BOOTSTRAP E BOOTSTRAP ITERADO

JORGE MIGUEL MILHAZES DE FREITAS

Resumo. Com olhos postos na inferência estatística, começamos por apresentar o método bootstrap introduzido por Efron (1979). Com base neste método, e numa abordagem semelhante à seguida por Beran e Ducharme (1991) e Diciccio e Romano (1988), estudamos a construção de intervalos de confiança e a respectiva consistência assimptótica. Posteriormente, introduzimos a pré-pivotagem, estabelecida por Beran (1987), como refinamento da metodologia bootstrap. Exibimos alguns resultados, expostos por Hall, acerca de expansões Edgeworth e a sua utilidade na demonstração de algumas propriedades dos métodos bootstrap. Seguidamente, debruçamonos sobre o problema da estimação do índice de cauda de Pareto, expondo um procedimento bootstrap de cauda, introduzido por J. Bacro e M. Brito.(1998). Por fim, exibimos um refinamento do boostrap de cauda, estabelecemos a sua consistência assimptótica e calculamos a sua ordem de convergência.

Conteúdo

Date: Janeiro de 2001.
Key words and phrases. Bootstrap, pré-pivotagem, intervalos de confian ça, consistência assimptótica, expansões Edgeworth, boostrap de cauda, prepivotagem de cauda.

1. Introdução- Algumas definições e notação

Seja X uma variável aleatória (v.a.) com função de distribuição contínua à direita (f.d.) F dada por $F(x) = P(X \leq x)$.

Consideremos $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ um conjunto de observações independentes de X. Dizemos que $\mathbb{X}_n$ é uma a.a. de v.a. independentes e identicamente distribuídas (i.i.d.) com f.d. F.

Com base na a.a. $\mathbb{X}_n$, contruímos a função de distribuição empírica (f.d.e.), F_n, da seguinte forma:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{X_i \leq x\},$$

onde $\mathbb{I}$ designa a função indicatriz.

Definimos a função quantil da f.d F do seguinte modo:

$$F^{-1}(\alpha) = \inf\{x : F(x) \geq \alpha\}.$$

Habitualmente, F é desconhecida e apenas é sabido que $F \in \mathbb{F}$, um conjunto de funções de distribuição.

Essencialmente existem dois contextos:

- paramétrico- em que $\mathbb{F}$ pode ser indexado por um parâmetro $\theta \in \mathbb{R}^n$, de tal forma que $F = F_\theta$;
- não paramétrico- em que $\mathbb{F}$ é demasiado grande para ser indexado.

Para um estimador de F vamos introduzir a notação $\widehat{F}_n$. No primeiro contexto deverá ser considerado como $\widehat{F}_n = F_{\widehat{\theta}_n}$, onde $\widehat{\theta}_n$ é um estimador de θ, por exemplo, um estimador de máxima verosimilhança. Já no segundo contexto, habitualmente, considera-se que $\widehat{F}_n = F_n$, ou seja, toma-se a f.d.e. como estimador não paramétrico de F. Realce-se que em ambos os casos o estimador $\widehat{F}_n$ de F é calculado com base na a.a. $\mathbb{X}_n$.

Definimos uma **amostra bootstrap**, $\mathbb{X}_n^* = (X_1^*, X_2^*, ..., X_n^*)$, como sendo uma a.a. de v.a. com f.d. $\widehat{F}_n$. Usaremos a notação * para indicar que nos referimos à amostra bootstrap. Quando $\widehat{F}_n = F_n$ (**naïve bootstrap**), a amostra bootstrap pode ser obtida, retirando aleatoriamente, um a um e com reposição, n elementos do conjunto de observações $\{X_1, X_2, ..., X_n\}$.

Denotemos, agora, por $\widehat{F}_n^*$ um estimador de F calculado com base na amostra $\mathbb{X}_n^*$. Consideremos uma nova a.a. de v.a. com f.d. $\widehat{F}_n^*$, e que denotamos por $\mathbb{X}_n^{**} = (X_1^{**}, X_2^{**}, ..., X_n^{**})$.

2. Problema Básico em Estatística

Seja X uma v.a. com f.d. F. Acerca desta distribuição apenas se sabe que pertence a uma família $\mathbb{F}$. O interesse reside numa função $\tau = T(F)$, e é desejável obter informações acerca de τ para além do facto de pertencer ao conjunto $\mathcal{T} = \{T(F) : F \in \mathbb{F}\}$.

