Universidade de Brasília - UnB
Instituto de Ciências Exatas - IE
Mestrado acadêmico em Matemática

Paulo Augusto Caixeta Borges

Propriedades Assintóticas de um Estimador baseado em Valores Extremos para o Momento Caudal Condicional

Brasília/DF
2023

Paulo Augusto Caixeta Borges

Propriedades Assintóticas de um Estimador baseado em Valores Extremos para o Momento Caudal Condicional

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática da Universidade de Brasília, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática.

Orientadora: Profa. Dra. Cátia Regina Gonçalves

Brasília/DF

2023

Paulo Augusto Caixeta Borges

Propriedades Assintóticas de um Estimador baseado em Valores Extremos para o Momento Caudal Condicional

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática da Universidade de Brasília, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática.

Aprovada em: 15/12/2023.

BANCA EXAMINADORA

Profa. Dra. Cátia Regina Gonçalves (Orientador)
Universidade de Brasília (UnB)

Profa. Dra. Cira Etheowalda Guevara Otiniano
Universidade de Brasília (UnB)

Prof. Dr. Marcelo Bourguignon Pereira
Universidade Federal do Rio Grande do Norte (UFRN)

AGRADECIMENTOS

Neste momento, gostaria de expressar minha profunda gratidão a todos que estiveram ao meu lado, tornando esta jornada possível e enriquecedora. Primeiramente, agradeço a Deus pela minha vida e por me permitir ultrapassar todos os obstáculos encontrados ao longo da realização deste mestrado.

À minha família, agradeço por ser o meu porto seguro, pelo amor, apoio incondicional e por acreditarem em mim nos momentos em que eu mesmo duvidava. Sem vocês, nada disso seria possível.

À minha orientadora, Cátia, minha gratidão por sua orientação perspicaz, paciência e pela maneira dedicada com que conduziu este trabalho. Suas contribuições foram inestimáveis e fundamentais para meu crescimento acadêmico e pessoal.

Aos professores que me introduziram ao mundo da Probabilidade, Wembesom e Adriana, sou grato pela inspiração e apoio que me impulsionaram na direção deste mestrado. Seus ensinamentos foram um farol que guiou meus passos até aqui.

Aos amigos que fiz durante o mestrado (UnB) e graduação (IFB), cuja amizade acrescentou riqueza e significado à minha vida, expresso minha profunda gratidão. Cada um de vocês deixou uma marca indelével em minha jornada acadêmica, compartilhando conhecimento, risadas e apoio nos momentos desafiadores. Particularmente, minha gratidão aos meus amigos do IFB que seguiram comigo para a UnB, Eliézer e Gabriel Miranda, trazendo mais esperança ao mestrado e oferecendo apoio diário em meio a uma pandemia. Muito obrigado pelos momentos de estudo e descontração, online e presencialmente.

À banca examinadora, expresso meu profundo agradecimento pelos esforços na revisão e correção deste trabalho. Agradeço pelas críticas construtivas e contribuições que aprimoraram significativamente este estudo.

A todos os que, de alguma maneira, contribuíram para este trabalho, meu sincero reconhecimento. Suas palavras, apoio e incentivo foram elementos essenciais para esta conquista.

Muito obrigado a cada um de vocês.

"A matemática é a linguagem a qual Deus escreveu o universo."
- Galileu Galilei

RESUMO

Nesta dissertação estudamos as propriedades assintóticas de um estimador, apresentado por Goegebeur et al. (2022), para a medida de risco conhecida como momento caudal condicional. A situação considerada corresponde à extrapolação fora do intervalo de dados e requer argumentos da teoria de valores extremos para a construção do estimador apropriado. Realizamos uma breve análise das principais medidas de risco encontradas na literatura, bem como suas relações com o momento caudal condicional e apresentamos, em detalhes, os resultados obtidos por Goegebeur et al. (2022), que estabelecem, sob condições desejáveis, a distribuição limite do estimador devidamente normalizado.

Palavras-chave: Medidas de Risco; Momento Caudal Condicional; Estimação; Estatísticas de Ordem; Teoria de Valores Extremos.

ABSTRACT

In this dissertation we study the asymptotic properties of an estimator, presented by Goegebeur et al. (2022), for the risk measure known as conditional tail moment. The situation considered corresponds to extrapolation outside the data range and requires arguments from extreme value theory for the construction of the appropriate estimator. We performed a brief analysis of the main risk measures found in the literature, as well as their relationships with conditional tail moment. The results obtained by Goegebeur et al. (2022), which establish under suitable conditions, the limit distribution of the properly normalised estimator are presented in detail.

Keywords: Risk Measures; Conditional Tail Moment; Estimation; Order statistics; Extreme Value Theory.

SUMÁRIO

Introdução

A gestão e mensuração de riscos, principalmente relacionados a eventos extremos, é de fundamental importância para instituições financeiras e de seguros.

Existem diversas definições na literatura para o conceito de risco, variando de acordo com o contexto em que é abordado. No âmbito da ciência atuarial, Denuit et al. (2006) define, de maneira abrangente, o seguro como um negócio destinado a transferir o impacto econômico associado a eventos imprevistos e, assim, um risco pode ser descrito como um evento extremo que ocorre ocasionalmente, podendo acarretar consequências financeiras desfavoráveis. Nesse cenário, sempre há um elemento de incerteza, seja em relação ao momento de ocorrência, à própria ocorrência ou à gravidade de suas implicações.

Assim, utilizando as ferramentas da teoria de probabilidade, o risco pode ser modelado como uma variável aleatória não negativa X, que, no contexto atuarial, representa a quantia aleatória de dinheiro desembolsada por uma seguradora para indenizar um segurado, um beneficiário e/ou um terceiro, conforme estabelecido em um contrato de seguro.

As chamadas *medidas de risco* são uma ferramenta extremamente importante nessa área e foram desenvolvidas com o objetivo de auxiliar na mensuração do risco de determinado seguro, empréstimo, aplicação, etc. De uma maneira geral, uma medida de risco é um funcional que associa a cada risco um número real. Por ser uma definição bastante ampla, uma questão que despertou interesse dos estudiosos no assunto foi a definição de propriedades teóricas desejáveis para que uma medida de risco específica se torne mais aplicável em situações do cotidiano. Um dos pioneiros nesse estudo foi Artzner et al. (1999) que propôs quatro axiomas (translatividade, homogeneidade positiva, monotonicidade e subaditividade), que caso sejam satisfeitos, definem o que se chama *medida de risco coerente*. Na literatura, esse conjunto de axiomas pode variar de autor para autor e produzir outras definições de medidas coerentes. Nesta dissertação, adotamos a definição de Denuit et al. (2006).

Uma das primeiras medidas de risco, e a mais popular, é o Valor em Risco (Value-at-Risk - VaR), introduzida pelo banco americano JP Morgan, no início da década de 1990. Ela se baseia no quantil da distribuição dos resultados e, de forma geral, quantifica, dentro de um nível de confiança especificado, a perda potencial que poderia ser sustentada por uma carteira inteira ou um banco como um todo, em um curto período de tempo, em condições normais de mercado. Especificamente, o VaR é definido para $p \in (0, 1)$, como

$$VaR_p(X) = U(1/p),$$

em que X é uma variável aleatória não negativa e U denota a função quantil caudal de X, isto é, $U(x) := \inf\{y : F(y) \geq 1 - 1/x\}$, $x > 1$, sendo F a função de distribuição de X. Para maiores informações sobre o VaR recomenda-se a leitura de Jorion (2007).

Para contornar algumas das desvantagens do VaR, como, por exemplo, não ser uma medida de risco coerente, uma medida alternativa também comumente usada é a Esperança Caudal Condicional (*Conditional Tail Expectation* - CTE). A CTE, representa a perda esperada condicional, dado que a perda excede seu VaR, ou seja, para $p \in (0, 1)$,

$$CTE_p(X) = E(X|X > U(1/p)),$$

desde que $E(X) < \infty$. Comparativamente ao VaR, a CTE é mais conservativa e satisfaz os quatro axiomas de Denuit et al. (2006) (caso o risco seja uma variável aleatória contínua, como mostraremos adiante) sendo uma medida de risco coerente. Mais informações sobre a CTE podem ser encontradas em Brazauskas et al. (2008) ou Cai e Tan (2007).

Uma generalização do CTE, introduzida recentemente por Methni, Gardes e Girard (2014) é o Momento Caudal Condicional (*Conditional Tail Moment* - CTM), que é definido, como

$$CTM_{\beta,p}(X) = E(X^\beta|X > U(1/p)),$$

para $p \in (0, 1)$ e $\beta > 0$ tal que $\mathbb{E}(X^\beta) < \infty$. A grande vantagem do CTM reside no fato de que estimá-lo permite estimar diversas outras medidas de risco, como por exemplo: a Variância Caudal Condicional (CTV), a Assimetria Caudal Condicional (CTS), o Valor em risco condicional (CVaR), a Esperança Caudal Condicional (CTE), dentre outras. Abordaremos mais sobre este assunto na Seção 2.4.

O principal objetivo desta dissertação é estudar propriedades assintóticas do estimador do CTM, apresentado recentemente por Goegebeur et al. (2022). Considerando uma amostra aleatória de tamanho n da variável risco X, com função de distribuição F, ou seja, n variáveis aleatórias independentes e identicamente distribuídas $X_1, ..., X_n$ com função de distribuição F. O foco do estudo é estimar $CTM_{\beta,p}(X)$ no caso em que p é muito pequeno, ou seja, $p < 1/n$. Para este caso, Goegebeur et al. (2022) propõem o uso das ferramentas da teoria de valores extremos para possibilitar a extrapolação além do intervalo de dados.

Dessa forma, assumindo que a conhecida condição de segunda ordem (SOC) é satisfeita, ou seja, para algum $\gamma > 0$, $\rho < 0$ e alguma função positiva ou negativa $A(t)$, com $\lim_{t\to\infty} A(t) = 0$, tem-se

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}, \qquad \forall x > 0,$$

então o seguinte estimador extrapolado é proposto por Goegebeur et al. (2022)

$$\hat{\theta}_{\beta,p} = \left(\frac{k}{np}\right)^{\beta\hat{\gamma}_k} \tilde{\theta}_{\beta,k/n},$$

em que $\hat{\gamma}_k$ é um estimador de γ devidamente escolhido e $\tilde{\theta}_{\beta,k/n}$ é o estimador do caso intermediário com $k, n \to \infty$ e $p = k/n \to 0$, que é definido por

$$\tilde{\theta}_{\beta,n} := \frac{1}{k}\sum_{j=1}^{k} X_{n-j+1,n}^{\beta},$$

em que $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ são as estatísticas de ordem associadas a amostra aleatória de X. Uma opção para a escolha do estimador $\hat{\gamma}_k$ que satisfaça as condições necessárias foi sugerida por Goegebeur et al. (2022). Eles mencionaram o estimador de Hill, apresentado em Hill (1975), que é dado por

$$\hat{\gamma}_k^H := \frac{1}{k}\sum_{i=1}^{k} \ln X_{n-i+1,n} - \ln X_{n-k,n}.$$

Em particular, a condição (SOC) implica que F está no domínio de atração da distribuição do valor extremo generalizada (GEV)

$$G_\gamma(x) = \exp\left[-(1+\gamma x)^{-1/\gamma}\right]$$

com índice extemal $\gamma > 0$. A classe de distribuições que satisfaz essa condição inclui as distribuições do tipo Pareto, cuja cauda é da forma

$$1 - F(x) = x^{-1/\gamma} l(x),\ x > 0,$$

em que l é uma função lentamente variante no infinito, ou seja,

$$\lim_{t\to\infty} \frac{l(tx)}{l(t)} = 1, \; \forall\, x > 0.$$

Assim, o parâmetro γ determina o peso da cauda superior de F, com valores grandes indicando caudas mais pesadas.

Os resultados centrais do artigo de Goegebeur et al. (2022) estabelecem, sob condições desejáveis, a distribuição assintótica dos estimadores $\tilde{\theta}_{\beta,n}$ e $\hat{\theta}_{\beta,p}$, em que $\beta > 0$ é tal que $E(X^{\beta}) < +\infty$ e $0 < \gamma < \frac{1}{2\beta}$.

Assim, no Capítulo 1 desta dissertação, apresentamos conceitos e resultados gerais preliminares referentes à teoria das funções de variação regular, teoria dos valores extremos e das estatísticas de ordem intermediárias.

No Capítulo 2 apresentamos inicialmente uma síntese sobre os conceitos de risco, medidas de risco e os axiomas que definem as medidas de riso coerentes. Além disso, nas Seções 2.3 e 2.4 apresentamos uma discussão sobre os axiomas de coerência para as medidas clássicas VaR e CTE, respectivamente. Finalizamos o capítulo com a introdução da medida de risco CTM e sua relação com outras medidas de risco comumente encontradas na literatura.

Por fim, no Capítulo 3, apresentamos em detalhes os resultados obtidos por Goegebeur et al. (2022) sobre a derivação das distribuições limites dos estimadores, devidamente normalizados, $\tilde{\theta}_{\beta,n}$ e $\hat{\theta}_{\beta,p}$ sob condições apropriadas.

Capítulo 1

Preliminares

O objetivo principal deste capítulo é apresentar diversos conceitos e resultados que serão utilizados ao longo de todo a dissertação.

Inicialmente, na Seção 1.1, introduzimos algumas terminologias e conceitos básicos envolvendo ordens de magnitude infinitesimais, inversa generalizada, função quantil, distribuição de Pareto, dentre outras.

Na Seção 1.2 apresentamos o conceito de função de variação regular e suas principais propriedades básicas. Uma síntese da teoria clássica dos valores extremos, com os resultados de interesse para esta dissertação, é apresentada na Seção 1.3. Por fim, na Seção 1.4 apresentamos o conceito e propriedades das estatísticas de ordem, especialmente a conhecida representação de Renyí e resultados sobre o comportamento assintótico das estatísticas de ordem, que serão importantes para o desenvolvimento do Capítulo 3.

As principais referências utilizadas neste capítulo são Cordeiro (1999), Embrechts e Hofert (2013), Resnick (1987), de Haan e Ferreira (2006), Bingham, Goldie e Teugels (1989), Rohatgi e Saleh (2001), Rényi (1953) e Seneta (1976).

1.1 Terminologia e Conceitos Básicos

Nesta Seção apresentamos diversos conceitos e resultados gerais básicos que serão utilizadas ao longo de toda a dissertação.

Iniciamos com definições das relações assintóticas de equivalência e ordem de magnitude infinitesimais.

Definição 1.1. Dizemos que uma função a valores reais f definida no intervalo (c, ∞)

para algum $c > 0$ é *assintoticamente equivalente* (no infinito) a outra função g e denotamos $f(x) \sim g(x)$ quando $x \to \infty$ se

$$\frac{f(x)}{g(x)} \to 1, \text{ quando } x \to \infty.$$

Definição 1.2. Sejam sequências de constantes $\{b_n\}$ e $\{c_n\}$. Dizemos que $\{b_n\}$ e $\{c_n\}$ são da *mesma ordem de magnitude*, quando $n \to \infty$, e denotamos por $b_n = o(c_n)$, se

$$\frac{b_n}{c_n} \to 0, \text{ quando } n \to \infty.$$

A ordem de magnitude acima é naturalmente generalizada para variáveis aleatórias, conforme descrevemos a seguir.

Definição 1.3. Sejam $\{Y_n\}$ uma sequência de variáveis aleatórias definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e $\{b_n\}$ uma sequência de constantes reais. Denotamos $Y_n = o_P(b_n)$ se

$$\frac{Y_n}{b_n} \xrightarrow{P} 0,$$

em que $\xrightarrow{P}$ indica convergência em probabilidade.

Na proposição a seguir, listamos algumas propriedades básicas da relação o_P.

Proposição 1.1. Sejam $X, \{X_n\}$ e $\{Y_n\}$ variáveis aleatórias definidas em um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$, $\{b_n\}$ e c constantes reais.

(a) $X_n \xrightarrow{P} X$ se, e só se, $X_n = X + o_P(1)$.

(b) Suponha que $Y_n = o_P(b_n)$ e $X_n = o_P(b_n)$, então

(b.1) $c \cdot Y_n = o_p(b_n)$.

(b.2) $a_n \cdot Y_n = o_P(a_n \cdot b_n)$

(b.3) $X_n + Y_n = o_p(b_n)$ e $X_n - Y_n = o_p(b_n)$.

Demonstração. (a), (b.1) e (b.2) seguem diretamente da definição de convergência em probabilidade.

(b.3) Das propriedades de convergência em distribuição (Teorema de Slutsky), como $\frac{X_n}{b_n} \xrightarrow{P} 0$ e $\frac{Y_n}{b_n} \xrightarrow{P} 0$, segue que

$$\frac{X_n \pm Y_n}{b_n} = \frac{X_n}{b_n} \pm \frac{Y_n}{b_n} \xrightarrow{D} 0,$$

em que $\xrightarrow{D}$ indica convergência em distribuição. Agora, como $Z_n \xrightarrow{D} c$ implica $Z_n \xrightarrow{P} c$, obtemos

$$\frac{X_n \pm Y_n}{b_n} \xrightarrow{P} 0,$$

ou seja, $X_n + Y_n = o_p(b_n)$ e $X_n - Y_n = o_p(b_n)$.

□

A seguir, apresentamos o conceito e propriedades básicas da inversa generalizada de funções não-decrescentes e, correspondentemente, a função quantil de uma função de distribuição de probabilidade.

Definição 1.4. Seja $f : \mathbb{R} \to \mathbb{R}$ uma função não-decrescente. A *inversa generalizada* $f^{\leftarrow} : \mathbb{R} \to \overline{\mathbb{R}}$ de f é definida por

$$f^{\leftarrow}(y) = \inf\{x \in \mathbb{R};\, f(x) \geq y\},\; y \in \mathbb{R},$$

com a convenção que $\inf \varnothing = \infty$.

Em especial, se $F : \mathbb{R} \to [0,1]$ é uma função de distribuição de probabilidade, então a inversa generalizada de F, $F^{\leftarrow} : [0,1] \to \overline{\mathbb{R}}$, é chamada de *função quantil* de F.

Definição 1.5. Dada uma função de distribuição F, chamamos de *função quantil caudal* à função definida por

$$U(x) = \inf\left\{y \in \mathbb{R};\, F(y) \geq 1 - \frac{1}{x}\right\} = F^{\leftarrow}\left(1 - \frac{1}{x}\right) = \left(\frac{1}{1-F}\right)^{\leftarrow}(x),\;\; x > 1. \quad (1.1)$$

Na próxima proposição listamos algumas das propriedades básicas da inversa generalizada que são de interesse para o nosso trabalho.

Proposição 1.2. Seja $f : \mathbb{R} \to \mathbb{R}$ uma função não-decrescente e $f^{\leftarrow}$ sua respectiva inversa generalizada.

(a) $f^{\leftarrow}$ é contínua pela esquerda.

(b) Se f é contínua pela direita, então

(b.1) o conjunto $A(y) = \{s : f(s) \geq y\}$ é fechado

(b.2) $f(f^{\leftarrow}(y)) \geq y$

(b.3) $f^{\leftarrow}(y) \leq t$ se, e só se, $y \leq f(t)$ e $t < f^{\leftarrow}(y)$ se, e só se, $y > f(t)$.

(c) Se f é inversível, então a inversa de f é $f^{\leftarrow}$.

Demonstração. Vide: Resnick (1987) - Seção 0.2. □

Como consequência direta da proposição anterior, obtemos as seguintes propriedades da função quantil.

Proposição 1.3. Seja F uma função de distribuição de probabilidade.

(a) Se U é uma variável aleatória com distribuição Uniforme sobre o intervalo $(0,1)$, então $F^{\leftarrow}(U)$ é uma variável aleatória com função de distribuição F.

(b) Se F é contínua, então $F(F^{\leftarrow}(x)) = x$.

As distribuições do tipo Pareto, que definimos a seguir, são utilizadas para modelar fenômenos de caudas pesadas em diversas aplicações no ramos de atuária (seguros não vida), finanças (retorno de ações), telecomunicações (tamanhos de arquivos, tempos de espera), geologia (terremotos), dentre outros.

Definição 1.6. Uma variável aleatória X tem distribuição Pareto se sua função de distribuição acumulada é dada por

$$F(x) = \begin{cases} 0, & \text{se } x < k \\ 1 - \left(\dfrac{k}{x}\right)^{\alpha}, & \text{se } x \geq k, \end{cases}$$

em que $k, \alpha > 0$. Notação: $X \sim \text{Pareto}(\alpha, k)$.

