Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Um Estimador com Estrutura de U-Estatística para o Índice Caudal de Distribuições de Cauda Pesada

por

Luciene Pinheiro Lopes

Brasília

2007

Resumo

No presente trabalho, estudamos o estimador proposto por Fan (2004), para o índice caudal de distribuições no domínio de atração de leis α-estáveis, com $0 < \alpha < 2$. Este estimador tem estrutura de U-estatística, é robusto e assintoticamente não viesado. Utilizando as ferramentas da teoria clássica de U-estatística, são demonstradas a consistência e normalidade assintótica do estimador.

Palavras-chave: Cauda pesada, U-estatística, distribuição estável, índice caudal, domínio de atração.

Abstract

In this work we study the estimator proposed by Fan (2004) for the tail index of distributions in the domain of attraction of a α-stable law with $0 < \alpha < 2$. This estimator has U-statistic structure and is robust and asymptotically unbiased. By using the classical tools theory of U-statistic, the consistency and the asymptotic normality of the estimator are proved.

Keywords: Heavy tail, U-statistics, stable law, tail index, domain of attraction.

Introdução

Distribuições de cauda pesada têm aparecido em aplicações nas mais diversas áreas, tais como: economia, finanças, atuária, engenharia elétrica, sociologia. Dentre os diversos tipos de distribuições de cauda pesada que são comumente usadas, as distribuições α-estáveis, com $0 < \alpha < 2$ (o caso $\alpha = 2$ refere-se à distribuição normal que tem cauda leve), merecem destaque.

As distribuições estáveis, por definição, são invariantes sob a adição e caracterizam-se por serem limites em distribuição de somas normalizadas de variáveis aleatórias (v.a's) independentes e identicamente distribuídas (i.i.d). Assim, variáveis empíricas que são somas de variáveis aleatórias ajustam-se bem a estas distribuições. Um exemplo simples de aplicação prática é a variação da taxa de câmbio. É naturalmente desejável que a variação da taxa de câmbio em um dia tenha a mesma distribuição que a taxa de câmbio na hora j para qualquer j.

O uso de distribuições α-estáveis não-gaussianas na modelagem de problemas aplicados, especialmente à Economia e Finanças, entre outras áreas, foi impulsionado pelo trabalho de Mandelbrot (1963) que verificou que o modelo α-estável, $0 < \alpha < 2$, por possuir segundo momento infinito, se ajustava melhor a dados financeiros que apresentavam alta variabilidade do que o modelo normal, que era comumente usado.

De acordo com Fan (2004), no entanto, do ponto de vista prático é mais realista modelar as observações com distribuições no domínio de atração de alguma lei α-estável $(0 < \alpha < 2)$.

Precisamente, se $X_1, X_2, ...$ são variáveis aleatórias i.i.d. com função de distribuição F para as quais existem sequências $\{a_n\}$ e $\{b_n\} > 0$ tais que

$$\frac{X_1 + ... + X_n - b_n}{a_n} \xrightarrow{d} X, \tag{1}$$

onde X é uma v.a. α-estável, então dizemos que F está no domínio de atração de uma lei α-estável e denotamos $F \in DA(\alpha)$.

As distribuições no domínio de atração de leis α-estáveis, com $0 < \alpha < 2$, caracterizam-se por possuírem cauda à direita ou cauda à esquerda regularmente variantes, com mesmo índice caudal α. Especificamente, $F \in DA(\alpha)$ com $0 < \alpha < 2$, se, e só se,

$$x^{\alpha}(1 - F(x)) \sim pL(x) \quad e \quad x^{\alpha}(F(-x)) \sim qL(x), \quad \text{quando} \quad x \to \infty,$$

onde $p, q \geq 0$, $p + q = 1$ e L é uma função lentamente variante no infinito.

Neste contexto, a estimação do índice caudal é um problema de interesse entre os pesquisadores da área. Como em muitas aplicações é impossível determinar, a priori, qual família de distribuições é mais apropriada, o desenvolvimento de estimadores robustos para o índice caudal tem merecido atenção especial.

Dentre os estimadores conhecidos na literatura, o mais popular é o estimador de Hill (1975), baseado nas estatísticas de ordem extremais. As propriedades deste estimador têm sido objeto de estudo de inúmeros pesquisadores. Sua consistência fraca foi provada por Mason (1983) para distribuições com cauda de variação regular. No entanto, a normalidade assintótica do estimador é considerada sob condições adicionais mais restritivas, do tipo Von Mises (vide por exemplo, de Haan e Resnick (1985). Além disso, muitos autores têm comprovado que o estimador de Hill apresenta melhor performance somente para distribuições com cauda do tipo Pareto. Assim, nos últimos anos, têm surgido na literatura diversos estimadores do índice caudal, na tentativa de se obter estimadores mais simples de serem usados na prática e de melhor eficiência, a despeito do tipo específico de distribuição de cauda pesada considerada. Dentre eles, merecem destaque: Pickands (1975), de Haan e Resnick (1980), Csörgö (1985), de Haan e Pereira (1999), Meerschaert e Scheffler (1998) dentre muitos outros.

Neste trabalho, nós estudamos um estimador proposto no artigo de Fan (2004) para a estimação do índice caudal de distribuições no domínio de atração de leis α-estáveis, $0 < \alpha < 2$. O estimador considerado tem uma estrutura de U-estatística e, desta maneira, tem uma forma mais simples, utiliza a informação da amostra inteira e possui menor variância. Além disso, as propriedades assintóticas do estimador são obtidas utilizando-se as ferramentas da teoria clássica de U-estatística, sem impor condições restritivas sobre a distribuição da população.

Assim, dividimos este trabalho em 3 capítulos.

No Capítulo 1, apresentamos os conceitos e resultados preliminares sobre a teoria de variação regular (Seção 1.2), distribuições estáveis e domínios de atração (Seção 1.3). Merecem destaque as demonstrações dos teoremas 1.3.9 e 1.3.10 que serão fundamentais na obtenção do não viés assintótico do estimador de Fan (Teorema 3.2.1).

O Capítulo 2 é dedicado à apresentação dos principais resultados da teoria clássica de U-estatística. A principal referência utilizada foi Lee (1990). Basicamente, uma U-estatística baseada numa amostra $(X_1, ..., X_n)$ de uma função de distribuição F, é caracterizada como sendo uma estatística da forma

$$U_n = U(X_1, ..., X_n) = \binom{n}{m}^{-1} \sum_{C_{n,m}} h^*(X_{\alpha_1}, ..., X_{\alpha_m}), \tag{2}$$

onde $m < n$, o somatório $C_{n,m}$ é sobre todas as combinações de m inteiros escolhidos sem reposição do conjunto de inteiros de 1 a n e h^* é uma função simétrica. Sob a hipótese que $h^*(X_1, ..., X_m)$ é um estimador não viesado de um parâmetro $\theta = \theta(F)$, de grau m, mostra-se que U_n é o estimador de menor variância dentre os estimadores de θ baseados na amostra $(X_1, ..., X_n)$. Estes resultados e outras propriedades básicas de U-estatística são apresentados na Seção 2.2. As outras duas seções deste capítulo são dedicadas à apresentação do Teorema da Decomposição de Hoeffding, e aplicações, (Seção 2.3) e à descrição do método de estimação de variância assintótica de uma U-estatística, extraído de Arversen (1969).

No Capítulo 3 apresentamos o estimador proposto por Fan (2004), bem como as propriedades assintóticas obtidas por ele.

Assumimos que $(X_1, ..., X_n)$ é uma amostra de v.a's aleatórias i.i.d. com função de distribuição $F \in DA(\alpha)$, com $0 < \alpha < 2$, (ou seja, satisfazendo (1)) e consideramos o seguinte estimador para α^{-1}

$$\widehat{\alpha_n^{-1}} = \frac{log\ (X_1^2 + ... + X_n^2)}{2log\ n}, \tag{3}$$

onde $X_1^2, X_2^2, ..., X_n^2$ estão no domínio de atração de uma lei $0 < \dfrac{\alpha}{2} < 1$. A motivação para a escolha de $\widehat{\alpha_n^{-1}}$ vem da propriedade da preservação da soma de leis estritamente estáveis, ou seja, $\dfrac{Z_1 + .. + Z_n}{n^{\frac{1}{\alpha}}} \stackrel{d}{=} Z_1$, para variáveis aleatórias $Z_1, .., Z_n$ estritamente α-estáveis.

Vale ressaltar que optamos por considerar em (3) as somas dos quadrados dos X_i's para facilitar as demonstrações das propriedades assintóticas, já que neste caso, como $0 < \frac{\alpha}{2} < 1$, as constantes normalizadoras b_n em (1) são nulas.

Seguindo as idéias de Meerschaert e Scheffler (1998), provamos, no Teorema 3.2.1, que α_n^{-1} é um estimador assintóticamente não viesado e consistente. Além disso, apresentamos, ainda na Seção 3.2, um princípio de grandes desvios obtido por Fan (2004), caracterizando o estimador α_n^{-1} como um estimador robusto para α^{-1}, mas com taxa de convergência muito lenta e de variância grande.

Assim, com o propósito de diminuir a variância do estimador $\widehat{\alpha_n^{-1}}$, Fan (2004), propõe o uso do estimador com estrutura de U-estatística, baseado no núcleo dado por (3), ou seja, para $m < n$ considera-se

$$\widehat{\alpha_U{}^{-1}} = \binom{n}{m}^{-1} \sum_{\alpha \in C_{n,m}} \frac{log\ (X_{\alpha_1}^2 + ... + X_{\alpha_m}^2)}{2log\ n},$$

onde $\alpha = \{\alpha_1, ..., \alpha_m\}$ e o somatório é considerado como em (2). Utilizando os resultados da teoria clássica de U-estatística, apresentados no Capítulo 2, provamos que , quando $m \to \infty$, $\widehat{\alpha_U{}^{-1}}$ é assintóticamente não viesado e fracamente consistente (Teorema 3.2.1). A consistência em média quadrática e a normalidade assintótica de $\widehat{\alpha_U{}^{-1}}$ são demonstradas sob a condição adicional que $m = o(n^{\frac{1}{2}})$, quando $n \to \infty$ (Teorema 3.3.1 e Teorema 3.4.1, respectivamente).

Para finalizar, apresentamos na Seção 3.5 uma breve análise das simulações numéricas, realizadas por Fan (2004), sobre a performance do estimador $\widehat{\alpha_U{}^{-1}}$, que indicam que o estimador com estrutura de U-estatística proposto é bastante competitivo em relação a outros estimadores do índice caudal considerados, em particular, o popular estimador de Hill (1975).

Capítulo 1

Preliminares

1.1 Introdução

Neste capítulo introduzimos conceitos e resultados preliminares que serão utilizados no desenvolvimento dos capítulos seguintes.

Na Seção 1.2, apresentamos a definição e alguns dos principais resultados da teoria de variação regular que serão úteis para caracterizar os domínios de atração de distribuições estáveis. Para mais detalhes indicamos Bingham, Goldie e Teugels (1987).

Veremos, na Seção 1.3, dedicada às distribuições estáveis, a Proposição 1.3.8 que garante que uma ou ambas as caudas de uma distribuição no domínio de atração de uma distribuição α-estável, $0 < \alpha < 2$ variam regularmente. Isto permite classificá-las como distribuições de cauda pesada.

Apresentamos na Seção 1.3 as definições e propriedades principais de distribuições estáveis e seus respectivos domínios de atração. Como referência básica citamos Feller (1971) e Samorodnitski e Taqqu (1987).

1.2 Variação regular

Definição 1.2.1. Uma função $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ é de *variação regular* de índice ρ se

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^\rho, \ \forall \lambda \in \mathbb{R}.$$

Neste caso, denotamos $f \in R_\rho$.

Quando $\rho = 0$ dizemos que a função é *lentamente variante.*

Proposição 1.2.2. *Seja* $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$*, uma função de variação regular de índice* ρ*. Então*
(i) $f(x) = x^\rho L(x)$ *onde* $L(x)$ *é lentamente variante.*
(ii) Se $f(x) \in R_\rho$*, então* $(f(x))^\alpha \in R_{\rho^\alpha}, \quad \alpha \in \mathbb{R}$.
(iii) Se $f_1 \in R_\rho$ *e* $f_2 \in R_\rho$*,* $f_2(x) \to \infty$ *quando* $x \to \infty$*, então* $f_1(f_2(x)) \in R_{\rho_1 \rho_2}$.
(iv) Se L *varia lentamente, então* $\dfrac{log\ L(x)}{log\ x} \to 0$ *quando* $x \to \infty$.

Proposição 1.2.3. *Se* f *varia regularmente com índice* $\rho \neq 0$*, então quando* $x \to \infty$

$$f(x) \to \infty, \quad se \quad \rho > 0$$

e

$$f(x) \to 0, \quad se \quad \rho < 0,$$

ou seja, se L *varia lentamente e* $\rho > 0$*,*

$$x^\rho L(x) \to \infty, \quad e \quad x^{-\rho} L(x) \to 0 \quad (x \to \infty).$$

.

Teorema 1.2.4. *(Teorema da convergência uniforme) Se* f *varia regularmente com índice* ρ*, então (no caso em que* $\rho > 0$*, assuma que* f *é limitada em cada intervalo* $(0, X]$*),*

$$\frac{f(\lambda x)}{f(x)} \to \lambda^\rho (x \to \infty) \ \ uniformemente \ \ em \ \ \lambda,$$

sobre cada $[a, b]$ $(0 < a \leq b < \infty)$ *se* $\rho = 0$*, sobre cada* $(0, b]$ $(0 < b < \infty)$ *se* $\rho > 0$ *e sobre cada* $[a, \infty)$ $(0 < a < \infty)$ *se* $\rho < 0$.

Teorema 1.2.5. *(Teorema de Potter) Se f é de variação regular de índice ρ então dadas constantes $A > 1$, $\delta > 0$ existe $X = X(A, \delta)$ tal que*

$$\frac{f(y)}{f(x)} \leq Amax\left\{\left(\frac{y}{x}\right)^{\delta}, \left(\frac{y}{x}\right)^{-\delta}\right\} \quad x \geq X \quad e \quad y \geq X.$$

1.3 Distribuições estáveis

Definição 1.3.1. Dizemos que uma variável aleatória X possui *distribuição estável* se para $X_1, X_2, ...$ cópias independentes de X existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$$X_1 + ... + X_n \stackrel{d}{=} a_n X + b_n,$$

onde $\stackrel{d}{=}$ indica convergência em distribuição.

No caso em que $b_n = 0$ então $X_1 + ... + X_n \stackrel{d}{=} a_n X$ e dizemos que X é estritamente estável.

Teorema 1.3.2. *Uma variável aleatória X possui distribuição não degenerada α-estável X se, e somente se, sua função característica é da forma*

$$\phi_X(t) = exp\left\{ist - c|t|^{\alpha}\left[1 + i\theta\frac{t}{|t|}u(t, \alpha)\right]\right\} \tag{1.1}$$

onde $s \in \mathbb{R}$, $c > 0$, $0 < \alpha \leq 2$, $-1 \leq \theta \leq 1$, são constantes e

$$\begin{aligned} u(t, \alpha) &= \tan\left(\frac{\pi\alpha}{2}\right), \; se\ \alpha \neq 1 \\ &= \frac{2log\ |t|}{2}, \quad se\ \alpha = 1. \end{aligned}$$

O número α é chamado de índice (ou expoente) de estabilidade e dizemos que X é α-estável.

Observações 1.3.3. (a) *Quando $\alpha = 2$, X tem distribuição normal. No caso $0 < \alpha < 2$ costuma-se dizer também que X tem distribuição Levy estável.*

(b) *De* (1.1) *podemos notar que $|\phi_X(t)| = exp\left[-c|t|^{\alpha}\right]$, com $c > 0$ e $0 < \alpha \leq 2$. Assim, como $|\phi_X(t)|$ é integrável na reta toda então das propriedades de função característica* *segue que a função de distribuição de X é diferenciável e sua densidade é contínua. Pode se provar ainda que a densidade de X é infinitamente diferenciável (vide por exemplo Nolan (2007)).*

(c) Se $0 < \alpha < 1$ então $\phi_X(t)$ não é diferenciável em $t = 0$ e assim $E|X| = \infty$. Se, além disso, X tem distribuição simétrica então EX não existe. Se $1 < \alpha < 2$, $\phi_X(t)$ é diferenciável em $t = 0$ uma vez, ou seja, EX é finita e $EX^2 = \infty$. Somente no caso normal, $\alpha = 2$, X tem média e variância finitas.

(d) Pode-se mostrar que se X é α-estável então $E|X|^r < \infty$, $\forall r \in (0, \alpha)$ (veja, por exemplo, Feller (1971) pag 215).

No Capítulo 3, para provarmos um princípio de grandes desvios (Teorema 3.2.2), necessitamos de uma representação alternativa à (1.1) para a função característica de distribuições α-estáveis, que apresentamos a seguir e cuja prova pode ser encontrada em Ibragimov e Linnik (1971).

Proposição 1.3.4. *Se X é uma distribuição estável, então a função característica $f(t)$ de X é*

$$f(t) = exp\left\{-c_0|t|^\alpha L\left(\frac{1}{|t|}\right)\left(1 - i\beta sign(t) tan\left(\frac{\pi\alpha}{2}\right)(1 + o(1))\right)\right\} \quad para \quad \alpha \neq 1.$$

onde dizemos que $f = o(g)$ se $\lim_{x\to\infty} \dfrac{f(x)}{g(x)} = 0$ e L é uma função lentamente variante.

Teorema 1.3.5. *Uma v.a. X tem distribuição α-estável se, e só se, X é o limite em distribuição de somas normalizadas de v.a's i.i.d., ou seja, existem v.a's i.i.d. tais que*

$$\frac{X_1 + ... + X_n - b_n}{a_n} \xrightarrow{d} X,$$

para escolhas apropriadas de constantes $a_n > 0$ e $b_n \in \mathbb{R}$.

Definição 1.3.6. Suponha que $X_1, X_2, ...$ é uma sequência de variáveis aleatórias i.i.d com distribuição F. Se existe uma sequência de números positivos $\{a_n\}$ e uma sequência de números reais $\{b_n\}$, tais que

$$\frac{X_1 + ... + X_n - b_n}{a_n} \xrightarrow{d} X$$

para alguma variável aleatória X α-estável, então dizemos que F está no domínio de atração de uma α-estável e denotamos por $F \in DA(\alpha)$.

