

4

A Teoria dos Valores Extremos

A Teoria dos Valores Extremos (TVE) vem sendo bastante utilizada em campos ligados a eventos raros. Sua estatística é aplicada na estimação de eventos climáticos, cálculo de seguros e eventos pouco comuns no mercado financeiro. A utilização da TVE teve impulso em 1953, quando barragens que protegem a Holanda do avanço do mar se romperam e causaram a inundação de boa parte do país, provocando a morte de 1800 pessoas. Após o desastre, o governo da Holanda criou um comitê que utilizava o ferramental ligado a Teoria dos Valores Extremos para estabelecer a altura das barragens.

4.1. As medidas de risco e a Teoria dos Valores Extremos

Observando-se as definições do VaR, fica evidente que tais medidas são extremamente dependentes de boas estimativas das caudas da distribuição dos retornos, dado que são calculados a partir dos retornos mais baixos da carteira ou ativo em questão.

Os métodos tradicionais de cálculo de VaR podem apresentar deficiências ao estimar as caudas das distribuições, além de se basearem na distribuição inteira dos dados. Sendo assim, novas metodologias estão sendo propostas, com base na Teoria dos Valores Extremos, pois esta tem a capacidade de estimar com razoável precisão probabilidades e quantis nos extremos da amostra, bem como fora desta.

Isso torna a TVE atraente para modelos de cálculo de VaR, dado que este está intimamente relacionado às observações extremas. No entanto, uma das hipóteses fundamentais da TVE é que os dados que compõem a amostra sejam independentes e identicamente distribuídos.

4.2. A distribuição dos mínimos

Assuma que r_t é o retorno de um ativo medido em um intervalo de tempo fixo. Considere a reunião dos n retornos como $\{r_1, \dots, r_n\}$. O retorno mínimo desse

conjunto é $r_{(1)}$, ou seja, quando estes retornos são ordenados levando-se em conta seu valor, $r_{(1)}$ é o menor deles. Aplicando-se o mesmo conceito, $r_{(n)}$ é o retorno de maior valor.

Essa abordagem enfoca as propriedades de $r_{(1)}$ porque esse valor mínimo é de extrema relevância para o cálculo do VaR em posições compradas no ativo, que é o objeto do presente estudo.

Assuma ainda que os retornos r são serialmente independentes e com uma função de distribuição acumulada $F(x)$. Os limites do retorno r são $[l, u]$. Para retornos logarítmicos, tem-se $l = -\infty$ e $u = \infty$. Então a função de distribuição acumulada (CDF) de $r_{(1)}$, chamada de $F_{n,1}(x)$, é dada por:

$$F_{n,1}(x) = P[r_{(1)} \leq x] = 1 - P[r_{(1)} > x]$$

$$F_{n,1}(x) = 1 - P[r_1 > x, r_2 > x, \dots, r_n > x]$$

Utilizando a premissa de que os retornos dos ativos são independentes:

$$F_{n,1}(x) = 1 - \prod_{j=1}^n P(r_j > x)$$

$$F_{n,1}(x) = 1 - \prod_{j=1}^n [1 - P(r_j \leq x)]$$

Como os retornos seguem uma mesma distribuição de probabilidade:

$$F_{n,1}(x) = 1 - \prod_{j=1}^n [1 - F(x)]$$

$$F_{n,1}(x) = 1 - [1 - F(x)]^n \quad (25)$$

Na prática a CDF $F(x)$ de r_i é desconhecida e, portanto, $F_{n,1}(x)$ de $r_{(1)}$ também é desconhecida. Entretanto, fazendo uma adaptação da demonstração para

máximos descrita em Embrechts et al (1997), se n cresce para o infinito, $F_{n,1}(x)$ torna-se degenerada, pois teria-se:

$$\lim_{n \rightarrow \infty} F_{n,1}(x) = 1 - \lim_{n \rightarrow \infty} [1 - F(x)]^n$$

Se $x \leq l$, $F(x) = 0$ e, portanto, $\lim_{n \rightarrow \infty} [1 - F(x)]^n = \lim_{n \rightarrow \infty} [1]^n = 1$, o que daria:

$$\lim_{n \rightarrow \infty} F_{n,1}(x) = 1 - 1 = 0$$

Se $x > l$, $F(x) > 0$ e, portanto, $1 - F(x) < 1$, fazendo com que $\lim_{n \rightarrow \infty} [1 - F(x)]^n = 0$.

Teria-se então:

$$\lim_{n \rightarrow \infty} F_{n,1}(x) = 1 - 0 = 1$$

Essa CDF degenerada não possui valores práticos.