Dizemos que G é uma *função de distribuição do tipo Pareto* se $G(x) = F(A + Bx)$, em que F é a f.d. Pareto e $A, B \in \mathbb{R}$ e $B > 0$.

Exemplo 1.1. São exemplos de distribuição do tipo Pareto:

(a) *Frechet.* Para $\sigma > 0$ e $\lambda > 0$ a função de densidade da distribuição Frechet é dada por:

$$f_{\sigma,\lambda}(x) = \lambda \sigma^{\lambda} x^{-(\lambda+1)} \exp\left[-\left(\frac{\sigma}{x}\right)^{\lambda}\right], \; x > 0;$$

(b) *F de Fisher-Snedecor (F_{v_1,v_2}).* Para $v_1 > 0$ e $v_2 > 0$ temos que:

$$f_{v_1,v_2}(x) = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} \frac{x^{\frac{v_1-2}{2}}}{\left(1 + \left(\frac{v_1}{v_2}\right)x\right)^{\frac{v_1+v_2}{2}}}, \; x > 0;$$

onde Γ é a função gama.

(c) *T-Student.*

$$f_v(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{x^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, \; x > 0;$$

em que v é o número de graus de liberdade e Γ é a função gama.

(d) *Burr Tipo XII.* Para $c > 0$ e $k > 0$ a função de densidade da distribuição Burr Tipo XII ou Burr é dada por:

$$f_{c,k}(x) = ck\frac{x^{c-1}}{(1+x^c)^{k+1}}, \; x > 0;$$

(e) *Pareto generalizada.* A função de densidade distribuição Generalizada de Pareto é denotada por:

$$f_{\xi,\sigma,\mu}(x) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-(1+1/\xi)}, & \text{para } 0 \leq x, \text{ se } \xi \neq 0 \\ \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\}, & \text{para } 0 \leq x \leq \frac{1}{|\xi|} \text{ se } \xi = 0. \end{cases}$$

Finalizamos esta Seção com uma propriedade de esperança condicional que será utilizada no Capítulo 2.

Proposição 1.4. Seja X uma variável aleatória com esperança finita e seja $x \in \mathbb{R}$ tal que $P(X > x) = 1 - F(x) > 0$. Então, para todo evento A tal que $P(A) = 1 - F(x)$, temos que

$$E(X|A) \leq E(X|X > x).$$

Demonstração. Seja um evento A tal que $P(A) = P(X > x) > 0$. Primeiramente, usando as propriedades básicas de esperança condicional, podemos escrever

$$E(X|X > x) \;\; = \;\; x + \frac{E[(X-x)I_{(X>x,A)}]}{P(X > x)} + \frac{E[(X-x)I_{(X>x,A^c)}]}{P(X > x)} \qquad (1.2)$$

Agora, por um lado, se $P(X > x, A) > 0$, então

$$E(X|X > x) \geq x + E(X - x|X > x, A)P(A|X > x).$$

Mas, como $P(A) = P(X > x) > 0$, temos que $P(A \mid X > x) = P(X > x \mid A)$ e, consequentemente,

$$\begin{aligned}
E(X|X>x) &\geq x + E(X-x|X>x,A)\cdot P(X>x|A) \\
&\geq x + E(X-x|X>x,A)P(X>x|A) + E(X-x|X\leq x,A)P(X\leq x|A) \\
&= x + \frac{E[(X-x)I_{(X>x,A)}]}{P(A)} + \frac{E[(X-x)I_{(X\leq x,A)}]}{P(A)} \\
&= x + E(X-x|A) = E(X|A).
\end{aligned}$$

Por outro lado, se $P(X > x, A) = 0$, então $P(A^c, X > x) = P(X > x) = P(A) > 0$. Convencionando que: para $P(C) = 0$, $P(B|C) = P(B)$, para todo evento B, temos definida $E(X \mid C) = E(X)$. Dessa forma, o resultado segue seguindo o mesmo raciocínio anterior. □

1.2 Variação Regular

A teoria das funções de variação regular ou regularmente variantes é uma ferramenta fundamental para o estudo de valores extremos, de cauda pesada e domínios de atração. Nesta Seção apresentamos os principais conceitos e resultados dessa teoria que serão úteis para o desenvolvimento dos próximos capítulos. Para maiores detalhes, veja por exemplo, de Haan e Ferreira (2006), Bingham, Goldie e Teugels (1989) ou Seneta (1976).

Definição 1.7. Uma função mensurável a Lebesgue $f : \mathbb{R}^+ \to \mathbb{R}$ é *regularmente variante*, ou *de variação regular, no infinito*, se para algum $\alpha \in \mathbb{R}$

$$\lim_{t\to\infty} \frac{f(tx)}{f(t)} = x^{\alpha},\ x > 0. \tag{1.3}$$

Neste caso, denotamos $f \in RV_\alpha$ e α é chamado de *índice de variação regular.* Se $\alpha = 0$ então dizemos que f é *de variação lenta* ou *lentamente variante.*

Teorema 1.1. Se $f \in RV_\alpha$, então a relação (1.3) é válida uniformemente para $x \in [a, b]$, com $0 < a < b < \infty$.

Demonstração. Veja de Haan e Ferreira (2006) Teorema B.1.4. □

Teorema 1.2. (**Teorema de Karamata**) Suponha $f \in RV_\alpha$. Então existe $t_0 > 0$ tal que $f(t)$ é positiva e localmente limitada para $t \geq t_0$. Se $\alpha \geq -1$ então

$$\lim_{t\to\infty} \frac{t\cdot f(t)}{\int_{t_0}^{t} f(s)\mathrm{d}s} = \alpha + 1. \tag{1.4}$$

Se $\alpha \leq -1$ e $\int_0^\infty f(s)\mathrm{d}s < \infty$, então

$$\lim_{t\to\infty} \frac{t \cdot f(t)}{\int_t^\infty f(s)\mathrm{d}s} = -\alpha - 1. \tag{1.5}$$

Reciprocamente, se (1.4) vale com $-1 < \alpha < \infty$, então $f \in RV_\alpha$; se (1.5) vale com $-\infty < \alpha < -1$, então $f \in RV_\alpha$.

Demonstração. Veja de Haan e Ferreira (2006) Teorema B.1.5. □

Teorema 1.3. (**Teorema da Representação**) Se $f \in RV_\alpha$, existem funções mensuráveis $a : \mathbb{R}^+ \to \mathbb{R}$ e $c : \mathbb{R}^+ \to \mathbb{R}$, com

$$\lim_{t\to\infty} c(t) = c_0,\ (0 < c_0 < \infty) \text{ e } \lim_{t\to\infty} a(t) = 0 \tag{1.6}$$

e $t_0 \in \mathbb{R}^+$ tal que para $t > t_0$,

$$f(t) = t^\alpha c(t) \cdot \exp\left(\int_{t_0}^t \frac{a(s)}{s}\mathrm{d}s\right), \tag{1.7}$$

ou seja,

$$f(t) = t^\alpha l(t), \tag{1.8}$$

em que $l \in RV_0$.

Reciprocamente, se (1.7) vale com a e c satisfazendo (1.6) então $f \in RV_\alpha$.

Demonstração. Veja de Haan e Ferreira (2006) Teorema B.1.6 □

Observação 1.1. Na fórmula (1.7) podemos considerar $t_0 \in [0, \infty)$ arbitrário, bastando alterar adequadamente as funções $c(t)$ e $a(t)$.

Demonstração. Veja Seneta (1976) Lema 1.4. □

Exemplo 1.2. As distribuições do tipo Pareto, dadas na Definição 1.6, possuem cauda de variação regular. Pelo Teorema 1.3, se F é uma função de distribuição do tipo Pareto, então pode ser representada por

$$1 - F(x) = x^{-1/\gamma} l(x),\ x > 0, \tag{1.9}$$

em que $\gamma > 0$ e $l \in RV_0$.

Finalizamos esta Seção, apresentando na proposição a seguir uma lista de algumas propriedades básicas das funções de variação regular.

Proposição 1.5. (a) Se $f \in RV_\alpha$, então $(\ln f(t))/(\ln t) \xrightarrow{t\to\infty} \alpha$. Isso implica que

$$\lim_{t\to\infty} f(t) = \begin{cases} 0, & \text{se } \alpha < 0 \\ \infty, & \text{se } \alpha > 0. \end{cases}$$

(b) Se $f_1 \in RV_{\alpha_1}$, $f_2 \in RV_{\alpha_2}$, então $f_1 + f_2 \in RV_{\max(\alpha_1,\alpha_2)}$

Se além disso, $\lim_{t\to\infty} f_2(t) = \infty$ então $f_1 \circ f_2 \in RV_{\alpha_1\alpha_2}$.

(c) Se $f \in RV_\alpha$ com $\alpha > 0$ (ou $\alpha < 0$) então f é assintoticamente equivalente a uma função diferenciável estritamente crescente (ou decrescente) g com derivada $g' \in RV_{\alpha-1}$ se $\alpha > 0$ e $-g$ com derivada $g' \in RV_{\alpha-1}$ se $\alpha < 0$.

(d) Se $f \in RV_\alpha$ e $g \in RV_\beta$ então $f \cdot g \in RV_{\alpha+\beta}$.

(e) Se $f \in RV_\alpha$, com $\alpha > 0$, e é limitada sobre intervalos finitos de $\mathbb{R}^+$, então $f^{\leftarrow} \in RV_{1/\alpha}$. Em particular, se $f \in RV_\alpha$, $\alpha > 0$, e f é estritamente crescente, então a função inversa $f^{-1} \in RV_{1/\alpha}$.

(f) (**Desigualdade de Potter**, 1942) Suponha que $f \in RV_\alpha$. Se $\delta_1, \delta_2 > 0$ são arbitrários, então existe $t_0 = t_0(\delta_1, \delta_2)$ tal que para $t \geq t_0$, $tx \geq t_0$,

$$(1-\delta_1)x^\alpha \min(x^{\delta_2}, x^{-\delta_2}) < \frac{f(tx)}{f(t)} < (1+\delta_1)x^\alpha \max(x^{\delta_2}, x^{-\delta_2}).$$

(g) Se $f \in RV_\rho$, com $\rho < 0$, e $\{a_n\}$ e $\{b_n\}$ são sequências de constantes tais que $a_n \to \infty$, $b_n \to \infty$ e $\frac{a_n}{b_n} \to \infty$, quando $n \to \infty$, então $\frac{f(a_n)}{f(b_n)} \to 0$, quando $n \to \infty$.

Demonstração. A demonstração das propriedades (a)-(f) podem ser encontradas em de Haan e Ferreira (2006), Seneta (1976) ou Bingham, Goldie e Teugels (1989). A propriedade (g) segue de (f). $\square$

1.3 Teoria dos Valores Extremos

A teoria dos valores extremos estuda o comportamento assintótico dos extremos amostrais $X_{n,n} = \max(X_1, X_2, ..., X_n)$ e $X_{1,n} = \min(X_1, X_2, ..., X_n)$, em que $X_1, X_2, ..., X_n$ são variáveis aleatórias i.i.d. com função de distribuição F.

Como $\min(X_1, X_2, ..., X_n) = -\max(-X_1, -X_2, ..., -X_n)$, o estudo sobre $X_{1,n}$ pode ser obtido a partir do estudo de $X_{n,n}$.

Nesta Seção, apresentamos uma síntese dos principais conceitos e resultados da teoria dos valores extremos, apresentados em de Haan e Ferreira (2006).

O objetivo central da teoria clássica dos valores extremos consiste em estudar o comportamento limite, quando $n \to \infty$, da distribuição do máximo $X_{(n,n)}$ estabilizado linearmente, ou seja, o comportamento limite de

$$\frac{\max(X_1, X_2, ..., X_n) - b_n}{a_n},$$

em que $\{a_n\}$ e $\{b_n\}$ são sequências de constantes reais, com $a_n > 0$.

Especificamente, o problema consiste em determinar condições para a existência de sequências de constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$$\lim_{n\to\infty} P\left(\frac{X_{n,n} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{1.10}$$

para cada ponto de continuidade x de G, e G é uma função de distribuição não degenerada. Essas distribuições limites são chamadas de *distribuições de valores extremos.* A classe de distribuições F satisfazendo (1.10) é chamada de *domínio de atração de* G e denotada por $D(G)$.

Vale observar que $X_{n,n}$ é um estimador natural para o ponto extremo superior da distribuição F, ou seja, $\omega(F) = sup\{x : F(x) < 1\}$ e $X_{n,n}$ converge quase certamente para $\omega(F)$. Assim, a obtenção de uma distribuição limite não degenerada em (1.10) fornece uma aproximação para a distribuição de $X_{n,n}$.

O teorema a seguir estabelece que as distribuições dos valores extremos podem ser representadas em uma única parametrização, obtida por Von Mises (1936) e Jenkinson (1955), também conhecida como *distribuição do valor extremo generalizada.*

Teorema 1.4. A classe de distribuições de valores extremos é $G_\gamma(ax + b)$ com $a > 0$, $b \in \mathbb{R}$, onde

$$G_\gamma(x) = \exp\left[-(1 + \gamma x)^{-1/\gamma}\right]. \tag{1.11}$$

Demonstração. Veja de Haan e Ferreira (2006), Teorema 1.1.3. □

O parâmetro γ em (1.11) é chamado de *índice de valor extremo* e determina o peso da cauda superior, com valores maiores indicando caudas mais pesadas.

A parametrização (1.11) inclui os três tipos básicos de distribuições do valor extremos, estabelecidas por Fisher e Tippett (1928) e Gnedenko (1943), conforme descrevemos a seguir.

(a) *Distribuição de Fréchet:* para $\gamma > 0$, use $G_\gamma((x-1)/\gamma)$ e considere $\alpha = 1/\gamma > 0$,

$$\Phi_\alpha(x) := \begin{cases} 0, & \text{se } x \leq 0 \\ \exp\left(-x^{-\alpha}\right), & \text{se } x > 0 \end{cases}$$

(b) *Distribuição de Gumbel:* para $\gamma = 0$,

$$G_0(x) = \exp\left(-\mathrm{e}^{-x}\right), \ \forall\, x \in \mathbb{R}.$$

(c) *Distribuição de Weibull (reversa):* para $\gamma < 0$ use $G_\gamma(-(1+x)/\gamma)$ e considere $\alpha = -1/\gamma > 0$,

$$\Psi_\alpha(x) := \begin{cases} \exp\left(-(-x)^{-\alpha}\right),, & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}.$$

A parametrização (1.11) possibilita a obtenção de expressões equivalentes à (1.10), descritas no teorema a seguir.

Teorema 1.5. Seja F uma função de distribuição e considere $U = \left(\dfrac{1}{1-F}\right)^{\leftarrow}$. Para $\gamma \in \mathbb{R}$ as seguintes afirmações são equivalentes:

(a) Existem constantes reais $a_n > 0$ e b_n, tais que

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x) = \exp\left(-(1+\gamma x)^{-1/\gamma}\right),$$

para todo x com $1 + \gamma x > 0$.

(b) Existe uma função positiva a tal que para $x > 0$,

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = D_\gamma(x) = \frac{x^\gamma - 1}{\gamma},$$

em que, para $\gamma = 0$ o lado direito é interpretado como $\log x$.

(c) Existe uma função positiva a, tal que

$$\lim_{t\to\infty} t\left(1 - F\left(a(t)x + U(t)\right)\right) = (1+\gamma x)^{-1/\gamma},$$

para todo x com $1 + \gamma x > 0$.

(d) Existe uma função positiva f, tal que

$$\lim_{t\uparrow x^*} \frac{1-F(t+xf(t))}{1-F(t)} = (1+\gamma x)^{-1/\gamma}$$

para todo x para o qual $1+\gamma x > 0$, onde $x^* = \sup\{x : F(x) < 1\}$.

Demonstração. Veja de Haan e Ferreira (2006) Teorema 1.1.6..

□

Uma questão estudada na teoria de valores extremos é estabelecer condições suficientes para que uma função F pertença ao domínio de atração de uma distribuição do valor extremo $G = G_\gamma$ para algum $\gamma \in \mathbb{R}$, que denotamos $F \in D(G)$, ou seja, para que F satisfaça a relação (1.10).

Vamos destacar os resultados referentes a duas dessas condições suficientes estabelecidas na teoria dos valores extremos, que são estudadas em de Haan e Ferreira (2006). Outros resultados com condições alternativas podem ser encontrados em de Haan e Ferreira (2006) e Resnick (1987), entre outros.

Iniciamos com o resultado referente à conhecida por *condição de Von Mises.*

Teorema 1.6. Seja F uma função de distribuição acumulada e x^* seu ponto extremo superior. Suponha que $F''(x)$ exista e $F'(x)$ seja positivo para todo x em alguma vizinhança esquerda de x^*. Se

$$\lim_{t\uparrow x^*} \left(\frac{1-F}{F'}\right)'(t) = \gamma, \tag{1.12}$$

ou, equivalentemente,

$$\lim_{t\uparrow x^*} \frac{(1-F(t))F''(t)}{(F'(t))^2} = -\gamma - 1, \tag{1.13}$$

então F está no domínio de atração de G_γ.

Demonstração. Veja de Haan e Ferreira (2006), Teorema 1.1.8. □

Apresentamos a seguir uma outra condição, conhecida como condição de segunda ordem, que será de fundamental importância especialmente para o desenvolvimento do Capítulo 3. Apresentaremos somente os resultados que serão de interesse para o desenvolvimento do restante desta dissertação. Um estudo mais detalhado pode ser encontrado em de Haan e Ferreira (2006).

Definição 1.8. Dizemos que a função de distribuição F, ou a função $U = \left(\dfrac{1}{1-F}\right)^{\leftarrow}$, satisfaz a *condição de segunda ordem* (SOC) se, para alguma função positiva a e para alguma função positiva ou negativa A, com $\lim_{t\to\infty} A(t) = 0$,

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^{\gamma}-1}{\gamma}}{A(t)} =: \; H(x),\ x > 0 \tag{1.14}$$

em que H é alguma função que não é múltipla da função $(x^{\gamma}-1)/\gamma$. Em particular, H não deve ser identicamente zero. As funções a e A são comumente chamadas *funções auxiliares* de primeira e segunda ordem, respectivamente.

Esta condição está relacionada com a condição apresentada no item (b) do **Teorema 1.5** e é possível provar que se F satisfaz a condição de segunda ordem, então $F \in D(G_{\gamma})$, para algum γ.

As distribuições do tipo Pareto, como as distribuições de Fréchet, F_{v_1,v_2}, t-Student, Burr e Pareto generalizada, são exemplos de distribuições que satisfazem a condição de segunda ordem.

Vale destacar que a função auxiliar A da condição SOC é de variação regular, conforme estabelece o teorema abaixo.

Teorema 1.7. Suponha que a condição de segunda ordem (1.14) seja válida. Então existem constantes $c_1, c_2 \in \mathbb{R}$ e algum parâmetro $\rho \leq 0$ tal que

$$H(x) = c_1 \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} \mathrm{d}u \mathrm{d}s + c_2 \int_1^x s^{\gamma+\rho-1} \mathrm{d}s.$$

Além disso, para $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{a(tx)}{a(t)} - x^{\gamma}}{A(t)} = c_1 x^{\gamma} \frac{x^{\rho}-1}{\rho}$$

e

$$\lim_{t\to\infty} \frac{A(tx)}{A(t)} = x^{\rho}.$$

Demonstração. Veja de Haan e Ferreira (2006), Teorema 2.3.3. □

Dentre os vários resultados relacionados com a condição de segunda ordem, enunciamos a seguir um teorema para o caso em que $\gamma > 0$, que é de nosso particular interesse.