Proposição 1.3.7. *Sejam $X_1, X_2, ...$ variáveis independentes e com distribuição $F \in DA(\alpha)$. Então as constantes a_n e b_n da Definição 1.3.6 anterior satisfazem:*

$$a_n = n^{\frac{1}{\alpha}} L(n), \text{ onde } L(n) \text{ é uma função lentamente variante, } \alpha \in (0, 2], \quad e$$

$$b_n = n \int_{|y| \leq a_n} y dF(y) = E\left(X I_{|X| \leq a_n}\right),$$

$$b_n = \begin{cases} nEX, & \alpha \in (1, 2] \\ 0, & \alpha \in (0, 1) \\ 0, & \alpha = 1 \quad e \quad F \quad \text{simétrica.} \end{cases}$$

O resultado a seguir permite classificar as distribuições no domínio de atração de uma α-estável com $0 < \alpha < 2$ como distribuições de cauda pesada devido à variação regular de uma ou ambas as caudas.

Proposição 1.3.8. *Uma distribuição F está no domínio de atração de uma distribuição α-estável com $0 < \alpha < 2$ se, e só se,*

$$x^{\alpha}(1 - F(x)) \sim pL(x) \tag{1.2}$$

e

$$x^{\alpha}(F(-x)) \sim qL(x), \tag{1.3}$$

quando $x \to \infty$. Onde p, $q \geq 0$, $p + q = 1$, $L(x)$ é de variação lenta e $f \sim g$ indica $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1$.

Note que (1.2) equivale a dizer que a cauda à direita de F é regularmente variante com expoente $-\alpha$ e da mesma forma, de (1.3), temos que a cauda à esquerda de F é regularmente variante com expoente $-\alpha$. Em particular, se F é uma função de distribuição α-estável então a função lentamente $L(x)$ em (1.2) e (1.3) é constante, $L(x) = c$ e assim, as respectivas caudas são do tipo Pareto.

Assim, como p, $q \geq 0$ e $p + q = 1$, concluímos da Proposição 1.3.8 acima que a família de distribuições no domínio de atração de uma lei Levy estável, $0 < \alpha < 2$, coincide com a família de distribuições que possuem caudas pesadas à direita ou caudas pesadas à esquerda (ou ambas), com o mesmo índice α, que mede a espessura das respectivas caudas. Neste caso, concluímos que $E|X|^{\delta} < \infty$ para $0 < \delta < \alpha$ e $E|X|^{\delta} = \infty$ para $\delta > \alpha$.

Finalizamos esta seção demonstrando as Proposições 1.3.10 e 1.3.11, que serão utilizadas no Capítulo 3, para provar as propriedades assintóticas do estimador a ser estudado naquele capítulo. Para estas demonstrações necessitamos do resultado a seguir.

Proposição 1.3.9. *Se $X_1, X_2, ...$ são variáveis aleatórias independentes e com distribuição F no domínio de atração de uma α-estável então as funções definidas por*

$$U_\zeta(t) = E\{|X|^\zeta I_{(|X|\leq t)}\} \ \ e \ \ V_\eta(t) = \{E|X|^\eta I_{(|X|>t)}\}$$

variam regularmente com índice $\zeta - \alpha$ e $\eta - \alpha$, respectivamente.

Proposição 1.3.10. *Seja Y uma variável aleatória α-estável. Então $E\{log\ |Y|\}^2 < \infty$.*

Demonstração: Para $A > 0$ arbitrário temos

$$\begin{aligned} E\{log\ |Y|\}^2 &= \int_{(|Y|\leq A)} (log\ |Y|)^2 dP + \int_{(|Y|>A)} (log\ |Y|)^2 dP \\ &\leq M_1 + \int_{(|Y|>A)} (log\ |Y|)^2 dP, \end{aligned}$$

onde $M_1 = (log\ A)^2 > 0$ constante.

Então, basta mostrar que

$$\lim_{A\to\infty} \int_{(|Y|>A)} (log\ |Y|)^2 dP = 0, \tag{1.4}$$

pois daí teremos que existe $M_2 > 0$ tal que $\displaystyle\int_{(|Y|>A)} (log\ |Y|)^2 dP \leq M_2$, $\forall A > 0$ e assim,

$$E\{log\ |Y|\}^2 \leq M_1 + M_2 < \infty. \tag{1.5}$$

Provemos (1.4). Para todo $A > 1$ podemos escrever

$$\begin{aligned} E\left\{(log\ |Y|)^2 I_{(|Y|>A)}\right\} &= E\left\{\int_1^{|Y|} \frac{2log\ s}{s} I_{(|Y|>A)} ds\right\} \\ &= E\left\{\int_1^{A} \frac{2log\ s}{s} I_{(|Y|>A)} ds\right\} + E\left\{\int_A^{\infty} \frac{2log\ s}{s} I_{(|Y|>A,\ |Y|>s)} ds\right\} \end{aligned}$$

e, usando o Teorema de Fubini, segue

$$
\begin{aligned}
E\left\{(log\ |Y|)^2 I_{(|Y|>A)}\right\} &= (log\ A)^2 E\{I_{(|Y|>A)}\} + \int_A^\infty \frac{2log\ s}{s} E\{I_{(|Y|>A,\ |Y|>s)}\} ds \\
&= (log\ A)^2 P(|Y|>A) + \int_A^\infty \frac{2log\ s}{s} P(|Y|>s) ds.
\end{aligned}
$$

Agora como Y é α-estável, então pela Proposição 1.3.8 $P(|Y|>s)$ é regularmente variante com expoente $-\alpha$, ou seja, $P(|Y|>s) = s^{-\alpha}L(s)$, onde $L(s)$ é uma função lentamente variante. Logo, para $A>1$

$$
E\left\{(log\ |Y|)^2 I_{(|Y|>A)}\right\} = (log\ A)^2 A^{-\alpha} L(A) + \int_A^\infty \frac{2log\ s}{s} s^{-\alpha} L(s) ds.
$$

Por um lado, como $log\ x,\ x>0$ é lentamente variante, então $(log\ x)^2$ também o é. Daí, temos

$$
(log\ A)^2 A^{-\alpha} L(A) = A^{-\alpha} L_1(A), \tag{1.6}
$$

onde $L_1(A)$ é uma função lentamente variante. Como $A^{-\alpha}L_1(A)$ é regularmente variante com expoente $-\alpha<0$ segue da Proposição 1.2.3 que

$$
\lim_{A\to\infty} (log\ A)^2 A^{-\alpha} L(A) = 0. \tag{1.7}
$$

Por outro lado, segue do Teorema de Potter (Teorema 1.2.5) que dado $\delta>0$, existe $A_0>1$ e uma constante $K(\delta, A_0)$ tal que

$$
L(s) \le K s^{\delta}, \qquad \forall s \ge A_0. \tag{1.8}
$$

Assim, tomando $0<\delta<\alpha$ temos para todo $A \ge A_0$,

$$
\begin{aligned}
0 \le \int_A^\infty \frac{2log\ s}{s} s^{-\alpha} L(s) ds &\le K \int_A^\infty \frac{2log\ s}{s} s^{-\alpha+\delta} ds \\
&= 2K \int_{log\ A}^\infty t e^{(-\alpha+\delta)t} dt \\
&= \frac{2K}{\alpha-\delta}\left[(log\ A)A^{-\alpha+\delta} - \frac{A^{-\alpha+\delta}}{\alpha-\delta}\right].
\end{aligned}
$$

Como $(-\alpha+\delta)<0$, temos da Proposição 1.2.3 que

$$\lim_{A\to\infty}(log\ A)A^{-\alpha+\delta}=0 \quad e \quad \lim_{A\to\infty}A^{-\alpha+\delta}=0.$$

Daí, segue que,

$$\int_A^\infty \frac{2log\ s}{s}s^{-\alpha}L(s)ds=0. \tag{1.9}$$

De (1.9) e (1.7) temos (1.4) e o resultado segue. □

Proposição 1.3.11. *Sejam $X_1, X_2, ...$ v.a's i.i.d com função de distribuição F no domínio de atração de uma distribuição α-estável. Seja $S_n^*=\dfrac{X_1+...+X_n-b_n}{a_n}$. Então*

$$E\{log\ |S_n^*|\}^k \longrightarrow E\{log\ |Y|\}^k, \quad quando \quad n\to\infty, \quad para \quad k=1,2.$$

Demonstração: Pela Proposição 1.3.10 temos que $E\{log\ |Y|\}^k<\infty$ para $k=1,2$.

Denote $Z_n=log\ S_n^*$, $\quad G_n(x)=P(Z_n\le x)$ e $G(x)=P(log\ |Y|\le x)$.

Primeiramente, temos da Definição 1.3.6 que $Z_n \stackrel{d}{\longrightarrow} log\ |Y|$, ou seja, $G_n(x)\to G(x), \forall x\in\mathbb{R}$ e das propriedades de convergência fraca segue que

$$E\{Z_n^kI_{(|Z_n|\le A)}\}=\int_{(|x|\le A)}x^kdG_n(x)\longrightarrow\int_{(|x|\le A)}x^kdG(x) \quad quando \quad n\to\infty, \tag{1.10}$$

pois $g(x)=x^k$ é contínua e limitada em $[-A,A]$. Assim, como

$$\begin{aligned}|EZ_n^k-E\{log\ |Y|\}^k| \quad &\le \quad |EZ_n^k-E\{Z_n^kI_{(|Z_n|\le A)}\}|+\left|E\{Z_n^kI_{(|Z_n|\le A)}\}-\int_{(|x|\le A)}x^kdG(x)\right| \\ &+ \quad \left|\int_{(|x|\le A)}x^kdG(x)-\int_{-\infty}^\infty x^kdG(x)\right|,\end{aligned}$$

segue que, para todo $A>0$

$$\limsup_{n\to\infty}|EZ_n^k-E\{log\ |Y|\}^k|\le\limsup_{n\to\infty}E\{|Z_n|^kI_{(|Z_n|>A)}\}. \tag{1.11}$$

Assim, basta mostrar que

$$\lim_{A\to\infty}\limsup_{n\to\infty}E\{|Z_n|^kI_{(|Z_n|>A)}\}=0. \tag{1.12}$$

Para isto vamos usar um raciocínio análogo ao utilizado na prova de (1.4) da Proposição 1.3.10. Para todo $A > 1$, podemos escrever para $k = 1, 2$,

$$
\begin{aligned}
E\{|Z_n|^k I_{(|Z_n|>A)}\} &= E\left\{\int_0^A k s^{k-1} I_{(|Z_n|>A)} ds\right\} + E\left\{\int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} I_{(|S_n^*|>e^A),|S_n^*|>e^s)} ds\right\} \\
&= A^k P(|Z_n| > A) + \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|S_n^*| > e^s) ds.
\end{aligned}
$$

Agora, como $Z_n \xrightarrow{d} log\ |Y|$ e $log\ |Y|$ é uma variável aleatória contínua, pois Y é α-estável, segue que $\forall A > 0$

$$\lim_{n\to\infty} A^k P(|Z_n| > A) = A^k P(|log\ |Y|| > A). \tag{1.13}$$

Mas, como $E\{|log\ |Y||^k\} < \infty, \quad k = 1, 2$, temos

$$\lim_{A\to\infty} A^k P(|log\ |Y|| > A) = 0 \tag{1.14}$$

e, assim,

$$\lim_{A\to\infty} \lim_{n\to\infty} A^k P(|Z_n| > A) = 0. \tag{1.15}$$

Por outro lado, como $\lim\limits_{n\to\infty} P(|S_n^*| \leq x) = P(|Y| \leq x)$ e $P(|Y| \leq x)$ é contínua, então a convergência é uniforme em $x \in \mathbb{R}$. Então,

$$\lim_{n\to\infty} \frac{P(|S_n^*| > x)}{P(|Y| > x)} = 1, \tag{1.16}$$

uniformemente em x. Ou seja, $\exists\ \ K_1 > 0$ tal que $\dfrac{P(|S_n^*| > x)}{P(|Y| > x)} \leq K_1$, $\forall x \in \mathbb{R}$ e n grande.

Daí, $\forall A > 1$ e n suficientemente grande obtemos

$$
\begin{aligned}
\int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|S_n^*| > e^s) ds &\leq K_1 \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|Y| > e^s) ds \\
&\leq K_1 \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|Y| > s) ds.
\end{aligned}
$$

Como Y é α-estável, pela Proposição 1.3.8, $P(|Y| > s)$ é regularmente variante com expoente $-\alpha$. Então para $A > 1$ e n suficientemente grande

$$\int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|S_n^*| > e^s) \leq K_1 \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} s^{-\alpha} L(s) ds.$$

Agora, utilizando raciocínio análogo ao utilizado para provar (1.9) na prova da Proposição 1.3.10, como $k = 1, 2$ e $0 < \alpha < 2$ podemos mostrar

$$\lim_{A\to\infty} \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} s^{-\alpha} L(s) ds = 0, \tag{1.17}$$

e daí segue

$$\lim_{A\to\infty} \limsup_{n\to\infty} K_1 \int_{e^A}^{\infty} k\frac{(log\ s)^{k-1}}{s} P(|S_n^*| > e^s) ds = 0. \tag{1.18}$$

Logo, usando (10) e (11) em (9) obtemos (8). Portanto, de (7) segue que $EZ_n^k \longrightarrow E\{log\ |Y|\}^k$, para k = 1,2, como queríamos. $\square$

Capítulo 2

U-Estatística

2.1 Introdução

Para subsidiar o desenvolvimento do próximo capítulo, apresentamos neste capítulo um estudo dos principais resultados da teoria clássica de U-estatística. A principal referência foi Lee (1990), citamos também como referência básica para esta teoria Handles e Wolfe (1979) e Korolyuk e Borovskich (1994).

A teoria de U-estatistica vem sendo desenvolvida há mais de 50 anos e o trabalho pioneiro mais importante é o de Hoeffding (1948).

Vários estimadores não paramétricos conhecidos na literatura pertencem à classe de U-estatísticas, os quais caracterizam-se por terem uma estrutura simples e serem não viesados (daí o significado da letra U: do inglês "unbiased "), e a aplicação da teoria existente possibilita a geração de novas estatísticas úteis em problemas de estimação, como é o caso dos estimadores a serem utilizados no Capítulo 3.

Assim, na Seção 2.2, caracterizamos uma U-estatística, baseada numa amostra de tamanho n de uma função de distribuição F, como sendo, toda estatística da forma

$$U_n = U(X_1, ..., X_n) = \binom{n}{m}^{-1} \sum_{C_{n,m}} h^*(X_{\alpha_1}, ..., X_{\alpha_m}),$$

onde o somatório $C_{n,m}$ é sobre todas as combinações de m inteiros escolhidos sem reposição do conjunto de inteiros de 1 a n e h^* é uma função simétrica. Daí, segue que uma U-estatística leva em consideração a amostra inteira.

Sob a hipótese que $h^*(X_1, ..., X_m)$ é um estimador não viesado do parâmetro $\theta = \theta(F)$ de grau m, mostramos que U_n é o estimador de menor variância dentre os estimadores não viesados de θ, baseados na amostra $(X_1, ..., X_n)$.

Apresentamos também, nessa seção, algumas caracterizações importantes da variância de U_n e, como consequência, apresentamos uma prova da consistência em média quadrática de U_n, Corolário 2.2.9.

Na Seção 2.3, provamos o Teorema da Decomposição de Hoeffding, (Teorema 2.3.1) que permite escrever uma U-estatística em termos de outras U-estatísticas não correlacionadas. Algumas propriedades decorrentes desta decomposição, que serão úteis para o desenvolvimento do próximo capítulo, também são apresentadas.

Finalizamos o capítulo apresentando, na Seção 2.4, a normalidade assintótica de U_n e, devido a dificuldade de se determinar a distribuição da U-estatística e sua variância assintótica, apresentamos um método consistente de aproximação dessa variância assintótica, extraído de Arversen (1969).

2.2 Definição e Propriedades

Considere um subconjunto qualquer $\mathcal{F}$ do conjunto de funções de distribuições sobre $\mathbb{R}$ e seja $\theta = \theta(F)$, $F \in \mathcal{F}$, um funcional definido sobre $\mathcal{F}$.

Seja $F \in \mathcal{F}$ um membro desconhecido de $\mathcal{F}$ e considere uma amostra $X_1, X_2, ..., X_n$ de v.a.'s independentes e identicamente distribuídas(i.i.d.) com função de distribuição F. Suponha que desejamos estimar o parâmetro $\theta = \theta(F)$ baseado nesta amostra.

Em geral, assumimos que $\mathcal{F}$ consiste de todas as distribuições absolutamente contínuas ou discretas ou subclasses destas. Note que nenhuma hipótese referente ao conhecimento de F é assumida, a menos que ela é uma distribuição da família $\mathcal{F}$.

Definição 2.2.1. Um parâmetro θ é *estimável de grau m* para uma família de distribuições $\mathcal{F}$ se m é o menor tamanho da amostra para o qual existe uma função $h(x_1, ..., x_m)$ tal que

$$E_F[h(X_1, ..., X_m)] = \theta, \quad \forall \ F \in \mathcal{F}, \tag{2.1}$$

onde $X_1, ..., X_m$ são v.a.'s independentes e com distribuição comum F.

Assim, podemos dizer que um parâmetro é estimável se existe um estimador não viesado para esse parâmetro.

A função $h(x_1, ..., x_m)$ que satisfaz (2.1) é chamada núcleo de θ.

Observação 2.2.2. *Observe que se θ_1 e θ_2 são parâmetros estimáveis de graus m_1 e m_2, respectivamente, então $\theta_1 + \theta_2$ é estimável de grau $m = max\{m_1, m_2\}$ e $\theta_1\theta_2$ é estimável de grau $m_1 + m_2$. Assim, um parâmetro estimável é um funcional linear. Costuma-se denominar um parâmetro estimável por funcional regular.*

Sem perda de generalidade, podemos supor que $h(x_1, ..., x_m)$ é simétrica em seus argumentos, isto é,

$$h(x_1, ..., x_m) = h(x_{\alpha_1}, ..., x_{\alpha_m}),$$

para qualquer permutação $\{\alpha_1, ..., \alpha_m\}$ do conjunto $\{1, ..., m\}$. Se $h(x_1, ..., x_m)$ não for simétrica, então podemos construir a partir de $h(x_1, ..., x_m)$ uma função que seja simétrica nos seus argumentos:

$$h^*(x_1, ..., x_m) = \frac{1}{m!} \sum_{P_m} h(x_{\alpha_1}, ..., x_{\alpha_m}), \tag{2.2}$$

onde P_m indica a soma sobre todas as permutações no conjunto $\{1, ..., m\}$.

