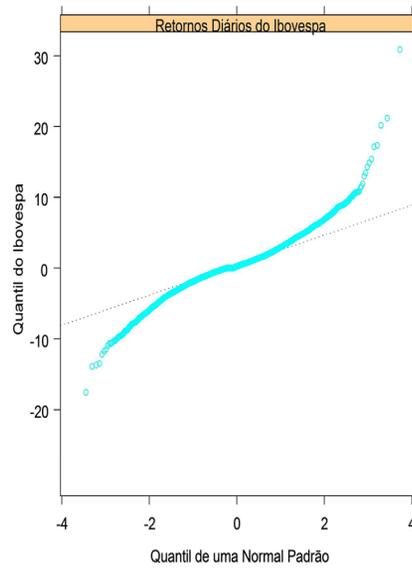


Figura 8 – Gráfico quantílico em relação à distribuição normal dos dados dos retornos do MSCI industrial