Teorema 1.8. Suponha que para algum γ positivo e para alguma função positiva ou negativa A,

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^{\gamma}}{A(t)} := K(x)$$

existe para todo $x > 0$ e K não é identicamente nulo. Então, para uma função possivelmente diferente A, positiva ou negativa,

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho} \tag{1.15}$$

para todo $x > 0$ com $\rho \leq 0$. Além disso, para qualquer $\varepsilon, \delta > 0$ existe $t_0 = t_0(\varepsilon, \delta) > 1$ tal que para todo $t, tx \geq t_0$,

$$\left| \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A_0(t)} - x^\gamma \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^{\gamma+\rho} \max(x^\gamma, x^{-\gamma}) \tag{1.16}$$

com

$$A_0(t) := \begin{cases} \rho \left(1 - \frac{\lim_{s\to\infty} s^{-\gamma} U(s)}{t^{-\gamma} U(t)}\right), & \text{se } \rho < 0 \\ 1 - \frac{\int_0^t s^{-\gamma} U(s) \mathrm{d}s}{t^{1-\gamma} U(t)}, & \text{se } \rho = 0. \end{cases}$$

A relação (1.15) é equivalente a

$$\lim_{t\to\infty} \frac{\frac{1-F(tx)}{1-F(t)} - x^{-1/\gamma}}{\alpha(t)} = x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\gamma\rho} \tag{1.17}$$

para todo $x > 0$ com $\rho \leq 0$ e $\alpha(t) := A\left(\frac{1}{1-F(t)}\right)$.

Finalizamos a Seção com uma proposição que estabelece uma propriedade assintótica da cauda de uma distribuição satisfazendo a condição (SOC), com $\gamma > 0$ e que será de importância para a obtenção dos resultados do Capítulo 3.

Proposição 1.6. Se X é uma variável aleatória com função de distribuição F satisfazendo a condição (SOC) com $0 < \gamma < 1/\beta$, então

$$\lim_{u\to\infty} u^\beta \cdot \overline{F}(u \cdot U(1/p)) = 0.$$

Demonstração. Primeiramente, podemos escrever:

$$\begin{aligned} u^\beta \cdot \overline{F}(u \cdot U(1/p)) &= \frac{u^\beta \cdot \overline{F}(u) \cdot \overline{F}(u \cdot U(1/p))}{\overline{F}(u)} \\ &= u^\beta \cdot \overline{F}(u) \cdot \left[\frac{\overline{F}(u \cdot U(1/p))}{\overline{F}(u)} - [U(1/p)]^{-1/\gamma} + [U(1/p)]^{-1/\gamma} \right] \\ &= u^\beta \cdot \overline{F}(u) \cdot \left\{ \left[\frac{\frac{\overline{F}(u \cdot U(1/p))}{\overline{F}(u)} - [U(1/p)]^{-1/\gamma}}{A(1/\overline{F}(u))} \right] \cdot A(1/\overline{F}(u)) + [U(1/p)]^{-1/\gamma} \right\}. \end{aligned}$$

Por um lado, pelo teorema anterior, a condição (SOC) com $\gamma > 0$ é equivalente a (1.17) e como $\lim_{t\to\infty} A(t) = 0$, obtemos que

$$\left[\frac{\frac{\overline{F}(u\cdot U(1/p))}{\overline{F}(u)} - [U(1/p)]^{-1/\gamma}}{A(1/\overline{F}(u))}\right] \cdot A(1/\overline{F}(u)) + [U(1/p)]^{-1/\gamma} \xrightarrow{u\to\infty} k, \text{ para algum } k \in \mathbb{R}.$$

Por outro lado, $u^\beta \in RV_\beta$ e $\overline{F(u)} \in RV_{-1/\gamma}$ (pela condição (SOC)), então, pelo item (d) da Proposição 1.5, segue que

$$u^\beta \cdot \overline{F(u)} \in RV_{\beta-1/\gamma}.$$

Mas, como $\beta < 1/\gamma$ temos, pelo item (a) da Proposição 1.5 que $\lim_{u\to\infty} u^\beta \cdot \overline{F}(u) = 0$. Portanto, $\lim_{u\to\infty} u^\beta \cdot \overline{F}(u \cdot U(1/p)) = 0$. □

1.4 Estatísticas de Ordem Intermediárias

As estatísticas de ordem estão entre as ferramentas fundamentais da estatística. Em particular, elas serão a base para a definição e derivação das propriedades assintóticas do estimador que é o objeto de estudo desta dissertação.

Definição 1.9. Seja uma amostra aleatória $(X_1, \ldots, X_n)$, ou seja, $X_1, \ldots, X_n$ v.a.'s independentes e identicamente distribuídas. As *estatísticas de ordem* da amostra aleatória $X_1, X_2, ..., X_n$ são os valores amostrais colocados em ordem crescente e denotados por $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$.

Em outras palavras, as estatísticas de ordem são as v.a.'s dadas por

$$\begin{aligned}
X_{1,n} &= \min(X_1, \ldots, X_n) = \text{ menor valor entre } X_1, \ldots, X_n \\
X_{2,n} &= \text{ segundo menor valor entre } X_1, \ldots, X_n \\
&\vdots \\
X_{k,n} &= k - \text{ésimo menor valor entre } X_1, \ldots, X_n \\
&\vdots \\
X_{n,n} &= \max(X_1, \ldots, X_n) = \text{ maior valor entre } X_1, \ldots, X_n.
\end{aligned}$$

As estatísticas $X_{2,n}, \ldots, X_{n-1,n}$ são também chamadas *estatísticas de ordem intermediárias.*

Apresentamos inicialmente resultados básicos sobre as distribuições conjuntas das estatísticas de ordem de variáveis aleatórias absolutamente contínuas.

Proposição 1.7. Sejam $X_1, X_2, ..., X_n$ variáveis aleatórias i.i.d. e com densidade f. Se

$X_{1,n}, X_{2,n}, ..., X_{n,n}$ são as suas respectivas estatísticas de ordem, então

(a) a densidade conjunta de $X_{1,n}, X_{2,n}, ..., X_{n,n}$ é dada por

$$f_{X_{1,n},...,X_{n,n}}(y_1, y_2, \dots, y_n) = \begin{cases} n! \prod_{i=1}^{n} f(y_i) & , \text{ se } y_1 < y_2 < \dots < y_n \\ 0 & , \text{ caso contrário;} \end{cases}$$

(b) a densidade marginal de $X_{k,n}$, para cada $k = 1, 2, \dots, n$ é dada por

$$f_{X_{k,n}}(y_k) = \frac{n!}{(k-1)!(n-k)!}[F(y_k)]^{k-1}[1 - F(y_k)]^{n-k} f(y_k),$$

em que F é a função de distribuição comum de $X_1, X_2, ..., X_n$;

(c) a densidade conjunta de $X_{j,n}$ e $X_{k,n}$, $1 \leq j < k \leq n$, é dada por

$$f_{X_{j,n},X_{k,n}}(y_j, y_k) = \begin{cases} \frac{n!}{(j-1)!(k-j-1)!(n-k)!}[F(y_j)]^{j-1}[F(y_k) - F(y_j)]^{k-j-1}\cdot & \\ \cdot[1 - f(y_k]^{n-k} f(y_j) f(y_k) & , \text{ se } y_j < y_k \\ 0 & , \text{ caso contrário.} \end{cases}$$

Demonstração. Veja Rohatgi e Saleh (2001), Seção 4.7. □

Descrevemos no próximo teorema um método, apresentado por Rényi (1953), para reescrever uma estatística de ordem como uma função da soma de variáveis aleatórias mutuamente independentes, que simplifica o tratamento das estatísticas de ordem.

Primeiramente, consideramos o caso particular das estatísticas de ordem associadas a uma amostra aleatória da distribuição exponencial, que será útil para a obtenção do resultado no caso geral.

Proposição 1.8. Sejam $\zeta_1, \zeta_2, ..., \zeta_n$ variáveis aleatórias i.i.d. com distribuição Exponencial(λ) e $\zeta_{1,n}, \zeta_{2,n}, ..., \zeta_{n,n}$ suas estatísticas de ordem. Então, para $k = 1, \dots, n-1$, as diferenças $\zeta_{k+1,n} - \zeta_{k,n}$ são variáveis aleatórias independentes e com distribuição exponencial($(n - k)\lambda$), respectivamente. Além disso, as estatísticas de ordem $\zeta_{k,n}$ podem ser expressas como

$$\zeta_{k,n} = \frac{\delta_1}{n} + \frac{\delta_2}{n-1} + ... + \frac{\delta_k}{n-k+1}, \; k = 1, 2, ..., n, \tag{1.18}$$

em que $\delta_1, \delta_2, ..., \delta_n$ são i.i.d. com distribuição Exponencial(λ).

Demonstração. As distribuições e a independência das diferenças $\zeta_{k+1,n} - \zeta_{k,n}$, podem ser obtidas usando os resultados da Proposição 1.7 para a distribuição exponencial. Assim, segue que as variáveis

$$\delta_{k+1} := (n-k)(\zeta_{k+1,n} - \zeta_{k,n}),\ k = 1, \ldots, n$$

são i.i.d. com distribuição exponencial(λ). Portanto, as estatísticas de ordem $\zeta_{k,n}$ podem ser expressas como em (1.18) $\square$

Teorema 1.9. (Representação de Rényi) Seja ξ uma variável aleatória contínua com função de distribuição acumulada $F(x)$ estritamente crescente no intervalo $(\alpha(F), \omega(F))$, em que $\alpha(F)$ e $\omega(F)$ são, respectivamente, os pontos extremos inferior e superior de F. Considere $(\xi_1, \xi_2, ..., \xi_n)$ uma amostra aleatória de tamanho n de ξ. Então, para cada $k = 1, 2, \ldots, n$ a estatística de ordem $\xi_{k,n}$ pode ser representada por

$$\xi_{k,n} \quad = \quad F^{-1}\left(\mathrm{e}^{-\left(\frac{\delta_1}{n}+\frac{\delta_2}{n-1}+\ldots+\frac{\delta_{n+1-k}}{k}\right)}\right), \tag{1.19}$$

em que as variáveis aleatórias $\delta_1, \delta_2, ..., \delta_n$ são i.i.d. com distribuição Exponencial(1).

Demonstração. Defina as variáveis aleatórias auxiliares

$$\eta_k = F_\xi(\xi_k) \text{ e } \zeta_k = \log\frac{1}{\eta_k}. \tag{1.20}$$

Como $\log\frac{1}{x}$ é uma função decrescente, temos que as respectivas estatísticas de ordem são dadas por

$$\eta_{k,n} = F_\xi(\xi_{k,n}) \text{ e } \zeta_{k,n} = \log\frac{1}{\eta_{n+1-k,n}}.$$

Por um lado, como F é estritamente crescente na região $\{x : 0 < F(x) < 1\}$, segue que, para $0 < x < 1$

$$P(\eta_k < x) = P(\xi_k < F^{-1}(x)) = F(F^{-1}(x)) = x, \tag{1.21}$$

ou seja, as variáveis η_k são uniformemente distribuídas sobre o intervalo $(0, 1)$. Além disso, como as v.a.'s ξ_k são independentes, segue que as variáveis $\eta_k,\ k \geq 1$, também são independentes.

Por outro lado, para $x \geq 0$, temos

$$\begin{aligned} P(\zeta_k < x) &= P\left(F(\xi_k) > \mathrm{e}^{-x}\right) \\ &= P(\xi_k > F^{-1}(\mathrm{e}^{-x})) \\ &= 1 - F(F^{-1}(\mathrm{e}^{-x})) \\ &= 1 - \mathrm{e}^{-x}, \end{aligned}$$

pois, $0 < \mathrm{e}^{-x} < 1$, para $x > 0$, e, por hipótese, F é estritamente crescente na região $\{x : 0 < F(x) < 1\}$.

Logo, as variáveis aleatórias $\zeta_1, ..., \zeta_n$ são i.i.d. com distribuição exponencial(1). Desta forma, pela Proposição 1.8, as estatísticas de ordem $\zeta_{k,n}$ podem ser representadas por

$$\zeta_{k,n} = \frac{\delta_1}{n} + \frac{\delta_2}{n-1} + ... + \frac{\delta_k}{n-k+1},\ k = 1, 2, ..., n, \tag{1.22}$$

em que $\delta_1, \delta_2, ..., \delta_n$ são i.i.d. com distribuição Exponencial(1).

Agora, por (1.20), podemos escrever

$$\begin{aligned} \xi_{k,n} &= F^{-1}(\eta_{k,n}) \\ &= F^{-1}\left(\mathrm{e}^{-\zeta_{n+1-k,n}}\right) \end{aligned} \tag{1.23}$$

e, usando (1.22), obtemos (1.19). □

Usando a representação de Rényi, obtemos o seguinte resultado sobre estatísticas de ordem da distribuição $Pareto(1,1)$, que será útil para o estudo das propriedades assintóticas do estimador a ser estudado no Capítulo 3.

Proposição 1.9. Sejam $Y_1, Y_2, ..., Y_n$ variáveis aleatórias i.i.d. com distribuição Pareto$(1,1)$ com estatísticas de ordem $Y_{1,n}, Y_{2,n}, ..., Y_{n,n}$, então para $\alpha \in \mathbb{R}$

$$\sum_{j=1}^{k} \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} \stackrel{D}{=} \sum_{j=1}^{k} Y_j^{\alpha},$$

em que $\stackrel{D}{=}$ indica igualdade em distribuição.

Demonstração. Por (1.23) temos:

$$\sum_{j=1}^{k} \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} = \sum_{j=1}^{k} \left(\frac{F_Y^{-1}(\mathrm{e}^{-\zeta_{j,n}})}{F_Y^{-1}(\mathrm{e}^{-\zeta_{k+1,n}})}\right)^{\alpha}$$

onde $\zeta_{j,n} := \ln\left(\frac{1}{\eta_{k_1-j,n}}\right)$ e $\eta_{k,n} := F_Y(Y_{k,n})$. Como $Y \sim$ Pareto$(1,1)$, temos que $F_Y^{-1}(x) = \dfrac{1}{1-x}$ e, consequentemente,

$$\begin{aligned} \sum_{j=1}^{k} \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} &= \sum_{j=1}^{k} \left(\frac{1-\mathrm{e}^{-\zeta_{k+1,n}}}{1-\mathrm{e}^{-\zeta_{j,n}}}\right)^{\alpha} \\ &= \sum_{j=1}^{k} \left(\frac{1-\eta_{n-k,n}}{1-\eta_{n+1-j,n}}\right)^{\alpha}. \end{aligned}$$

Mas, por (1.21), temos que as variáveis $\eta_k, k \geq 1$ têm distribuição uniforme no intervalo $(0,1)$. Logo, $1-\eta_k,\ k=1,2,...$, também possuem distribuição uniforme no intervalo (0,1) e, além disso, são independentes por serem funções de Y_k. Assim, definindo $1-\eta_k =: \tilde{\eta}_k$ e $1-\eta_{k,n} = \tilde{\eta}_{n-k+1,n}$ podemos escrever

$$\begin{aligned}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} &= \sum_{j=1}^{k}\left(\frac{\tilde{\eta}_{k+1,n}}{\tilde{\eta}_{j,n}}\right)^{\alpha}\\ &= \sum_{j=1}^{k}\left(\frac{F_{\tilde{\eta}}(\tilde{\eta}_{k+1,n})}{F_{\tilde{\eta}}(\tilde{\eta}_{j,n})}\right)^{\alpha}.\end{aligned}$$

Agora, utilizando (1.19) obtemos

$$\begin{aligned}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} &= \sum_{j=1}^{k}\left(\frac{F_{\tilde{\eta}}\left(F_{\tilde{\eta}}^{-1}\left(\mathrm{e}^{-\left(\frac{\tilde{\delta}_1}{n}+...+\frac{\tilde{\delta}_{n-k}}{k+1}\right)}\right)\right)}{F_{\tilde{\eta}}\left(F_{\tilde{\eta}}^{-1}\left(\mathrm{e}^{-\left(\frac{\tilde{\delta}_1}{n}+...+\frac{\tilde{\delta}_{n-j+1}}{j}\right)}\right)\right)}\right)^{\alpha}\\ &= \sum_{j=1}^{k}\left(\frac{\mathrm{e}^{-\left(\frac{\tilde{\delta}_1}{n}+...+\frac{\tilde{\delta}_{n-k}}{k+1}\right)}}{\mathrm{e}^{-\left(\frac{\tilde{\delta}_1}{n}+...+\frac{\tilde{\delta}_{n-j+1}}{j}\right)}}\right)^{\alpha}\\ &= \sum_{j=1}^{k}\left(\frac{1}{\mathrm{e}^{-\left(\frac{\tilde{\delta}_{n-k+1}}{k}+...+\frac{\tilde{\delta}_{n-j+1}}{j}\right)}}\right)^{\alpha},\end{aligned}$$

em que as variáveis aleatórias $\tilde{\delta}_1, \tilde{\delta}_2, ..., \tilde{\delta}_n$ são i.i.d com distribuição Exponencial(1). E como $Y \sim$ Pareto$(1,1)$, segue que

$$\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} = \sum_{j=1}^{k}\left(F_Y^{-1}\left(1-\mathrm{e}^{-\left(\frac{\tilde{\delta}_{n-k+1}}{k}+...+\frac{\tilde{\delta}_{n-j+1}}{j}\right)}\right)\right)^{\alpha}.$$

Por outro lado, como $\tilde{\eta}_k \sim$ Uniforme$(0,1)$, novamente por (1.19) obtemos que

$$\tilde{\eta}_{j,k} = F_{\tilde{\eta}}(\tilde{\eta}_{j,k}) = \mathrm{e}^{-\left(\frac{\tilde{\delta}_1}{k}+...+\frac{\tilde{\delta}_{k-j+1}}{j}\right)} \stackrel{D}{=} \mathrm{e}^{-\left(\frac{\tilde{\delta}_{n-k+1}}{k}+...+\frac{\tilde{\delta}_{n-j+1}}{j}\right)}.$$

Substituindo esse resultado em (1.24), chegamos que

$$\begin{aligned}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} &\stackrel{D}{=} \sum_{j=1}^{k}\left(F_Y^{-1}(1-\tilde{\eta}_{j,k})\right)^{\alpha}\\ &= \sum_{j=1}^{k}\left(F_Y^{-1}(\eta_{k-j+1,k})\right)^{\alpha}\\ &= \sum_{j=1}^{k}\left(F_Y^{-1}(\mathrm{e}^{-\zeta_{j,k}})\right)^{\alpha},\end{aligned}$$

pois $\zeta_{j,k} := \ln\left(\frac{1}{\eta_{k-j+1,k}}\right)$. Usando novamente (1.23), obtemos o resultado desejado

$$\begin{aligned}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\alpha} &\stackrel{D}{=} \sum_{j=1}^{k} Y_{k+1-j,k}^{\alpha}\\ &= \sum_{j=1}^{k} Y_j^{\alpha}.\end{aligned}$$

□

Nesta última parte desta Seção, apresentamos alguns resultados importantes sobre o comportamento assintótico das estatísticas de ordem intermediárias (ou seja, as estatísticas de ordem quando $k, n \to \infty$ e $p = k/n \to 0$) que estão relacionados com os resultados para as estatísticas de ordem extremais estudados pela Teoria dos Valores Extremos.

Iniciamos com a propriedade de Normalidade Assintótica.

Teorema 1.10. Suponha que a condição de Von Mises para o domínio de atração de uma distribuição de valores extremos G_γ seja válida. Se $k = k(n) \to \infty$, $k/n \to 0$, quando $n \to \infty$, então

$$\sqrt{k}\frac{X_{n-k,n} - U(n/k)}{\frac{n}{k}U'(n/k)} \xrightarrow{D} \mathcal{N}(0,1),$$

em que $\xrightarrow{D}$ indica convergência em distribuição.

Demonstração. Veja de Haan e Ferreira (2006), Teorema 2.2.1. □

Aplicando o teorema anterior para uma distribuição tipo Pareto(1,1), temos o seguinte resultado.

Corolário 1.1. Para a f.d. tipo Pareto(1,1) $F_Y(y) = 1-1/y$, $y \geq 1$, quando $n \to \infty$, $k \to$

∞, $k/n \to 0$, temos

$$\sqrt{k}\left(\frac{k}{n}Y_{n-k,n} - 1\right) \xrightarrow{D} \mathcal{N}(0,1).$$

No Teorema 1.10, a normalidade assintótica das estatísticas de ordem intermediárias é obtida sob a condição de Von Mises. O próximo teorema estabelece que sob a condição SOC, as estatísticas de ordem intermediárias são também assintoticamente normais.

Teorema 1.11. Sejam $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ as n estatísticas de ordem de uma amostra i.i.d. com função de distribuição F. Suponha que a condição de segunda ordem (SOC), dada em (1.14), é válida para algum $\gamma \in \mathbb{R}$, $\rho \leq 0$. Então,

$$\sqrt{k}\frac{X_{n-k,n} - U(n/k)}{a(n/k)} \xrightarrow{D} \mathcal{N}(0,1),$$

desde que a sequência $k = k(n)$ seja tal que $k(n) \to \infty$, $n \to \infty$, e $\sqrt{k}A(n/k) \to \lambda \in \mathbb{R}$. Além disso, é importante ressaltar que $a \in RV_\gamma$, pelo Teorema 1.7.