Sob a família $\mathcal{F}$ contendo as funções de distribuições de suporte finito ou contendo as funções de distribuição absolutamente contínuas é possível mostrar que existe um único estimador de θ que é não viesado e simétrico (vide Lee (1990)). Este estimador chamado de U-estatística (onde U refere-se a "unbiased"), é definido a seguir:

Definição 2.2.3. Sejam $X_1, X_2, ..., X_n$ v.a's i.i.d. com f.d. F e seja $h(X_1, ..., X_m)$ um estimador não viesado de um parâmetro estimável θ de grau $m \leq n$. Definimos a *U-estatística* para a amostra como sendo

$$U_n(h) = U(X_1, ..., X_n) = \binom{n}{m}^{-1} \sum_{\alpha \in C_{n,m}} h^*(X_{\alpha_1}, ..., X_{\alpha_m}), \tag{2.3}$$

onde $\alpha = \{\alpha_1, ..., \alpha_m\}$, o somatório, sobre $\alpha \in C_{n,m}$, é sobre todas as combinações de m inteiros escolhidos sem reposição do conjunto de inteiros de 1 a n e h^* é o núcleo simétrico correspondente a h definido em (2.2).

Observação 2.2.4. *Observe que a condição não viesado sobre $h^*(X_1, ..., X_m)$ garante que a U-estatística seja também não viesada, pois aplicando a esperança em ambos os lados de (2.3) e usando o fato das variáveis aleatórias serem i.i.d. temos que $EU(X_1, ..., X_n) = Eh^*(X_1, ..., X_m) = \theta$. Pelo mesmo argumento, se $h^*(X_1, ..., X_m)$ for um estimador assintoticamente não viesado então o estimador U-estatística criado a partir dele também é assintoticamente não viesado.*

Uma das principais vantagens da U-estatística é que ela é o estimador de menor variância dentre todos os estimadores não viesados de θ. Isto pode ser mostrado observando que a U-estatística U_n é uma função da estatística de ordem $(X_{(1)}, ..., X_{(n)})$ que é suficiente e completa. Daí segue do Teorema 8.4.6 em Rohatgi (2001) que U_n é o estimador não viesado de variância mínima do seu valor esperado que é θ. Uma prova mais simples pode ser obtida se assumirmos que $\mathcal{F}$ é a família de funções de distribuição de suporte finito ou a família de funções de distribuição absolutamente contínuas, e usarmos a unicidade do estimador não viesado e simétrico. Esta última prova é apresentada a seguir.

Teorema 2.2.5. *Seja θ um parâmetro estimável de grau m com núcleo h, definido sobre o conjunto $\mathcal{F}$ das funções de distribuição absolutamente contínuas ou das funções de distribuição de suporte finito. Então a U-estatística $U_n(h)$ é o estimador de variância mínima na classe de todos os estimadores não viesados de θ, baseado numa amostra de tamanho $n \geq m$.*

Demonstração: Seja $f = f(X_1, ..., X_n)$ um estimador não viesado de θ baseado numa amostra de tamanho n, $\{X_1, ..., X_n\}$, de $F \in \mathcal{F}$.

Seja f^* o estimador simétrico correspondente a f, ou seja,

$$f^*(x_1, ..., x_n) = \frac{1}{n!} \sum_{P_n} f(x_{\alpha_1}, ..., x_{\alpha_n}),$$

onde a soma é considerada sobre todas as permutações P_n de $\{1, ..., n\}$.

Então f^* é um estimador não viesado e simétrico de θ. Logo, como observamos anteriormente, f^* coincide com U_n sobre $\mathbb{R}$.

Assim, aplicando a desigualdade de Cauchy Schwartz, podemos escrever

$$\begin{aligned} U_n^2 &= \left(\sum_{P_n} \frac{1}{n!} f(x_{\alpha_1}, ..., x_{\alpha_n})\right)^2 \\ &\leq \sum_{P_n} \left(\frac{1}{n!}\right)^2 \sum_{P_n} f^2(x_{\alpha_1}, ..., x_{\alpha_n}) \\ &= \frac{1}{n!} \sum_{P_n} f^2(x_{\alpha_1}, ..., x_{\alpha_n}). \end{aligned}$$

Logo,

$$\begin{aligned} EU_n^2 &\leq \frac{1}{n!} \sum_{P_n} Ef^2(X_{\alpha_1}, ..., X_{\alpha_n}) \\ &= \frac{1}{n!} \sum_{P_n} Ef^2(X_1, ..., X_n) \\ &= Ef^2(X_1, ..., X_n). \end{aligned}$$

Como $EU_n = Ef$, seque que $Var\ U_n \leq Var\ f$. □

A seguir apresentamos algumas propriedades importantes sobre a variância de uma U-estatística U_n.

Primeiramente, obtemos uma expressão da variância de U_n em termos de certas esperanças condicionais.

Teorema 2.2.6. *Seja $U_n(h)$ uma U-estatística com núcleo $h(X_1, ..., X_m)$ de grau m. Então*

$$Var\ U_n(h) = \binom{n}{m}^{-1} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \sigma_c^2, \tag{2.4}$$

onde

$$\sigma_c^2 = Var\ h_c(X_1, ..., X_c) \tag{2.5}$$

e

$$h_c(x_1, ..., x_c) = E(h(X_1, ..., X_m)|X_1 = x_1, ..., X_c = x_c),\ c = 1, ..., m. \tag{2.6}$$

Para demonstrar o teorema acima necessitamos de alguns resultados preliminares.

O primeiro deles é uma consequência imediata das propriedades de esperança condicional.

Lema 2.2.7. *Seja* $h_c(x_1, ..., x_c) = E\{h(X_1, ..., X_m)|X_1 = x_1, ..., X_c = x_c\}$, $c = 1, ..., m$. *Então*

(*i*) $h_c(x_1, ..., x_c) = E\{h_d(X_1, ..., X_d)|X_1 = x_1, ..., X_c = x_c\}$ *para* $1 \leq c < d \leq m$

(*ii*) $Eh_c(X_1, ..., X_c) = Eh(X_1, ..., X_m)$.

No lema a seguir obtemos uma expressão para a variância $\sigma_c^2 = Var\ h_c(X_1, ..., X_c)$ como uma covariância.

Lema 2.2.8. *Seja* $\sigma_c^2 = Var\ h_c(X_1, ..., X_c)$, *onde* $h_c(X_1, ..., X_c)$ *foi definido por (2.6). Então*

$$\zeta_c = Cov\ \{h(X_{\alpha_{1_c}}, ..., X_{\alpha_{m_c}}), h(X_{\beta_{1_c}}, ..., X_{\beta_{m_c}})\} = \sigma_c^2, \tag{2.7}$$

onde $\{\alpha_{1_c}, ..., \alpha_{m_c}\}$ *e* $\{\beta_{1_c}, ..., \beta_{m_c}\}$ *são subconjuntos de* m *elementos escolhidos de* $\{1, ..., n\}$ *com* c *elementos em comum.*

Demonstração: Por definição, $\sigma_c^2 = Eh_c^2(X_1, ..., X_c) - [Eh_c(X_1, ..., X_c)]^2$.

Pelo Lema 2.2.7, item (*ii*), temos que para c arbitrário, $Eh_c(X_1, ..., X_c) = Eh(X_1, ..., X_m)$. Assim,

$$\sigma_c^2 = Eh_c^2(X_1, ..., X_c) - [Eh(X_1, ..., X_m)]^2.$$

Por outro lado, como $X_1, ..., X_n$ são identicamente distribuídas e $\{\alpha_{1_c}, ..., \alpha_{m_c}\}$ e $\{\beta_{1_c}, ..., \beta_{m_c}\}$ têm c elementos em comum, podemos escrever

$$\begin{aligned}
Cov\{h(X_{\alpha_{1_c}}, ..., X_{\alpha_{m_c}}), h(X_{\beta_{1_c}}, ..., X_{\beta_{m_c}})\} &= E\{h(X_1, ..., X_m)h(X_1, ..., X_c, X_{m+1}, ..., X_{2m-c})\} \\
&- Eh(X_1, ..., X_m)Eh(X_1, ..., X_c, X_{m+1}, ..., X_{2m-c}) \\
&= E\{h(X_1, ..., X_m)h(X_1, ..., X_c, X_{m+1}, ..., X_{2m-c})\} \\
&- [Eh(X_1, ..., X_m)]^2.
\end{aligned}$$

Comparando as expressões para $Cov\{h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}}),h(\beta_{1_c},...,X_{\beta_{m_c}})\}$ e σ_c^2, é suficiente mostrar que

$$E\{h(X_1,...,X_m)h(X_1,...,X_c,X_{m+1},...,X_{2m-c})\}=Eh_c^2(X_1,...,X_c).$$

Mas, como $X_1,...,X_n$ são i.i.d. temos

$$\begin{aligned}
&E\{h(X_1,...,X_m)h(X_1,...,X_c,X_{m+1},...,X_{2m-c})\}\\
&= \int ... \int h(x_1,...,x_m)h(x_1,...,x_c,x_{m+1},...,x_{2m-c})\prod_{i=1}^{2m-c}dF(x_i)\\
&= \int ... \int \left\{\int ... \int \ h(x_1,...,x_m)\prod_{i=c+1}^{m}dF(x_i)\right\}\\
&\ \ .\ \ \left\{\int ... \int h(x_1,...,x_c,x_{m+1},...,x_{2m-c})\prod_{i=m+1}^{2m-c}dF(x_i)\right\}\prod_{i=1}^{c}dF(x_i)\\
&= \int ... \int h_c^2(x_1,...,x_c)\prod_{i=1}^{c}dF(x_i)\\
&= Eh_c^2(X_1,...,X_c),
\end{aligned}$$

e o resultado segue. □

Demonstração do Teorema 2.2.6: Temos

$$\begin{aligned}
Var\ U_n(h) &= Var\ \left\{\binom{n}{m}^{-1}\sum_{C_{n,m}}h(X_{\alpha_1},...,X_{\alpha_m})\right\}\\
&= Cov\left\{\binom{n}{m}^{-1}\sum_{C_{n,m}}h(X_{\alpha_1},...,X_{\alpha_m}),\binom{n}{m}^{-1}\sum_{C_{n,m}}h(X_{\beta_1},...,X_{\beta_m})\right\}\\
&= \binom{n}{m}^{-2}\sum_{C_{n,m}}\sum_{C_{n,m}}Cov\{h(X_{\alpha_1},...,X_{\alpha_m}),h(X_{\beta_1},...,X_{\beta_m})\}.
\end{aligned}$$

A covariância é zero entre os conjuntos $\{h(\alpha_{1_0},...,\alpha_{m_0})\}$ e $\{h(\beta_{1_0},...,\beta_{m_0})\}$, pois não há elementos em comum e então há independência entre eles. Assim, precisamos considerar a covariância entre os pares $\{h(\alpha_{1_c}, ..., \alpha_{m_c})\}$ e $\{h(\beta_{1_c}, ..., \beta_{m_c})\}$ para $c \neq 0$, $c = 1, ..., m$. Daí,

$$Var\ U_n(h) = \binom{n}{m}^{-2} \sum_{c=1}^{m} k\ Cov\ \{h(X_{\alpha_{1_c}}, ..., X_{\alpha_{m_c}}), h(X_{\beta_{1_c}}, ..., X_{\beta_{m_c}})\}.$$

Vamos determinar k. O primeiro membro do par $\{h(\alpha_{1_c}, ..., \alpha_{m_c})\}$ e $\{h(\beta_{1_c}, ..., \beta_{m_c})\}$ pode ser escolhido de $\binom{n}{m}$ maneiras. O segundo membro tem que ter c elementos em comum com o primeiro membro, assim existem $\binom{m}{c}$ maneiras de escolher c. Como não queremos ter mais do que c elementos em comum com o primeiro membro, os $m - c$ elementos restantes devem ser escolhidos de $\binom{n-m}{m-c}$ maneiras. Assim $k = \binom{n}{m}\binom{m}{c}\binom{n-m}{m-c}$. Então,

$$\begin{aligned} Var\ U_n(h) &= \binom{n}{m}^{-2} \sum_{c=1}^{m} \binom{n}{m}\binom{m}{c}\binom{n-m}{m-c}\zeta_c \\ &= \binom{n}{m}^{-1} \sum_{c=1}^{m} \binom{m}{c}\binom{n-m}{m-c}\zeta_c. \end{aligned}$$

Agora, pelo Lema 2.2.8 temos que $\zeta_c = \sigma_c^2$ e (2.4) está provado. □

Como consequência do Teorema 2.2.6 obtemos um resultado sobre o comportamento assintótico da variância de uma U-estatística, quando o tamanho da amostra cresce, e a partir daí, obtemos a consistência em média quadrática do estimador.

Corolário 2.2.9. *Seja $X_1, .., X_n$ uma amostra aleatória com distribuição $F \in \mathcal{F}$ e θ um parâmetro estimável de grau m. Para $n \geq m$ seja $U_n(h) = U(X_1, ..., X_n)$ uma U-estatística com núcleo simétrico h. Se $Eh^2(X_1, ..., X_m) < \infty$ então*

$$\lim_{n\to\infty} nVar\ U(X_1, ..., X_n) = m^2\zeta_1, \tag{2.8}$$

onde ζ_1 é a covariância dada em (2.7), e consequentemente obtemos

$$U(X_1, ..., X_n) \xrightarrow{(2)} \theta, \tag{2.9}$$

onde $\xrightarrow{(2)}$ indica convergência em média quadrática.

Demonstração: Para provar (2.8), temos do Teorema 2.2.6 que

$$nVar\ U(X_1, ..., X_n) = \frac{\displaystyle\sum_{c=1}^{m} n\binom{m}{c}\binom{n-m}{m-c}\zeta_c}{\dbinom{n}{m}}.$$

Desenvolvendo o termo geral do somatório, obtemos para cada $c = 1, ..., m$

$$\begin{aligned}
\frac{n\dbinom{m}{c}\dbinom{n-m}{m-c}}{\dbinom{n}{m}}\zeta_c &= \frac{nm!(n-m)!m!(n-m)!}{n!c!(m-c)!(m-c)!(n-2m+c)!}\zeta_c \\
&= K_c\frac{(n-m)!(n-m)!n}{(n-2m+c)!n!}\zeta_c \\
&= K_c\frac{(n-m)(n-m+1)...(n-2m+c+1)}{(n-1)...(n-m+1)}\zeta_c,
\end{aligned} \tag{2.10}$$

onde $K_c = \dfrac{[m!]^2}{c![(m-c)]^2}$.

Quando $c = 1$, temos em (2.10) o mesmo número de termos envolvendo n no numerador e no denominador. Assim,

$$\frac{(n-m)(n-m+1)...(n-2m+c+1)}{(n-1)...(n-m+1)} \rightarrow 1, \quad \text{quando} \quad n \rightarrow \infty.$$

Quando $c > 1$, temos mais termos envolvendo n no denominador do que no numerador de (2.10). Daí,

$$\frac{(n-m)(n-m+1)...(n-2m+c+1)}{(n-1)...(n-m+1)} \rightarrow 0, \quad \text{quando} \quad n \rightarrow \infty.$$

Consequentemente, no somatório sobra um único termo referente a $c = 1$, e portanto,

$$n\ Var\ U(X_1, ..., X_n) \xrightarrow{p} K_1\zeta_1 = m^2\zeta_1, \text{ quando } n \rightarrow \infty.$$

Resta provar (2.9). Para isso, podemos escrever

$$E[U(X_1,...,X_n)-\theta]^2 = Var\ U(X_1,...,X_n),$$

pois $EU(X_1,...,X_n) = Eh(X_1,...,X_m)$. Assim, como $Eh(X_1,...,X_m) = \theta$ temos que provar que:

$$Var\ U(X_1,...,X_n) \to 0 \ \text{ quando } \ n \to \infty.$$

Mas, por (2.8) temos que o limite $nVar\ U(X_1,...,X_n)$ existe e é finito. Logo,

$$\lim_{n\to\infty} Var\ U(X_1,...,X_n) = \lim_{n\to\infty} \frac{nVar\ U(X_1,...,X_n)}{n} = 0.$$

□

2.3 Decomposição de Hoeffding e Aplicações

Nessa seção apresentaremos uma representação obtida por Hoeffding (1961), de uma U-estatística de grau m como soma de U-estatísticas não correlacionadas de grau $1,2,...,m$.

Para obter esta decomposição, definimos recursivamente os núcleos $h^{(1)},...,h^{(m)}$ correspondentes ao núcleo simétrico h, da seguinte forma:

$$h^{(1)} = h^1(x_1) = h_1(x_1) - Eh(X_1,...,X_m) \tag{2.11}$$

e para $c = 2,...,m$,

$$h^{(c)} = h^c(x_1,...,x_c) = h_c(x_1,...,x_c) - \sum_{j=1}^{c-1}\sum_{C_{c,j}} h^j(x_{\alpha_1},...,x_{\alpha_j}) - Eh(X_1,...,X_m). \tag{2.12}$$

Observe que os núcleos h^c definidos em (2.11) e em (2.12) são simétricos em virtude de $h(x_1,...,x_m)$ ser simétrica. É fundamental que $h^{(c)}$ seja simétrica pois veremos que $h^{(c)}$ vai ser o núcleo de uma U-estatística.