Dada as circunstâncias, recorre-se a resultados de convergência fraca para mínimos centrados e normalizados, fornecidos pelo teorema de Fisher-Tippett (anunciado adiante). Tal teorema argumenta que podem ser obtidas duas seqüências $\{\beta_n\}$ e $\{\alpha_n\}$ onde $\alpha_n > 0$, de tal forma que a distribuição de

$r_{(1)^*} = \frac{(r_{(1)} - \beta_n)}{\alpha_n}$ converge para uma distribuição não degenerada quando n vai

para infinito. A seqüência $\{\beta_n\}$ é uma série local e $\{\alpha_n\}$ é uma série de fatores escalares. Seguindo ainda a ótica do Teorema de Fisher-Tippett sobre o fato de assumirem-se retornos independentes, a distribuição limite dos mínimos normalizados $r_{(1)^*}$ é dada por:

$$F_*(x) = \begin{cases} 1 - \exp\left[-(1 + kx)^{\frac{1}{k}}\right] & \text{se } k \neq 0 \\ 1 - \exp[-\exp(x)] & \text{se } k = 0 \end{cases} \quad (26)$$

para $x < -\frac{1}{k}$ se $k < 0$ e para $x > -\frac{1}{k}$ se $k > 0$, onde o subscrito $*$ significa mínimo. No caso de $k = 0$, aplicou-se $\lim_{k \rightarrow 0} \left((1 + kx)^{\frac{1}{k}} \right) = \exp(x)$. O parâmetro k é chamado de parâmetro de formato que governa o comportamento da cauda da distribuição limite.

A distribuição limite na eq. (26) é a distribuição generalizada do valor extremo de Jenkinson para mínimo. Ela compreende os três tipos de distribuições limites de Gnedenko:

Tipo 1: $k = 0$, a família Gumbel. A CDF é:

$$F_*(x) = 1 - \exp[-\exp(x)], \quad -\infty < x < \infty \quad (27)$$

Tipo 2: $k < 0$, a família Fréchet. A CDF é:

$$F_*(x) = \begin{cases} 1 - \exp\left[-(1 + kx)^{\frac{1}{k}}\right] & \text{se } x < -\frac{1}{k} \\ 1 & \text{caso contrário} \end{cases} \quad (28)$$

Tipo 3: $k > 0$, a família Weibull. A CDF é:

$$F_*(x) = \begin{cases} 1 - \exp\left[-(1 + kx)^{\frac{1}{k}}\right] & \text{se } x > -\frac{1}{k} \\ 0 & \text{caso contrário} \end{cases} \quad (29)$$

A figura abaixo mostra as CDFs para as três distribuições: Gumbel, Fréchet e Weibull.

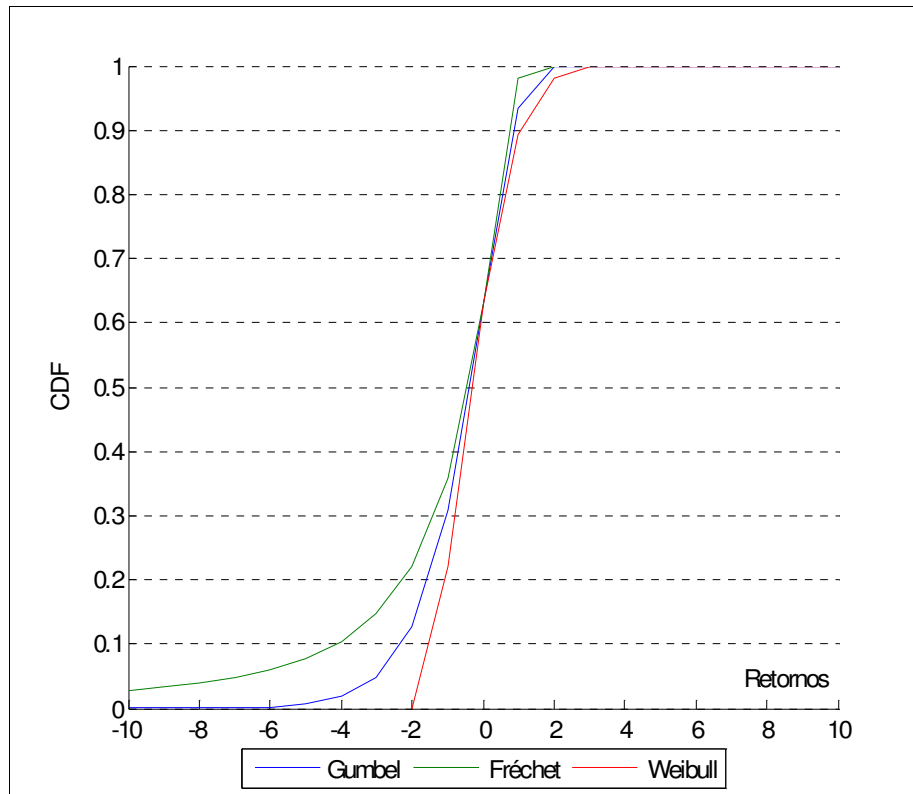


Figura 2 - Função de distribuição acumulada das três distribuições para mínimos.