Seja $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ uma a.a. de observações de X. Consideremos $\widehat{F}_n$ um estimador de F e consequentemente $\widehat{\tau}_n = T\left(\widehat{F}_n\right)$ um estimador de τ.

Usando a nomenclatura introduzida por Beran(1987), definimos agora uma **raíz**, como sendo uma v.a. $R_n(\mathbb{X}_n, T(F))$. É conveniente escolher esta raíz de forma a que a sua distribuição tenha um limite fraco não degenerado. Designaremos por $H_n(x, F) = P[R_n(\mathbb{X}_n, \tau) \leq x]$ a função de distribuição da raíz $R_n(\mathbb{X}_n, \tau)$, e por $H_A(\cdot, F)$ o limite fraco não degenerado da f.d. supra referida. A partir daqui, supomos sempre que $H_A(\cdot, F)$ é não degenerada.

Muitas informações (do ponto de vista probabilístico) acerca de τ podem ser retiradas de funcionais apropriadas de $H_n(\cdot, F)$, como por exemplo, a variância de $H_n(\cdot, F)$ como medida de dispersão, ou os seus quantis de forma a construir intervalos de confiança. É aqui que surge o problema. Se $H_n(\cdot, F)$ depender da f.d. desconhecida F, nada do atrás referido pode ser calculado. Em raras ocasiões, porém, $H_n(\cdot, F)$ é independente de F , o que significa que as referidas funcionais podem ser calculadas. Nestes casos, em que o problema fica resolvido, dizemos que $R_n(\mathbb{X}_n, \tau)$ é um **pivot verdadeiro**.

Exemplo 2.0.1. *Suponhamos que a v.a.* $X \sim N\left(\mu, \sigma^2\right)$, *onde* $\theta = \left(\mu, \sigma^2\right)$.

Neste caso, $F = F_\theta$, *donde* $\tau = T(F) = T(\theta) = \mu = \int x dF(x)$.

Tomemos

$$R_n(\mathbb{X}_n, \tau) = \sqrt{n}\left(\overline{X}_n - \mu\right) \tag{2.0.1}$$

onde $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

É um resultado clássico que $H_n(x, F) = H_n(x, \theta) = \Phi\left(\frac{x}{\sigma}\right)$ *onde* Φ *representa a f.d. da distribuição normal* $N(0,1)$. *Conclui-se então que* $H_A(x,\theta) = \Phi\left(\frac{x}{\sigma}\right)$. *Logo* $R_n(\mathbb{X}_n, \tau)$ *é raíz mas não é um pivot verdadeiro.*

Exemplo 2.0.2. *Sob as mesmas hipóteses do exemplo anterior suponhamos, agora, que*

$$R_n(\mathbb{X}_n, \tau) = \sqrt{n}\frac{\overline{X}_n - \mu}{S'_n}$$

onde $S'_n = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}_n\right)^2$.

Neste caso, $R_n \sim t_{n-1}$ *(isto é,* R_n *segue uma distribuição t de Student com n-1 graus de liberdade), logo* $H_n(\cdot, \theta) = t_{n-1}$, *consequentemente, neste contexto,* R_n *é um pivot verdadeiro.*

Exemplo 2.0.3. *Suponhamos que* X *tem f.d.* $F \in \mathbb{F}$ *que consiste no conjunto de todas as funções de distribuição em* $\mathbb{R}$. *Tomemos como parâmetro de interesse* $\tau = T(F) = F$. *Considere-se a estatística de Kolmogorov:*

$$R_n(\mathbb{X}_n, \tau) = \sqrt{n}\left\|F_n - F\right\|_\infty, \text{ onde } F_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}\left\{X_i \leq x\right\}.$$

Se F *for contínua então* $H_n(\cdot, F)$ *é independente de* F *e converge para* $H_A(\cdot, F)$, *que também é independente de* F. *Conclui-se então que* $R_n(\mathbb{X}_n, \tau)$ *é, neste contexto, um pivot verdadeiro.*

Como pivots verdadeiros são raridades usam-se métodos para aproximar $H_n(\cdot, F)$.