Teorema 2.3.1. *(Teorema da decomposição de Hoeffding) Seja $U(X_1, ..., X_n)$ a U-estatística correspondente a um núcleo simétrico $h(x_1, ..., x_m)$ de grau m. Para $c = 1, ..., m$, seja H_n^j a U-Estatística de grau j, $j = 1, ..., m$, baseada no núcleo $h^{(j)}$ definido em* (2.12) *e em* (2.11)*, ou seja,*

$$H_n^j = \binom{n}{j}^{-1} \sum_{C_{n,j}} h^j(X_{\alpha_1}, ..., X_{\alpha_j}). \tag{2.13}$$

Então

$$U(X_1, ..., X_n) = Eh(X_1, ..., X_m) + \sum_{j=1}^{m} \binom{m}{j} H_n^j. \tag{2.14}$$

Provaremos (2.14) em vários passos, que serão apresentados na forma de lemas. Primeiramente, defina para $j < c < m$

$$S_j(i_1, ..., i_m) = \sum_{C_{m,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}). \tag{2.15}$$

Lema 2.3.2. *Temos*

$$\sum_{C_{n,m}} S_j(i_1, ..., i_m) = \binom{n-j}{m-j} \sum_{C_{n,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_m}).$$

Demonstração: Por (2.15) temos

$$\sum_{C_{n,m}} S_j(i_1, ..., i_m) = \sum_{C_{n,m}} \sum_{C_{m,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}). \tag{2.16}$$

Em (2.16), para cada conjunto de m elementos formado a partir do conjunto $\{1, ..., n\}$, estamos considerando todos os conjuntos de j elementos que é possível formar a partir deles. Assim, temos em (2.16) todos os conjuntos de j elementos que é possível obter a partir do conjunto $\{1, ..., n\}$, cada um deles aparece um número k de vezes em (2.16). Por exemplo, se considerarmos os conjuntos $\{1, ..., m\}$ e $\{1, ..., m-1, m+1\}$ formados a partir de $\{1, ..., n\}$, podemos retirar de $\{1, ..., m\}$ e de $\{1, ..., m-1, m+1\}$ o mesmo conjunto $\{1, ..., j\}$. Dessa discussão segue,

$$\sum_{C_{n,m}} S_j(i_1, ..., i_m) = k \sum_{C_{n,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_m}).$$

Vamos determinar k. Note que para ocorrer uma situação similar ao exemplo dado, ou seja, para que conjuntos de m elementos distintos gerem o mesmo conjunto de j elementos, é necessário que eles tenham j elementos em comum, assim o número k de vezes que cada conjunto de j elementos formados a partir do conjunto $\{1, ..., n\}$ aparece em (2.16) é $\binom{n-j}{m-j}$. □

Lema 2.3.3. $h(x_1, ..., x_m) = \sum_{j=1}^{m} S_j(i_1, ..., i_m) + Eh(X_1, ..., X_m)$.

Demonstração:
Da definição de $h^{(m)}$ em (2.12) temos

$$h^m(x_1, ..., x_m) = h_m(x_1, ..., x_m) - \sum_{j=1}^{m-1} \sum_{C_{m,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}) - Eh(X_1, ..., X_m). \quad (2.17)$$

Pela definição de $S_j(i_1, ..., i_m)$ em (2.15), podemos reescrever (2.17),

$$h^m(x_1, ..., x_m) = h_m(x_1, ..., x_m) - \sum_{j=1}^{m-1} S_j(i_1, ..., i_m) - Eh(X_1, ..., X_m). \quad (2.18)$$

Isolando $\sum_{j=1}^{m-1} S_j(i_1, ..., i_m)$ em (2.18), obtemos

$$\sum_{j=1}^{m-1} S_j(i_1, ..., i_m) = h_m(x_1, ..., x_m) - h^m(x_1, ..., x_m) - Eh(X_1, ..., X_m). \quad (2.19)$$

Temos também que

$$\sum_{j=1}^{m} S_j(i_1, ..., i_m) + Eh(X_1, ..., X_m) = \sum_{j=1}^{m-1} S_j(i_1, ..., i_m) + S_m(i_1, ..., i_m) + Eh(X_1, ..., X_m). \quad (2.20)$$

Daí, substituindo (2.19) em (2.20) concluímos,

$$\sum_{j=1}^{m} S_j(i_1, ..., i_m) + Eh(X_1, ..., X_m) = h_m(x_1, ..., x_m) - h^m(x_1, ..., x_m) + S_m(i_1, ..., i_m). \quad (2.21)$$

Novamente pela definição de $S_j(i_1, ..., i_m)$ em (2.15), obtemos

$$S_m(i_1, ..., i_m) = h^m(x_1, ..., x_m). \tag{2.22}$$

Daí,

$$\sum_{j=1}^{m} S_j(i_1, ..., i_j) + Eh(X_1, ..., X_m) = h_m(x_1, .., x_m). \tag{2.23}$$

Mas,

$$h_m(x_1, ..., x_m) = E\{h(X_1, ..., X_m)|X_1 = x_1, ..., X_m = x_m\} = h(x_1, ..., x_m),$$

e assim o resultado está provado. □

Demonstração do Teorema 2.3.1: Dos lemas 2.3.2 e 2.3.3 e de (2.15) segue que

$$\begin{aligned} U(X_1, ..., X_m) &= \binom{n}{m}^{-1} \sum_{C_{n,m}} h(X_{\alpha_1}, ..., X_{\alpha_m}) \\ &= \binom{n}{m}^{-1} \sum_{C_{n,m}} \left[\sum_{j=1}^{m} S_j(i_1, ..., i_m) + Eh(X_1, ..., X_m) \right] \\ &= \binom{n}{m}^{-1} \left[\sum_{j=1}^{m} \sum_{C_{n,m}} S_j(i_1, ..., i_m) \right] + Eh(X_1, ..., X_m) \\ &= \binom{n}{m}^{-1} \sum_{j=1}^{m} \binom{n-j}{m-j} \sum_{C_{n,j}} h^j(X_{\alpha_1}, ..., X_{\alpha_m}) + Eh(X_1, ..., X_m) \\ &= \sum_{j=1}^{m} \binom{n}{m}^{-1} \binom{n-j}{m-j} \sum_{C_{n,j}} h^j(X_{\alpha_1}, ..., X_{\alpha_m}) + Eh(X_1, ..., X_m). \end{aligned}$$

Agora, podemos verificar que:

$$\binom{n}{m}^{-1} \binom{n-j}{m-j} = \binom{m}{j} \binom{n}{j}^{-1}. \tag{2.24}$$

Assim,

$$\begin{aligned} U(X_1,...,X_n) &= \sum_{j=1}^{m} \binom{m}{j} \binom{n}{j}^{-1} \sum_{C_{n,j}} h^j(X_{\alpha_1},...,X_{\alpha_m}) + Eh(X_1,...,X_m). \\ &= \sum_{j=1}^{m} \binom{m}{j} H_n^j + Eh(X_1,...,X_m), \end{aligned}$$

onde H_n^j é uma U-estatística de grau j baseado no núcleo simétrico $h^{(j)}$. □

As componentes U-Estatísticas que aparecem na decomposição de Hoeffding também podem ser escritas como uma U-estatística de grau m e núcleo $S_j(i_1,...,i_m)$ pois da definição (2.13) de H_n^j, de (2.24) e do Lema 2.3.2 podemos escrever:

$$\begin{aligned} \binom{m}{j} H_n^j &= \binom{m}{j} \binom{n}{j}^{-1} \sum_{C_{n,j}} h^j(X_{\alpha_1}...,X_{\alpha_m}) \\ &= \binom{n}{m}^{-1} \binom{n-j}{m-j} \sum_{C_{n,j}} h^j(X_{\alpha_1}...,X_{\alpha_m}) \qquad (2.25) \\ &= \binom{n}{m}^{-1} \sum_{C_{n,m}} S_j(i_1,...,i_m). \end{aligned}$$

Observe também que se considerarmos o termo R_n^c obtido do truncamento da decomposição H depois de c termos, então R_n^c é uma U-estatística de grau m com núcleo $\sum_{j=c+1}^{m} S_j(i_1,...,i_m)$, pois por (2.25) segue

$$\begin{aligned} R_n^c = \sum_{j=c+1}^{m} \binom{m}{j} H_n^j &= \sum_{c+1}^{m} \binom{n}{m}^{-1} \sum_{C_{n,m}} S_j(i_1,...,i_m) \\ &= \binom{n}{m}^{-1} \sum_{C_{n,m}} \sum_{j=c+1}^{m} S_j(i_1,...,i_m). \end{aligned}$$

Falta provar que as componentes U-estatísticas que aparecem na decomposição de Hoeffding são não correlacionadas, mas para isso precisamos do seguinte lema cuja prova será omitida e pode ser vista em detalhes em Lee ((1990), página 28).

Lema 2.3.4. *Sejam $h_c(x_1,...,x_c)$ e $h^c(x_1,...,x_c)$ como foram definidas em (2.6), (2.11) e (2.12),respectivamente. Então*
(i) Para $c = 1,..,j-1$ e $j = 1,...,m$ temos $h^j_c(x_1,...,x_c) = 0$
(ii) $Eh^j(X_1,...,X_j) = 0$.

Teorema 2.3.5. *Seja H^j_n como no Teorema 2.3.1.*
(i) Seja $j < j^{'}$ e sejam $\{\alpha_1,...,\alpha_{j'}\}$ e $\{\beta_1,...,\beta_{j'}\}$ subconjuntos de j e $j^{'}$ elementos respectivamente de $\{1,...,n\}$. Então

$$Cov\ \{h^j(X_{\alpha_1},...,X_{\alpha_j}), h^{j^{'}}(X_{\beta_1},...,X_{\beta_{j'}})\} = 0, \tag{2.26}$$

e, assim,

$$Cov\ (H^j_n, H^{j^{'}}_n) = 0. \tag{2.27}$$

(ii) Sejam $\{\alpha_1,...,\alpha_j\}$ e $\{\beta_1,...,\beta_j\}$ subconjuntos de j elementos distintos de $\{1,...,n\}$. Então

$$Cov\ \{h^j(X_{\alpha_1},...,X_{\alpha_j}), h^j(X_{\beta_1},...,X_{\beta_j})\} = 0. \tag{2.28}$$

(iii) Temos

$$Var\ H^j_n = \binom{n}{j}^{-1} \delta^2_j, \tag{2.29}$$

onde $\delta^2_j = Var\ h^j(X_1,...,X_j)$.

Demonstração:
(i) Para provar (2.26) temos pelo Lema 2.3.4, item (ii), que $Eh^j(X_1,...,X_j) = 0\ \ \forall\ j = 1,...,m$. Assim,

$$Cov\ \{h^j(X_{\alpha_1},...,X_{\alpha_j}), h^{j^{'}}(X_{\beta_1},...,X_{\beta_{j'}})\} = E\{h^j\left(X_{\alpha_1},...,X_{\alpha_j}\right) h^{j^{'}}(X_{\beta_1},...,X_{\beta_{j'}})\}.$$

Como $j < j^{'}$, certamente há elementos de $\{\beta_1,...,\beta_{j'}\}$ que não estão em $\{\alpha_1,...,\alpha_{j'}\}$. Daí, há pelo menos uma variável aleatória $X_{j'}$ que aparece em $h^j(X_{\beta_1},...,X_{\beta_{j'}})$, mas não aparece em $h^j(X_{\alpha_1},...,X_{\alpha_j})$. Portanto, $h^j(X_{\alpha_1},...,X_{\alpha_j})$ é independente de $X_{j'}$. Então, podemos escrever

$$
\begin{aligned}
Cov \ \{h^j(X_{\alpha_1}, ..., X_{\alpha_j})), h^{j'}(X_{\beta_1}, ..., X_{\beta_{j'}})\} = & \\
&= E\{E\ (h^j(X_{\alpha_1}, ..., X_{\alpha_j})h^{j'}(X_{\beta_1}, ..., X_{\beta_{j'}})\ |X_{j'}\} \\
&= E\{h^j(X_{\alpha_1}, ..., X_{\alpha_j})E(h^{j'}(X_{\beta_1}, ..., X_{\beta_{j'}})|X_{j'})\} \\
&= E\{h^j(X_{\alpha_1}, ..., X_{\alpha_j})h_1^{j'}(X_{j'})\} \\
&= 0,
\end{aligned}
$$

pois pelo Lema 2.3.4, item (i), temos que $h_1^{j'}(X_{j'}) = 0$.

Daí, (2.27) segue imediatamente de (2.26), pois

$$
Cov\ (H_n^j, H_n^{j'}) = \binom{n}{j}^{-1}\binom{n}{j'}^{-1}\sum_{C_{n,j}}\sum_{C_{n,j'}} Cov\ \{h^j(X_{\alpha_1}, ..., X_{\alpha_j})), h^{j'}(X_{\beta_1}, ..., X_{\beta_{j'}})\} = 0.
$$

(ii) Se $\{\alpha_1, ..., \alpha_j\} \cap \{\beta_1, ..., \beta_j\}$ é vazia, então $h^j(X_{\alpha_1}, ..., X_{\alpha_j})$ e $h^j(x_{\beta_1}, ..., x_{\beta_j})$ são independentes e portanto a $Cov\ \{(h^j(X_{\alpha_1}, ..., X_{\alpha_j}), h^j(X_{\beta_1}, ..., X_{\beta_j})\} = 0$. Senão, sem perda de generalidade suponha que $\{1, ..., c\}$ sejam os índices das vaviáveis aleatórias que não estão em $\{\alpha_1, ..., \alpha_j\}$. Daí,

$$
\begin{aligned}
Cov\ \{h^j(X_{\alpha_1}, ..., X_{\alpha_j})), h^j(X_{\beta_1}, ..., X_{\beta_j})\} = & \\
&= E\{Eh^j(X_{\alpha_1}, ..., X_{\alpha_j})h^j(X_{\beta_1}, ..., X_{\beta_j})|X_1, ..., X_c\} \\
&= E\{h^j(X_{\alpha_1}, ..., X_{\alpha_j})E\{h^j(X_{\beta_1}, ..., X_{\beta_j})|X_1, ..., X_c\}\} \\
&= Eh^j(X_{\alpha_1}, ..., X_{\alpha_j})h_c^j(X_1, ..., X_c) \\
&= 0
\end{aligned}
$$

pois $c < j$ e, pelo Lema 2.3.4 (i), $h_c^j(X_1, ..., X_c) = 0$.

(iii) Como H_n^j é uma U-estatística baseada no núcleo $h^{(j)}$, podemos usar o Teorema 2.2.6

e obter

$$Var\ H_n^j \;=\; \binom{n}{j}^{-1}\sum_{c=1}^{j}\binom{j}{c}\binom{n-j}{j-c}Var\ h_c^j(X_1,...,X_c).$$

Mas, pelo Lema 2.3.4, $h_c^j(X_1,...,X_c)=0$ para $c=1,...,j-1$ e daí

$$\begin{aligned}Var\ H_n^j &= \binom{n}{j}^{-1}Var\ h_j^j(X_1,...,X_j)\\ &= \binom{n}{j}^{-1}Var\ h^j(X_1,...,X_j)\\ &= \binom{n}{j}^{-1}\delta_j^2.\end{aligned}$$

□

Como aplicação da decomposição de Hoeffding podemos obter uma expressão alternativa para a variância de uma U-estatística em termos das esperanças das variâncias de $h^{(j)}$.

Corolário 2.3.6. *A variância de uma U-estatística é dada por:*

$$Var\ U(X_1,...,X_n)=\sum_{j=1}^{m}\binom{m}{j}^2\binom{n}{j}^{-1}\delta_j^2,$$

onde $\delta_j^2 = Var\ h^j(X_1,...,X_j)$, com h^j definido em (2.12).

Demonstração: Dos teoremas 2.3.1 e 2.3.5 segue

$$\begin{aligned}Var\ U(X_1,...,X_n) &= Var\left[\sum_{j=1}^{m}\binom{m}{j}H_n^j + Eh(X_1,...,X_m)\right]\\ &= Var\sum_{j=1}^{m}\binom{m}{j}H_n^j\\ &= \sum_{j=1}^{m}\binom{m}{j}^2 Var\ H_n^j\end{aligned}$$

$$= \sum_{j=1}^{m} \binom{m}{j}^2 \binom{n}{j}^{-1} \delta_j^2.$$

□

Observe que no Teorema 2.2.6 obtivemos uma expressão para a variância de uma U-estatística em termos de $\sigma_c^2 = Var\ h_c(X_1, ..., X_c)$. Na proposição a seguir obtemos uma relação entre σ_c^2 e δ_j^2 que aparece na expressão da variância obtida no corolário acima.

Proposição 2.3.7. *Temos*

$$\sigma_c^2 = \sum_{j=1}^{c} \binom{c}{j} \delta_j^2,$$

onde σ_c^2 definido em (2.5) e δ_j^2 em (2.29).

Demonstração: Em (2.12), definimos

$$h_c(x_1, ..., x_c) = h^c(x_1, ..., x_c) - \sum_{j=1}^{c-1} \sum_{C_{c,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}) + Eh(X_1, ..., X_m).$$

Por outro lado,

$$\begin{aligned}
\sum_{j=1}^{c} \sum_{C_{c,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}) + Eh(X_1, ..., X_m) &= \sum_{j=1}^{c-1} \sum_{C_{c,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}) + \sum_{C_{c,c}} h^c(x_{\alpha_1}, ..., x_{\alpha_c}) \\
&+ Eh(X_1, ..., X_m) \\
&= \sum_{j=1}^{c-1} \sum_{C_{c,j}} h^j(x_{\alpha_1}, ..., x_{\alpha_j}) + h^c(x_1, ..., x_c) \\
&+ Eh(X_1, ..., X_m) \\
&= h_c(x_1, ..., x_c).
\end{aligned}$$

Assim,

$$h_c(x_1, ..., x_c) = \sum_{j=1}^{c} \sum_{C_{c,j}} h^j(X_{\alpha_1}, ..., X_{\alpha_j}) + Eh(X_1, ..., X_m). \tag{2.30}$$

Daí, calculando a variância de ambos os lados de (2.30) e usando o Teorema 2.3.5, obtemos

$$
\begin{aligned}
Var\ h_c(X_1,...,X_c) &= Var\left[\sum_{j=1}^{c}\sum_{C_{c,j}} h^j(X_{\alpha_1},...,X_{\alpha_j}) + Eh(X_1,...,X_m)\right] \\
&= Var\sum_{j=1}^{c}\sum_{C_{c,j}} h^j(X_{\alpha_1},...,X_{\alpha_j}) \\
&= \sum_{j=1}^{c} Var\ \sum_{C_{c,j}} h^j(X_{\alpha_1},...,X_{\alpha_j}) \\
&= \sum_{j=1}^{c} Var\left[\frac{\binom{c}{j}}{\binom{c}{j}}\ \sum_{C_{c,j}} h^j(X_{\alpha_1},...,X_{\alpha_j})\right] \\
&= \sum_{j=1}^{c} \binom{c}{j}^2 Var\left[\binom{c}{j}^{-1}\ \sum_{C_{c,j}} h^j(X_{\alpha_1},...,X_{\alpha_j})\right] \\
&= \sum_{j=1}^{c} \binom{c}{j}^2 Var\ H_c^j \\
&= \sum_{j=1}^{c} \binom{c}{j}^2 \binom{c}{j}^{-1} \delta_j^2 \\
&= \sum_{j=1}^{c} \binom{c}{j} \delta_j^2.
\end{aligned}
$$

□

Como consequência da proposição anterior obtemos uma relação de ordem entre as variâncias σ_c^2, $0 \leq c \leq m$, que utilizaremos no próximo capítulo.