Demonstração. Veja de Haan e Ferreira (2006), Teorema 2.4.1. $\square$

Como consequência do Teorema 1.11, na proposição a seguir mostramos que $X_{n-k,n}$ pode ser usada como aproximação de $U(n/k)$, quando $n \to \infty$, $k \to \infty$, $k/n \to 0$.

Proposição 1.10. Seja X uma variável aleatória que satisfaça a condição (SOC) e $U = \left(\frac{1}{1-F}\right)^{\leftarrow}$ a função quantil caudal. Então,

$$\frac{X_{n-k,n}}{U(n/k)} \xrightarrow{P} 1, \text{ quando } n \to \infty,\, k \to \infty,\, k/n \to 0. \tag{1.24}$$

Demonstração. Pelo Teorema 1.11, temos que, para n suficientemente grande,

$$P\left(\frac{X_{n-k,n}}{U(n/k)} - 1 \leq x\right) \approx F_{N(0,1)}\left(x\sqrt{k}\cdot\frac{U(n/k)}{a(n/k)}\right).$$

Agora, como X satisfaz a condição (SOC), então $U \in RV_\gamma$ e, além disso, $a \in RV_\gamma$, pelo Teorema 1.7. Logo, pela Proposição 1.5, $\frac{U(x)}{a(x)} \in RV_0$. Consequentemente, utilizando a mesma proposição, é possível mostrar que $\sqrt{k}\cdot\frac{U(n/k)}{a(n/k)} \xrightarrow{n\to\infty} \infty$.

Assim, para n suficientemente grande

$$
\begin{aligned}
P\left(\frac{X_{n-k,n}}{U(n/k)}-1\leq x\right) &\approx F_{N(0,1)}\left(x\sqrt{k}\cdot\frac{U(n/k)}{a(n/k)}\right)\\
&\approx \begin{cases} 1, & \text{se } x>0\\ 1/2, & \text{se } x=0\\ 0, & \text{se } x<0.\end{cases}
\end{aligned}
$$

Como 0 é um ponto de descontinuidade da distribuição limite, pela definição de convergência em distribuição, segue que

$$\frac{X_{n-k,n}}{U(n/k)}-1 \xrightarrow{D} 0.$$

Consequentemente, $\dfrac{X_{n-k,n}}{U(n/k)} \xrightarrow{P} 1.$ □

Finalizamos a Seção com um resultado que estabelece um tipo de Lei dos Grandes Números para as estatísticas de ordem intermediárias, no caso especial da distribuição tipo $Pareto(1,1)$.

Proposição 1.11. Sejam $Y_1,...,Y_n$ variáveis aleatórias i.i.d com distribuição tipo $Pareto(1,1)$ e $Y_{1,n},...,Y_{n,n}$ suas estatísticas de ordem. Então, para $k=k(n)=o(n)$, quando $n\to\infty$, temos

$$\lim_{n\to\infty} Y_{n-k,n}=\infty, \tag{1.25}$$

e

$$\frac{k}{n}Y_{n-k,n}\xrightarrow{P} 1. \tag{1.26}$$

Demonstração. (a) Veja de Haan e Ferreira (2006) Lema 3.2.1.

(b) Pelo Corolário 1.1, para n suficientemente grande temos

$$P\left(\frac{k}{n}Y_{n-k,n}-1\leq x\right)\approx F_{N(0,1)}(\sqrt{k}x)\approx \begin{cases} 1, & \text{se } x>0\\ 1/2, & \text{se } x=0\\ 0, & \text{se } x<0\end{cases}$$

pois $\sqrt{k}\to\infty$ quando $n\to\infty$. Como 0 é um ponto de descontinuidade da distribuição limite, segue que

$$\frac{k}{n}Y_{n-k,n}-1\xrightarrow{D} 0.$$

Consequentemente, $\dfrac{k}{n}Y_{n-k,n}\xrightarrow{P} 1.$ □

Capítulo 2

Medidas de Risco

2.1 Introdução

O conceito de risco é bastante amplo e existem diversas definições encontradas na literatura, dependendo do contexto em que está inserido. No contexto da ciência atuarial, Denuit et al. (2006) descreve, de maneira ampla, um seguro como um negócio destinado a transferir (total ou parcialmente) o impacto econômico relativo a contratempos imprevistos. Nesse sentido, em linhas gerais, um *risco* pode ser descrito como um evento extremo que ocorre ocasionalmente e que pode acarretar algumas consequências financeiras adversas. Neste caso, sempre há um elemento de incerteza (ou o momento de sua ocorrência, ou sua própria ocorrência ou a gravidade das suas consequências). Por exemplo, as companhias de seguro eventualmente enfrentam sinistros extremos que podem comprometer a solvência de uma carteira, ou ainda, podem comprometer parte substancial da companhia.

Assim, usando as ferramentas da teoria de probabilidade, um risco é modelado por uma variável aleatória não negativa, que, no contexto atuarial, representa a quantia aleatória de dinheiro paga por uma seguradora para indenizar um tomador de seguro, um beneficiário e/ou um terceiro na execução de um contrato de seguro.

As *medidas de risco* são uma ferramenta extremamente importante para auxiliar na gestão do risco relacionado a eventos extremos, bastante utilizadas por seguradoras, bancos, instituições financeiras, entre outros e têm o papel de auxiliar na mensuração do risco de um determinado investimento, seguro, empréstimo, etc.

Neste capítulo, apresentamos alguns conceitos e propriedades básicas relacionadas a medidas de risco e destacamos algumas medidas comumente estudadas na literatura, em especial aquelas utilizadas por companhias seguradoras. Para um estudo mais detalhado e aprofundado sobre as medidas de risco utilizadas nas seguradoras, sugerimos a leitura de Denuit et al. (2006), Cai e Tan (2007) e Goegebeur et al. (2022).

Na Seção 2.1, apresentamos as definições preliminares de risco e medida de risco e os axiomas que caracterizam as chamadas *medidas de risco coerentes.* Na Seção 2.2, apresentamos a medida de risco clássica mais amplamente utilizada pelas empresas reguladoras, conhecida como *Valor em Risco* (VaR) e discutimos brevemente os axiomas relativos à coerência desta medida. Na Seção 2.3, analisamos uma medida popular alternativa ao VaR, conhecida por *Esperança Caudal Condicional* ou *Esperança de Cauda Condicional* (CTE). Finalmente, na Seção 2.4, introduzimos a medida de risco chamada *Momento Caudal Condicional* ou *Momento de Cauda Condicional* (CTM), que permite unificar as definições de diversas medidas de risco e cuja estimação será objeto de estudo no próximo capítulo.

As principais referências bibliográficas utilizadas neste capítulo são Artzner et al. (1999), Cardoso (2008), Denuit et al. (2006), Goegebeur et al. (2022), Hardy (2006), Oliveira (2009), Righi e Ceretta (2014), entre outras.

2.2 Conceitos Preliminares

Nesta Seção apresentamos uma síntese de conceitos básicos relacionados as medidas de risco, baseados no livro do Denuit et al. (2006).

Iniciamos com uma definição geral do que chamamos de risco.

Definição 2.1. Denominamos um *risco* como sendo uma variável aleatória não negativa X, que está associada às perdas ocasionadas por eventos adversos.

Como citamos anteriormente, em atuária, por exemplo, X representa a quantia aleatória de dinheiro paga por uma seguradora para indenizar um tomador de seguro, um beneficiário e/ou um terceiro na execução de um contrato de seguro. Assim, a possível perda da seguradora é dada por X menos o valor do prêmio pago pelo segurado. Logo, neste contexto, é natural, em geral, modelar os riscos por variáveis aleatórias não negativas.

Dessa forma, em linhas gerais, uma medida de risco pode ser definida como um funcional que associa a cada risco um número real, que quantifica o grau de "perigo" de X.

Definição 2.2. Uma *medida de risco* é um funcional $g : \mathcal{G} \to \mathbb{R}$, em que $\mathcal{G}$ é o conjunto de todos os riscos.

Seguindo o raciocínio anterior, podemos interpretar que, grandes valores de $g[X]$ indicam que X é "perigoso". Não existe um consenso sobre qual é a medida de risco mais adequada, pois nenhuma delas consegue abranger todo o quadro de risco presente em situações do dia-a-dia. Cada uma delas se concentra em um aspecto particular do risco. Conforme Denuit et al. (2006) descreve, podemos dizer que há um paralelo com a estatística matemática, em que determinadas características das distribuições podem ter significados e usos bastante diferentes, como, por exemplo, a média para medir a tendência central, a variância para medir a dispersão, entre outros.

A definição de medida de risco é bastante ampla, por isso não é tão útil para fins práticos. Por conta disso, são definidos alguns axiomas que, se satisfeitos, tornam a medida de risco mais aplicável em situações reais. A medida de risco que satisfizer os axiomas de translatividade, homogeneidade positiva, monotonicidade e subaditividade será chamada de *medida de risco coerente*. Um dos trabalhos pioneiros no estudo de medidas de risco coerentes foi Artzner et al. (1999). Vale a pena destacar que esse conjunto de axiomas não é universalmente aceito, então eles podem variar de autor para autor. Modificar o conjunto de axiomas leva a outras definições de medidas de risco "coerentes".

Listamos abaixo alguns dos axiomas descritos por Denuit et al. (2006).

Para motivar o primeiro axioma, seguindo a nossa linha de raciocínio, uma medida de risco pode ser interpretada como uma função que fornece a quantidade de capital que deve ser adicionado para tornar o risco aceitável e, então, qualquer aumento no passivo por um valor constante c deve resultar no mesmo aumento no capital.

Definição 2.3. (Translatividade). Uma medida de risco g é *translativa* se para todas as variáveis aleatórias X e cada constante $c \in \mathbb{R}$.

$$g[X + c] = g[X] + c.$$

O segundo axioma garante que um risco deve ser independente da unidade monetária utilizada.

Definição 2.4. (Homogeneidade Positiva). Uma medida de risco g é *homogênea positiva* se para todas as variáveis aleatórias X e qualquer constante positiva c

$$g[cX] = cg[X].$$

O próximo axioma é bastante natural em termos práticos, pois se um risco Y pode levar a perdas maiores do que outro risco X, então a quantidade de capital que deve ser adicionado para tornar Y aceitável deve ser maior que a quantidade que deve ser adicionada para tornar X aceitável.

Definição 2.5. (Monotonicidade). Uma medida de risco g é *monôtona* se para todas as variáveis aleatórias X e Y tais que $P(X \leq Y) = 1$ então

$$g[X] \leq g[Y].$$

O último axioma reflete a ideia de que o risco pode ser diminuído pela diversificação.

Definição 2.6. (Sub-aditividade). Uma medida de risco g é *subaditiva* se para todas as variáveis aleatórias X e Y

$$g[X+Y] \leq g[X] + g[Y].$$

Definição 2.7. Uma medida de risco que é translativa, homogênea positiva, subaditiva e monótona é chamada de *coerente*.

De forma prática, é interessante que a medida de risco seja coerente. Mas, como já comentamos anteriormente, cada medida irá se basear em um aspecto particular do risco. Por isso, as vezes é conveniente usar medidas de risco mesmo que elas não cumpram todos os axiomas definidos anteriormente.

Neste trabalho, estamos interessados nas medidas de risco que são utilizadas para determinar provisões e requisitos de capital com o objetivo de evitar a insolvência da empresa. Assim, vamos nos concentrar nas medidas de risco que medem as caudas superiores das funções de distribuição.

Nas próximas seções apresentamos algumas dessas medidas de interesse.

2.3 Valor em Risco (VaR)

O *Valor em Risco* é a medida de risco clássica mais popular. No setor financeiro ela é usada para determinar a quantidade de ativos necessários para cobrir possíveis perdas.

Definição 2.8. Definimos o *Valor em Risco*, denotado VaR (*Value-at-Risk*), por

$$VaR_p(X) = U(1/p),\, p \in (0,1), \tag{2.1}$$

em que $X \in \mathcal{G}$ e U denota a função quantil caudal de X, isto é, $U(x) := F^{\leftarrow}(1-1/x) = \inf\{y : F(y) \geq 1 - 1/x\}$, $x > 1$. Ou seja,

$$VaR_p(X) = \inf\{y : F(y) \geq 1-p\} = F^{\leftarrow}(1-p),\, p \in (0,1). \tag{2.2}$$

Vale observar que como o VaR é definido em termos da função quantil, então todas as propriedades listadas na Proposição (1.2) são diretamente aplicáveis ao VaR.

Na próxima proposição mostramos que o VaR satisfaz os axiomas de translatividade, homogeneidade positiva e monotonicidade, mas de maneira geral, não satisfaz a subatividade. Assim, exceto para alguns casos especiais o VaR não é uma medida de risco coerente. Para maiores informações sobre o VaR veja, por exemplo, Jorion (2007).

Proposição 2.1. Para $0 < p < 1$, seja $VaR_p(X)$, $X \in \mathcal{G}$ definido como em (2.1). Então,

(a) o VaR é uma medida de risco translativa, homogênea positiva e monótona

(b) o VaR não satisfaz a subaditividade.

Demonstração.

(a) Sejam X e Y variáveis aleatórias não negativas, definidas no mesmo espaço de probabilidade.

(a.1) Para $c \in \mathbb{R}$, segue da definição do VaR e das propriedades de ínfimo que

$$\begin{aligned} VaR_p(X+c) &= \inf\{y : P(X+c \leq y) \geq 1-p\} \\ &= \inf\{z+c : P(X \leq z) \geq 1-p\} \\ &= \inf\{z : P(X \leq z) \geq 1-p\} + c \\ &= VaR_p(X) + c. \end{aligned}$$

Logo, o VaR satisfaz o axioma da translatividade.

(a.2) Para provar a homogeneidade positiva, consideramos dois casos separadamente. *Caso 1:* $\lambda = 0$. Por um lado, se $VaR_p(X) < \infty$, então

$$VaR_p(\lambda \cdot X) = \inf\{y : P(0.X \leq y) \geq 1-p\} = \inf\{y : y \geq 0\} = 0 = 0 \cdot VaR_p(X)$$

Por outro lado, se $VaR_p(X) = \infty$, basta convencionarmos $0. + \infty = 0$.

Caso 2: $\lambda > 0$. Neste caso, da definição do VaR e das propriedades de ínfimo segue que

$$\begin{aligned} VaR_p(\lambda \cdot X) &= \inf\{y : P(\lambda \cdot X \leq y) \geq 1-p\} \\ &= \inf\{z \cdot \lambda : P(X \leq z) \geq 1-p\} \\ &= \lambda \cdot \inf\{z : P(X \leq z) \geq 1-p\} \\ &= \lambda \cdot VaR_p(X). \end{aligned}$$

(a.3) Para provar o axioma da monotonicidade, suponha, sem perda de generalidade, que o espaço de probabilidade considerado é completo. Assim, se $P(X \leq Y) = 1$, segue que $P(Y \leq y) \leq P(X \leq y),\ \forall\, y \in \mathbb{R}$ e, consequentemente,

$$\{y;\, P(Y \leq y) \geq 1 - p\} \subseteq \{y;\, P(X \leq y) \geq 1 - p\}.$$

Logo, pelas propriedades do ínfimo, temos

$$VaR_p(Y) = \inf\{y;\, P(Y \leq y) \geq 1-p\} \geq \inf\{y;\, P(X \leq y) \geq 1-p\} = VaR_p(X).$$

(b) Para comprovar que o VaR não satisfaz o axioma da subaditividade considere o seguinte exemplo: sejam X_1, X_2 variáveis aleatórias i.i.d, com distribuição

$$P(X_i = 100) = 0,01 \text{ e } P(X_i = 2) = 0,99,\ i \in \{1, 2\}$$

e distribuição conjunta dada por

$$P(X_1 = 100, X_2 = 100) = 0,0001$$

$$P(X_1 = 100, X_2 = 2) = P(X_1 = 2, X_2 = 100) = 0,0099$$

$$P(X_1 = 2, X_2 = 2) = 0,9801.$$

Assim, podemos obter que a função de distribuição acumulada da soma $X_1 + X_2$ é

$$F_{X_1+X_2}(x) = \begin{cases} 1, & \text{se } x \geq 200 \\ 0,9999, & \text{se } 102 \leq x < 200 \\ 0,9801, & \text{se } 4 \leq x < 102 \\ 0, & \text{se } x < 4 \end{cases}$$

Agora, considerando $p = 0,0198$ obtemos que

$$VaR_{0,0198}(X_1 + X_2) = 102 > 4 = VaR_{0,0198}(X_1) + VaR_{0,0198}(X_2).$$

□

Embora tenhamos apresentado acima um exemplo que comprova a não subatidi-tividade do VaR para o caso de riscos discretos, no exemplo a seguir mostramos que também para o caso contínuo o VaR, em geral, não é subaditivo. Para outros exemplos, veja Artzner et al. (2001).

Exemplo 2.1. Sejam X e Y variáveis aleatórias i.i.d. com distribuição Exponencial(1). Neste caso, $X + Y \sim \text{Gama}(2, 1)$ e as respectivas funções de distribuição acumulada são dadas por

$$F_X(x) = F_Y(x) = \begin{cases} 1 - e^{-x}, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases}$$

e

$$\begin{aligned} F_{X+Y}(x) &= \begin{cases} \displaystyle\int_0^x ze^{-z}\mathrm{d}z, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases} \\ &= \begin{cases} 1 - e^{-x} - xe^{-x}, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0. \end{cases} \end{aligned}$$

Agora, considerando $p = 0,99$, obtemos

$$\begin{aligned} VaR_{0,99}(X) = VaR_{0,99}(Y) &= \inf\{z : F_X(z) \geq 1 - 0,99\} \\ &= \inf\left\{z : z > 0 \text{ e } 1 - e^{-z} \geq 0,01\right\} \\ &= \inf\{z : z > 0 \text{ e } z \geq -\ln(0,99)\} \\ &\approx 0,01 \end{aligned}$$

e

$$\begin{aligned} VaR_{0,99}(X+Y) &= \inf\{z : F_{X+Y}(z) \geq 1 - 0,99\} \\ &= \inf\left\{z : z > 0 \text{ e } 1 - e^{-z} - ze^{-z} \geq 0,01\right\} \\ &\approx 0,145. \end{aligned}$$

Portanto,

$$VaR_{0,99}(X+Y) > VaR_{0,99}(X) + VaR_{0,99}(Y).$$

✓

2.4 Esperança Caudal Condicional (CTE)

Como vimos na Seção anterior uma das desvantagens do VaR é que ele calcula a probabilidade de uma perda ocorrer, mas não captura a magnitude de tal perda caso ela ocorra. A *Esperança Caudal Condicional* ou *Esperança de Cauda Condicional* (Conditional Tail Expectation), denotada por CTE, é uma medida de risco que pretende resolver esta e outras possíveis desvantagens do VaR.

Definição 2.9. Definimos a *Esperança Caudal Condicional* ou *Esperança de Cauda Con-*

dicional, denotada CTE, de uma variável aleatória não negativa X, por

$$CTE_p(X) = E(X|X > U(1/p)), \tag{2.3}$$

desde que $E|X| < \infty$, para $0 < p < 1$. Em que U denota a função quantil caudal de X, isto é, $U(x) := \inf\{y : F(y) \geq 1 - 1/x\}$, $x > 1$. Quando $P(X = 0) = 1$ assumimos $CTE_p(X) = 0$.

Ou seja, como $U(1/p) = VaR_p(X)$, a CTE representa o risco esperado condicional, dado que o risco excede seu VaR.

Obviamente, $CTE_p(X) \geq VaR_p(X)$ e, assim, a CTE é uma medida de risco mais conservadora em relação ao VaR. Além disso, ela leva em consideração toda a informação contida na cauda superior da distribuição e é frequentemente aplicada em investimentos financeiros ou na indústria de seguros.

Para maiores informações sobre esta medida de risco recomendamos a leitura de Brazauskas et al. (2008) ou Cai e Tan (2007).