A função de densidade de probabilidade (pdf) da distribuição limite generalizada da eq. (26) pode ser obtida por diferenciação e aplicação da regra da cadeia:

Se $k \neq 0$:

$$f_*(x) = \frac{dF_*(x)}{dx} = -\exp\left[-(1+kx)^{\frac{1}{k}}\right] \frac{d}{dx}\left[-(1+kx)^{\frac{1}{k}}\right] = -\exp\left[-(1+kx)^{\frac{1}{k}}\right] \left(-\frac{1}{k}(1+kx)^{\left(\frac{1}{k}-1\right)}k\right)$$

$$f_*(x) = (1+kx)^{\left(\frac{1}{k}-1\right)} \exp\left[-(1+kx)^{\frac{1}{k}}\right] \quad (30)$$

Se $k = 0$:

$$f_*(x) = \frac{dF_*(x)}{dx} = -(-\exp(x))\exp(-\exp(x)) = \exp(x)\exp(-\exp(x))$$

$$f_*(x) = \exp(x - \exp(x)) \quad (31)$$

Resumindo:

$$f_*(x) = \begin{cases} (1+kx)^{\left(\frac{1}{k}-1\right)} \exp\left[-(1+kx)^{\frac{1}{k}}\right] & \text{se } k \neq 0 \\ \exp(x - \exp(x)) & \text{se } k = 0 \end{cases} \quad (32)$$

onde $-\infty < x < \infty$ para $k = 0$;

$$x < -\frac{1}{k} \text{ para } k < 0;$$

$$x > -\frac{1}{k} \text{ para } k > 0.$$

A figura a seguir mostra o formato da função de densidade de probabilidade para as três distribuições. Resumindo, o comportamento da cauda de $F(x)$ determina a distribuição limite de mínimos $F_*(x)$. Analisando a cauda esquerda das três distribuições, nota-se que, para a família Gumbel, ela declina exponencialmente; para a Fréchet, ela declina seguindo uma função potência; na família Weibull ela é finita. Para gerenciamento de risco, na maior parte dos casos o interesse recai sobre a família Fréchet, que inclui distribuições estáveis e t-Student. A família Gumbel consiste em distribuições com caudas leves que mais se aproximam às distribuições normal e log-normal.

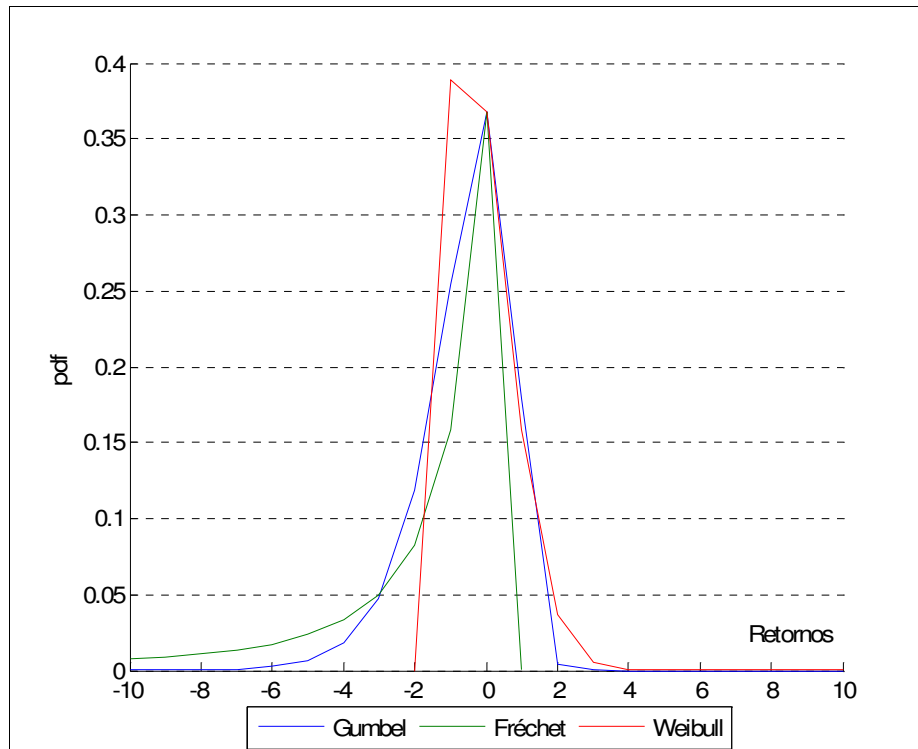


Figura 3 – Funções densidade de probabilidade para as três distribuições para mínimos.