Proposição 2.3.8. *Para* $0 \leq c \leq d \leq m$

$$\frac{\sigma_c^2}{c} \leq \frac{\sigma_d^2}{d}.$$

Demonstração: Temos que provar que $c\sigma_d^2 - d\sigma_c^2 \geq 0$. Pela Proposição 2.3.7, obtemos

$$\begin{aligned} c\sigma_d^2 - d\sigma_c^2 &= c\sum_{j=1}^{d}\binom{d}{j}\delta_j^2 - d\sum_{j=1}^{c}\binom{c}{j}\delta_j^2 \\ &= \sum_{j=1}^{c}\left[c\binom{d}{j} - d\binom{c}{j}\right]\delta_j^2 + \sum_{c+1}^{d} c\binom{d}{j}\delta_j^2. \end{aligned}$$

O segundo termo da soma é positivo pois para c arbitrário $\delta_c^2 \geq 0$. Precisamos verificar que o primeiro termo também é positivo. Para isso basta verificar que $c\binom{d}{j} - d\binom{c}{j} \geq 0$ para $d \geq c \geq j \geq 1$.

Mas,

$$\begin{aligned} c\binom{d}{j} - d\binom{c}{j} &= c\frac{d!}{j!(d-j)!} - d\frac{c!}{j!(c-j)!} \\ &= \frac{cd}{j!}(d-1)(d-2)...(d-j+1) - \frac{cd}{j!}(c-1)(c-2)...(c-j+1) \geq 0, \end{aligned}$$

pois $d \geq j$. □

2.4 Normalidade Assintótica e Estimação da Variância

Como aplicação da decomposição de Hoeffding (Teorema 2.3.1), podemos obter a normalidade assintótica do estimador U-estatística.

Teorema 2.4.1. *Seja $U_n = U(X_1, ..., X_n)$ a U-Estatística baseada em um núcleo simétrico $h(X_1, ..., X_m)$ correspondente a um parâmetro $\theta = \theta(F)$, $F \in \mathcal{F}$, de grau m.*

Se $Eh^2(X_1, ..., X_m) < \infty$ então $n^{\frac{1}{2}}(U_n - EU_n)$ é assintoticamente normal, com média zero e variância assintótica $m^2\zeta_1 = m^2\sigma_1^2$.

Demonstração: Da decomposição de Hoeffding (Teorema 2.3.1), podemos escrever

$$\begin{aligned}
\sqrt{n}\{U_n - EU_n\} &= \sqrt{n}\left\{Eh(X_1,...,X_m) + \sum_{j=1}^{m}\binom{m}{j}H_n^j - Eh(X_1,...,X_m)\right\} \\
&= \sqrt{n}\left\{\binom{m}{1}H_n^1 + \sum_{j=2}^{m}\binom{m}{j}H_n^j\right\} \\
&= \sqrt{n}\left\{\binom{m}{1}H_n^1 + R_n\right\} \\
&= \sqrt{n}\left\{m\binom{n}{1}^{-1}\sum_{C_{n,1}} h^1(X_i) + \ R_n\right\} \\
&= \sqrt{n}\left\{\frac{m}{n}\sum_{i=1}^{n} h^1(X_i) + \ R_n\right\},
\end{aligned}$$

onde $R_n = \sum_{j=2}^{m}\binom{m}{j}H_n^j$.

Do Teorema 2.3.5 seguem as seguintes igualdades,

$$\begin{aligned}
Var\ \sqrt{n}R_n = nVar\ R_n &= nVar\ \sum_{j=2}^{m}\binom{m}{j}H_n^j \\
&= n\sum_{j=2}^{m} Var\ \binom{m}{j}H_n^j \\
&= \sum_{j=2}^{m} n\binom{m}{j}^2 Var\ H_n^j \\
&= \sum_{j=2}^{m} n\binom{m}{j}^2\binom{n}{j}^{-1}\delta_j^2.
\end{aligned}$$

Então $Var\ \sqrt{n}R_n = O(n^{-1})$ e portanto temos que o comportamento assintótico de

$\sqrt{n}\{U_n - EU_n)\}$ é o mesmo de $\dfrac{m}{\sqrt{n}}\sum_{i=1}^{n} h(X_i)$. Agora,

$Eh^1(X_1) = Eh_1(X_1) - Eh(X_1, ..., X_m) = 0$ e $Var\ h^1(X_1) = Var\ h_1(X_1) = \sigma_1^2 = \zeta_1 < \infty$,

então, como $h^1(X_1)$, $i \geq 1$ são i.i.d, segue do Teorema do Limite Central

$$\frac{\sum_{i=1}^{n} h^1(X_i)}{\sqrt{n}} \xrightarrow{d} N(0,1),\ n \to \infty,$$

o que conclui a prova do teorema. □

Em virtude da dificuldade de se obter a variância assintótica do estimador U_n, finalizamos esta seção apresentando um método consistente de aproximação desta variância assintótica. Para detalhes desse método, veja Arversen (1969).

Considere $X_1, ..., X_n$ uma amostra de tamanho n de v.a's i.i.d. com distribuição $F \in \mathcal{F}$ e seja $h(x_1, ..., x_m)$ um núcleo simétrico de $\theta = \theta(F)$. Suponha que se consiga dividir essa amostra em m grupos de k observações de maneira que $n = mk$. Caso isso não seja possível deletamos alguns dados da amostra.

Defina

$U_n^0 = U_n =$ U-estatística baseada no núcleo h, e para $i = 1, ..., n$,

U_{n-1}^i é o estimador U-estatística obtido deletando o i-ésimo grupo de observações,

e

$$\widehat{U_n} = \frac{\sum_{i=1}^{n} U_{n-1}^i}{n}. \tag{2.31}$$

Teorema 2.4.2. *Seja* $\widehat{U_n}$ *definido em (2.31). Se* $Eh^2(X_1, ..., X_m) < \infty$*, então*

$$s_n^2 = (n-1)\sum_{i=1}^{n}(U_{n-1}^i - \widehat{U}_n)^2 \xrightarrow{p} m^2\zeta_1,\ \text{quando } n \to \infty, \tag{2.32}$$

onde ζ_1 *é dado em (2.7).*

Para demonstrar o Teorema 2.4.2 necessitamos do resultado a seguir, (Teorema 2.4.3), cuja demonstração será omitida e pode ser encontrada em Lee (1990).

Teorema 2.4.3. *Seja S_n uma sequência de estatísticas simétricas e suponha que para $n > m$,*

$$S_n(X_1, ..., X_n) = \frac{\sum_{i=1}^{n} S_{n-1}(X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)}{n}. \tag{2.33}$$

Então S_n é uma U-estatística de grau m. Em contrapartida, qualquer U-estatística de grau m satisfaz (2.33).

Demonstração do Teorema 2.4.2:

Sem perda de generalidade suponha que a U-estatística tem média zero. Vamos considerar que $k = 1$, ou seja, que dividimos a amostra de tamanho n em n grupos. Assim U_{n-1}^i é o estimador U-estatística deletando a observação X_i. Pelo Teorema 2.4.3 temos que $U_n = \widehat{U}_n$. Daí, podemos obter

$$\begin{aligned}
(n-1)\sum_{i=1}^{n}(U_{n-1}^i - \widehat{U}_n)^2 &= (n-1)\sum_{i=1}^{n}(U_{n-1}^i - U_n)^2 \\
&= (n-1)\left[\sum_{i=1}^{n}(U_{n-1}^i)^2 - 2U_n\sum_{i=1}^{n}U_{n-1}^i + \sum_{i=1}^{n}{U_n}^2\right] \\
&= (n-1)\left[\sum_{i=1}^{n}(U_{n-1}^i)^2 - nU_n^2\right] \\
&= (n-1)\sum_{i=1}^{n}\binom{n-1}{m}^{-2}\sum_{i} h(S)h(T) \\
&- n(n-1)\binom{n}{m}^{-2}\sum_{C_{n,m}} h(X_{\alpha_1}, ..., X_{\alpha_m})h(X_{\beta_1}, ..., X_{\beta_m}),
\end{aligned}$$

onde $\sum_{i}$ indica a soma sobre todos os conjuntos S e T de m elementos que podem ser escolhidos no conjunto $\{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$.

Se S e T tiverem c elementos em comum então esses dois conjuntos totalizam $2m - c$

elementos distintos, logo $h(S)h(T)$ aparece em $n-2m+c$ termos do somatório

$$\sum_{c=0}^{m} h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}).$$

Consequentemente,

$$\sum_{i=1}^{n}\sum_{i} h(S)h(T) = \sum_{c=0}^{m}(n-2m+c)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}).$$

Logo,

$$\begin{aligned}(n-1)\sum_{i=1}^{n}&(U_{n-1}^{i}-\widehat{U}_n)^2 = \\ = \quad &(n-1)\binom{n-1}{m}^{-2}\sum_{c=0}^{m}(n-2m+c)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}) - \\ - \quad &n(n-1)\binom{n}{m}^{-2}\sum_{c=0}^{m}h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_1},...,X_{\beta_{m_c}}).\end{aligned} \tag{2.34}$$

Desenvolvendo o primeiro termo da diferença em (2.34), obtemos

$$(n-1)\binom{n-1}{m}^{-2}\sum_{c=0}^{m}(n-2m+c)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}})\frac{n}{n} =$$

$$= \frac{n-1}{n}\binom{n-1}{m}^{-2}\sum_{c=0}^{m}(n^2-2mn+cn)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}).$$

Desenvolvendo o segundo termo da diferença em (2.34), também obtemos,

$$n(n-1)\binom{n}{m}^{-2}\sum_{c=0}^{m}h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_m}},...,X_{\beta_{m_c}}) =$$

$$= \frac{(n-1)}{n}\binom{n-1}{m}^{-2}\sum_{c=0}^{m}(n^2-2mn+m^2)h(X_{\alpha_{1_m}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}).$$

Logo, substituindo em (2.34), temos

$$(n-1)\sum_{i=1}^{n}(U_{n-1}^{i}-U_n)^2 =$$

$$= \frac{n-1}{n}\binom{n-1}{m}^{-2}\left\{\sum_{c=0}^{m}(n^2-2mn+cn)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}) \right. -$$

$$\left. - \sum_{c=0}^{m}(n^2-2mn+m^2)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}})\right\} =$$

$$= \frac{n-1}{n}\binom{n-1}{m}^{-2}\sum_{c=0}^{m}(cn-m^2)h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}).$$

Denotando,

$$U_c = \binom{n}{c}^{-1}\binom{n-c}{m-c}^{-1}\binom{n-m}{m-c}^{-1}h(X_{\alpha_{1_c}},...,X_{\alpha_{m_c}})h(X_{\beta_{1_c}},...,X_{\beta_{m_c}}),$$

podemos escrever,

$$(n-1)\sum_{i=1}^{n}(U_{n-1}^{i}-U_n)^2 = \sum_{c=0}^{m}\frac{n-1}{n}\binom{n-1}{m}^{-2}\binom{n}{c}\binom{n-c}{m-c}\binom{n-m}{m-c}(cn-m^2)U_c. \tag{2.35}$$

Desenvolvendo o termo geral do somatório em (2.35), segue para $c = 0, 1, ..., m$:

$$\frac{n-1}{n}\binom{n-1}{m}^{-2}\binom{n}{c}\binom{n-c}{m-c}\binom{n-m}{m-c}(cn-m^2)U_c =$$

$$= \frac{n-1}{n}\left[\frac{m!}{(m-c)!}\right]^2\frac{n...(n-2m+c+1)}{[(n-1)...(n-m)]^2}(cn-m^2)U_c.$$

Seja

$$A_n(c) = \frac{n...(n-2m+c+1)}{[(n-1)...(n-m)]^2}. \tag{2.36}$$

No produto $A_n(c)$ há $2m-c$ e $2m$ fatores envolvendo n no numerador e no denominador, respectivamente.

Quando $c = 1$, o termo geral de (2.35) se reduz a

$$m^2\frac{(n-1)}{n}(nA_n(1)-m^2A_n(1)), \tag{2.37}$$

onde em $nA_n(1)$ há o mesmo número de termos envolvendo n no numerador e denominador. Assim $nA_n(1) \to 1$ e $m^2A_n(1) \to 0$ quando $n \to \infty$. Daí, o termo geral quando $c = 1$ tende para m^2 quando $n \to \infty$.

Quando $c > 1$, em $nA_n(c)$ temos $2m - c + 1$ termos envolvendo n no numerador e no denominador sempre há $2m$ termos envolvendo n, então $nA_n(c) \to 0$ e $m^2A_n(c) \to 0$. Assim, o termo geral para $c > 1$ reduz a

$$\left[\frac{m!}{(m-c)!}\right]^2 \frac{(n-1)}{n}(cnA_n(c) - m^2A_n(c)) \to 0, \text{ quando } n \to \infty.$$

E portanto, no somatório sobra um único termo referente a $c = 1$, daí, se provarmos que $U_c \xrightarrow{p} \zeta_c$ concluímos a prova do teorema.

Observe que U_c é uma U-estatística com núcleo simétrico dado por:

$K(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\beta_1}, ..., X_{\beta_{m-c}}, X_{\gamma_1}, ..., X_{\gamma_{m-c}}) =$

$$= \binom{(2m-c)}{c(m-c)(m-c)}^{-1} \sum_{P_{m_c}} h(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\beta_1} ..., X_{\beta_{m-c}})h(X_{\alpha_1}, ..., X_{\alpha_c}, X_{\gamma_1}, ..., X_{\gamma_{m-c}}),$$

onde $\sum_{P_{m_c}}$ indica a soma sobre todas as $\binom{(2m-c)}{c(m-c)(m-c)}$ permutações do conjunto $\{\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_{m-c}, \gamma_1, ..., \gamma_{m-c}\}$.

Agora,

$EK(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\beta_1}, ..., X_{\beta_{m-c}}, X_{\gamma_1}, ..., X_{\gamma_{m-c}}) =$

$= \zeta_c - Eh(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\beta_1}, ..., X_{\beta_{m-c}})Eh(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\gamma_1}, ..., X_{\gamma_{m-c}})$.

Logo,

$$EU_c = \zeta_c - Eh(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\beta_1}, ..., X_{\beta_{m-c}})Eh(X_{\alpha_1}, ..., X_{\alpha_m}, X_{\gamma_1}, ..., X_{\gamma_{m-c}}).$$

Como estamos assumindo que o parâmetro estimável é zero segue que $EU_c = \zeta_c$ e sob a condição que $Eh^2(X_1, ..., X_m) < \infty$ segue de (2.9) do Corolário 2.2.9 do Teorema 2.2.6 que $U_c \xrightarrow{p} \zeta_c$. Portanto o resultado está provado. $\square$

Capítulo 3

Estimador do Índice Caudal com Estrutura de U-Estatística

3.1 Introdução

Consideramos uma função de distribuição F no domínio de atração de alguma distribuição Levy estável, ou seja, $F \in DA(\alpha)$ com $0 < \alpha < 2$. Pela Proposição 1.3.8 temos que

$$x^{\alpha}(1 - F(x)) \sim pL(x) \quad e \quad x^{\alpha}(F(-x)) \sim qL(x), \text{ quando } \quad x \to \infty,$$

onde $p, q \geq 0$, $p + q = 1$ e L é lentamente variante.

Ou seja, F possui cauda à direita ou cauda à esquerda regularmente variantes com o mesmo expoente $-\alpha$, que mede a espessura das respectivas caudas.

Assim, as funções de distribuição no domínio de atração de leis estáveis, $0 < \alpha < 2$, são distribuições de cauda pesada com mesmo índice caudal α.

Nós estamos interessados no problema de estimação do índice caudal α, a partir de uma amostra $(X_1, ..., X_n)$ de variáveis aleatórias i.i.d. com função de distribuição F.

Neste capítulo estudamos o estimador robusto, com estrutura de U-estatística, proposto por Fan (2004).

Na Seção 3.2 apresentamos a motivação e os princípios básicos para a escolha do núcleo do estimador. Especificamente, consideramos

$$\widehat{\alpha_n^{-1}} = \frac{log\ (X_1^2 + ... + X_n^2)}{2\ log\ n}, \tag{3.1}$$

onde, $X_1^2, ..., X_n^2$ estão no domínio de atração de uma lei $\frac{\alpha}{2}$-estável, com $0 < \frac{\alpha}{2} < 1$, pois $F \in DA(\alpha)$. Usando as idéias Meerschaert e Scheffler (1998) provamos que $\widehat{\alpha_n^{-1}}$ é um estimador de α^{-1} assintóticamente não viesado e consistente (Teorema 3.2.1). Além disso, apresentamos um princípio dos grandes desvios obtido por Fan (2004).

Na Seção 3.3 consideramos o estimador com estrutura de U-estatística, baseado no núcleo dado em (3.1), ou seja, para $m < n$ seja

$$\widehat{\alpha_U{}^{-1}} = \binom{n}{m}^{-1} \sum_{\alpha \in C_{n,m}} \frac{log\ (X_{\alpha_1}^2 + ... + X_{\alpha_m}^2)}{2\ log\ n},$$

onde $\alpha = \{\alpha_1, ..., \alpha_m\}$ e o somatório é sobre todas as combinações de m inteiros escolhidos sem reposição do conjunto de inteiros 1 a n. Em virtude da estrutura de U-estatistica, a vantagem de α_U^{-1} é que ele tem variância menor do que $\widehat{\alpha_n^{-1}}$.

Ainda nesta seção, provamos que, quando $m \to \infty$, α_U^{-1} é assintoticamente não viesado e fracamente consistente e a consistência em média quadrática é obtida sob a condição adicional que $m = o(n^{\frac{1}{2}})$ quando $n \to \infty$. Vale observar que tendo em vista que as provas apresentadas no Capítulo 2, referentes à teoria clássica de U-estatísticas, permanecem válidas quando o núcleo considerado é assintóticamente não viesado, foi possível utilizar os resultados lá obtidos nas demonstrações das propriedades assintóticas de α_U^{-1} descritas nas seções 3.3 e 3.4.

Finalizamos o capítulo apresentando na Seção 3.5 uma breve análise das simulações numéricas realizadas por Fan (2004) sobre a performance do estimador com estrutura de U-estatística em comparação com outros estimadores conhecidos, dentre eles o estimador de Hill (1975).

3.2 Estimação do Índice Caudal

Sejam $X_1, ..., X_n, ...$ variáveis aleatórias i.i.d. com função de distribuição F e suponha que $F \in DA(\alpha)$. Ou seja, existem sequências de constantes $\{a_n\}$ e $\{b_n\}$, com $a_n > 0$, tais que:

$$\frac{X_1 + ... + X_n - b_n}{a_n} \xrightarrow{d} X, \tag{3.2}$$

onde X é uma v.a. com distribuição α-estável, onde $\alpha \in (0, 2)$.