Antes de analisarmos os axiomas de coerência para a CTE, vamos fazer uma breve descrição da sua relação com duas outras medidas de risco comumente encontradas na literatura: o *Valor em Risco Caudal* (TVaR - *Tail Value at Risk*) e o *Déficit Esperado* (ES - *Expected Shortfall*), definidos, respectivamente, por

$$TVaR_p(X) = \frac{1}{p}\int_0^p VaR_\varepsilon(X)\mathrm{d}\varepsilon, \ \forall\, 0 < p < 1 \tag{2.4}$$

e

$$ES_p(X) = E[(X - VaR_p(X))^+], \ \forall\, 0 < p < 1. \tag{2.5}$$

em que $Y^+ = Y.I_{Y>0}$, denota a parte positiva da v.a. Y. Vale observar que, no contexto da atuária, o ES, também chamado *Stop-loss Premium* (SP), é uma medida de risco de resseguro que possibilita identificar os casos "perigosos".

Mais relações com outras medidas de risco, podem ser encontradas em Denuit et al. (2006).

Na proposição a seguir mostramos que no caso de distribuições de risco contínuas, o $TVaR$ coincide com a CTE e o ES é proporcional à diferença entre a CTE e o VaR.

Proposição 2.2. Se X é uma variável aleatória não negativa, com função de distribuição contínua F, então para todo $0 < p < 1$

$$CTE_p(X) = \frac{1}{p}\int_0^p VaR_\varepsilon(X)\mathrm{d}\varepsilon = TVaR_p(X) \tag{2.6}$$

e

$$CTE_p(X) = VaR_p(X) + \frac{1}{p}E[(X - VaR_p(X))^+] = VaR_p(X) + \frac{1}{p}ES_p(X), \qquad (2.7)$$

em que $Y^+ = Y.I_{Y>0}$, denota a parte positiva da v.a. Y.

Demonstração. Primeiramente, por (2.2), temos

$$P(X > U(1/p)) = P\left(X > VaR_p(X)\right) = 1 - F(F^{\leftarrow}(1-p)).$$

Agora, como X é uma variável aleatória contínua, pelo item (b) da Proposição 1.3, segue que $F(F^{\leftarrow}(1-p)) = 1-p$. Assim, $P(X > U(1/p)) = p > 0$ e pelas propriedades de esperança condicional, segue que

$$\begin{aligned} CTE_p(X) = E(X|X > U(1/p)) &= \frac{1}{P(X > VaR_p(X))}\int_{(X>VaR_p(X))} X\,\mathrm{d}P \\ &= \frac{1}{p}\int_{(X>VaR_p(X))} X\,\mathrm{d}P. \end{aligned}$$

Mas, pelo item (a) da Proposição 1.3, temos que X tem a mesma distribuição de $F^{\leftarrow}(1-\xi)$, em que ξ é uma v.a. com distribuição Uniforme sobre $(0,1)$, pois $U = 1-\xi$ também tem distribuição Uniforme sobre $(0,1)$. Logo, podemos escrever

$$\begin{aligned} CTE_p(X) &= \frac{1}{p}\int_{(F^{\leftarrow}(1-\xi)>F^{\leftarrow}(1-p))} F^{\leftarrow}(1-\xi)\,\mathrm{d}P \\ &= \frac{1}{p}\int_0^p F^{\leftarrow}(1-\varepsilon)\mathrm{d}\varepsilon \\ &= \frac{1}{p}\int_0^p VaR_\varepsilon(X)\mathrm{d}\varepsilon \end{aligned}$$

e obtemos (2.6). Para provar (2.7), basta observar que

$$\begin{aligned} ES_p(X) &= E[(X - VaR_p(X))^+] \\ &= \int_{(X-VaR_p(X)>0)} [X - VaR_p(X)]\,\mathrm{d}P \\ &= E(X - VaR_p(X) \mid X > VaR_p(X))P(X - VaR_p(X) > 0) \\ &= [E(X \mid X > VaR_p(X)) - VaR_p(X)][1 - F(F^{\leftarrow}(1-p))] \\ &= (CTE_p(X) - VaR_p(X))\; p \end{aligned}$$

e obtemos (2.7). $\square$

Observação 2.1. Pela proposição anterior, $CTE_p(X)$ coincide com $TVaR_p(X)$, para todo $0 < p < 1$, no caso em que X é variável aleatória contínua, pois, neste caso, temos que $F(F^{\leftarrow}(1-p)) = 1-p$. Entretanto, para o caso geral temos as seguintes relações, que podem ser verificadas de maneira análoga ao caso contínuo:

$$CTE_p(X) = VaR_p(X) + \frac{1}{1 - F(VaR_p(X))} ES_p(X)$$

e

$$TVaR_p(X) = CTE_p(X) + \left(\frac{1}{p} - \frac{1}{1 - F(VaR_p(X))} \right) ES_p(X).$$

$\Diamond$

Finalizamos a Seção com a análise dos axiomas de coerência para a CTE, que é uma medida coerente no caso em que os riscos são variáveis aleatórias contínuas.

Proposição 2.3. Para $0 < p < 1$, seja $CTE_p(X)$, com $X \in \mathcal{G}$, definida como em (2.3).

(a) A CTE é uma medida de risco translativa e homogênea positiva.

(b) Se X é uma v.a. contínua, então a CTE é monótona e sub-aditiva. Consequentemente, neste caso, é uma medida coerente.

Demonstração. (a) Primeiramente, vamos provar o axioma da translatividade. Para $c \in \mathbb{R}$, pela Proposição 2.1, como o VaR é translativo, segue da linearidade da esperança condicional que

$$\begin{aligned} CTE_p(X+c) &= E(X+c|X+c > VaR_p(X+c)) \\ &= E(X \mid X+c > VaR_p(X)+c) + c \\ &= E(X|X > VaR_p(X)) + c \\ &= CTE_p(X) + c. \end{aligned}$$

Para provar o axioma da homogêneidade positiva, seja $\lambda \geq 0$. Com a convenção: $CTE_p(X) = 0$, quando $P(X = 0) = 1$, assumida na Definição 2.9, é imediato a validade do axioma para $\lambda = 0$. Agora, para $\lambda > 0$, pela Proposição 2.1, como o VaR é homogêneo positivo, segue da linearidade da esperança condicional que

$$\begin{aligned} CTE_p(\lambda X) &= E(\lambda X|\lambda X > VaR_p(\lambda X)) \\ &= \lambda E(X \mid \lambda X > \lambda VaR_p(X)) \\ &= \lambda E(X|X > VaR_p(X)) \\ &= \lambda . CTE_p(X). \end{aligned}$$

(b) Se X é uma v.a. contínua, então pela Proposição 2.6, temos

$$CTE_p(X) = \frac{1}{p}\int_0^p VaR_\varepsilon(X)\mathrm{d}\varepsilon.$$

Logo, a monotonicidade da CTE segue da monotonicidade do VaR que foi provada na Proposição 2.1.

Agora, para provar a subaditividade da CTE, sejam X e Y variáveis aleatórias contínuas, com funções de distribuição F_X e F_Y, respectivamente. Então, $X+Y$ também possui função de distribuição contínua, que denotamos F_{X+Y}. Logo, como vimos anteriormente, das propriedades da inversa generalizada, como F_X, F_Y e F_{X+Y} são funções contínuas segue que, para todo $0 < p < 1$, $F_X(F_X^{\leftarrow}(1-p)) = F_Y(F_Y^{\leftarrow}(1-p)) = F_{X+Y}(F_{X+Y}^{\leftarrow}(1-p)) = 1-p$ e, consequentemente,

$$P(X > VaR_p(X)) = P(Y > VaR_p(Y)) = P(X+Y > VaR_p(X+Y)) = p \in (0,1).$$

Assim, usando a Proposição 1.4, podemos obter que

$$\begin{aligned}
CTE_p(X+Y) &= E(X+Y|X+Y > VaR_p(X+Y)) \\
&= E(X|X+Y > VaR_p(X+Y)) + E(Y|X+Y > VaR_p(X+Y)) \\
&\leq E(X|X > VaR_p(X)) + E(Y|Y > VaR_p(Y)) \\
&\leq CTE_p(X) + CTE_p(Y).
\end{aligned}$$

Portanto, no caso em que os riscos são variáveis aleatórias contínuas então a CTE é uma medida de risco coerente. □

Apresentamos a seguir um exemplo no caso contínuo, em que vale a subaditividade do CTE, mas não não vale a do VaR, ressaltando, assim, uma vantagem do CTE em comparação ao VaR.

Exemplo 2.2. Considere como no Exemplo 2.1, X e Y variáveis aleatórias i.i.d com distribuição Exponencial(1). Vimos, que para $p = 0,99$, $VaR_{0,99}(X) = VaR_{0,99}(Y) \approx 0,01$ e $VaR_{0,99}(X+Y) \approx 0,145$ e, assim,

$$VaR_{0,99}(X+Y) > VaR_{0,99}(X) + VaR_{0,99}(Y).$$

Agora, calculando para a CTE, obtemos

$$\begin{aligned} CTE_{0,99}(X) = CTE_{0,99}(Y) &= E(X|X > VaR_{0,99}(X)) \\ &\approx \frac{\int_{0,01}^{\infty} x f_X(x)\mathrm{d}x}{1 - P(X \le VaR_{0,99}(X))} \\ &\approx \frac{\int_{0,01}^{\infty} x e^{-x}\mathrm{d}x}{1 - F_X(F^{\leftarrow}(1-0,99))} \\ &\approx 1,01, \end{aligned}$$

e, de maneira análoga,

$$\begin{aligned} CTE_{0,99}(X+Y) &= E(X+Y|X+Y > VaR_{0,99}(X+Y)) \\ &\approx \frac{\int_{0,145}^{\infty} x^2 e^{-x}(x)\mathrm{d}x}{1 - P(X+Y \le VaR_{0,99}(X+Y))} \\ &\approx 2,02. \end{aligned}$$

Portanto,

$$CTE_{0,99}(X+Y) \le CTE_{0,99}(X) + CTE_{0,99}(Y).$$

✓

2.5 Momento Caudal Condicional (CTM)

O *Momento Caudal Condicional* ou *Momento de Cauda Condicional*, denotado por CTM (*Conditional Tail Moment*), é uma medida de risco que foi introduzida por Methni, Gardes e Girard (2014), como uma nova ferramenta para unificar a estimação de diversas medidas de risco, incluindo a CTE.

Nesta Seção, vamos apresentar a definição do CTM e a sua relação com outras medidas de risco. No próximo capítulo estudaremos em detalhes o comportamento assintótico de um estimador natural para o CTM baseado na Teoria dos Valores Extremos, proposto por Goegebeur et al. (2022).

Definição 2.10. Definimos o *Momento Caudal Condicional* ou *Momento de Cauda Condicional*, denotado CTM, de uma variável aleatória não-negativa X, por

$$CTM_{\beta,p}(X) = E(X^{\beta}|X > U(1/p)) = E(X^{\beta} \mid X > VaR_p(X)), \tag{2.8}$$

desde que $\mathbb{E}(X^{\beta}) < \infty$, em que $p \in (0,1)$ e $\beta > 0$. Ou seja, o CTM é definido como o momento condicional de ordem β da distribuição de risco (ou perda) acima do $p-$quantil superior.

Não é difícil verificar que para $\beta \neq 1$ de uma maneira geral, CTM não cumpre a translatividade e a homogeneidade positiva. Ou seja, não é uma medida de risco coerente. Mas, como mencionamos anteriormente, a grande utilidade do CTM é possibilitar a unificação da estimação de várias medidas de risco por meio da sua estimação.

É óbvio que o CTM de ordem um ($\beta = 1$) coincide com a CTE. Listamos abaixo, outras medidas de risco, destacadas por Methni, Gardes e Girard (2014), que podem ser definidas em função do CTM.

Definição 2.11. Seja $X \in \mathcal{G}$ e $p \in (0,1)$.

(a) A *Variância Caudal Condicional* ou *Variância de Cauda Condicional*, denotada por CTV (*Conditional Tail Variance*), é definida como

$$\mathrm{CTV}_p(X) := E[(X - CTE_p(X))^2 | X > VaR_p(X)].$$

Ou seja, ela mede a variabilidade condicional de X acima de $VaR_p(X)$ e indica a dispersão dos valores de X em relação ao $CTE_p(X)$.

(b) A *Assimetria Caudal Condicional* ou a *Assimetria de Cauda Condicional*, denotada por CTS (*Conditional Tail Skewness*), é definido como o coeficiente de assimetria das leis de perdas sabendo que as perdas são maiores que o VaR. Ou seja,

$$\mathrm{CTS}_p(X) := \frac{E(X^3 | X > VaR_p(X))}{\mathrm{CTV}_p(X)^{3/2}}.$$

(c) O *Valor em risco condicional*, denotado por CVaR (*Conditional-Value-at-Risk*), é definido por

$$\mathrm{CVaR}_{p,\lambda}(X) = \lambda VaR_p(X) + (1-\lambda) CTE_p(X)$$

com $0 \leq \lambda \leq 1$. Em particular, $\mathrm{CVaR}_{p,1}(Y) = VaR_p(Y)$ e $\mathrm{CVaR}_{p,0}(Y) = CTE_p(Y)$.

Podemos verificar facilmente que as medidas CTE, ES (definido em (2.5)), CTV, CTS e CVaR podem ser reescritas na forma

$$\Phi\left(VaR_p(X), CTM_{1,p}(X), CTM_{2,p}(X), CTM_{3,p}(X)\right),$$

em que $\Phi : (t_0, t_1, t_2, t_3) \to \mathbb{R},\ (t_0, t_1, t_2, t_3) \in \mathbb{R}^4$, é uma função dada, em cada caso, da seguinte forma:

$$
\begin{aligned}
\mathrm{CTE}_p &: \Phi(t_0, t_1, t_2, t_3) = t_1 \\
\mathrm{ES}_p &: \Phi(t_0, t_1, t_2, t_3) = p(t_1 - t_0) \\
\mathrm{CTV}_p &: \Phi(t_0, t_1, t_2, t_3) = t_2 - t_1^2 \\
\mathrm{CTS}_p &: \Phi(t_0, t_1, t_2, t_3) = \frac{t_3}{(t_2 - t_1^2)^{3/2}} \\
\mathrm{CVaR}_{p,\lambda} &: \Phi(t_0, t_1, t_2, t_3) = \lambda t_0 + (1 - \lambda) t_1,\ \lambda \in [0, 1].
\end{aligned}
$$

Assim, a grande vantagem do CTM reside no fato que estimá-lo permite estimar as medidas de risco baseadas em momentos condicionais, de ordens arbitrárias, da distribuição de risco acima do VaR.

Capítulo 3

Estimação do CTM

3.1 Introdução

Neste capítulo, abordamos o problema da estimação do CTM, que permite a estimação de outras medidas de risco obtidas como funções do CTM, conforme observamos no capítulo anterior.

Para facilitar a notação, vamos denotar o momento caudal condicional (CTM) de uma variável aleatória não-negativa X, com função de distribuição F e função quantil caudal U, conforme definido na Definição 2.10, por

$$\theta_{\beta,p} := CTM_{\beta,p}(X) = E(X^{\beta}|X > U(1/p)) = E(X^{\beta} \mid X > VaR_p(X)), \tag{3.1}$$

para $p \in (0,1)$ e $\beta > 0$ tal que $\mathbb{E}(X^{\beta}) < \infty$.

Nosso objetivo é apresentar em detalhes os resultados obtidos por Goergebeur et al. (2022) sobre as propriedades assintóticas do estimador proposto por eles, baseado em uma amostra $X_1, ..., X_n$ de variáveis aleatórias i.i.d. com função de distribuição F, no caso especial em que p é muito pequeno, ou seja, $p < 1/n$. Lembrando que

$$U(1/p) = \inf \{y \in \mathbb{R};\ F(y) \geq 1 - p\}.$$

Para este caso, Goergebeur et al. (2022) fazem uso da teoria de valores extremos, a qual permite extrapolar além do intervalo de dados. Especificamente, assume-se a condição de segunda ordem apresentada na Definição 1.8, com $\rho < 0$ e $\gamma > 0$, que especifica a velocidade de convergência de $\dfrac{U(tx)}{U(t)}$, quando $t \to \infty$, ou seja,

Condição de Segunda Ordem (SOC). Para algum $\gamma > 0$, $\rho < 0$ e alguma função positiva ou negativa $A(t)$, com $\lim_{t\to\infty} A(t) = 0$, temos

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^{\gamma}}{A(t)} = x^{\gamma}\frac{x^{\rho}-1}{\rho}, \qquad \forall x > 0. \tag{3.2}$$

Pelo Teorema 1.5, podemos concluir que a condição (SOC) implica que $F \in D(G_{\gamma})$, com índice extremal $\gamma > 0$, em que

$$G_{\gamma}(x) = \exp\left[-(1+\gamma x)^{-1/\gamma}\right]$$

é a distribuição do valor extremo generalizada e, pelo Teorema 1.7, concluímos que a função $A(.)$ é de variação regular, com índice ρ.

Alternativamente, a condição (3.2) pode ser expressa, em termos da função de distribuição $\overline{F}$ da seguinte forma

$$\lim_{t\to\infty} \frac{\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma}}{\alpha(t)} = x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho}, \qquad \forall x > 0, \tag{3.3}$$

em que $\alpha(t) := A(1/\overline{F}(t))$. Neste caso, o parâmetro γ determina o peso da cauda superior de F, com valores grandes indicando caudas mais pesadas.

Exemplos de distribuições que satisfazem a condição (SOC) são as distribuições tipo Pareto descritas no Exemplo 1.7.

Para a estimação de $\theta_{\beta,p}$, considera-se, primeiramente, o caso intermediário: $k, n \to \infty$ e $p = k/n \to 0$.

Neste caso, considerando as estatísticas de ordem da amostra, $X_{1,n}, ..., X_{n,n}$, e assumindo a condição de segunda ordem (SOC), temos pela Proposição 1.10, que $U(1/p) = U(n/k)$ pode ser estimado por $X_{n-k,n}$ e, consequentemente, $\theta_{\beta,p}$ pode ser estimado por uma média empírica dos dados acima de $X_{n-k,n}$.

Com isso, para este caso intermediário, Goergebeur et al. (2022) propõem o seguinte estimador para $\theta_{\beta,k/n}$

$$\tilde{\theta}_{\beta,n} := \frac{1}{k}\sum_{j=1}^{k} X_{n-j+1,n}^{\beta} \tag{3.4}$$

e obtém a distribuição assintótica de $\tilde{\theta}_{\beta,n}$, devidamente normalizado, para o caso em que $0 < \gamma < 1/(2\beta)$ (Teorema 3.2).

Dessa forma, através de uma expansão assintótica do CTM (Teorema 3.1), provase que para $p \downarrow 0$

$$\theta_{\beta,p} \sim \left(\frac{(U(1/p))}{(U(n/k))}\right)^{\beta} \theta_{\beta,k/n} \sim \left(\frac{k}{np}\right)^{\beta\gamma} \theta_{\beta,k/n}.$$

Com base nesses resultados, Goergebeur et al. (2022), propõem um estimador extrapolado para $\theta_{\beta,p}$, definido por

$$\hat{\theta}_{\beta,p} = \left(\frac{k}{np}\right)^{\beta\hat{\gamma_k}} \tilde{\theta}_{\beta,k/n}, \tag{3.5}$$

em que $\tilde{\theta}_{\beta,k/n}$ é o estimador do caso intermediário, dado em (3.4), e $\hat{\gamma_k}$ é um estimador de γ devidamente escolhido. O principal resultado do artigo de Goergebeur et al. (2022) é a obtenção da normalidade assintótica de $\hat{\theta}_{\beta,p}$

Assim, para estabelecermos os resultados obtidos por Goergebeur et al. (2022), apresentamos inicialmente na Seção 3.2 uma lista de resultados auxiliares que serão utilizados para a derivação dos resultados principais a serem demonstrados na Seção 3.3, os quais estabelecem, sob condições desejáveis, a normalidade assintótica dos estimadores descritos acima.

3.2 Lemas Auxiliares

O seguinte lema foi apresentado por Goergebeur et al. (2022) e contém uma propriedade útil das estatísticas de ordem associadas a uma amostra aleatória da distribuição $Pareto(1,1)$ e que será utilizado para obter a normalidade assintótica dos estimadores descritos anteriormente.