A Teoria dos Valores Extremos possui duas importantes implicações. Primeiro, o comportamento da cauda da CDF $F(x)$ de r_t determina a distribuição limite (normalizada) de mínimos $F_*(x)$. Depois, a teoria é geralmente aplicável a um amplo número de distribuições de retornos r_t .

4.2.1. O teorema de Fisher-Tippett

Seja r_t , $t = 1, 2, \dots, n$ uma seqüência de variáveis aleatórias iid. Se existirem constantes $\alpha_n > 0$ e β_n e alguma distribuição não degenerada H tais que

$$\alpha_n^{-1}(r_{(1)} - \beta_n) \xrightarrow{d} H, \quad n \rightarrow \infty$$

então H pertence a um dos três tipos de distribuição:

Fréchet:

$$\Phi_{\alpha}(x) = \begin{cases} 0 & x \leq 0 \\ \exp\{-x^{-\alpha}\} & x > 0 \end{cases} \quad \alpha > 0 \quad (33)$$

Weibull:

$$\Psi_{\alpha}(x) = \begin{cases} \exp\{-(-x)^{\alpha}\} & x \leq 0 \\ 1 & x > 0 \end{cases} \quad \alpha > 0 \quad (34)$$

Gumbel:

$$\Lambda(x) = \exp\{-e^{-x}\} \quad x \in \Re \quad (35)$$

A referência da demonstração do teorema pode ser encontrada em Embrechts et al (1997).

Embora tais distribuições sejam diferentes, do ponto de vista matemático elas estão intimamente relacionadas, como pode ser verificado através das equivalências abaixo (Embrechts et al (1997)):

$$X \sim \Phi_{\alpha} \Leftrightarrow \ln X^{\alpha} \sim \Lambda \Leftrightarrow -X^{-1} \sim \Psi_{\alpha}$$

O teorema de Fisher-Tippett é válido para qualquer variável aleatória independente e identicamente distribuída, tal que a transformação referida produza uma variável aleatória não degenerada. Dizemos então que a variável aleatória pertence à classe das distribuições max-estáveis.

4.3. Métodos de estimação

Apresentaremos agora alguns dos métodos de estimação das caudas de distribuições utilizando-se a Teoria dos Valores Extremos. O método baseado no estimador de Hill será utilizado no presente trabalho.

A distribuição de valores extremos contém três parâmetros: k , β_n e α_n . Eles podem ser estimados usando-se tanto métodos paramétricos como não-paramétricos.

Para uma amostra dada, só existem um único máximo e um único mínimo. E sendo assim, não se consegue estimar os três parâmetros com apenas uma observação extrema. Usa-se então uma idéia alternativa, que consiste em dividir a amostra em sub-períodos e aplicar a Teoria dos Valores Extremos a esses sub-períodos.

Assuma que existem T retornos possíveis. Divide-se a amostra em g sub-períodos não repetidos, cada qual com n observações, de tal forma que $T = ng$. Se n for suficientemente grande, supõe-se que a Teoria dos Valores Extremos se aplica para cada sub-amostra.

Considere $r_{n,i}$ como o retorno mínimo da i -ésima sub-amostra, onde n significa o tamanho da sub-amostra. Quando n é suficientemente grande, $x_{n,i} = \frac{(r_{n,i} - \beta_n)}{\alpha_n}$ deve seguir uma distribuição de valor extremo, e o conjunto dos mínimos das sub-amostras $\{r_{n,i} \mid i = 1, \dots, g\}$ pode então ser considerada como uma amostra de g observações que segue uma distribuição de valores extremos. Especificamente, define-se:

$$r_{n,i} = \min_{1 \leq j \leq n} \{r_{(i-1)n+j}\}, \quad i = 1, \dots, g$$

Essa definição dá a exata localização dos mínimos das sub-amostras dentro da amostra como um todo. Esse conjunto dos mínimos das sub-amostras serão os dados utilizados na estimativa dos parâmetros da distribuição de valores extremos.

4.3.1. Abordagem Paramétrica

Nesta seção serão apresentadas duas abordagens paramétricas: o método de máxima verossimilhança e o da regressão.