Da Proposição 1.3.7, podemos escolher $a_n = n^{\frac{1}{\alpha}} L(n)$, onde L é lentamente variante.

Como dissemos anteriormente, o índice de estabilidade α da v.a. X corresponde ao índice caudal da f.d. F, o qual deseja-se estimar.

Por facilidade, lembremos que se $F \in DA(\alpha)$ então a distribuição comum de $X_1^2, X_2^2, ...$ está no domínio de atração de uma distribuição $\dfrac{\alpha}{2}$ -estável. Como $0 < \dfrac{\alpha}{2} < 1$ então temos da Proposição 1.3.7,

$$\frac{X_1^2 + ... + X_n^2}{a_n} \xrightarrow{d} Y, \tag{3.3}$$

onde Y é uma variável aleatória $\dfrac{\alpha}{2}$ -estável e $a_n = n^{\frac{2}{\alpha}} L(n)$, com $L(n)$ uma função lentamente variante.

Como motivação para a construção do estimador do índice caudal α, baseado em Fan (2004), consideremos primeiramente o caso particular em que as variáveis aleatórias $X_1^2, ..., X_n^2$ são estáveis, ou seja,

$$\frac{X_1^2 + ... + X_n^2}{n^{\frac{2}{\alpha}}} \stackrel{d}{=} X_1^2.$$

Assim, temos

$$\frac{log\ (X_1^2 + ... + X_n^2)}{log\ n} - \frac{2}{\alpha} \stackrel{d}{=} \frac{log\ (X_1^2)}{log\ n}, \tag{3.4}$$

Se definirmos

$$\widehat{\alpha_n^{-1}} = \frac{log\ (X_1^2 + ... + X_n^2)}{2\ log\ n}, \tag{3.5}$$

então por (3.4) temos para $\epsilon > 0$

$$P\left(\left|\widehat{\alpha_n^{-1}} - \frac{1}{\alpha}\right| > \epsilon\right) = P\left(\left(\log\ X_1^2\right) > 2\epsilon\ log\ n\right) \longrightarrow 0,\ n \to \infty.$$

Assim, neste caso particular, temos que $\widehat{\alpha_n^{-1}}$ é um estimador fracamente consistente de $\frac{1}{\alpha}$.

No teorema a seguir mostramos que para o caso mais geral em que F está no domínio de atração de uma distribuição α-estável, o estimador definido em (3.5) é também fracamente consistente. Além disso, também obtemos a consistência em média quadrática. A demonstração segue as idéias de Meerschaert e Scheffler (1998).

Teorema 3.2.1. *Sejam $X_1, X_2, ...$ variáveis aleatórias i.i.d. com função de distribuição $F \in DA(\alpha)$, $0 < \alpha < 2$ e seja $\widehat{\alpha_n^{-1}}$ definido em (3.5). Então*
(i) existe uma sequência de constantes $\{c_n\} > 0$, com $c_n > 0$ e $c_n \to 0$, quando $n \to \infty$, tais que

$$2log\ n\left(\widehat{\alpha_n^{-1}} - \frac{1}{\alpha} + c_n\right) \xrightarrow{d} log\ |Y|, \tag{3.6}$$

onde Y é uma v.a. $\frac{\alpha}{2}$-estável.

(ii) $\widehat{\alpha_n^{-1}} \xrightarrow{p} \frac{1}{\alpha}$, quando $n \to \infty$.

(iii) $E(\widehat{\alpha_n^{-1}}) \to \frac{1}{\alpha}$ e $E\left(\widehat{\alpha_n^{-1}} - \frac{1}{\alpha}\right)^2 \longrightarrow 0$ quando $n \to \infty$.

(iv) $Var\ [2log\ n\widehat{\alpha_n^{-1}}] \longrightarrow Var\ \ [log\ |Y|] < \infty$, quando $n \to \infty$.

Demonstração: Da hipótese segue o limite (3.3) e, consequentemente,

$$log\ \left(\frac{X_1^2 + ... + X_n^2}{a_n}\right) \xrightarrow{d} log\ |Y|, \tag{3.7}$$

onde Y é α-estável, $0 < \frac{\alpha}{2} < 1$ e $a_n = n^{\frac{2}{\alpha}} L(n)$, $L(n)$ lentamente variante.

(i) Podemos escrever

$$\begin{aligned} log\ \left(\frac{X_1^2 + ... + X_n^2}{a_n}\right) &= log\ (X_1^2 + ... + X_n^2) - log\ n^{\frac{2}{\alpha}} L(n) \\ &= log\ (X_1^2 + ... + X_n^2) - \frac{2}{\alpha} log\ n - log\ L(n). \end{aligned} \tag{3.8}$$

Note que, como $L(n)$ é lentamente variante, segue da Proposição 1.2.2, item (iv) que $\dfrac{log\ L(n)}{log\ n} \to 0$ quando $n \to \infty$. Então de (3.7) e (3.8) temos

$$(2log\ n)\left(\widehat{\alpha_n^{-1}} - \frac{1}{\alpha} - \frac{log\ L(n)}{2log\ n}\right) = log\ \left(\frac{X_1^2 + ... + X_n^2}{a_n}\right) \xrightarrow{d} log\ |Y|.$$

Assim (i) está provado com $c_n = \dfrac{log\ L(n)}{2log\ n}$.

(ii) Para facilitar, denote $Z_n = 2\ log\ n\left(\widehat{\alpha_n^{-1}} - \dfrac{1}{\alpha} - c_n\right)$.

Temos

$$\widehat{\alpha_n^{-1}} - \frac{1}{\alpha} = \frac{Z_n}{2log\ n} + c_n \xrightarrow{p} 0, \tag{3.9}$$

pois por (3.6) $Z_n \xrightarrow{d} log\ |Y|$ e $c_n \to 0$.

(iii) Como Y é α-estável pela Proposição 1.3.10, seque que $E\{log\ |Y|\} < \infty$ e como $Z_n = log\ \left|\dfrac{X_1^2 + ... + X_n^2}{a_n}\right| \xrightarrow{d} log\ |Y|$, segue da Proposição 1.3.11 que

$$EZ_n \longrightarrow E\{log\ |Y|\} < \infty\ e\ EZ_n^2 \longrightarrow E\{(log\ |Y|)^2\} < \infty.$$

Assim, como $\widehat{\alpha_n^{-1}} - \dfrac{1}{\alpha} = \dfrac{Z_n}{2log\ n} + c_n$, com $c_n \to 0$, o resultado segue.

(iv) Basta observar que

$$Var\ [2log\ n\ \widehat{\alpha_n^{-1}}] = Var\ \left(Z_n + \frac{log\ n}{\alpha} + c_n log\ n\right) = Var\ Z_n,$$

e por (iii) $Var\ Z_n \longrightarrow Var\ \ [log\ |Y|] < \infty$. $\square$

Além da consistência provada acima, um princípio de grandes desvios para o estimador

$\widehat{\alpha_n^{-1}}$ foi provado por Fan (2004), o qual apresentamos a seguir.

Teorema 3.2.2. *Sejam $X_1, ..., X_n$ variáveis aleatórias i.i.d no domínio de atração de uma distribuição α-estável e seja $\widehat{\alpha_n^{-1}}$ definido em (3.5). Então, para cada $\epsilon > 0$,*

$$(i) \lim_{n\to\infty} \frac{log\ P(\widehat{\alpha_n^{-1}} - \alpha^{-1} > \epsilon)}{log\ n} = -\alpha\epsilon.$$

$$(ii) \lim_{n\to\infty} \frac{log\ P(\widehat{\alpha_n^{-1}} - \alpha^{-1} < -\epsilon)}{log\ n} = -\epsilon.$$

Para demonstrar este teorema, necessitaremos de um resultado auxiliar, cuja demonstração será omitida e pode ser encontrada em Heyde (1968).

Lema 3.2.3. *(Lema de Heyde) Suponha que $S_{(n)} = X_1 + ... + X_n$, com $X_1, ..., X_n$ i.i.d. no domínio de atração de uma distribuição α-estável X e $c_n \to \infty$, quando $n \to \infty$. Então*

$$\lim_{n\to\infty} \frac{P(|S_{(n)}| > a_n c_n)}{nP(|X| > a_n c_n)} = 1.$$

Demonstração do Teorema 3.2.2:
(i) Da hipótese temos (3.3) com $a_n = n^{\frac{1}{\alpha}} L(n)$, onde $L(n)$ é uma função lentamente variante.

Denotando $S_n = X_1^2 + ... + X_n^2$, da definição (3.5) de $\widehat{\alpha_n^{-1}}$ podemos escrever:

$$\begin{aligned}
P(\widehat{\alpha_n^{-1}} - \alpha^{-1} > \epsilon) &= P\left(\frac{log\ (X_1^2 + ... + X_n^2)}{2log\ n} - \frac{1}{\alpha} > \epsilon\right) \\
&= P\left(\frac{log\ (X_1^2 + ... + X_n^2)}{2log\ n} - \frac{1}{\alpha} - \frac{log\ L(n)}{log\ n} > \epsilon - \frac{log\ L(n)}{log\ n}\right) \\
&= P\left(log\ |S_n| - log\ n^{\frac{2}{\alpha}} - log\ (L(n))^2 > log\ n^{2\epsilon} - log\ (L(n))^2\right) \\
&= P\left(log\ \left(\frac{|S_n|}{n^{\frac{2}{\alpha}}(L(n))^2}\right) > log\ \frac{n^{2\epsilon}}{(L(n))^2}\right) \\
&= P\left(|S_n| > \frac{a_n^2 n^{2\epsilon}}{(L(n))^2}\right) \\
&= P\left(|S_n| > a_n^2 c_n\right),
\end{aligned}$$

onde $c_n = n^{2\epsilon}(L(n))^{-2}$ e $a_n = n^{\frac{1}{\alpha}}L(n)$.

Agora, como $c_n = n^{2\epsilon}(L(n))^{-2}$ é regularmente variante com expoente $\rho = 2\epsilon > 0$, então da Proposição 1.2.2, temos que $c_n \to \infty$ quando $n \to \infty$. Assim, por (3.3) e pelo Lema de Heyde, segue que para $n \to \infty$

$$P(\widehat{\alpha_n^{-1}} - \alpha^{-1} > \epsilon) \sim nP\left(|Y| > a_n^2 c_n\right) = nP\left(|Y| > n^{2\epsilon + \frac{2}{\alpha}}\right). \tag{3.10}$$

Mas, como Y é $\dfrac{\alpha}{2}$-estável então da Proposição 1.3.8 temos que $P(|Y| > x) \sim Cx^{\frac{-\alpha}{2}}$, quando $x \to \infty$, onde C é uma constante. Logo, de (3.10) obtemos

$$\begin{aligned} P(\widehat{\alpha_n^{-1}} - \alpha^{-1} > \epsilon) &\sim Cn(n^{\frac{2}{\alpha}} + 2\epsilon)^{-\frac{\alpha}{2}} \\ &\sim Cn^{-\alpha\epsilon}. \end{aligned}$$

Daí, segue $log\ P(\widehat{\alpha_n^{-1}} - \alpha^{-1} > \epsilon) \sim log\ C - \alpha\epsilon log\ n,\ quando\ n \to \infty$ e assim (i) está provado.

(ii) Sejam $\phi_{S_n}(t)$ e $\phi_{X_1^2}(t)$ as funções característica de $S_n = X_1^2 + ... + X_n^2$ e de X_1^2 respectivamente. Então, como $X_1^2, ..., X_n^2$ são i.i.d., da fórmula da inversão de funções características podemos escrever

$$\begin{aligned} P(\widehat{\alpha_n^{-1}} - \alpha^{-1} < -\epsilon) &= P\left(\frac{|S_n|}{n^{\frac{2}{\alpha}}} < n^{-\epsilon}\right) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{exp(itn^{-\epsilon}) - exp(-in^{-\epsilon}t)}{it}\phi_{S_n}(\frac{t}{n^{2/\alpha}})dt \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{sen(n^{-\epsilon}t)}{t}\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt \\ &= \frac{n^{-\epsilon}}{\pi}\int_{-\infty}^{\infty} \frac{sen(n^{-\epsilon}t)}{n^{-\epsilon}t}\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt \\ &= \frac{n^{-\epsilon}}{\pi}\int_{-\infty}^{0} \frac{sen(n^{-\epsilon}t)}{n^{-\epsilon}t}\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt \end{aligned}$$

$$
\begin{aligned}
&+ \quad \frac{n^{-\epsilon}}{\pi}\int_0^\infty \frac{sen(n^{-\epsilon}t)}{n^{-\epsilon}t}\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt \\
&= \quad \frac{n^{-\epsilon}}{\pi}\int_0^\infty \frac{sen(n^{-\epsilon}t)}{n^{-\epsilon}t}\left[\phi_{X_1^2}\left(\frac{-t}{n^{2/\alpha}}\right)\right]^n dt \\
&+ \quad \frac{n^{-\epsilon}}{\pi}\int_0^\infty \frac{sen(n^{-\epsilon}t)}{n^{-\epsilon}t}\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt.
\end{aligned}
$$

Como $\lim_{x\to 0}\frac{senx}{x} = 1$, então quando $n \to \infty$ segue que

$$
P(\widehat{\alpha_n^{-1}}-\alpha^{-1} < -\epsilon) \sim \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]\left[\phi_{X_1^2}\left(\frac{-t}{n^{2/\alpha}}\right)\right]^n dt + \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]\left[\phi_{X_1^2}\left(\frac{t}{n^{2/\alpha}}\right)\right]^n dt.
$$

Como a f.d. de X_1^2 está no domínio de atração de uma distribuição $\frac{2}{\alpha}$-estável, com $0 < \frac{\alpha}{2} < 1$, então da Proposição 1.3.2 segue que sua função característica é da forma

$$
\phi_{X_1^2}(t) = exp\left\{-c_0|t|^{\frac{\alpha}{2}}L\left(\frac{1}{|t|}\right)\left(1 - i\beta sinal(t)tan\left(\frac{\pi\alpha}{2}\right)(1+o(1))\right)\right\},
$$

onde L é uma função lentamente variante. Assim, obtemos

$$
\begin{aligned}
&P(\widehat{\alpha_n^{-1}} - \alpha^{-1} < -\epsilon) \sim \\
&\sim \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0 n\left|\frac{t}{n^{2/\alpha}}\right|^{\alpha/2} L\left(\frac{n^{2/\alpha}}{t}\right)\left(1 - i\beta sinal(t)\tan\left(\frac{\pi\alpha}{4}\right)[1+o(1)]\right)\right\} dt \\
&+ \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0 n\left|\frac{t}{n^{2/\alpha}}\right|^{2/\alpha} L\left(\frac{n^{2/\alpha}}{t}\right)\left(1 - i\beta sinal(-t)\tan\left(\frac{\pi\alpha}{4}\right)[1+o(1)]\right)\right\} dt \\
&\sim \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0 t^{\frac{\alpha}{2}}L(n^{2/\alpha})[1+o(1)]\left(1 - i\beta\tan\left(\frac{\pi\alpha}{4}\right)[1+o(1)]\right)\right\} dt \\
&+ \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)])exp\left\{-c_0 t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\left(1 + i\beta\tan\left(\frac{\pi\alpha}{4}\right)[1+o(1)]\right)\right\} dt \\
&= \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0 t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\left(1 - i\beta K[1+o(1)]\right)\right\} dt
\end{aligned}
$$

$$+\ \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\left(1+i\beta K[1+o(1)]\right)\right\}dt$$

$$\sim\ \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\right\}exp\left\{c_0t^{\alpha/2}L(n^{2/\alpha})i\beta K[1+o(1)]\right\}dt$$

$$+\ \frac{n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\right\}exp\left\{-c_0t^{\alpha/2}L(n^{2/\alpha})i\beta K[1+o(1)]\right\}dt,$$

onde $K=\tan\left(\frac{\pi\alpha}{4}\right)$. Mas, usando que $e^{ix}=\cos x+i$ sen x obtemos,

$$P(\widehat{\alpha_n^{-1}}-\frac{1}{\alpha}<-\epsilon)=$$

$$=\frac{2n^{-\epsilon}}{\pi}\int_0^\infty [1+o(1)]exp\left\{-c_0t^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\right\}\cos\left\{\beta Kc_ot^{\alpha/2}L(n^{2/\alpha})[1+o(1)]\right\}dt.$$

Fazendo a mudança de variável $u=t\left[L(n^{2/\alpha})\right]^{\frac{2}{\alpha}}$, seque que

$$P(\widehat{\alpha_n^{-1}}-\frac{1}{\alpha}<-\epsilon)\ \sim\ K_1n^{-\epsilon}\left[L(n^{2/\alpha})\right]^{\frac{-2}{\alpha}}\int_0^\infty exp\{-c_2u^{\alpha/2}\}\cos\{c_1u^{\alpha/2}du\}.$$

Agora, considerando a mudança $s=c_2u^{\alpha/2}$, temos

$$\begin{aligned}\int_0^\infty exp\{-c_2u^{\alpha/2}\}\cos\{\ c_1u^{\alpha/2}\}du &\leq \int_0^\infty exp\{-c_2u^{\alpha/2}\}du\\ &= \frac{2c_2}{\alpha}\int_0^\infty s^{1-\frac{2}{\alpha}}e^{-s}ds\\ &= \frac{2c_2}{\alpha}\int_0^\infty s^{\left(2-\frac{2}{\alpha}\right)-1}e^{-s}ds\\ &= \frac{2c_2}{\alpha}\Gamma(2-\frac{2}{\alpha})<\infty,\end{aligned}$$

onde $2-\dfrac{2}{\alpha}>0$ e $\Gamma(x)$ denota a função gama. Assim,

$$P(\widehat{\alpha_n^{1}}-\frac{1}{\alpha}<-\epsilon)\sim Cn^{-\epsilon}\left[L(n^{2/\alpha})\right]^{\frac{-2}{\alpha}},$$

onde C é uma constante.