Lema 3.1. Sejam $Y_1, ..., Y_n$ variáveis aleatórias independentes Pareto$(1,1)$, com estatísticas de ordem $Y_{1,n} \leq ... \leq Y_{n,n}$. Considere, para $\beta_i > 0,\ i = 1,2$ e $\gamma > 0$,

$$P_{n,\beta_i} = \sqrt{k}\left(\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta_i} - \frac{1}{1-\gamma\beta_i}\right),\ i = 1,2 \tag{3.6}$$

e

$$Q_n = \sqrt{k}\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right]. \tag{3.7}$$

Então, se $\gamma < \min(\frac{1}{2\beta_1}, \frac{1}{2\beta_2})$, temos para $k, n \to \infty$ com $\frac{k}{n} \to 0$

$$\begin{pmatrix} P_{n,\beta_1} \\ P_{n,\beta_2} \\ Q_n \end{pmatrix} \xrightarrow{D} \begin{pmatrix} P_{\beta_1} \\ P_{\beta_2} \\ Q \end{pmatrix} \sim \mathcal{N}_3\left(\mathbf{0},\ \mathbf{\Sigma}\right), \tag{3.8}$$

em que $\mathcal{N}_3\left(\mathbf{0},\ \mathbf{\Sigma}\right)$ indica a distribuição Normal 3−dimensional, com vetor média $\mathbf{0}$ e matriz de covariâncias $\mathbf{\Sigma}$ dados por

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad \mathbf{\Sigma} := \begin{pmatrix} \frac{(\gamma\beta_1)^2}{(1-\gamma\beta_1)^2(1-2\gamma\beta_1)} & \frac{\gamma^2\beta_1\beta_2}{(1-\gamma\beta_1)(1-\gamma\beta_2)(1-\gamma(\beta_1+\beta_2))} & 0 \\ \frac{\gamma^2\beta_1\beta_2}{(1-\gamma\beta_1)(1-\gamma\beta_2)(1-\gamma(\beta_1+\beta_2))} & \frac{(\gamma\beta_2)^2}{(1-\gamma\beta_2)^2(1-2\gamma\beta_2)} & 0 \\ 0 & 0 & \gamma^2 \end{pmatrix}. \tag{3.9}$$

Demonstração. Primeiramente, para facilidade de notação, considere

$$\Lambda_n := \begin{pmatrix} P_{n,\beta_1} \\ P_{n,\beta_2} \end{pmatrix},$$

em que P_{n,β_1} e P_{n,β_2} são dados em (3.6). Assim, pela Proposição 1.9 segue que

$$\begin{aligned} \Lambda_n &\stackrel{D}{=} \sqrt{k}\begin{pmatrix} \frac{1}{k}\sum_{j=1}^{k} Y_j^{\gamma\beta_1} - \frac{1}{1-\gamma\beta_1} \\ \frac{1}{k}\sum_{j=1}^{k} Y_j^{\gamma\beta_2} - \frac{1}{1-\gamma\beta_2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{k}}\sum_{j=1}^{k}\left(Y_j^{\gamma\beta_1} - \frac{1}{1-\gamma\beta_1}\right) \\ \frac{1}{\sqrt{k}}\sum_{j=1}^{k}\left(Y_j^{\gamma\beta_2} - \frac{1}{1-\gamma\beta_2}\right) \end{pmatrix}. \end{aligned}$$

Mas, como $Y_j,\ j = 1, \ldots, n$ têm distribuição Pareto(1,1), podemos obter que para todo $0 < \alpha < \frac{1}{2}$

$$E(Y_j^{\alpha}) = \int_1^{\infty} x^{\alpha}\frac{1}{x^2}dx = \frac{1}{1-\alpha}.$$

Agora, por hipótese temos que $0 < \gamma < \min(\frac{1}{2\beta_1}, \frac{1}{2\beta_2})$, então

$$E(Y_j^{\gamma\beta_i}) = \frac{1}{1-\gamma\beta_i} \quad \text{e} \quad E(Y_j^{\gamma\beta_i})^2 = \frac{1}{1-2\gamma\beta_i} < \infty,\ i \in \{1,2\} \tag{3.10}$$

Logo, utilizando o Teorema do Limite Central Multivariado, obtemos

$$\Lambda_n \xrightarrow{D} \mathcal{N}(\widetilde{\mathbf{0}}, \tilde{\mathbf{\Sigma}}), \tag{3.11}$$

em que $\widetilde{\mathbf{0}} = (0,0)^t$ e $\tilde{\mathbf{\Sigma}}$ é a submatriz formada pelas duas primeiras linhas e colunas da matriz $\mathbf{\Sigma}$ dada em (3.9). Além disso, pelo Corolário 1.1

$$\sqrt{k}\left(\frac{k}{n}Y_{n-k,n} - 1\right) \xrightarrow{D} \mathcal{N}(0,1).$$

Logo, segue do conhecido método Delta (veja Teorema 1, Seção 7.5 do livro Rohatgi e Saleh (2001)), que decorre das propriedades da convergência em distribuição, que

$$\sqrt{k}\frac{\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1^{\gamma}\right]}{|\gamma|\cdot 1} = \sqrt{k}\frac{\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right]}{\gamma} \xrightarrow{D} \mathcal{N}(0,1),$$

ou seja,

$$Q_n := \sqrt{k}\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right] \xrightarrow{D} \mathcal{N}(0,\gamma^2). \tag{3.12}$$

Mas, pelo Teorema 1.9, da representação de Rényi, temos que os dois componentes de Λ_n (P_{n,β_1} e P_{n,β_2}) são independentes de Q_n e assim, a função característica conjunta de Λ_n e Q_n é o produto das funções caraterísticas de Λ_n e Q_n. Logo, por (3.11) e (3.12), segue da versão multivariada do Teorema de Levy, que para $\mathbf{t} = (t_0, t_1, t_2) \in \mathbb{R}^3$,

$$\begin{aligned}
\Phi_{\binom{\Lambda_n}{Q_n}}(\mathbf{t}) &= E(e^{i[(t_0,t_1)\cdot\Lambda_n + t_2\cdot Q_n]}) \\
&= \Phi_{\Lambda_n}(t_0,t_1)\cdot\Phi_{Q_n}(t_2) \\
&\xrightarrow{D} \Phi_{\mathcal{N}(\mathbf{0},\tilde{\mathbf{\Sigma}})}(t_0,t_1)\cdot\Phi_{\mathcal{N}(0,\gamma^2)}(t_2) \\
&= \exp\left(-\frac{(t_0,t_1)\cdot\tilde{\mathbf{\Sigma}}\cdot\binom{t_0}{t_1} + t_2^2\gamma^2}{2}\right) \\
&= \exp\left(-\frac{\mathbf{t}\cdot\mathbf{\Sigma}\cdot\mathbf{t}^t}{2}\right)
\end{aligned}$$

em que $\mathbf{t}^t$ denota a matriz transposta de $\mathbf{t}$. Portanto, novamente como consequência do Teorema de Levy multivariado obtemos (3.8). □

O próximo lema também estabelece uma propriedade assintótica de estatísticas de ordem associadas a distribuição $Pareto(1,1)$ e será utilizado para a obtenção da distribuição limite do estimador intermediário $\tilde{\theta}_{\beta,n}$.

Lema 3.2. Sejam $Y_1, Y_2, ..., Y_n$ variáveis aleatórias i.i.d com distribuição $Pareto(1,1)$ e $Y_{1,n} \leq ... \leq Y_{n,n}$ suas estatísticas de ordem. Sejam, também, $A(t)$ uma função positiva (ou negativa) tal que $A \in RV_\rho$, com $\rho < 0$, e $A_0(t)$ uma função tal que $A_0(t) \sim A(t)$, quando $t \to \infty$. Se $k = k(n) \to \infty$, $k = o(n)$ e $\sqrt{k}A(n/k) \to \lambda \in \mathbb{R}$, quando $n \to \infty$, então

$$\sqrt{k}\cdot A_0(Y_{n-k,n}) = \lambda + o_P(1)$$

e, consequentemente, $\sqrt{k}\cdot o_P(A_0(Y_{n-k,n})) = o_P(1)$

Demonstração. Podemos escrever

$$\sqrt{k}\cdot A_0(Y_{n-k,n}) = \sqrt{k}A(n/k)\cdot\frac{A_0(Y_{n-k,n})}{A_0(n/k)}\cdot\frac{A_0(n/k)}{A(n/k)}. \tag{3.13}$$

Por um lado, por hipótese, temos $\sqrt{k}A(n/k) \to \lambda \in \mathbb{R}$ e $A_0(t) \sim A(t)$, então

$$\sqrt{k}A(n/k)\,\frac{A_0(n/k)}{A(n/k)} \longrightarrow \lambda, \text{ quando } n \to \infty. \tag{3.14}$$

Por outro lado, utilizando a Proposição 1.5 (desigualdade de Potter), com $f(t) = |A_0(t)| \in$

RV_ρ, obtemos que para $\delta_1, \delta_2 > 0$ arbitrários, existe $t_0 = t_0(\delta_1, \delta_2)$ tal que para $n/k \geq t_0$ e $Y_{n-k,n} \geq t_0$ temos

$$\left|\frac{A_0(Y_{n-k,n})}{A_0(n/k)}\right| > (1-\delta_1)\left(\frac{k}{n}Y_{n-k,n}\right)^\rho \min\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\delta_2}, \left(\frac{k}{n}Y_{n-k,n}\right)^{-\delta_2}\right] \tag{3.15}$$

e

$$\left|\frac{A_0(Y_{n-k,n})}{A_0(n/k)}\right| < (1+\delta_1)\left(\frac{k}{n}Y_{n-k,n}\right)^\rho \max\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\delta_2}, \left(\frac{k}{n}Y_{n-k,n}\right)^{-\delta_2}\right]. \tag{3.16}$$

Mas, pela Proposição 1.11 temos $\frac{k}{n}Y_{n-k,n} \xrightarrow{\mathbb{P}} 1$ e como δ_1 é arbitrário, de (3.15) e (3.16) segue que

$$\frac{A_0(Y_{n-k,n})}{A_0(n/k)} \xrightarrow{P} 1. \tag{3.17}$$

Portanto, usando (3.14) e (3.17) em (3.13), concluímos que

$$\sqrt{k} \cdot A_0(Y_{n-k,n}) \xrightarrow{P} \lambda$$

ou seja, $\sqrt{k} \cdot A_0(Y_{n-k,n}) = \lambda + o_P(1)$. □

Para facilitar a apresentação da demonstração do Teorema 3.2, que estabelece a distribuição limite do estimador intermediário $\tilde{\theta}_{\beta,n}$, destacamos uma parte dessa demonstração e apresentamos no lema a seguir, no qual obtém-se uma decomposição assintótica relacionada às estatísticas de ordem associadas a uma amostra de uma distribuição $Pareto(1,1)$.

Lema 3.3. Sejam $Y_1, Y_2, ..., Y_n$ variáveis aleatórias i.i.d com distribuição $Pareto(1,1)$ e $Y_{1,n} \leq ... \leq Y_{n,n}$ suas estatísticas de ordem. Se F é uma função de distribuição, com respectiva função quantil caudal U, satisfazendo a condição (SOC), para algum $\gamma > 0$, $\rho < 0$ e alguma função positiva ou negativa $A(t)$, então, para $k, n \to \infty$ com $k/n \to 0$ e $\sqrt{k}A(\frac{n}{k}) \to \lambda \in \mathbb{R}$, temos

$$\begin{aligned}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^\beta &= \frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta} + \beta A_0(Y_{n-k,n})\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^\rho - 1}{\rho} \\ &\quad + o_{\mathbb{P}(A_0(Y_{n-k,n}))}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}.\end{aligned} \tag{3.18}$$

Demonstração. Sejam γ e ρ dados pela condição (SOC) e $A_0(t) \sim A(t)$, quando $t \to \infty$, podemos escrever

$$
\begin{aligned}
\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})} &= \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} + A_0(Y_{n-k,n}) \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \\
&\quad + \frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})} - \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} - A_0(Y_{n-k,n}) \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \\
&= \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \left\{ 1 + A_0(Y_{n-k,n}) \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \right\} \\
&\quad + A_0(Y_{n-k,n}) \left\{ \frac{\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})} - \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma}}{A_0(Y_{n-k,n})} - \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \right\} \\
&=: \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \left\{ 1 + A_0(Y_{n-k,n}) \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \right\} + A_0(Y_{n-k,n}) R_{j,k},
\end{aligned}
$$

em que

$$
R_{j,k} = \left\{ \frac{\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})} - \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma}}{A_0(Y_{n-k,n})} - \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma} \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} \right\}.
$$

Logo,

$$
\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} = \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta} \left\{ 1 + A_0(Y_{n-k,n}) \frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1}{\rho} + A_0(Y_{n-k,n}) \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{-\gamma} R_{j,k} \right\}^{\beta}. \tag{3.19}
$$

Agora, usando a desigualdade (1.16), do Teorema 1.8, considerando $x = \frac{Y_{n-j+1,n}}{Y_{n-k,n}}$ e $t = Y_{n-k,n}$ e como, pela Proposição 1.11, temos $Y_{n-k,n} \xrightarrow{P} \infty$, então $\forall \varepsilon, \delta > 0$ e $j \in \{1, ..., k\}$, existe n suficientemente grande tal que, com probabilidade arbitrariamente grande, temos

$$
|R_{j,k}| \leq \varepsilon \cdot \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma+\rho} \max\left\{ \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\delta}, \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{-\delta} \right\} = \varepsilon \cdot \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma+\rho+\delta}.
$$

Mas, $\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right) > 1$, então tomando $0 < \delta < -\rho$, obtemos que

$$
\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{-\gamma} |R_{j,k}| \leq \varepsilon.
$$

Assim, de (3.19), segue que

$$\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} = \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\left\{1 + A_0(Y_{n-k,n})\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho} + A_0(Y_{n-k,n})o_{\mathbb{P}}(1)\right\}^{\beta},$$

com o termo $o_{\mathbb{P}}(1)$ uniforme em $j \in \{1, ..., k\}$. Utilizando a expansão de Taylor com resto infinitesimal, segue que

$$\begin{aligned}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} &= \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\left\{1 + \beta A_0(Y_{n-k,n})\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho} + \beta A_0(Y_{n-k,n})o_{\mathbb{P}}(1)\right. \\ &\left. \quad + o_{\mathbb{P}}\left(A_0(Y_{n-k,n})\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho} + A_0(Y_{n-k,n})o_{\mathbb{P}}(1)\right)\right\}\end{aligned}$$

e, como $\frac{Y_{n-j+1,n}}{Y_{n-k,n}} \geq 1$ e $\rho < 0$, podemos escrever

$$\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} = \left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\left\{1 + \beta A_0(Y_{n-k,n})\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho} + o_{\mathbb{P}}(A_0(Y_{n-k,n}))\right\}.$$

Logo, como $o_{\mathbb{P}}(1)$ é um termo uniforme em $j \in \{1, ..., k\}$, obtemos (3.18). □

Finalizamos esta Seção, com um lema auxiliar para a demonstração do principal resultado deste capítulo (Teorema 3.3.), que estabelece a normalidade assintótica do estimador $\hat{\theta}_{\beta,p}$, que foi obtida, sob condições desejáveis, por Goergebeur et al. (2022).

Lema 3.4. Seja F uma função de distribuição, com respectiva função quantil caudal U, satisfazendo a condição (SOC), para algum $\gamma > 0$, $\rho < 0$ e alguma função positiva ou negativa $A(t)$. Suponha que $k = k(n) \to \infty$, $k = o(n)$ e $\sqrt{k}A(\frac{n}{k}) \to \lambda \in \mathbb{R}$, quando $n \to \infty$.

(a) Se $\hat{\gamma_k}$ um estimador para γ, tal que $\sqrt{k}(\hat{\gamma_k} - \gamma) \xrightarrow{D} \Gamma$, com Γ tendo distribuição não degenerada, e p é tal que $\frac{k}{np} \to \infty$ e $\frac{\ln(k/(np))}{\sqrt{k}} \to 0$, quando $n \to \infty$, então

$$\frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta} - 1\right] \xrightarrow{D} \beta\Gamma \tag{3.20}$$

(b) Se p é tal que $\frac{k}{np} \to \infty$, então

$$\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta} - 1\right\} \xrightarrow{n\to\infty} \frac{\beta}{\rho}\lambda \tag{3.21}$$

Demonstração. (a) Para provar (3.20), podemos escrever

$$\begin{aligned}
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}-1\right] &= \sqrt{k}\left(\frac{\exp\left\{(\hat{\gamma}_k-\gamma)\beta\cdot\ln(k/np)\right\}-1}{\ln\frac{k}{np}}\right)\\
&= \sqrt{k}(\hat{\gamma}_k-\gamma)\beta\left(\frac{\exp\left\{\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot\frac{\ln(k/np)}{\sqrt{k}}\right\}-1}{\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot\frac{\ln(k/np)}{\sqrt{k}}}\right).
\end{aligned}$$

Agora, por hipótese, temos que $\sqrt{k}(\hat{\gamma}_k-\gamma)\xrightarrow{D}\Gamma$ e $\frac{\ln(k/(np))}{\sqrt{k}}\to 0$, então

$$\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\frac{\ln(k/np)}{\sqrt{k}}\xrightarrow{P}0.$$

Com isso, utilizando a seguinte função contínua

$$g(x)=\begin{cases}\dfrac{e^x-1}{x}, & \text{se } x\neq 0\\ 1, & \text{se } x=0\end{cases}$$

e o fato que funções contínuas preservam a convergência em probabilidade, obtemos que

$$\frac{\exp\left\{\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot\frac{\ln(k/np)}{\sqrt{k}}\right\}-1}{\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot\frac{\ln(k/np)}{\sqrt{k}}}=g\left(\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\frac{\ln(k/np)}{\sqrt{k}}\right)\xrightarrow{P}g(0)=1.$$

Logo, como $\sqrt{k}(\hat{\gamma}_k-\gamma)\xrightarrow{D}\Gamma$, segue das propriedades de convergência em distribuição que

$$\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}-1\right)=\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot g\left(\sqrt{k}(\hat{\gamma}_k-\gamma)\beta\cdot\frac{\ln(k/np)}{\sqrt{k}}\right)\xrightarrow{D}\beta\Gamma.$$

(b) Para provar (3.21), primeiramente escrevemos

$$\begin{aligned}
\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\} &= \sqrt{k}\left\{\left\{1+\cdot\frac{U(1/p)}{U(n/k)}\left(\frac{np}{k}\right)^{\gamma}-1\right\}^{-\beta}-1\right\}\\
&= \sqrt{k}\left\{\left\{1+\left(\frac{np}{k}\right)^{\gamma}\cdot\left[\frac{U(1/p)}{U(n/k)}-\left(\frac{k}{np}\right)^{\gamma}\right]\right\}^{-\beta}-1\right\}\\
&= \sqrt{k}\left\{\left\{1+\left(\frac{np}{k}\right)^{\gamma}A(n/k)\cdot\left[\frac{\frac{U(1/p)}{U(n/k)}-\left(\frac{k}{np}\right)^{\gamma}}{A(n/k)}\right]\right\}^{-\beta}-1\right\}.
\end{aligned}$$

Agora, utilizando (1.15) da condição (SOC), com $t=n/k$ e $x=k/np$, obtemos que

$$\frac{\frac{U(1/p)}{U(n/k)}-\left(\frac{k}{np}\right)^{\gamma}}{A(n/k)}=\left(\frac{k}{np}\right)^{\gamma}\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o(1)$$

e, consequentemente,

$$\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\}=\sqrt{k}\left\{\left[1+A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(A\left(\frac{n}{k}\right)\right)\right]^{-\beta}-1\right\}.$$

Usando a fórmula de Taylor com resto infinitesimal, segue que

$$\begin{aligned}\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\} &= \sqrt{k}\left\{-\beta\left[A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(A\left(\frac{n}{k}\right)\right)\right]+\right.\\ &\left.+\,o\left(A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(A\left(\frac{n}{k}\right)\right)\right)\right\}.\end{aligned}$$

Mas como $k/np \to \infty$ e $\rho < 0$ podemos verificar que

$$o\left(A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(A\left(\frac{n}{k}\right)\right)\right)=o\left(A\left(\frac{n}{k}\right)\right).$$

Logo,

$$\begin{aligned}\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\} &= \sqrt{k}\left\{-\beta A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(A\left(\frac{n}{k}\right)\right)\right\}\\ &= -\beta\sqrt{k}A\left(\frac{n}{k}\right)\frac{\left(\frac{k}{np}\right)^{\rho}-1}{\rho}+o\left(\sqrt{k}A\left(\frac{n}{k}\right)\right).\end{aligned}$$

Finalmente, como, por hipótese, $\sqrt{k}A(n/k) \to \lambda$, $k/np \to \infty$ e $\rho < 0$, podemos concluir que

$$\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\}\xrightarrow{n\to\infty}\frac{\beta}{\rho}\lambda.$$

□

3.3 Propriedades Assintóticas do Estimador

Nesta Seção, apresentamos os resultados centrais de Goergebeur et al. (2022) que estabelecem, sob condições apropriadas, as distribuições limites dos estimadores $\tilde{\theta}_{\beta,n}$ e $\hat{\theta}_{\beta,p}$, definidos em (3.4) e (3.5), respectivamente.