2.1. **Aproximação Assimptótica.** Como $H_n(\cdot, F)$ converge fracamente para $H_A(\cdot, F)$, então, para n grande, podemos tomar $H_A(\cdot, F)$ como aproximação de $H_n(\cdot, F)$.

Em alguns casos importantes acontece que a dependência da f.d. F em $H_n(\cdot, F)$ desaparece à medida que n aumenta, de tal forma que, $H_A(\cdot, F)$ é uma f.d. independente de F. Nestes casos, $R_n(\mathbb{X}_n, \tau)$ diz-se um **pivot assimptótico**.

Quando a f.d. $H_A(\cdot, F)$ não é independente de F, e, por conseguinte, não é possível o cálculo da mesma, usa-se $H_A\left(\cdot, \widehat{F}_n\right)$ como aproximação assimptótica da f.d. $H_n(\cdot, F)$.

2.2. **O Método de Bootstrap.** Este método, introduzido por Efron (1979), consiste em aproximar $H_n(\cdot, F)$ usando $H_n\left(\cdot, \widehat{F}_n\right)$.

A parte interessante do método consiste em constatar que o processo que nos leva a $H_n\left(\cdot, \widehat{F}_n\right)$ é uma mímica do processo que nos conduz a $H_n(\cdot, F)$. Verifiquemos mais cuidadosamente o conteúdo desta afirmação.

O ponto de partida é a distribuição F que podemos imaginar como sendo imposta pela natureza. Assim, obtém-se $X_1, X_2, ..., X_n$ a.a.de v.a. i.i.d. Construímos então $R_n(\mathbb{X}_n, \tau)$ com f.d. $H_n(\cdot, F)$ desconhecida e um estimador consistente de F. Damos então início ao processo de mímica. Para tal, começamos por "copiar" a natureza criando uma a.a. $\mathbb{X}_n^* = (X_1^*, X_2^*, ..., X_n^*)$ da função de distribuição $\widehat{F}_n$, que permitirá construir uma aproximação de $H_n\left(\cdot, \widehat{F}_n\right)$.

Notemos que a utilização do bootstrap permite ultrapassar o entrave inicial de dispormos apenas de uma única amostra $\mathbb{X}_n$, eventualmente pequena, e por conseguinte um único valor de $R_n(\mathbb{X}_n, \tau)$. A criação das "novas" amostras por reamostragem é, como veremos, motivada pelo Teorema de Glivenko- Cantelli.

Cálculo da Distribuição Bootstrap. Basicamente, existem duas formas de fazê-lo: a primeira, calculando explicitamente de forma analítica; a segunda, por meio das aproximações de Monte Carlo. Pese embora o cálculo analítico seja preferível à segunda forma, esta última tem um horizonte de aplicação muito mais vasto.

Cálculo Analítico das Distribuição de Bootstrap. Este cálculo é possível, sobretudo, no caso paramétrico onde $H_n(\cdot, F) = H_n(\cdot, \theta)$ e consiste em substituir θ por $\widehat{\theta}_n$ na expressão de $H_n(\cdot, \theta)$, já que $H_n\left(\cdot, \widehat{F}_n\right) = H_n\left(\cdot, F_{\widehat{\theta}_n}\right) = H_n\left(\cdot, \widehat{\theta}_n\right)$

Existem dois casos importantes onde é possível obter uma expressão analítica para $H_n\left(\cdot, \widehat{\theta}_n\right)$: o caso de θ ser um parâmetro num espaço de dimensão finita e o caso dos quantis.

Exemplo 2.2.1. *Suponhamos que $X \sim N(\mu, \sigma^2)$ onde $\theta = (\mu, \sigma^2)$. Considere-se o seguinte estimador centrado de $\theta : \widehat{\theta}_n = \left(\overline{X}_n, S_n'^2\right)$. Se tomarmos $R_n(\mathbb{X}_n, \tau) = \sqrt{n}(\overline{X}_n - \mu)$ como em (2.0.1), então $H_n(x, \theta) = \Phi\left(\frac{x}{\sigma}\right)$, donde imediatamente se conclui que $H_n\left(x, \widehat{\theta}_n\right) = \Phi\left(\frac{x}{S_n'}\right)$.*

Cálculo das Aproximações de Monte Carlo da Distribuição de Bootstrap. O método analítico raramente se aplica, uma vez que em muito poucos casos $R_n(\mathbb{X}_n, \tau)$ é suficientemente simples de forma a permitir o cálculo explícito de $H_n(\cdot, F)$.