Pela Proposição 1.2.2, item (iii), temos que $L(n^{2/\alpha})$ é lentamente variante, pelos itens

(ii) e (iv), respectivamente da mesma proposição, segue que $h(n) = L(n^{2/\alpha})^{-\frac{\alpha}{2}}$ é lentamente variante e $\dfrac{log\ h(n)}{log\ n} \to 0$ quando $n \to \infty$, daí, obtemos,

$$\lim_{n\to\infty} \frac{log\ P\left(\widehat{\alpha_n^{-1}} - \dfrac{1}{\alpha} < -\epsilon\right)}{log\ n} = -\epsilon + C \lim_{n\to\infty} \frac{log\ L(n^{2/\alpha})^{\frac{-\alpha}{2}}}{log\ n} = -\epsilon.$$

$\square$

Dos resultados anteriores podemos observar que o estimador $\widehat{\alpha_n^{-1}}$ definido em (3.5), embora seja robusto e consistente, pode ter variância alta. Como vimos no capítulo anterior, um caminho natural para diminuir a variância do estimador, preservando suas propriedades assintóticas, é introduzir uma estrutura de U-estatística. Assim, na próxima seção definimos um estimador com estrutura de U-estatística baseado no estimador $\widehat{\alpha_n^{-1}}$.

3.3 Estimador com Estrutura de U-Estatística

Considere $X_1, ..., X_n$ uma amostra i.i.d. de uma função de distribuição F no domínio de atração de uma distribuição α-estável, $0 < \alpha < 2$. Ou seja, temos as relações (3.3) e (3.5) descritas na seção anterior.

Para $m < n$, defina a função

$$h(x_1, ..., x_m) = \frac{log\ (x_1^2 + ... + x_n^2)}{2log\ n}, \tag{3.11}$$

ou seja, $h(X_1, ..., X_m) = \widehat{\alpha_m^{-1}}$, onde $\widehat{\alpha_m^{-1}}$ é o estimador definido em (3.5).

Temos que $h(x_1, ..., x_m)$ é uma função simétrica nos seus argumentos e do Teorema 3.2.1 seque que, quando $m \to \infty$,

$$\widehat{\alpha_m^{-1}} \longrightarrow \frac{1}{\alpha} \tag{3.12}$$

e

$$E\left(\widehat{\alpha_m^{-1}} - \frac{1}{\alpha}\right)^2 \longrightarrow 0, \tag{3.13}$$

ou seja, $\widehat{\alpha_m^{-1}}$ é um estimador assintóticamente não viesado de α^{-1} e seu erro médio quadrático tende a zero, quando $m \to \infty$.

Baseado no núcleo h, definimos o seguinte estimador de α^{-1}, com estrutura de U-estatística

$$\widehat{\alpha_U^{-1}} = U_n(h) = \binom{n}{m}^{-1} \sum_{\alpha \in C_{n,m}} \frac{log\ (X_{\alpha_1}^2 + ... + X_{\alpha_m}^2)}{2log\ n}, \tag{3.14}$$

onde $\alpha = \{\alpha_1, ..., \alpha_m\} \in C_{n,m}$ e o somatório é sobre todas as combinações de m inteiros escolhidos sem reposição do conjunto de inteiros 1 a n.

Assim, $\widehat{\alpha_U^{-1}}$ é uma média dos estimadores $h(X_{\alpha_1}, ..., X_{\alpha_m})$ considerados sobre todas as possíveis sub-amostras de tamanho m extraídas da amostra $(X_1, ..., X_n)$.

Cabe observar que, embora o núcleo $h(X_1, ..., X_m)$ não seja não viesado, como requerido na definição de U-estatística dada no Capítulo 2, as propriedades que usaremos nesta e na próxima seção referentes à teoria de U-estatística, permanecem válidas sob a condição de que $h(X_1, ..., X_m)$ seja assintoticamente não viesado, conforme pode ser comprovado analisando-se as respectivas demonstrações apresentadas no capítulo anterior.

Desta forma, segue imediatamente que $E\widehat{\alpha_U^{-1}} = Eh(X_1, ..., X_m)$ e, daí, $\widehat{\alpha_U^{-1}}$ é um estimador de α^{-1} assintoticamente não viesado e fracamente consistente, desde que $m \to \infty$. Além disso, podemos obter a consistência em média quadrática de $\widehat{\alpha_U^{-1}}$, sob certas restrições sobre m. É o que provamos no próximo teorema.

Teorema 3.3.1. *Sejam $X_1, ..., X_n$ observações i.i.d. com distribuição F, que está no domínio de atração de uma lei α-estável, $0 < \alpha < 2$, e seja $\widehat{\alpha_U^{-1}}$ definido em (3.14).*
(i) Se $m \to \infty$ então

$$E\widehat{\alpha_U^{-1}} \longrightarrow \alpha^{-1} \quad e \quad \widehat{\alpha_U^{-1}} \xrightarrow{p} \alpha^{-1}. \tag{3.15}$$

(ii) Se $m \to \infty$ e $m = o(n^{\frac{1}{2}})$, quando $n \to \infty$ então

$$\widehat{\alpha_U^{-1}} \xrightarrow{(2)} \frac{1}{\alpha}. \tag{3.16}$$

Demonstração:
(i) Temos $E\widehat{\alpha_U^{-1}} = Eh(X_1, ..., X_m)$ e (3.15) segue imediatamente de (3.12) e (3.13).
(ii) De (3.13) temos que $Eh^2(X_1, .., X_m) < \infty$ para m suficientemente grande, então do Corolário 2.2.9 podemos concluir que

$$\lim_{n \to \infty} nVar\ \widehat{\alpha_U^{-1}} = m^2 \zeta_1, \tag{3.17}$$

onde $\zeta_1 = Var\ E(h(X_1, ..., X_m)|X_1)$.

Agora, como $E\widehat{\alpha_U{}^{-1}} = Eh(X_1, ..., X_m)$, temos

$$E(\widehat{\alpha_U{}^{-1}} - \alpha^{-1})^2 = Var\ \widehat{\alpha_U{}^{-1}} + \{Eh(X_1, ..., X_m) - \alpha^{-1}\}^2. \tag{3.18}$$

Mas por (3.12) temos $Eh(X_1, ..., X_m) - \alpha^{-1} \to 0$ quando $m \to \infty$ e por (3.16), como $m = o(n^{\frac{1}{2}})$ quando $n \to \infty$ e $\zeta_1 < \infty$ para m suficientemente grande, segue que

$$Var\ \widehat{\alpha_U{}^{-1}} = \frac{nVar\widehat{\alpha_U{}^{-1}}}{m^2}\frac{m^2}{n} \to 0, \quad \text{quando} \quad n \to \infty\ e\ m \to \infty.$$

Daí segue (3.15). □

Vale ressaltar que, em virtude das propriedades decorrentes da estrutura de U-estatística, o estimador $\widehat{\alpha_U{}^{-1}}$ definido em (3.14) tem menor variância e melhora o estimador $\widehat{\alpha_n{}^{-1}}$ dado por (3.5), na seção anterior, no sentido que ele possui menor erro médio quadrático, pois por (3.18) segue

$$E[\widehat{\alpha_U{}^{-1}} - \alpha^{-1}]^2 < Var\ \widehat{\alpha_n^{-1}} + [E\widehat{\alpha_n^{-1}} - \alpha^{-1}]^2 = E[\widehat{\alpha_n^{-1}} - \alpha^{-1}]^2.$$

3.4 Normalidade Assintótica

Nessa seção, apresentaremos a demonstração da normalidade assintótica do estimador $\widehat{\alpha_U{}^{-1}}$, sob a restrição $m \to \infty$ e $m = o(n^{\frac{1}{2}})$ quando $n \to \infty$.

Para isto, podemos observar que é muito difícil obter a distribuição de $\widehat{\alpha_U{}^{-1}}$, bem como a sua variância assintótica que, por (3.17), é dada em termos de ζ_1, ou seja, $nVar\ \widehat{\alpha_U{}^{-1}} \sim m\zeta_1$.

Assim, vamos utilizar o método do cálculo da variância assintótica descrito na seção 2.4. Se $X_1, X_2, ...$ são v.a's i.i.d. com função de distribuição $F \in DA(\alpha)$, considere o estimador com estrutura de U-estatística $\widehat{\alpha_U{}^{-1}}$ definido em (3.14), baseado no núcleo h dado por 3.11.

Defina

$U_n^0 = \widehat{\alpha_U{}^{-1}}$

e para $i = 1, ..., n$,

U^i_{n-1} = a U-estatística correspondente a h, baseada em todas as observações exceto X_i, isto é,

$$X_1, ..., X_{i-1}, X_{i+1}, ..., X_n.$$

Denote

$$\widehat{U_n} = \frac{\sum_{i=1}^{n} U^i_{n-1}}{n}$$

e defina

$$s^2_n = (n-1)\sum_{i=1}^{n}(U^i_{n-1} - \widehat{U}_n)^2. \tag{3.19}$$

Então, pelo Teorema 2.4.2, temos que

$$s^2_n \xrightarrow{p} m^2\zeta_1, \text{ quando } n \to \infty.$$

Ou seja,

$$\frac{s^2_n}{nVar\ \widehat{\alpha^{-1}_U}} \xrightarrow{p} 1, \text{ quando } n \to \infty. \tag{3.20}$$

Podemos, então, enunciar o teorema que garante a normalidade assintótica de $\widehat{\alpha_U{}^{-1}}$, sob condições impostas sobre m.

Teorema 3.4.1. *Sejam $X_1, ..., X_n$ com distribuição F no domínio de atração de uma lei α-estável, $0 < \alpha < 2$. Seja $\widehat{\alpha^{-1}_U}$ definido em (3.14) e (3.11). Se $m \to \infty$ e $m = o(n^{\frac{1}{2}})$ quando $n \to \infty$ então*

$$\sqrt{n}s^{-1}_n(\widehat{\alpha_U{}^{-1}} - E\widehat{\alpha^{-1}_U}) \xrightarrow{d} N(0,1), \ quando\ n \to \infty,$$

onde s^2_n é dado por (3.19).

Observe que, sob as condições do Teorema 3.4.1, temos $E\widehat{\alpha^{-1}_U} = Eh(X_1, ..., X_m) = E\widehat{\alpha^{-1}_m}$ e da prova do Teorema 3.2.1 (ii), temos $E\widehat{\alpha^{-1}_m} = \dfrac{1}{\alpha} + E\left(\dfrac{Z_m}{2log\ m} + c_m\right)$, onde $Z_m = log\ \dfrac{X^2_1 + ... + X^2_m}{a_m}$, $a_m = n^{\frac{2}{\alpha}}L(m)$ e $E\left(\dfrac{Z_m}{2log\ m}\right) + c_m \longrightarrow 0$, quando $m \to \infty$.

Assim, segue que

$$\sqrt{n}s_n^{-1}\left(\widehat{\alpha_U{}^{-1}} - \frac{1}{\alpha} + d_m\right) \xrightarrow{d} N(0,1), \text{ quando } n \to \infty, \tag{3.21}$$

onde $d_m \to 0$ quando $m \to \infty$ e $\dfrac{m^2}{n} \to 0$ quando $n \to \infty$. Portanto, $\widehat{\alpha_U^{-1}}$ é um estimador com muito pouco viés quando o tamanho da amostra é grande.

Para a demonstração do Teorema 3.4.1 necessitaremos de alguns lemas auxiliares.

Lema 3.4.2. *Seja $X_1, ..., X_m$ variáveis aleatórias i.i.d. no domínio de atração de uma lei α-estável, $0 < \alpha < 2$. Para $m < n$, defina*

$$Z = E\left(log\ \left(1 + \frac{X_1^2}{X_2^2 + ... + X_m^2}\right)\middle| X_1\right). \tag{3.22}$$

Então

$$\lim_{m\to\infty} mEZ = \frac{2}{\alpha} \tag{3.23}$$

e

$$\liminf_{m\to\infty} mEZ^2 \geq d_1, \tag{3.24}$$

onde d_1 é uma constante.

Demonstração:
(i) Para provar (3.23), denote $S_m = X_1^2 + ... + X_m^2$. Então, usando propriedade de esperança condicional, obtemos

$$\begin{aligned} EZ &= E\left(log\ \left(1 + \frac{X_1^2}{X_2^2 + ... + X_m^2}\right)\right) \\ &= E\left(log\ \frac{S_m}{S_{m-1}}\right) \\ &= E\ log\ S_m - E\ log\ S_{m-1}. \end{aligned}$$

Pelo Teorema 3.2.1, item (iii), temos $Eh(X_1, ..., X_m) = E\left(\dfrac{log\ S_m}{2log\ m}\right) \to \dfrac{1}{\alpha}$ quando

$m \to \infty$. Assim quando $m \to \infty$ temos

$$\begin{aligned}
EZ &= E\ log\ S_m - E\ log\ S_{m-1} \\
&\sim 2\left[\frac{log\ m}{\alpha} - \frac{log\ (m-1)}{\alpha}\right] \\
&= \frac{2}{\alpha} log\ \frac{m}{m-1} \\
&= \frac{2}{\alpha} log\ \left(1 - \frac{1}{m}\right)^{-1} \\
&\sim \frac{2}{\alpha m}.
\end{aligned}$$

(ii) Provemos (3.24). Da desigualdade de Chebyschev temos que $P(|Z| \geq m|EZ|) \leq \dfrac{EZ^2}{m^2|EZ|^2}$. Então, por (3.23) temos para $m \to \infty$

$$\begin{aligned}
EZ^2 &\geq P(|Z| \geq m|EZ|)m^2(EZ)^2 \\
&\sim \frac{4}{\alpha^2} P(|Z| \geq \frac{2}{\alpha}).
\end{aligned}$$

Denote

$$Y_{m-1} = \frac{X_2^2 + ... + X_m^2}{L(m-1)(m-1)^{\frac{2}{\alpha}}}. \tag{3.25}$$

Então, como $X \geq a \Rightarrow E(X|Y) \geq a$, temos

$$P(|Z| \geq \frac{2}{\alpha}) \geq P\left(\left|log\left(1 + \frac{X_1^2}{X_2^2 + ... + X_m^2}\right)\right| \geq \frac{2}{\alpha}\right). \tag{3.26}$$

Logo para $P\left(\left(log\left|1 + \frac{X_1^2}{X_2^2 + ... + X_m^2}\right)\right| \geq \frac{2}{\alpha}\right)$, temos as seguintes possibilidades:

$$ou\ log\left(1 + \frac{X_1^2}{L(m-1)(m-1)^{\frac{2}{\alpha}} Y_{m-1}}\right) \geq \frac{2}{\alpha}\ ou\ log - \left(1 + \frac{X_1^2}{L(m-1)(m-1)^{\frac{2}{\alpha}} Y_{m-1}}\right) \geq \frac{2}{\alpha}$$

$$ou\ log\left(1+\frac{X_1^2}{L(m-1)(m-1)^{\frac{2}{\alpha}}Y_{m-1}}\right)\leq\frac{-2}{\alpha}\ ou\ log-\left(1+\frac{X_1^2}{L(m-1)(m-1)^{\frac{2}{\alpha}}Y_{m-1}}\right)\leq\frac{-2}{\alpha}.$$

Assim, de (3.26) segue

$$\begin{aligned}
EZ^2 &\geq \frac{4}{\alpha^2}P\left(\frac{X_1^2}{Y_{m-1}}\geq(e^{\frac{2}{\alpha}}-1)L(m-1)(m-1)^{\frac{2}{\alpha}}\right)+\\
&+ \frac{4}{\alpha^2}P\left(\frac{X_1}{Y_{m-1}}\leq-(1+e^{\frac{-2}{\alpha}})L(m-1)(m-1)^{\frac{2}{\alpha}}\right)+\\
&+ \frac{4}{\alpha^2}P\left(\frac{X_1^2}{Y_{m-1}}\leq-(1-e^{\frac{2}{\alpha}})L(m-1)(m-1)^{\frac{2}{\alpha}}\right)+ \qquad (3.27)\\
&+ \frac{4}{\alpha^2}P\left(\frac{X_1^2}{Y_{m-1}}\geq-(1+e^{\frac{2}{\alpha}})L(m-1)(m-1)^{\frac{2}{\alpha}}\right)+\\
&= I_1+I_2+I_3+I_4.
\end{aligned}$$

Basta calcular I_1, pois $I_2+I_3+I_4$ são da mesma ordem que I_1.

Denotando $A_m=e^{\frac{2}{\alpha}}-1)L(m-1)(m-1)^{\frac{2}{\alpha}}$, podemos escrever para todo $\epsilon>0$

$$\begin{aligned}
\frac{\alpha^2 I_1}{4} &= P\left(\frac{X_1^2}{Y_{m-1}}\geq A_m\right)\\
&= \int_0^\infty P\left(X_1^2\geq A_m s\right)dF_{Y_{m-1}}(s)+\int_{-\infty}^0 P\left(X_1^2\leq A_m s\right)dF_{Y_{m-1}}(s) \qquad (3.28)\\
&\geq \int_\epsilon^\infty P\left(X_1^2\geq A_m s\right)dF_{Y_{m-1}}(s)+\int_{-\infty}^\epsilon P\left(X_1^2\leq A_m s\right)dF_{Y_{m-1}}(s).
\end{aligned}$$

Agora, como $X_1^2,...,X_n^2$ estão no domínio de atração de uma distribuição $\dfrac{\alpha}{2}$-estável, com $0<\dfrac{\alpha}{2}<1$, da Proposição 1.3.8 seque que quando $m\to\infty$

$$P(X_1^2\geq A_m s)\sim pA_m^{-\alpha/2}s^{-\alpha/2}L(A_m s)$$

e

$$P(X_1^2 \leq A_m s) \sim q A_m^{-\alpha/2} |s|^{-\alpha/2} L(A_m s),$$

onde $p + q = 1$.

Substituindo a expressão de A_m, segue

$$P(X_1^2 \geq A_m s) \sim \frac{c_1 s^{-\alpha/2}}{m-1}, \quad c_1 = pL(m-1)^{-\alpha/2}(e^{\frac{1}{\alpha}} - 1)^{-\alpha/2} > 0 \tag{3.29}$$

$$P(X_1^2 \leq A_m s) \sim \frac{c_2 |s|^{-\alpha/2}}{m-1}, \quad c_2 = qL(m-1)^{-\frac{\alpha}{2}}(e^{\frac{1}{\alpha}} - 1)^{-\alpha/2} > 0. \tag{3.30}$$

Assim, $P(X_1^2 \geq A_m s)$ e $P(X_1^2 \leq A_m s)$ variam regularmente com expoente negativo $\frac{-\alpha}{2}$. Então pelo Teorema 1.2.4 segue que o limite em (3.29) é uniforme sobre $s \in [\epsilon, \infty)$ e o limite em (3.30) é uniforme sobre $|s| \in [\epsilon, \infty)$ e daí, particularmente para $s \in (-\infty, \epsilon]$.