4.3.1.1. Método de Máxima Verossimilhança

Assuma que o conjunto dos mínimos das sub-amostras segue uma distribuição de valores extremos de tal modo que a pdf de $x_{n,i} = \frac{(r_{n,i} - \beta_n)}{\alpha_n}$ é dada pela eq. (32). Pode-se obter a pdf de $r_{n,i}$ realizando uma simples substituição:

$$f(r_{n,i}) = \begin{cases} \left(1 + \frac{k_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{\left[\left(\frac{1}{k_n}\right)-1\right]} \exp\left[-\left(1 + \frac{k_n(r_{n,i} - \beta_n)}{\alpha_n}\right)^{\frac{1}{k_n}}\right] & \text{se } k_n \neq 0 \\ \exp\left[\frac{(r_{n,i} - \beta_n)}{\alpha_n} - \exp\left(\frac{(r_{n,i} - \beta_n)}{\alpha_n}\right)\right] & \text{se } k_n = 0 \end{cases} \quad (36)$$

onde se entende que $1 + \frac{k_n(r_{n,i} - \beta_n)}{\alpha_n} > 0$ se $k_n \neq 0$. O subscrito n adicionado ao parâmetro k significa que sua estimativa depende da escolha de n . Como se assumiu que os retornos são independentes, a função de verossimilhança do conjunto dos mínimos das sub-amostras é:

$$l(r_{n,1}, \dots, r_{n,g} | k_n, \beta_n, \alpha_n) = \prod_{i=1}^g f(r_{n,i}) \quad (37)$$

$$l(r_{n,1}, \dots, r_{n,g} | k_n, \beta_n, \alpha_n) = \prod_{i=1}^g \left\{ \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right]^{\left(\frac{-1+k_n}{k_n} \right)} \right\} \exp \left\{ \sum_{i=1}^g \left[- \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right]^{\frac{1}{k_n}} \right] \right\} \quad (38)$$

Logo, o logaritmo da função de verossimilhança é dado por:

$$L(k_n, \beta_n, \alpha_n) = \ln[l(r_{n,1}, \dots, r_{n,g} | k_n, \beta_n, \alpha_n)] \quad (39)$$

$$\begin{aligned}
&= -\left(\frac{-1+k_n}{k_n}\right) \sum_{i=1}^g \ln \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right] - \sum_{i=1}^g \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right]^{\frac{1}{k_n}} \\
&= \sum_{i=1}^g \left\{ \left[-\left(\frac{-1+k_n}{k_n}\right) \right] \ln \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right] - \left[1 + k_n \left(\frac{r_{n,i} - \beta_n}{\alpha_n} \right) \right]^{\frac{1}{k_n}} \right\} \quad (40)
\end{aligned}$$

Os estimadores de máxima verossimilhança de k_n , β_n e α_n são obtidos pela solução do sistema de equações não lineares formada pelas derivadas parciais da equação acima, em relação a cada parâmetro e igualadas a zero, ou seja:

$$\frac{\partial L(k_n, \beta_n, \alpha_n)}{\partial k_n} = 0$$

$$\frac{\partial L(k_n, \beta_n, \alpha_n)}{\partial \beta_n} = 0$$

$$\frac{\partial L(k_n, \beta_n, \alpha_n)}{\partial \alpha_n} = 0$$

Essa estimativa é não viesada, assintoticamente normal e de mínima variância sob premissas apropriadas. Devemos levar em conta que, ao ajustar a distribuição generalizada do valor extremo aos mínimos de mínimos disjuntos, podemos estar perdendo informações relevantes contidas em outras observações que não os mínimos. O estimador de Hill, o qual será discutido adiante, contorna tal problema, e por isso será utilizado no presente trabalho para estimação do índice de cauda da distribuição generalizada do valor extremo. Para referências sobre tal método de estimação, ver Mendes (2000) e McNeil(1998).

4.3.1.2. Método da Regressão

Esse método assume que o conjunto dos mínimos das sub-amostras são amostras aleatórias da distribuição de valores extremos generalizada e faz uso de propriedades de ordenação estatística. Como resultado, têm-se a regressão:

$$\ln \left[-\ln \left(\frac{g+1-i}{g+1} \right) \right] = \begin{cases} \frac{1}{k_n} \ln \left(1 + k_n \frac{(r_{n(i)} - \beta_n)}{\alpha_n} \right) + e_i & \text{se } k \neq 0 \\ \frac{r_{n(i)}}{\alpha_n} - \frac{\beta_n}{\alpha_n} + e_i & \text{se } k = 0 \end{cases} \quad (41)$$

onde: $i = 1, \dots, g$;

e_i é o erro associado a regressão.