Efron sugere então o seguinte método para obter uma aproximação de $H_n\left(\cdot, \widehat{F}_n\right)$:

(1) Uma vez observado $\mathbb{X}_n$, calcular $\widehat{F}_n$ e $\widehat{\tau}_n$;
(2) Obter uma amostra bootstrap $\mathbb{X}^*_{n,1} = (X^*_1, X^*_2, ..., X^*_n)$ de v.a. i.i.d. com f.d. $\widehat{F}_n$ e calcular $R^*_{n,1} = R_n\left(\mathbb{X}^*_{n,1}, \widehat{\tau}_n\right)$
(3) Repetir o passo 2. M vezes de forma a obter $R^*_{n,i}$, $i = 1, 2, ..., M$.
(4) Construir a f.d.e. baseada nos M $R^*_{n,i}$,
$$\widetilde{H}_B(x) = \widetilde{H}_n\left(x, \widehat{F}_n\right) = \frac{1}{M}\sum_{i=1}^{M} \mathbb{I}\left\{R^*_{n,i} \leq x\right\},$$
e usá-la como aproximação de $H_n\left(x, \widehat{F}_n\right)$.

Observação 2.1. *Para n fixo, a proximidade entre $\widetilde{H}_B(x)$ e $H_n\left(x, \widehat{F}_n\right)$ depende do número de réplicas M. De facto, pelo Teorema de Glivenko- Cantelli temos que:*
$$\left\|\widetilde{H}_B - H_n\left(\cdot, \widehat{F}_n\right)\right\|_\infty \underset{M\to\infty}{\longrightarrow} 0 \ q.c.$$

3. Intervalos de confiança

No texto subsequente, de forma a simplificar a nomenclatura, não faremos qualquer distinção entre intervalo de confiança aleatório e concretização do mesmo, designando ambos por intervalo de confiança.

Seja $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ uma a.a. obtida de uma f.d. F, como na secção 2. O que se pretende é a construção de um intervalo de confiança para o parâmetro de interesse $\tau = T(F)$. Seja $\mathcal{T} = \{T(F) : F \in \mathbb{F}\}$. Uma região de confiança para τ, com grau de confiança $1-\alpha$, é um subconjunto C (aleatório) de $\mathcal{T}$ tal que $P[\tau \in C] = 1 - \alpha$.

A construção de C será feita com base na raíz $R_n(\mathbb{X}_n, \tau)$, com f.d.
$$H_n(x, F) = P[R_n(\mathbb{X}_n, \tau) \leq x]$$
que tem limite fraco não degenerado $H_A(x, F)$. O intervalo de confiança natural é $C_\tau = \{\tau : R_n(\mathbb{X}_n, \tau) \leq c_{1-\alpha}\}$. O problema reside na determinação do nível crítico $c_{1-\alpha}$ de forma a que $P[\tau \in C_\tau]$ esteja perto de $1-\alpha$. Teoricamente a solução exacta é $c_{1-\alpha} = H_n^{-1}(1-\alpha, F)$. Porém, na maior parte das situações, $H_n(\cdot, F)$ depende da f.d. F que é desconhecida. Assim sendo, e à semelhança da secção 2, veremos aproximações diferentes para a resolução do problema.

3.1. Intervalo de Confiança Assimptótico.

Neste caso, a construção do intervalo de confiança é feita usando $H_A\left(\cdot, \widehat{F}_n\right)$ como aproximação de $H_n(\cdot, F)$, de forma que, substituindo

$H_n^{-1}(1-\alpha, F)$ por $H_A^{-1}\left(1-\alpha, \widehat{F}_n\right)$ obtém-se o seguinte intervalo de confiança:

$$\begin{aligned} C_A &= \left\{\tau : R_n \leq H_A^{-1}\left(1-\alpha, \widehat{F}_n\right)\right\} \\ &\simeq \left\{\tau : H_A\left(R_n, \widehat{F}_n\right) \leq 1-\alpha\right\}. \end{aligned}$$

3.1.1. *Consistência assimptótica.* Uma questão importante é saber a eficiência deste método na construção de intervalos de confiança com um grau de confiança $1-\alpha$.