Por outro lado, como por hipótese $Y_{m-1} \xrightarrow{d} Y$, onde Y é uma variável aleatória $\frac{\alpha}{2}$-estável, então das propriedades de convergência em distribuição (vide por exemplo Chung (1974), pag 98) segue de (3.28) que, quando $m \to \infty$,

$$\frac{\alpha^2 I_1}{4} \sim \frac{c_1}{m-1} \int_{\epsilon}^{\infty} s^{-\alpha/2} dF_Y(s) + \frac{c_2}{m-1} \int_{-\epsilon}^{-\infty} |s|^{-\alpha/2} dF_Y(s).$$

Como cada uma das integrais acima são finitas e positivas, quando $m \to \infty$ segue que

$$\frac{\alpha^2 I_1}{4} \sim \frac{d_1}{m},$$

onde d_1 é uma constante positiva. Portanto, de (3.27) segue o resultado. $\square$

Lema 3.4.3. *Sob as hipóteses do Lema 3.4.2, temos que*

$$\zeta_1 = Var\ E(h(X_1, ..., X_m)|X_1) \sim \frac{d}{m\ (2log\ m)^2},\ \textit{quando}\ m \to \infty,$$

onde h é definido como em (3.11) e $d > 0$ constante.

Demonstração: Podemos escrever

$$
\begin{aligned}
h(X_1,...,X_m) &= \frac{log\ \sum_{i=1}^{m} X_i^2}{2log\ m} \\
&= \frac{1}{2log\ m}\left[log\ \left(1+\frac{X_1^2}{X_2^2+...+X_m^2}\right)+log\ (X_1^2+...+X_m^2)\right]
\end{aligned}
$$

Substituindo a expressão de Z dada em (3.22) no Lema 3.4.2, pela linearidade da esperança condicional e pela independência de X_1 e $(X_2^2,...,X_m^2)$ segue

$$
\begin{aligned}
\zeta_1 &= Var\ E\,(h(X_1,...,X_m)|X_1) \\
&= \frac{Var\ Z}{(2log\ m)^2}+Var\ E\left(\left.\frac{log\ (X_2^2+...+X_m^2)}{2log\ m}\right| X_1\right) \\
&= \frac{Var\ Z}{(2log\ m)^2} \\
&= \frac{EZ^2-(EZ)^2}{(2log\ m)^2}.
\end{aligned}
$$

Por outro lado, pelo Lema 3.4.2, segue que

$$
\begin{aligned}
\liminf_{m\to\infty} m(2log\ m)^2\zeta_1 &= \liminf_{m\to\infty}\left[mEZ^2-\frac{(mEZ)^2}{m}\right] \\
&\geq d_1.
\end{aligned} \tag{3.31}
$$

Por outro lado, pela Proposição 2.3.8, temos

$$
\begin{aligned}
\zeta_1 &\leq \frac{\zeta_m}{m} \\
&= \frac{Var\ h(X_1,...,X_m)}{m}.
\end{aligned}
$$

Mas, como $h(X_1,...,X_m) = \widehat{\alpha_m^{-1}}$, do Teorema 3.2.1 (iv), segue que

$$Var\ h(X_1,...,X_m) \sim \frac{Var\ log\ |Y|}{(2log\ m)^2},\ \text{quando } m \to \infty,$$

onde Y é $\frac{\alpha}{2}$-estável, quando $m \to \infty$ e $Var\ (\log\ |Y|) < \infty$. Assim, temos

$$\limsup_{m\to\infty} m(2log\ m)^2 \zeta_1 \leq Var\ \ [log\ |Y|] = d_2 < \infty. \tag{3.32}$$

Assim, de (3.31) e (3.32), podemos concluir

$$\zeta_1 \sim \frac{d}{m(2log\ m)^2},\ \text{quando } m \to \infty$$

para alguma constante $d > 0$. □

Finalmente, podemos demonstrar a normalidade assintótica de $\widehat{\alpha_U^{-1}}$.

Demonstração do Teorema 3.4.1: Temos que $\widehat{\alpha_U{}^{-1}} = U_n(h) = \binom{n}{m}^{-1} \sum_{C_{n,m}} h(X_{\alpha_1},...,X_{\alpha_m})$ e $h(X_1,...,X_m)$ definido por (3.11). Então usando o Teorema 2.3.1 podemos obter a decomposição de Hoeffding para $\widehat{\alpha_U{}^{-1}}$ dada por:

$$\begin{aligned}\widehat{\alpha_U{}^{-1}} &= Eh(X_1,...,X_m) + \sum_{c=1}^{m} \binom{m}{c} H_n^c \\ &= Eh(X_1,...,X_m) + mH_n^1 + \sum_{c=2}^{m} \binom{m}{c} H_n^c.\end{aligned}$$

onde $H_n^c = \binom{n}{c}^{-1} \sum_{C_{n,c}} h^c(X_{\alpha_1},...,X_{\alpha_c})$ e $h^{(c)}$ definidos como em (2.11) e (2.12).

Seja $R_n = \sum_{c=2}^{m} \binom{m}{c} H_n^c$. Como $H_n^1 = \frac{\sum_{i=1}^{n} h^1(X_i)}{n}$ e $\widehat{\alpha_U{}^{-1}} = Eh(X_1,...,X_m)$ podemos escrever

$$\widehat{\alpha_U{}^{-1}} = E\widehat{\alpha_U{}^{-1}} + \frac{m}{n} \sum_{i=1}^{n} h^1(X_i) + R_n,$$

e assim,

$$s_n^{-1}\sqrt{n}(\widehat{\alpha_U^{-1}} - E\widehat{\alpha_U^{-1}}) = \frac{m}{s_n\sqrt{n}}\sum_{i=1}^{n} h^1(X_i) + s_n^{-1}\sqrt{n}R_n. \tag{3.33}$$

(i) Vamos provar primeiramente que para m suficientemente grande

$$\frac{m}{s_n\sqrt{n}}\sum_{i=1}^{n} h^1(X_i) \xrightarrow{d} N(0,1), \text{ quando } n \to \infty. \tag{3.34}$$

Para isto, note que $h^1(X_1) = E(h(X_1, ..., X_m)|X_i) - Eh(X_1, ..., X_m)$.

Assim, $h^1(X_1), ..., h^1(X_n)$ são v.a's i.i.d com $Eh^1(X_i) = 0$ e $Var\ h_1(X_i) = \zeta_1 < \infty$, pois para m suficientemente grande $Eh^2(X_1, ..., X_m) < \infty$. Então temos

$$Var\ \sum_{i=1}^{n} h^1(X_i) = n\zeta_1, \tag{3.35}$$

e do Teorema 2.4.2, segue

$$\frac{ns_n^2}{m^2 Var\ \sum_{i=1}^{n} h^1(X_i)} \xrightarrow{p} 1, \text{ quando } n \to \infty.$$

Assim, (3.34) segue do Teorema do Limite Central.

(ii) Provaremos que, quando $m \to \infty$ e $m = o(n^{1/2})$ quando $n \to \infty$, temos

$$\sqrt{n}s_n^{-1}R_n \xrightarrow{p} 0, \text{ quando } n \to \infty. \tag{3.36}$$

De fato, temos $R_n = \sum_{c=2}^{m}\binom{m}{c}H_n^c$, onde H_n^c é definido como na decomposição de Hoeffding. Então, pelo Teorema 2.3.4 os H_n^j's são não correlacionados e

$$Var\ H_n^c = \binom{n}{c}^{-1}\delta_c^2,$$

onde $\delta_c^2 = Var\ h^c(X_1, ..., X_c)$, segue que

$$Var\ R_n = \sum_{c=2}^{m} \binom{m}{c}^2 \binom{n}{c}^{-1} \delta_c^2.$$

Por outro lado, do Corolário 2.2.9 da decomposição de Hoeffding de $\widehat{\alpha_U^{-1}}$, segue que

$$\begin{aligned} Var\ \widehat{\alpha_U^{-1}} &= \sum_{c=1}^{m} \binom{m}{c}^2 \binom{n}{c}^{-1} \delta_c^2 \\ &= \frac{m^2 \zeta_1}{n} + Var\ R_n. \end{aligned} \tag{3.37}$$

Mas, usando a expressão alternativa para $Var\ \widehat{\alpha_U^{-1}}$ obtida do Teorema 2.2.6, podemos escrever

$$Var\ \widehat{\alpha_U^{-1}} = \binom{n}{m}^{-1} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \sigma_c^2,$$

onde $\zeta_c = \sigma_c^2$ pelo Lema 2.2.8. Então

$$Var\ \widehat{\alpha_U^{-1}} = \binom{n}{m}^{-1} \binom{m}{1} \binom{n-m}{m-1} \zeta_1 + \binom{n}{m}^{-1} \sum_{c=2}^{m} \binom{m}{c} \binom{n-m}{m-c} \zeta_c, \tag{3.38}$$

temos também

$$\binom{n}{m}^{-1} \binom{m}{1} \binom{n-m}{m-1} = \frac{m^2}{n} \left(1 - \frac{m-1}{n-1}\right) \left(1 - \frac{m-1}{n-2}\right) ... \left(1 - \frac{m-1}{n-m+1}\right)$$

Como $m = o(n^{\frac{1}{2}})$ então $\dfrac{m-1}{n-k} \to 0$, $k = 1, ..., m-1$. e daí temos

$$\binom{n}{m}^{-1} \binom{m}{1} \binom{n-m}{m-1} \sim \frac{m^2}{n}.$$

Assim, comparando (3.37) e (3.38) temos que

$$Var\ R_n \sim \frac{\displaystyle\sum_{c=2}^{m} \binom{m}{c} \binom{n-m}{m-c} \zeta_c}{\dbinom{n}{m}}.$$

Pelo Teorema 2.3.8, temos que $\zeta_1 \leq \frac{\zeta_2}{2} \leq ... \leq \frac{\zeta_m}{m}$. Assim, obtemos

$$\begin{aligned}
Var\ R_n &\sim \frac{\sum_{c=2}^{m} \binom{m}{c}\binom{n-m}{m-c}\zeta_c}{\binom{n}{m}} \\
&\leq \frac{\zeta_m}{m}\frac{\sum_{c=2}^{m} c\binom{m}{c}\binom{n-m}{m-c}}{\binom{n}{m}} \\
&\leq \frac{\zeta_m}{m}\frac{\sum_{c=1}^{m} c\binom{m}{c}\binom{n-m}{m-c}}{\binom{n}{m}} - o(1),
\end{aligned}$$

pois $\frac{\binom{m}{1}\binom{n-m}{m-1}}{\binom{n}{m}} = o(1)$ quando $\frac{m^2}{n} \to 0$.

$$Var\ R_n \leq \frac{m^2}{n}\frac{\zeta_m}{m}(1 + o(1)).$$

Agora, como vimos na prova do Lema 3.4.2, segue do Teorema 3.2.1 que

$$(2log\ m)^2\zeta_m = (2log\ m)^2 Var\ h(X_1, ..., X_m) \sim Var\ log|Y| < \infty,$$

quando $m \to \infty$, ou seja, para m suficientemente grande

$$\frac{\zeta_m}{m} \sim \frac{Var\ log|Y|}{m(2log\ m)^2},$$

e pelo Lema 3.4.2 segue que $\frac{\zeta_m}{m} \sim K\zeta_1$, para uma certa constante $K > 0$. Logo, podemos concluir que, se $\frac{m^2}{n} \longrightarrow 0$ quando $n \to \infty$, então

$$Var\ R_n \leq \frac{m^2 K\zeta_1}{n}(1 + o(1)),$$

e assim,

$$\frac{Var\ R_n}{\frac{m^2\zeta_1}{n}} \to 0 \ \ se \ \ \frac{m^2}{n} \to 0 \text{ quando } n \to \infty.$$

Daí, segue da Desigualdade de Chebyschev que

$$\frac{R_n}{\frac{m\sqrt{\zeta_1}}{\sqrt{n}}} \xrightarrow{p} 0,$$

e, como $s_n^2 \xrightarrow{p} m^2\zeta_1$, então podemos obter (3.36). Finalmente, de (3.33), (3.34), (3.36) obtemos (3.36) como queríamos demonstrar. □

3.5 Comparação com outros estimadores

No artigo de Fan (2004), são apresentados os resultados de simulações numéricas para analisar a performance do estimador $\widehat{\alpha_U^{-1}}$ definido em (2.3). Nesta seção apresentamos uma breve análise destes resultados.

Primeiramente, observemos que, embora o estimador com estrutura de U-estatística tenha propriedades muito boas, na prática quando o tamanho da amostra n ou o grau m é grande, o cálculo de $\widehat{\alpha_U^{-1}}$ inclui a média de $\binom{n}{m}$ termos, podendo ser um processo muito demorado. Uma solução, apresentada por Fan (2004), para esse problema é inicialmente dividir a amostra em sub-amostras de igual tamanho, podendo-se deletar alguns termos, caso seja necessário, sem aumentar devidamente a variância. Para cada sub-amostra, considera-se o estimador com estrutura de U-estatística e a média desses estimadores é um estimador denominado U-estatística incompleta. Assim, por motivos computacionais, no caso de amostras grandes, as simulações realizadas no artigo de Fan (2004) são feitas para o estimador de U-estatística incompleta $\widehat{\alpha_{U_o^{-1}}}$ obtido através do estimador $\widehat{\alpha_U^{-1}}$.

Para o caso de estimação do índice caudal de distribuições de cauda pesada no domínio de atração de leis α-estáveis, foram feitas simulações comparando-se a performance do estimador $\widehat{\alpha_{U_o^{-1}}}$ com os estimadores de Hill (1975) e o estimador de Haan e Resnick (1980). Foram consideradas amostras de tamanho $20.50 = 1000$ e foram realizadas 1000 repetições de cada experimento.

Os resultados computacionais obtidos por Fan (2004) mostram que, em termos do erro médio quadrático, o estimador com estrutura de U-estatística tem performance melhor do que os outros dois estimadores analisados.

Cabe ainda ressaltar que, quando os estimadores são usados para estimar o índice caudal de uma distribuição normal, somente o estimador de Fan tem performance muito boa.

Foram feitas simulações para o caso de amostras de distribuições α-estáveis e a performance do estimador de Fan foi comparada com os estimadores de Press (1972) e Zolotarev (1986). Neste caso, quando $\alpha < 1$ os estimadores de Fan tem performance tão boa quanto o de Press e ambos são melhores do que o de Zolotarev. Já para $\alpha > 1$ o estimador de Fan é tão bom quanto o de Zolotarev e ambos são melhores do que o de Press. No entanto, o estimador de Fan é muito mais simples de implementar, enquanto os outros dois apresentam certas restrições de implementação.

Embora não tenha sido objeto de estudo dessa dissertação, no mesmo artigo de Fan (2004) são apresentados alguns procedimentos alternativos para melhorar a performance do estimador principal. Um deles baseia-se num procedimento de re-amostragem e é indicado por Fan na situação de amostras de tamanho não muito grande. Para mais detalhes, quanto aos aspectos teóricos e resultados computacionais, vide o referido artigo.

Bibliografia

[1] Arversen, J., *Jackknifing U-statistics*, Ann. Math. Statist 40, (6), 2076-2100., (1969).

[2] Bingham, N.H., Goldie, C.M., Teugels, J.L., *Regular Variation*, Cambridge University Press., 1987.

[3] Chung, K.L., *A Course in Probability Theory*, 2nd Ed, Academic Press, New York, U.S.A., 1974.

[4] Csörgö, S., Deuvels, P., Mason, D., *Kernel estimates of the tail index of a distribution*, Ann. Statist. 13, 1050-1077., (1985).

[5] De Haan, L., Resnick, S., *On Asymptotic Normality of the Hill Estimator*, Ann. Statist, 743-756., (1985).

[6] De Haan, L., Resnick, S., *A simple asymptotic estimate for the index of a stable distribution*, J. Roy. Statist. Soc. Ser. B 42, 83-88., (1980).

[7] De Haan, L., T. Pereira., *Estimation the index of a stable distribution*, Statistics & Probability Letters 41, 39-55., (1999).

[8] Embrechts, P., Kliippelberg C., Mikosch T., *Modeling Extremal Events for Insurance and Finance*, Springer., 1997.

[9] Fan, Z., *Estimation Problems for distributions with heavy tails*, J. Statist. Plann. Inference, 13-40., (2004).

[10] Feller, W., *An Introduction to Probability theory and its Applications II*, Wiley, New York., 1971.

[11] Heyde, C.C., *On large deviation probabilities in the case of attraction to a non-normal stable law*, Ann. Statist, 253-255., (1968).

[12] Hill, B.M., *A simple general approach to inference about the tail of a distribution*, Ann. Probab(3)., 1163-1174., (1975).

[13] Hoeffding, W., *A class of U-statistics with asymptotically normal distribution*, Ann. Math.Statist 19., 293-325., (1948).

[14] Hoeffding, W., *The strong law of large numbers for U-statistics*, Institute of Statistics Mimeo Series Nº 302., University of North Carolina., 1961.

[15] Ibragimov, I.A., Linnik, Y. V., *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen., 1971.

[16] Korolyuk, V.S., Borovskich, Y.V., *Theory of U-Statistics*, Kluwer Academic Publishers, Dordrecht., 1994.

[17] Lee, A. J., *U-statistics Theory and Practice*, New York and Basel., 1990.

[18] Mason, D., *A strong invariance theorem for the tail empirical process*, Ann. Inst. Henri Poincaré, 24, 491-506., (1988).

[19] Meerschaert M.M., Scheffler, H.P., *A simple robust estimation method for the thickness of heavy tails*, J. Statist. Plann. Inference (71), 19-34., (1998).

[20] Nolan, J., *Stable Distributions Models for Heavy Tailed Data*, Birkhauser., 2007. **Nota:** Em progresso, Capítulo 1 online em http://academic2.american.edu/ jpnolan/stable/chap1.pdf.

[21] Pickands J., *Statistical inference using extreme order statistics*, Ann. Statist.(3), 119-131., (1975).

[22] Press, S. J., *Estimation in univariate and multivariate stable distributions*, J. Amer. Statist. Assoc. 67 340, 842-846., (1972).

[23] Randles, R.H., Wolfe, D.A., *Introduction to the Theory of Non Parametric Statistics*, Wiley, New York., 1979.

[24] Rohatgi, V.K., Saleh, A.K., *An Introduction to Probability and Statistics*, Wiley, New York., 2001.

[25] Samorodnitski, G., Taqqu, M. S., *Stable non-Gaussian Random Processes. Stochastic Models with Infinite Variance*, Chapman and Hall, London Paris Tokyo., 1987.