A estimativa pelo método da regressão é consistente, mas menos eficiente do que os estimadores de máxima verossimilhança.

4.3.2. Abordagem Não Paramétrica

O parâmetro de forma k pode ser estimado por alguns métodos não paramétricos. Aqui, dois são mencionados, um proposto por Pickands (estimador Pickands) e outro por Hill (estimador Hill). Ambos estimadores usam diretamente os retornos, não necessitando dividi-los em sub-amostras.

Chamando a ordenação estatística de $r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(r)}$ e chamando de q um número inteiro positivo, os dois estimadores de k são definidos como:

$$\hat{k}_p(q) = -\frac{1}{\ln(2)} \ln \left[\frac{-r_{(q)} + r_{(2q)}}{-r_{(2q)} + r_{(4q)}} \right] \quad (42)$$

$$\hat{k}_h(q) = -\frac{1}{q} \sum_{i=1}^q \{ \ln[-r_{(i)}] - \ln[-r_{(q+1)}] \} \quad (43)$$

onde o argumento (q) é usado para enfatizar que os estimadores dependem de q .

4.3.2.1. Propriedades do estimador de Hill

Segundo o teorema 6.4.6 em Embrechts et al (1997), é possível obter consistência fraca quando (r_i) é iid: se $q \rightarrow \infty$, $q/n \rightarrow 0$ para $n \rightarrow \infty$ então

$$\hat{k}_h(q) \xrightarrow{p} k.$$

Consistência forte e normalidade assintótica também podem ser obtidas sob hipóteses adicionais. Neste caso temos $\sqrt{q}(\hat{k}_h(q) - k) \xrightarrow{d} N(0, k^2)$.

Vale aqui lembrar que o estimador de Hill é derivado com base na hipótese de uma amostra de dados iid. No entanto, o trabalho de Resnick e Starica (1998) mostra que é possível se obter consistência para dados que apresentam dependência.

Como podemos notar na sua definição, o estimador de Hill depende do número de estatísticas de ordem a serem utilizadas. Definir tal número é um problema não muito trivial, o qual será tratado com detalhes a seguir.

O estimador de Hill é aplicado apenas para a distribuição Fréchet, porém, quando aplicável, é mais eficiente que o estimador Pickands.

O estimador de Hill utiliza a amostra de dados como um todo, não havendo a necessidade de divisão da base de dados em sub-amostras para posterior obtenção da distribuição de valores extremos.

4.3.2.2. Escolha do número q de estatísticas de ordem a serem utilizadas

Uma das dificuldades encontradas na utilização do estimador de Hill diz respeito à escolha de q , ou, equivalentemente, à escolha da última estatística de ordem a ser utilizada na estimação. Escolher um valor particular para q significa escolher um ponto de corte onde as caudas da distribuição efetivamente começam, e a partir do qual a TVE pode ser aplicada eficientemente.

O problema é que não existe um consenso sobre qual a melhor maneira de determinar este ponto de corte, e o estimador de Hill é altamente sensível a este valor, principalmente em amostras pequenas. Esta situação é normalmente enfrentada quando tentamos modelar dados do mercado financeiro brasileiro, e portanto, devemos ser cuidadosos ao utilizar o estimador de Hill.

Como dito anteriormente, podemos mostrar que, sob determinadas hipóteses, $\sqrt{q}(\hat{k}_h(q) - k) \xrightarrow{d} N(0, k^2)$. Isso significa que, quanto maior o número q

de estatísticas de ordem, menor a variância do estimador. Poderíamos então pensar em escolher um valor suficientemente alto para q de forma a manter a variância reduzida. No entanto, é possível detectar a existência de um viés em amostras finitas, crescente com o valor de q , o que nos priva de escolher tal elevado valor. Isso nos mostra que existe um “trade-off” entre viés e variância no estimador de Hill, o que nos sugere que deve existir um valor ótimo de q onde viés e variância se estabilizam.

Para resolver tal problema, alguns métodos foram propostos na literatura, dentre eles um método baseado em *bootstrap*, desenvolvido por Danielsson e de Vries (1997a). Preferimos aqui utilizar um método gráfico, chamado de Hillplot (Embrechts et al (1997)).

Método gráfico – “Hillplot”

O Hillplot consiste em traçar o gráfico $\{(q, \hat{k}_h(q)): q = 1, 2, \dots, n\}$, e procurar a região onde o estimador de Hill se estabiliza, ou seja, onde viés e variância devem estar equilibrados. Vale aqui notar que nem sempre é fácil identificar tal patamar no gráfico.