A proposição seguinte (*vide* Beran e Ducharme (1991)) mostra que C_A é assimptoticamente consistente, isto é,

$$P[\tau \in C_A] \underset{n\to\infty}{\longrightarrow} 1-\alpha.$$

Proposição 3.1. *Seja d uma métrica em $\mathbb{F}$. Suponhamos que para quaisquer $\varepsilon > 0$ e $F \in \mathbb{F}$,*

(1) $\lim_{n\to\infty} P\left[d\left(\widehat{F}_n, F\right) > \varepsilon\right] = 0$

(2) *O limite fraco $H_A(\cdot, F)$ é contínuo em x para todo o $F \in \mathbb{F}$ e em F para todo o $x \in \mathbb{R}$.*

Então para quaisquer $\varepsilon > 0$ e $F \in \mathbb{F}$, temos que:

a): $\lim_{n\to\infty} P\left[\left\|H_A\left(\cdot, \widehat{F}_n\right) - H_A(\cdot, F)\right\|_\infty > \varepsilon\right] = 0$

e

b): $\lim_{n\to\infty} P[\tau \in C_A] = 1-\alpha$

Omitimos aqui a prova desta proposição já que pode ser consultada na referência, já citada, Beran e Ducharme (1991).

3.2. **Intervalo de Confiança Bootstrap.** A solução bootstrap para o problema consiste em aproximar a distribuição exacta $H_n(\cdot, F)$ pela distribuição bootstrap

$$\widehat{H}_B(x) = H_n\left(x, \widehat{F}_n\right) = P[R_n^* \leq x \mid \mathbb{X}_n]$$

onde $R_n^* = R_n(\mathbb{X}_n^*, \widehat{\tau}_n)$ e, como anteriormente, $\widehat{\tau}_n = T\left(\widehat{F}_n\right)$. Como esta distribuição bootstrap não envolve elementos desconhecidos, podemos usar $\widehat{H}_B(\cdot)$ como estimador de $H_n(\cdot, F)$ e calcular $\widehat{H}_B^{-1}(1-\alpha)$ como aproximação de $H_n^{-1}(1-\alpha, F)$. Deste modo obtém-se o intervalo de confiança bootstrap:

$$\begin{aligned} C_B &= \left\{\tau : R_n \leq \widehat{H}_B^{-1}(1-\alpha)\right\} && (3.2.1) \\ &\simeq \left\{\tau : H_n\left(R_n, \widehat{F}_n\right) \leq 1-\alpha\right\}. && (3.2.2) \end{aligned}$$

3.2.1. *Consistência assimptótica.* A proposição seguinte encontra-se enunciada na referência Beran e Ducharme (1991), e permite concluir que C_B é assimptoticamente consistente, isto é,

$$P[\tau \in C_B] \underset{n\to\infty}{\longrightarrow} 1-\alpha.$$

Proposição 3.2. *Seja d uma métrica em $\mathbb{F}$. Suponhamos que para quaisquer $\varepsilon > 0$ e $F \in \mathbb{F}$,*

(1) $\lim_{n\to\infty} P\left[d\left(\widehat{F}_n, F\right) > \varepsilon\right] = 0$;

(2) *(**Triangular Array Convergence Condition (TAC)**): Para qualquer sequência $\{G_n\}$ em $\mathbb{F}$ tal que $\lim_{n\to\infty} d(G_n, F) = 0$, temos que $H_n(\cdot, G_n)$ converge fracamente para $H_A(\cdot, F)$;*

(3) *O limite fraco $H_A(\cdot, F)$ é contínuo em x para todo o $F \in \mathbb{F}$.*

Então para quaisquer $\varepsilon > 0$ e $F \in \mathbb{F}$, temos que:

a): $\lim_{n\to\infty} P\left[\left\|\widehat{H}_B(\cdot) - H_A(\cdot, F)\right\|_\infty > \varepsilon\right] = 0$

e

b): $\lim_{n\to\infty} P[\tau \in C_B] = 1-\alpha$.