Primeiramente, apresentamos o teorema que obtém, sob a condição de segunda ordem (SOC), uma expansão assintótica do CTM, que não se restringe somente ao caso intermediário, sendo válida para qualquer $p \downarrow 0$.

Teorema 3.1. Se X satisfaz a condição (SOC), com $0 < \gamma < 1/\beta$ e $\rho < 0$, então para $p \downarrow 0$ temos

$$\theta_{\beta,p} = \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta} + \frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-p)}A\left(\frac{1}{p}\right) + o\left(A\left(\frac{1}{p}\right)\right)\right\}.$$

Demonstração. Primeiramente, das propriedades de esperança condicional e da integral de Lebesgue-Stieltjes, podemos escrever

$$\begin{aligned}
\theta_{\beta,p} &= \mathbb{E}\left(X^{\beta}|X > U(1/p)\right)\\
&= \frac{\int_{\Omega} X^{\beta}\cdot I_{X>U(1/p)}\mathrm{d}P}{1-P(X\leq U(1/p))}\\
&= \frac{\displaystyle\int_{U(1/p)}^{\infty} x^{\beta}\mathrm{d}F(x)}{\overline{F}(U(1/p))}\\
&= \frac{-1}{\overline{F}(U(1/p))}\cdot\int_{U(1/p)}^{\infty} x^{\beta}\mathrm{d}\overline{F}(x).
\end{aligned}$$

Agora, fazendo a mudança de variáveis $u = \frac{x}{U(1/p)}$ e usando a propriedade de integração por partes para a integral de Lebesgue-Stieltjes obtemos

$$\begin{aligned}
\theta_{\beta,p} &= \frac{-[U(1/p)]^{\beta}}{\overline{F}(U(1/p))}\cdot\int_{1}^{\infty} u^{\beta}\cdot\mathrm{d}\overline{F}(u\cdot U(1/p))\\
&= \frac{-[U(1/p)]^{\beta}}{\overline{F}(U(1/p))}\cdot\left[\lim_{u\to\infty} u^{\beta}\cdot\overline{F}(u\cdot U(1/p)) - \overline{F}(U(1/p)) - \right.\\
&\quad\left. - \beta\int_{1}^{\infty} u^{\beta-1}\cdot\overline{F}(u\cdot U(1/p))\mathrm{d}u\right].
\end{aligned}$$

Mas, pela Proposição 1.6, temos que $\lim_{u\to\infty} u^{\beta}\cdot\overline{F}(u\cdot U(1/p)) = 0$, então segue que

$$\theta_{\beta,p} = \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{1+\beta\int_{1}^{\infty} u^{\beta-1}\frac{\overline{F}(uU(1/p))}{\overline{F}(U(1/p))}\mathrm{d}u\right\}.$$

Agora, pelo Teorema 1.8, podemos escrever

$$\begin{aligned}
\theta_{\beta,p} &= \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{1+\beta\int_1^{\infty}u^{\beta-1}u^{-1/\gamma}\left\{1+\alpha_0(U(1/p))\left(\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right)\right.\right.\\
&\quad\left.\left.+\ \frac{\alpha_0(U(1/p))}{u^{-1/\gamma}}\left[\frac{\frac{\overline{F}(uU(1/p))}{\overline{F}(U(1/p))}-u^{-1/\gamma}}{\alpha_0(U(1/p))}-u^{-1/\gamma}\left(\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right)\right]\right\}\mathrm{d}u\right\}\\
&= \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{1+\beta\int_1^{\infty}u^{\beta-\frac{1}{\gamma}-1}\mathrm{d}u+\beta\alpha_0(U(1/p))\int_1^{\infty}u^{\beta-\frac{1}{\gamma}-1}\left(\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right)\mathrm{d}u\right.\\
&\quad\left.+\ \beta\alpha_0(U(1/p))\int_1^{\infty}u^{\beta-1}\left[\frac{\frac{\overline{F}(uU(1/p))}{\overline{F}(U(1/p))}-u^{-1/\gamma}}{\alpha_0(U(1/p))}-u^{-1/\gamma}\left(\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right)\right]\mathrm{d}u\right\},
\end{aligned}$$

em que $\alpha_0(.)$ é um função tal que $\alpha_0(.)\sim\alpha(.)$. Agora, como por hipótese, $0<\gamma<1/\beta$, ou seja, $\beta-1/\gamma<0$, segue que

$$\theta_{\beta,p}=\left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-\rho)}\alpha_0(U(1/p))+\alpha_0(U(1/p))R_p\right\},$$

com

$$R_p:=\beta\int_1^{\infty}u^{\beta-1}\left[\frac{\frac{\overline{F}(uU(1/p))}{\overline{F}(U(1/p))}-u^{-1/\gamma}}{\alpha_0(U(1/p))}-u^{-1/\gamma}\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right]\mathrm{d}u$$

Mas, por (1.16) para $\varepsilon>0$, $0<\delta<\frac{1-\rho}{\gamma}-\beta$ e $p<p_0(\varepsilon,\delta)$, podemos obter que

$$\begin{aligned}
|R_p| &\leq \beta\int_1^{\infty}u^{\beta-1}\left|\frac{\frac{\overline{F}(uU(1/p))}{\overline{F}(U(1/p))}-u^{-1/\gamma}}{\alpha_0(U(1/p))}-u^{-1/\gamma}\frac{u^{\rho/\gamma}-1}{\gamma\rho}\right|\mathrm{d}u\\
&\leq \beta\int_1^{\infty}u^{\beta-1}\cdot\varepsilon\cdot u^{-\frac{1}{\gamma}+\frac{\rho}{\gamma}}\cdot\max(u^{\delta},u^{-\delta})\mathrm{d}u\\
&\leq \varepsilon\beta\int_1^{\infty}u^{\beta-1}u^{-\frac{1}{\gamma}+\frac{\rho}{\gamma}+\delta}\mathrm{d}u\\
&\leq \varepsilon\frac{\gamma\beta}{1-\rho-\gamma\beta-\gamma\delta}.
\end{aligned}$$

Assim, $R_p=o(1)$, para $p\downarrow 0$ e, consequentemente, temos

$$\theta_{\beta,p}=\left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-p)}\alpha_0(U(1/p))+o(\alpha_0(U(1/p)))\right\}. \tag{3.22}$$

Agora, como, por definição temos $\alpha(t):=A\left(\frac{1}{\overline{F}(t)}\right)$ e $U(1/p)=F^{\leftarrow}(1-p))$, podemos escrever

$$\begin{aligned}\alpha_0(U(1/p)) &= \frac{\alpha_0(U(1/p))}{\alpha(U(1/p))}\cdot A\left(\frac{1}{1-F(F^{\leftarrow}(1-p))}\right)\\ &= \frac{\alpha_0(U(1/p))}{\alpha(U(1/p))}\cdot A\left(\frac{1}{p}\right).\end{aligned}$$

Mas, como α e U são funções regularmente variantes e, além disso, $\alpha_0 \sim \alpha$, temos que $\frac{\alpha_0(U(1/p))}{\alpha(U(1/p))} = 1 + o(1)$. Logo,

$$\alpha_0(U(1/p)) = A\left(\frac{1}{p}\right) + o\left(A\left(\frac{1}{p}\right)\right).$$

Finalmente, substituindo este resultado em (3.22) e utilizando a Proposição 1.1, concluímos que

$$\begin{aligned}\theta_{\beta,p} &= \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-p)}\left[A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)\right]\right.\\ &\quad\left.+o\left[A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)\right]\right\}\\ &= \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-p)}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)\right.\\ &\quad\left.+o\left(A\left(\frac{1}{p}\right)\right)\right\}\\ &= \left[U\left(\frac{1}{p}\right)\right]^{\beta}\left\{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-p)}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)\right\}.\end{aligned}$$

□

Como consequência do teorema anterior, podemos obter uma relação assintótica entre o caso intermediário e o caso geral.

Corolário 3.1. Sejam $X_1, ..., X_n$ variáveis aleatórias i.i.d. com função de distribuição satisfazendo a condição (SOC), com $0 < \gamma < 1/(2\beta)$ e $\rho < 0$. Então para $k, n \to \infty$ com $k/n \to 0$ e $\sqrt{k}A(\frac{n}{k}) \to \lambda \in \mathbb{R}$, e p tal que $\frac{k}{np} \to \infty$, temos

$$\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,k/n}}{\theta_{\beta,p}} \xrightarrow{n\to\infty} 1 \tag{3.23}$$

Demonstração. Podemos escrever

$$
\begin{aligned}
\left(\frac{k}{np}\right)^{\gamma\beta} \cdot \frac{\theta_{\beta,k/n}}{\theta_{\beta,p}} &= \left(\frac{k}{np}\right)^{\gamma\beta} \left(\frac{U(n/k)}{U(1/p)}\right)^{\beta} \cdot \frac{\theta_{\beta,k/n}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}} \\
&= \left\{ \frac{\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta} - 1\right]}{\sqrt{k}} + 1 \right\} \cdot \frac{\theta_{\beta,k/n}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}.
\end{aligned}
$$

Então, utilizando o Teorema 3.1 e o Lema 3.4 é possível mostrar que

$$
\frac{\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta} - 1\right]}{\sqrt{k}} \xrightarrow{n\to\infty} 0 \ \text{ e } \ \frac{\theta_{\beta,k/n}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}} \xrightarrow{n\to\infty} 1
$$

pois $p < 1/n$. Portanto,

$$
\left(\frac{k}{np}\right)^{\gamma\beta} \cdot \frac{\theta_{\beta,k/n}}{\theta_{\beta,p}} \xrightarrow{n\to\infty} 1
$$

□

O resultado principal para o estimador intermediário do CTM é dado no teorema abaixo.

Teorema 3.2. Sejam $X_1, ..., X_n$ variáveis aleatórias i.i.d. com função de distribuição F satisfazendo a condição (SOC), com $0 < \gamma < 1/(2\beta)$ e $\rho < 0$. Então para $k, n \to \infty$ com $k/n \to 0$ e $\sqrt{k}A(\frac{n}{k}) \to \lambda \in \mathbb{R}$, temos

$$
\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,k/n}} - 1\right) \xrightarrow{D} (1-\gamma\beta)P_{\beta} + \beta Q \sim \mathcal{N}\left(0, \frac{2(\gamma\beta)^2(1-\gamma\beta)}{1-2\gamma\beta}\right). \tag{3.24}
$$

Demonstração. Primeiramente, por definição, temos

$$
\tilde{\theta}_{\beta,n} = \frac{1}{k}\sum_{j=1}^{k} X_{n-j+1,n}^{\beta} = \frac{1}{k} X_{n-k,n}^{\beta} \sum_{j=1}^{k} \left(\frac{X_{n-j+1,n}}{X_{n-k,n}}\right)^{\beta}. \tag{3.25}
$$

Seja $U(x) = \left(\dfrac{1}{1-F}\right)^{\leftarrow}(x) = F^{\leftarrow}\left(1 - \frac{1}{x}\right)$, $x > 1$, a função quantil caudal de F e considere $Y_1, Y_2, ..., Y_n$ variáveis aleatórias i.i.d., com distribuição Pareto(1,1), ou seja, $F_Y(x) = 1 - \frac{1}{x}, x > 1,$. Então, como $F_Y^{-1}(y) = \frac{1}{1-y}$, temos que $X_k \stackrel{D}{=} U(Y_k)$, $1 \leq k \leq n$, pois

$$\begin{aligned} P(U(Y_k) \leq x) &= P\left(F^{\leftarrow}\left(1-\frac{1}{Y_k}\right) \leq x\right) = P\left(1-\frac{1}{Y_k} \leq F(x)\right) \\ &= P\left(Y_k \leq F_Y^{-1}(F(x))\right) \\ &= F(x). \end{aligned}$$

Assim, pela independência das variáveis aleatórias é possível concluir que

$$(X_{n-k,n}, X_{n-k+1,n}, ..., X_{n,n}) \stackrel{D}{=} (U(Y_{n-k,n}), U(Y_{n-k+1,n}), ..., U(Y_{n,n})). \tag{3.26}$$

Logo, por (3.25) obtemos que

$$\tilde{\theta}_{\beta,n} \stackrel{D}{=} \frac{1}{k} U(Y_{n-k,n})^{\beta} \sum_{j=1}^{k} \left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}. \tag{3.27}$$

Para calcular o limite do Teorema 3.2, definimos

$$\breve{\theta}_{\beta,\frac{k}{n}} := \frac{1}{1-\gamma\beta}\left[U\left(\frac{n}{k}\right)\right],$$

e podemos decompor

$$\begin{aligned} \sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right) &= \frac{\sqrt{k}\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}} - \sqrt{k} \\ &= \sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) + \frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}\sqrt{k}\left(\frac{\breve{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right). \end{aligned} \tag{3.28}$$

Por um lado, o primeiro termo da decomposição pode ser reescrito da seguinte forma

$$\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) \stackrel{d}{=} \sqrt{k}(1-\gamma\beta)\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}\right)^{\beta} \frac{1}{k} \cdot \sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} - \sqrt{k}$$

e expandindo $\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}\right)^{\beta}$ pela fórmula de Taylor com resto infinitesimal, obtemos

$$\begin{aligned} \sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) \stackrel{D}{=}\; & \sqrt{k}(1-\gamma\beta)\frac{1}{k} \cdot \sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} - \sqrt{k} \\ & +\beta\sqrt{k}\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}-1\right)(1-\gamma\beta)\frac{1}{k} \cdot \sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} \\ & +\sqrt{k}o_{\mathbb{P}}\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}-1\right)(1-\gamma\beta)\frac{1}{k} \cdot \sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}. \end{aligned} \tag{3.29}$$

Mas, podemos escrever

$$\begin{aligned}\sqrt{k}\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}-1\right) &= \sqrt{k}A\left(\frac{n}{k}\right)\frac{\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}\right)}{A\left(\frac{n}{k}\right)}-\sqrt{k}\left(\frac{k}{n}\cdot Y_{n-k,n}\right)^{\gamma}+\sqrt{k}\left(\frac{k}{n}\cdot Y_{n-k,n}\right)^{\gamma}-\sqrt{k}\\ &= \sqrt{k}A\left(\frac{n}{k}\right)\left(\frac{\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}-\left(\frac{k}{n}\cdot Y_{n-k,n}\right)^{\gamma}}{A\left(\frac{n}{k}\right)}\right)+\sqrt{k}\left[\left(\frac{k}{n}\cdot Y_{n-k,n}\right)^{\gamma}-1\right]\end{aligned}$$

e, usando a condição (SOC) em (1.15), a hipótese que $\sqrt{k}A\left(\frac{n}{k}\right)\to\lambda$ e o fato que $\frac{k}{n}Y_{n-k,n}\xrightarrow{\mathbb{P}}1$ (Proposição 1.11), chegamos que

$$\begin{aligned}\sqrt{k}\left(\frac{U(Y_{n-k,n})}{U(\frac{n}{k})}-1\right) &\overset{D}{=} \sqrt{k}A\left(\frac{n}{k}\right)\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\frac{\left(\frac{k}{n}Y_{n-k,n}\right)^{\rho}-1}{\rho}+Q_n+o_{\mathbb{P}}(1)\\ &\overset{D}{=} Q_n+o_{\mathbb{P}}(1),\end{aligned}\tag{3.30}$$

em que $Q_n:=\sqrt{k}\left[\left(\frac{k}{n}\cdot Y_{n-k,n}\right)^{\gamma}-1\right]$. Substituindo (3.30) em (3.29) obtemos

$$\begin{aligned}\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) &\overset{D}{=} (1-\gamma\beta)\sqrt{k}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}-\frac{1}{1-\gamma\beta}\right]\\ &\quad+\beta\left[Q_n+o_{\mathbb{P}}(1)\right]+o_{\mathbb{P}}(1)\\ &\overset{D}{=} (1-\gamma\beta)\sqrt{k}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}-\frac{1}{1-\gamma\beta}\right]+\beta Q_n+o_{\mathbb{P}}(1).\end{aligned}$$

Agora, utilizando (3.18) obtido no Lema 3.3, podemos reescrever esta igualdade como

$$\begin{aligned}\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) &\overset{D}{=} (1-\gamma\beta)\sqrt{k}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}-\frac{1}{1-\gamma\beta}\right]\\ &\quad+\beta(1-\gamma\beta)\sqrt{k}A_0(Y_{n-k,n})\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho}\\ &\quad+(1-\gamma\beta)\sqrt{k}o_{\mathbb{P}}(A_0(Y_{n-k,n}))+\beta Q_n+o_{\mathbb{P}}(1).\end{aligned}$$

Mas, pelo Lema 3.2 e pela Proposição 1.1, obtemos

$$\begin{aligned}\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) &\overset{D}{=} (1-\gamma\beta)\sqrt{k}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\beta}-\frac{1}{1-\gamma\beta}\right]+\\ &\quad+\beta(1-\gamma\beta)\left(\lambda+o_P(1)\right)\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho}+\beta Q_n+o_{\mathbb{P}}(1).\end{aligned}\tag{3.31}$$

Agora, pela Proposição 1.9 e pela Lei Forte dos Grandes Números de Kolmogorov, como $Y_j\sim Pareto(1,1)$, $0<\gamma\beta<1/2$ e $\rho<0$, podemos concluir que

$$\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\overset{D}{=}\frac{1}{k}\sum_{j=1}^{k}Y_j^{\gamma\beta}\xrightarrow{\mathbb{P}}\mathbb{E}(Y_j^{\gamma\beta})=\frac{1}{1-\gamma\beta}\tag{3.32}$$

e

$$\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\gamma\beta}\frac{\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho}-1}{\rho} \stackrel{D}{=} \frac{\beta}{\rho}\left[\frac{1}{k}\sum_{j=1}^{k}Y_j^{\gamma\beta+\rho}-\frac{1}{k}\sum_{j=1}^{k}Y_j^{\gamma\beta}\right]$$
$$\stackrel{\mathbb{P}}{\to} \frac{\beta}{(1-\gamma\beta-\rho)(1-\gamma\beta)}. \tag{3.33}$$

Além disso, por (3.32), (3.33) e como $A_0 \sim A$, com $\lim_{t\to\infty} A(t) = 0$, segue de (3.18) que

$$\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta} \stackrel{\mathbb{P}}{\to} \frac{1}{1-\gamma\beta} \tag{3.34}$$

Assim, usando (3.33) em (3.31), segue que

$$\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right) \stackrel{D}{=} (1-\gamma\beta)P_{n,\beta}+\beta Q_n+\frac{\beta}{1-\gamma\beta-\rho}\lambda+o_{\mathbb{P}}(1), \tag{3.35}$$

em que $P_{n,\beta} := \sqrt{k}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\beta}-\frac{1}{1-\gamma\beta}\right]$.