Para finalizar, vale ressaltar que a escolha do q ótimo depende das características de cada série, e que dados dependentes, com maior ocorrência de valores extremos sugerem valores mais elevados para q . Maiores referências podem ser encontradas em Souza (1999) e Danielson e de Vries (1997b).

4.4. Valores Extremos aplicados ao Value at Risk

Nesta seção será discutida uma abordagem para o cálculo do VaR usando a Teoria dos Valores Extremos. Essa discussão será dividida em duas subseções: a primeira será destinada a estimar os parâmetros da distribuição de valor extremo e a segunda enfoca o cálculo do VaR usando as probabilidades de interesse associadas a diferentes intervalos de tempo.

4.4.1. Estimativa dos Parâmetros

Assuma que existem T observações dos retornos de um ativo realizadas no período amostral. Como abordado anteriormente, dividiu-se o período em g sub-períodos não sobrepostos de tamanho n tal que $T = ng$. Se $T = ng + m$, com $1 \leq m \leq n$, então descartam-se as m primeiras observações da amostra. Os parâmetros k_n , β_n e α_n podem ser obtidos usando-se a teoria já abordada anteriormente. Substituindo $x_n = \frac{(r_n - \beta_n)}{\alpha_n}$ na CDF da eq. (26) e chamando de p^*

a menor probabilidade que indica perdas potenciais e r_n^* o p^* -ésimo quantil do sub-período sob a distribuição limite de valores extremos, tem-se:

$$p^* = \begin{cases} 1 - \exp \left[- \left(1 + \frac{k_n (r_n^* - \beta_n)}{\alpha_n} \right)^{\frac{1}{k_n}} \right] & \text{se } k \neq 0 \\ 1 - \exp \left[- \exp \left(\frac{(r_n^* - \beta_n)}{\alpha_n} \right) \right] & \text{se } k = 0 \end{cases}$$

$$-1 + p^* = \begin{cases} - \exp \left[- \left(1 + \frac{k_n (r_n^* - \beta_n)}{\alpha_n} \right)^{\frac{1}{k_n}} \right] & \text{se } k \neq 0 \\ - \exp \left[- \exp \left(\frac{(r_n^* - \beta_n)}{\alpha_n} \right) \right] & \text{se } k = 0 \end{cases}$$

$$1 - p^* = \begin{cases} \exp \left[- \left(1 + \frac{k_n (r_n^* - \beta_n)}{\alpha_n} \right)^{\frac{1}{k_n}} \right] & \text{se } k \neq 0 \\ \exp \left[- \exp \left(\frac{(r_n^* - \beta_n)}{\alpha_n} \right) \right] & \text{se } k = 0 \end{cases}$$

$$\ln(1 - p^*) = \begin{cases} -\left(1 + \frac{k_n(r_n^* - \beta_n)}{\alpha_n}\right)^{\frac{1}{k_n}} & \text{se } k \neq 0 \\ -\exp\left(\frac{(r_n^* - \beta_n)}{\alpha_n}\right) & \text{se } k = 0 \end{cases}$$

O quantil r_n^* será:

Para $k \neq 0$

$$-\ln(1 - p^*) = \left(1 + \frac{k_n r_n^* - k_n \beta_n}{\alpha_n}\right)^{\frac{1}{k_n}}$$

$$[-\ln(1 - p^*)]^{k_n} = \left(1 + \frac{k_n r_n^* - k_n \beta_n}{\alpha_n}\right)$$

$$-\frac{k_n r_n^*}{\alpha_n} = -[-\ln(1 - p^*)]^{k_n} + 1 - \frac{k_n \beta_n}{\alpha_n}$$

$$r_n^* = -\frac{\alpha_n}{k_n} \left(-[-\ln(1 - p^*)]^{k_n} + 1 - \frac{k_n \beta_n}{\alpha_n}\right)$$

$$r_n^* = \beta_n - \frac{\alpha_n}{k_n} \left(1 - [-\ln(1 - p^*)]^{k_n}\right)$$

Para $k = 0$:

$$-\ln(1 - p^*) = \exp\left(\frac{(r_n^* - \beta_n)}{\alpha_n}\right)$$

$$\ln[-\ln(1 - p^*)] = \frac{r_n^*}{\alpha_n} - \frac{\beta_n}{\alpha_n}$$

$$\frac{r_n^*}{\alpha_n} = \frac{\beta_n}{\alpha_n} + \ln[-\ln(1-p^*)]$$

$$r_n^* = \beta_n + \alpha_n \ln[-\ln(1-p^*)]$$

Resumindo:

$$r_n^* = \begin{cases} \beta_n - \frac{\alpha_n}{k_n} \left(1 - [-\ln(1-p^*)]^{k_n}\right) & \text{se } k \neq 0 \\ \beta_n + \alpha_n \ln[-\ln(1-p^*)] & \text{se } k = 0 \end{cases} \quad (44)$$

Como já havia sido dito anteriormente, em finanças, na maioria das vezes, estamos interessados no caso em que $k \neq 0$.

Caso a opção para cálculo do parâmetro de forma da cauda k seja o estimador de Hill, algumas mudanças devem ser efetuadas. Segundo Embrechts et al (1997), o estimador de Hill é uma particularização do estimador de máxima verossimilhança e depois de algum algebrismo, pode-se obter a função de distribuição acumulada abaixo:

$$F(r_i) = \frac{q}{n} \left(\frac{r_i}{r_{(q+1)}} \right)^{\frac{1}{k_h(q)}}$$

O estimador para o quantil torna-se:

$$r = \left[\frac{n}{q} (1-p) \right]^{k_h(q)} r_{(q+1)} \quad (45)$$

Como dito anteriormente, o estimador de Hill utiliza a amostra de dados como um todo, não havendo a necessidade de divisão da base de dados em sub-amostras para posterior obtenção da distribuição de valores extremos.

4.5. Cálculo do VaR

Para uma dada probabilidade p^* (cauda esquerda), o quantil r_n^* da eq. (44) é o VaR baseado na Teoria dos Valores Extremos aplicada aos mínimos dos sub-

períodos da amostra. Agora será explicitada a relação entre os mínimos dos sub-períodos da amostra e as séries de retornos observadas.

Como a maioria dos retornos dos ativos são serialmente não correlacionados ou são fracamente serialmente correlacionados, pode-se usar a relação na eq. (44) e obter-se:

$$\begin{aligned} p^* &= P(r_{n,i} \leq r_n^*) = 1 - [1 - P(r_{n,i} \leq r_n^*)]^n \\ 1 - p^* &= [1 - P(r_{n,i} \leq r_n^*)]^n \end{aligned} \quad (46)$$

Essa relação entre probabilidades permite que se obtenha o VaR através da série de retornos do ativo original. Mais precisamente, para uma pequena probabilidade p , o p -ésimo quantil de r_t é r_n^* se a probabilidade p^* for escolhida com base na eq. (46), onde $p = P(r_t \leq r_n^*)$. Consequentemente, para uma dada probabilidade pequena p , partindo da eq. (44), o VaR será:

$$\text{VaR} = \begin{cases} \beta_n - \frac{\alpha_n}{k_n} \left(1 - [-\ln(1 - p^*)]^{k_n} \right) & \text{se } k \neq 0 \\ \beta_n + \alpha_n \ln[-\ln(1 - p^*)] & \text{se } k = 0 \end{cases} \quad (47)$$

Substituindo a eq. (46) na eq. (44), temos que:

$$\begin{aligned} \text{VaR} &= \begin{cases} \beta_n - \frac{\alpha_n}{k_n} \left(1 - [-\ln(1 - p)]^{k_n} \right) & \text{se } k \neq 0 \\ \beta_n + \alpha_n \ln[-\ln(1 - p)] & \text{se } k = 0 \end{cases} \\ \text{VaR} &= \begin{cases} \beta_n - \frac{\alpha_n}{k_n} \left(1 - [-n \ln(1 - p)]^{k_n} \right) & \text{se } k \neq 0 \\ \beta_n + \alpha_n \ln[-n \ln(1 - p)] & \text{se } k = 0 \end{cases} \end{aligned} \quad (48)$$

Voltando ao caso particular do estimador de Hill, seguindo o resultado da eq. (45), o VaR ficaria:

$$\text{VaR} = \left[\frac{n}{q} (1-p) \right]^{k_h(q)} r_{(q+1)} \quad (49)$$

4.6. VaR Multiperiódico

A regra da raiz quadrada do tempo da metodologia RiskMetrics torna-se um caso especial sob a ótica da Teoria dos Valores Extremos. A relação entre os horizontes de ℓ dias e de 1 dia pode ser dada por:

$$\text{VaR}(\ell) = \ell^{\frac{1}{\alpha}} \text{VaR}(1) = \ell^{-k} \text{VaR}(1) \quad (50)$$

onde:

α é o índice de cauda;

k é o parâmetro de forma da distribuição de valor extremo.

Essa relação é conhecida como a regra da raiz α do tempo. Têm-se que

$\alpha = -\frac{1}{k}$, ou seja, não se deve confundir-lo com o parâmetro α_n .