Por outro lado, pelo Teorema 3.1, o segundo termo da decomposição (3.28) pode ser escrito da seguinte maneira

$$\begin{aligned}
\sqrt{k}\left(\frac{\breve{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right) &= \sqrt{k}\left[\frac{\frac{[U(n/k)]^{\beta}}{1-\gamma\beta}}{[U(n/k)]^{\beta}\left(\frac{1}{1-\gamma\beta}+\frac{\beta\cdot A(n/k)}{(1-\gamma\beta)(1-\gamma\beta-\rho)}+o(A(n/k))\right)}-1\right]\\
&= \sqrt{k}\left[\frac{1-\gamma\beta-\rho}{(1-\gamma\beta-\rho)+\beta A(n/k)+(1-\gamma\beta)(1-\gamma\beta-\rho)o(A(n/k))}-1\right]\\
&= \frac{-\beta\sqrt{k}A(n/k)-(1-\gamma\beta)(1-\gamma\beta-\rho)\sqrt{k}\cdot o(A(n/k))}{(1-\gamma\beta-\rho)+\beta A(n/k)+(1-\gamma\beta)(1-\gamma\beta-\rho)o(A(n/k))}.
\end{aligned}$$

Agora, como $\sqrt{k}o\left(A(n/k)\right) = \sqrt{k}A(n/k)\frac{o(A(n/k))}{A(n/k)} \to 0$ e $\lim_{t\to\infty} A(t) = 0$ concluímos que

$$\sqrt{k}\left(\frac{\breve{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right) = \frac{-\beta}{1-\gamma\beta-\rho}\cdot\lambda+o(1) \tag{3.36}$$

Assim, usando (3.35) e (3.36) podemos reescrever (3.28) como

$$\begin{aligned}
\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right) &= \sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}-1\right)+\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}\sqrt{k}\left(\frac{\breve{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right) \\
&\stackrel{D}{=} (1-\gamma\beta)P_{n,\beta}+\beta Q_n+\frac{\beta}{1-\gamma\beta-\rho}\cdot\lambda+o_{\mathbb{P}}(1) \\
&\quad +\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}}\left(\frac{-\beta}{1-\gamma\beta-\rho}\cdot\lambda+o(1)\right) \\
&\stackrel{D}{=} (1-\gamma\beta)P_{n,\beta}+\beta Q_n+\frac{\beta}{1-\gamma\beta-\rho}\cdot\lambda+o_{\mathbb{P}}(1) \\
&\quad +\left[(1-\gamma\beta)\left(\frac{U(Y_{n-k,n})}{U(n/k)}\right)^{\beta}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}\right]\left(\frac{-\beta}{1-\gamma\beta-\rho}\cdot\lambda+o(1)\right),
\end{aligned}$$

pois $\frac{\tilde{\theta}_{\beta,n}}{\breve{\theta}_{\beta,\frac{k}{n}}} \stackrel{D}{=} (1-\gamma\beta)\left(\frac{U(Y_{n-k,n})}{U(n/k)}\right)^{\beta}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}$. Mas, pelos resultados (3.26), (3.34) e pela Proposição 1.10, temos que

$$\left[(1-\gamma\beta)\left(\frac{U(Y_{n-k,n})}{U(n/k)}\right)^{\beta}\frac{1}{k}\sum_{j=1}^{k}\left(\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}\right)^{\beta}\right]\left(\frac{-\beta}{1-\gamma\beta-\rho}\cdot\lambda+o(1)\right)\xrightarrow{P}\frac{-\beta}{1-\gamma\beta-\rho}\cdot\lambda.$$

Portanto,

$$\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right)\xrightarrow{D}(1-\gamma\beta)P_{\beta}+\beta Q,$$

em que Q e P_β são independentes, pela representação de Rényi, $Q\sim\mathcal{N}(0,\gamma^2)$ e $P_\beta\sim\mathcal{N}\left(0,\frac{(\gamma\beta)^2}{(1-2\gamma\beta)(1-\gamma\beta)^2}\right)$, pelo Lema 3.1. Ou seja,

$$\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right)\xrightarrow{D}(1-\gamma\beta)P_{\beta}+\beta Q\sim\mathcal{N}\left(0,\frac{2(\gamma\beta)^2(1-\gamma\beta)}{1-2\gamma\beta}\right).$$

□

Finalmente, conforme mencionamos na introdução deste capítulo, pelo Teorema 3.1 temos que para $p\downarrow 0$

$$\frac{\theta_{\beta,p}/(U(1/p))^{\beta}}{\theta_{\beta,k/n}/(U(n/k))^{\beta}}\to 1.$$

Mas, como, pela condição (SOC) temos que $U\in RV_\gamma$, então $\lim_{n\to\infty}\frac{U(1/p)}{U(n/k)}=\left(\frac{k}{np}\right)^{\gamma}$ e, consequentemente, segue que

$$\theta_{\beta,p}\sim\left(\frac{(U(1/p))}{(U(n/k))}\right)^{\beta}\theta_{\beta,k/n}\sim\left(\frac{k}{np}\right)^{\beta\gamma}\theta_{\beta,k/n}.$$

Assim, Goegebeur et al. (2022) propuseram o estimador extrapolado do CTM, $\hat{\theta}_{\beta,p}$, definido em (3.5) e que foi obtido da relação acima, estimando-se $\theta_{\beta,k/n}$ pelo estimador do caso intermediário $\tilde{\theta}_{\beta,n}$ e substituindo-se γ por um estimador $\hat{\gamma}_k$.

Dessa forma, escolhendo-se um estimador $\hat{\gamma}_k$ de tal forma $\sqrt{k}(\hat{\gamma}_k - \gamma)$ tenha distribuição limite Γ, é possível obter uma distribuição limite para $\hat{\theta}_{\beta,p}$ a partir da distribuição Γ, conforme veremos no próximo teorema.

Uma opção para a escolha do estimador $\hat{\gamma}_k$ que satisfaça as condições necessárias foi sugerida por Goegebeur et al. (2022). Eles mencionaram o estimador de Hill, apresentado em Hill (1975), que é dado por

$$\hat{\gamma}_k^H := \frac{1}{k}\sum_{i=1}^{k} \ln X_{n-i+1,n} - \ln X_{n-k,n}.$$

Neste caso, sob a condição (SOC) e para $k, n \to \infty$, com $k/n \to 0$ e $\sqrt{k}A(n/k) \to \lambda$, a seguinte convergência é válida

$$\sqrt{k}(\hat{\gamma}_k^H - \gamma) \xrightarrow{D} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right).$$

Um outro estimador de viés-corrigido para γ, que foi também sugerido por Goegebeur et al. (2022), é o estimador Beirlant et al. (1999), cujos detalhes serão omitidos nesta dissertação.

Com isso, temos, a seguir, o resultado principal que determina a distribuição assintótica do estimador $\hat{\theta}_{\beta,p}$ devidamente padronizado.

Teorema 3.3. Sejam $X_1, ..., X_n$ variáveis aleatórias i.i.d. com função de distribuição satisfazendo a condição (SOC), com $0 < \gamma < 1/(2\beta)$ e $\rho < 0$. Então, para $k, n \to \infty$ com $k/n \to 0$ e $\sqrt{k}A(\frac{n}{k}) \to \lambda \in \mathbb{R}$, $\hat{\gamma}_k$ um estimador para γ com $\sqrt{k}(\hat{\gamma}_k - \gamma) \xrightarrow{D} \Gamma$, e p tal que $\frac{\ln(k/(np))}{\sqrt{k}} \to 0$ e $\frac{k}{np} \to \infty$, temos

$$\frac{\sqrt{k}}{\ln(k/(np))}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}} - 1\right) \xrightarrow{D} \beta\Gamma.$$

Demonstração. Por definição, temos

$$\hat{\theta}_{\beta,p} = \left(\frac{k}{np}\right)^{\beta\hat{\gamma}_k} \cdot \tilde{\theta}_{\beta,n}.$$

Com isso, nós temos a seguinte decomposição

$$
\begin{aligned}
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right) &= \frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{\hat{\gamma}_k\beta}\cdot\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,p}}-1\right]\\
&= \frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{\hat{\gamma}_k\beta}\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-\left(\frac{k}{np}\right)^{\hat{\gamma\beta}_k}\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}+\left(\frac{k}{np}\right)^{\hat{\gamma}_k\beta}\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}\right.\\
&\quad\left.-\left(\frac{k}{np}\right)^{\gamma\beta}\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}+\left(\frac{k}{np}\right)^{\gamma\beta}\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-1\right]\\
&= \frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}\right.\\
&\quad\left.+\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}+\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-1\right]\\
&= \frac{1}{\ln\frac{k}{np}}\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right)\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}} \qquad (3.37)\\
&\quad+\frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}-1\right]\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-1\right].
\end{aligned}
$$

Mas, escrevendo

$$
\begin{aligned}
\frac{1}{\ln\frac{k}{np}}\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta} &= \frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\frac{1}{\sqrt{k}}\\
&= \frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}-1\right]\cdot\frac{1}{\sqrt{k}}+\frac{1}{\ln\frac{k}{np}}
\end{aligned}
$$

e utilizando o Lema 3.4 e o fato de $k\to\infty$ quando $n\to\infty$, segue que

$$
\frac{1}{\ln\frac{k}{np}}\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\xrightarrow{P}0.
$$

Assim, usando este resultado, o Teorema 3.2 e o Lema 3.1, concluímos que

$$
\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left(\frac{\tilde{\theta}_{\beta,n}}{\theta_{\beta,\frac{k}{n}}}-1\right)\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}=o_{\mathbb{P}}(1).
$$

e substituindo este resultado em (3.37) obtemos

$$
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right)=\frac{\sqrt{k}}{\ln\frac{k}{np}}\left[\left(\frac{k}{np}\right)^{(\hat{\gamma}_k-\gamma)\beta}-1\right]\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}}{\theta_{\beta,p}}-1\right]+o_P(1).
$$

Agora, podemos usar o Lema 3.4 e o Lema 3.1 para obter

$$\begin{aligned}
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right) &= \beta\sqrt{k}(\hat{\gamma}_k-\gamma)+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}\cdot\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}-1\right]+o_P(1)\\
&= \beta\sqrt{k}(\hat{\gamma}_k-\gamma)+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left[\left(\frac{k}{np}\right)^{\gamma\beta}\cdot\frac{\theta_{\beta,\frac{k}{n}}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}\cdot\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}-\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\right.\\
&\quad\left.+\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right]+o_P(1)\\
&= \beta\sqrt{k}(\hat{\gamma}_k-\gamma)+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left[\frac{\theta_{\beta,\frac{k}{n}}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}-1\right]\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}\\
&\quad+\frac{1}{\ln\frac{k}{np}}\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\}+o_{\mathbb{P}}(1).
\end{aligned}$$

Assim, denotando

$$T_{1,k}=\sqrt{k}\left[\frac{\theta_{\beta,\frac{k}{n}}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}-1\right]\ \text{ e }\ T_{2,k}=\sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\},$$

podemos escrever

$$\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right)=\beta\sqrt{k}(\hat{\gamma}_k-\gamma)+\frac{1}{\ln\frac{k}{np}}T_{1,k}\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}+\frac{1}{\ln\frac{k}{np}}T_{2,k}+o_{\mathbb{P}}(1). \tag{3.38}$$

Por um lado, pelo Teorema 3.1, nós temos

$$\begin{aligned}
T_{1,k} &:= \sqrt{k}\left[\frac{\theta_{\beta,\frac{k}{n}}/\left[U(n/k)\right]^{\beta}}{\theta_{\beta,p}/\left[U(1/p)\right]^{\beta}}-1\right]\\
&= \sqrt{k}\left[\frac{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-\rho)}A\left(\frac{n}{k}\right)+o\left(A\left(\frac{n}{k}\right)\right)}{\frac{1}{1-\gamma\beta}+\frac{\beta}{(1-\gamma\beta)(1-\gamma\beta-\rho)}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)}-1\right]
\end{aligned}$$

e utilizando a Proposição 1.1 obtemos

$$\begin{aligned}
T_{1,k} &= \sqrt{k}\left[\frac{1+\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{n}{k}\right)+o\left(A\left(\frac{n}{k}\right)\right)}{1+\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)}-1\right]\\
&= \sqrt{k}\left[\frac{\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{n}{k}\right)+o\left(A\left(\frac{n}{k}\right)\right)-\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)}{1+\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)}\right].
\end{aligned}$$

Mas, como $\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right)+o\left(A\left(\frac{1}{p}\right)\right)=o(1)$, já que $1/p\xrightarrow{n\to\infty}\infty$ e $\lim_{t\to\infty}A(t)=0$, temos

$$T_{1,k} = \sqrt{k}\left[\frac{\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{n}{k}\right) + o\left(A\left(\frac{n}{k}\right)\right) - \frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right) + o\left(A\left(\frac{1}{p}\right)\right)}{1+o(1)}\right]. \tag{3.39}$$

Agora, temos

$$\frac{\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right) + o\left(A\left(\frac{1}{p}\right)\right)}{A\left(\frac{n}{k}\right)} = \frac{\beta}{1-\gamma\beta-\rho}\cdot\frac{A\left(\frac{1}{p}\right)}{A\left(\frac{n}{k}\right)} + \frac{o\left(A\left(\frac{1}{p}\right)\right)}{A\left(\frac{1}{p}\right)}\cdot\frac{A\left(\frac{1}{p}\right)}{A\left(\frac{n}{k}\right)}, \tag{3.40}$$

e como $A \in RV_\rho$, podemos usar a Proposição 1.5 (desigualdade de Potter), com δ_1 e δ_2 arbitrários, para obter que

$$(1-\delta_1)\left(\frac{k}{np}\right)^{\rho-\delta_2} < \frac{A(1/p)}{A(n/k)} < (1+\delta_1)\left(\frac{k}{np}\right)^{\rho+\delta_2}$$

como $\rho < 0$ e $\frac{k}{np} \to \infty$, por hipótese, podemos tomar $0 < \delta_2 < -\rho$ para obter que

$$\frac{A(1/p)}{A(n/k)} \xrightarrow{n\to\infty} 0$$

utilizando este resultado em (3.40), verificamos que

$$\frac{\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{1}{p}\right) + o\left(A\left(\frac{1}{p}\right)\right)}{A\left(\frac{n}{k}\right)} \xrightarrow{n\to\infty} 0.$$

Assim, como $\sqrt{k}A\left(\frac{n}{k}\right) \to \lambda \in \mathbb{R}$, por hipótese, de (3.39), segue que

$$\begin{aligned} T_{1,k} &= \sqrt{k}\left[\frac{\frac{\beta}{1-\gamma\beta-\rho}A\left(\frac{n}{k}\right) + o\left(A\left(\frac{n}{k}\right)\right)}{1+o(1)}\right] \\ &= \frac{\frac{\beta}{1-\gamma\beta-\rho}\sqrt{k}A\left(\frac{n}{k}\right) + o\left(\sqrt{k}A\left(\frac{n}{k}\right)\right)}{1+o(1)} \\ &\xrightarrow{n\to\infty} \frac{\beta}{1-\gamma\beta-\rho}\lambda. \end{aligned}$$

Ou seja,

$$T_{1,k} = \frac{\beta}{1-\gamma\beta-\rho}\lambda + o(1). \tag{3.41}$$

Por outro lado, pelo Lema 3.4, temos

$$
\begin{aligned}
T_{2,k} &:= \sqrt{k}\left\{\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right\} \\
&= \frac{\beta}{\rho}\lambda+o(1).
\end{aligned} \tag{3.42}
$$

Assim, utilizando (3.41) e (3.42) em (3.38), obtemos

$$
\begin{aligned}
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right) &= \beta\sqrt{k}(\hat{\gamma}_k-\gamma)+\frac{1}{\ln\frac{k}{np}}\left(\frac{\beta}{1-\gamma\beta-\rho}\lambda+o(1)\right)\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta} \\
&\quad +\frac{1}{\ln\frac{k}{np}}\left(\frac{\beta}{\rho}\lambda+o(1)\right)+o_{\mathbb{P}}(1).
\end{aligned} \tag{3.43}
$$

Mas, utilizando novamente o Lema 3.4, podemos obter

$$
\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}=\frac{\sqrt{k}\left[\left(\frac{U(n/k)}{U(1/p)}\right)^{\beta}\left(\frac{k}{np}\right)^{\gamma\beta}-1\right]}{\sqrt{k}}+1 \xrightarrow{n\to\infty} 1
$$

e como , por hipótese, $\sqrt{k}(\hat{\gamma}_k-\gamma)\to\Gamma$ e $k/np\to\infty$ de (3.43) concluímos que

$$
\frac{\sqrt{k}}{\ln\frac{k}{np}}\left(\frac{\hat{\theta}_{\beta,p}}{\theta_{\beta,p}}-1\right)\xrightarrow{D}\beta\Gamma
$$

□

Vale observar que escolhendo-se o estimador $\hat{\gamma}_k$ de tal forma que Γ seja a distribuição Normal, como é o caso dos dois estimadores citados acima e sugeridos por Goegebeur et al. (2022), obtemos, pelo Teorema 3.3, a normalidade assintótica do estimador $\hat{\theta}_{\beta,p}$.

Bibliografia

[1] ARTZNER, Philippe et al. Coherent measures of risk. **Mathematical finance**, v. 9, n. 3, p. 203-228, 1999.

[2] ASH, Robert B.; DOLÉANS-DADE, Catherine A. **Probability and measure theory**. Academic press, 2000.

[3] BARRY, James R. **Probabilidade: um curso em nível intermediário**. Rio de Janeiro, Brasil: IMPA, 2008.

[4] BINGHAM, Nicholas H.; GOLDIE, Charles M.; TEUGELS, Jef L. **Regular variation**. Cambridge university press, 1989.

[5] BRAZAUSKAS, Vytaras et al. Estimating conditional tail expectation with actuarial applications in view. **Journal of Statistical Planning and Inference**, v. 138, n. 11, p. 3590-3604, 2008.

[6] CAI, Jun; TAN, Ken Seng. Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. **ASTIN Bulletin: The Journal of the IAA**, v. 37, n. 1, p. 93-112, 2007.

[7] CARDOSO, Petrusca Arrieiro. **Uma metodologia para estimação do capital econômico**: incorporação de dependência entre riscos via cópulas. 2008. Dissertação (Mestrado em Ciências Atuariais) - Instituto de Gestão de Riscos Financeiros e Atuariais, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

[8] CASELLA, George; BERGER, Roger L. **Statistical inference**. Cengage Learning, 2002.

[9] CORDEIRO, Gauss Moutinho. **Introdução à teoria Assintótica**. IMPA, 1999.

[10] de HAAN, Laurens; FERREIRA, Ana. **Extreme value theory**: an introduction. New York: Springer, 2006.

[11] DENUIT, Michel et al. **Actuarial theory for dependent risks**: measures, orders and models. John Wiley & Sons, 2006.

[12] EMBRECHTS, Paul; HOFERT, Marius. A note on generalized inverses. **Mathematical Methods of Operations Research**, v. 77, p. 423-432, 2013.

[13] FISHER, Ronald Aylmer; TIPPETT, Leonard Henry Caleb. Limiting forms of the frequency distribution of the largest or smallest member of a sample. **Mathematical proceedings of the Cambridge philosophical society**. Cambridge University Press, p. 180-190, 1928.

[14] GNEDENKO, Boris. Sur la distribution limite du terme maximum d'une serie aleatoire. **Annals of mathematics**, p. 423-453, 1943.

[15] GOEGEBEUR, Yuri et al. Extreme-value based estimation of the conditional tail moment with application to reinsurance rating. **Insurance: Mathematics and Economics**, v. 107, p. 102-122, 2022.

[16] HARDY, Mary R. An introduction to risk measures for actuarial applications. **SOA Syllabus Study Note**, v. 19, 2006.

[17] JENKINSON, Arthur F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 81, n. 348, p. 158-171, 1955

[18] JORION, Philippe. **Value at risk**: the new benchmark for managing financial risk. The McGraw-Hill Companies, Inc., 2007.

[19] MARKOWITZ, Harry. Portfolio Selection. **The Journal of Finance**, v. 7, p. 77-91. 1952.

[20] METHNI, Jonathan El; GARDES, Laurent; GIRARD, Stéphane. Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions. **Scandinavian Journal of Statistics**, v. 41, n. 4, p. 988-1012, 2014.

[21] OLIVEIRA, Elsio Paiva. **Medidas coerentes de risco**. 2009. Dissertação - Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2009.

[22] RÉNYI, Alfréd. On the theory of order statistics. **Acta Mathematica Academiae Scientiarum Hungaricae**, v. 4, n. 2, p. 191-231, 1953.

[23] RESNICK, S.I. **Extreme Values, Regular Variation, and Point Processes**. Springer-Verlag, New York, 1987.

[24] RIGHI, Marcelo Brutti; CERETTA, Paulo Sergio. Teoria de medidas de risco: Uma revisão abrangente. **Revista Brasileira de Finanças**, v. 12, n. 3, p. 411-464, 2014.

[25] ROHATGI, Vijay K.; SALEH, A. K. M. D. **An introduction to probability and statistics**. John Wiley & Sons, 2001.

[26] SENETA, Eugene. **Regularly varying functions**. Springer, 1976.

[27] VON MISES, Richard. La distribution de la plus grande de n valuers. **Rev. math. Union interbalcanique**, v. 1, p. 141-160, 1936.