Demonstração. a)

Seja $\{G_n\} \in \mathbb{F}$ tal que o $\lim_{n\to\infty} d(G_n, F) = 0$. Então por 2. temos que $\lim_{n\to\infty} H_n(x, G_n) = H_A(x, F)$. Pelo Teorema de Polya este limite é uniforme em x, isto é,

$$\lim_{n\to\infty} \|H_n(\cdot, G_n) - H_A(\cdot, F)\|_\infty = 0,$$

Então usando 1. vem que $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} P\left[\left\|H_n\left(\cdot, \widehat{F}_n\right) - H_A(\cdot, F)\right\|_\infty > \varepsilon\right] = 0. \tag{3.2.3}$$

Por (3.2.3) temos que

$$\forall \varepsilon > 0, \ \lim_{n\to\infty} P\left[\left\|H_n\left(R_n, \widehat{F}_n\right) - H_A(R_n, F)\right\|_\infty > \varepsilon\right] = 0. \tag{3.2.4}$$

Seja R uma v.a. com f.d. $H_A(\cdot, F)$. Como $R_n \xrightarrow{D} R$ e $H_A(\cdot, F)$ é contínua em x para todo o $F \in \mathbb{F}$ (hipótese 3.) então

$$H_A(R_n, F) \xrightarrow{D} H_A(R, F). \tag{3.2.5}$$

Escrevendo

$$H_n\left(R_n, \widehat{F}_n\right) = \underbrace{H_n\left(R_n, \widehat{F}_n\right) - H_A(R_n, F)}_{V} + \underbrace{H_A(R_n, F)}_{W}$$

em que $V \xrightarrow{P} 0$, por (3.2.4), e $W \xrightarrow{D} H_A(R, F)$, por (3.2.5), resulta, pelo Teorema de Slutsky, que

$$H_n\left(R_n, \widehat{F}_n\right) \xrightarrow{D} H_A(R, F).$$

Posto isto, uma vez que $H_A(R, F) = U \sim \mathcal{U}_{[0,1]}$ (distribuição uniforme em $[0, 1]$) e lembrando que $C_B = \left\{\tau : H_n\left(R_n, \widehat{F}_n\right) \leq 1 - \alpha\right\}$ então

$$\lim_{n\to\infty} P[\tau \in C_B] = P[U \leq 1 - \alpha] = 1 - \alpha$$

. □

A conclusão a) é como uma "licença" de utilização do bootstrap. Do ponto de vista teórico, importa averiguar quão razoável é a aproximação bootstrap. Uma forma de o fazer é mostrar que

$$H_n\left(\cdot, \widehat{F}_n\right) - H_A(\cdot, F) \underset{n\to\infty}{\longrightarrow} 0$$

ou equivalentemente, usando o Terema de Polya, mostrar que

$$\left\|H_n\left(\cdot, \widehat{F}_n\right) - H_A(\cdot, F)\right\|_\infty \underset{n\to\infty}{\longrightarrow} 0 \tag{3.2.6}$$

onde a convergência pode ser fraca ou forte.

Nesta proposição, mostramos (3.2.6) em probabilidade, sob certas condições. No entanto, noutras circunstâncias, é possível mostrar a convergência forte. Bickel e Freedman (1981) apresentam vários destes exemplos. Entre eles destacamos o seguinte:

Seja $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ uma a.a. de uma f.d. F de média μ e variância σ^2. Considere-se $R_n(\mathbb{X}_n, \tau) = \sqrt{n}\left(\overline{X}_n - \mu\right)$ e $\widehat{F}_n = F_n$. Pelo Teorema do limite central temos que a distribuição de $R_n(\mathbb{X}_n, \tau)$ converge fracamente para $N\left(0, \sigma^2\right)$. Neste contexto, os referidos autores mostraram o seguinte Teorema.

Teorema 3.3. *(Bickel e Freedman (1981)) Suponhamos que X_1, X_2, ... é uma sucessão de v.a. i.i.d. com variância finita σ^2. Então para quase todas as sucessões amostrais X_1, X_2, ..., dado $(X_1, X_2, ..., X_n)$, quando $n \to \infty$ temos que a distribuição condicional de*

$$R_n^* = R_n(\mathbb{X}_n^*, \tau) = \sqrt{n}\left(\overline{X}_n^* - \overline{X}_n\right)$$

converge fracamente para $N\left(0, \sigma^2\right)$.

4. Refinamentos dos intervalos de confiança bootstrap

4.1. **Introdução e Raízes na Forma Normalizada.** Os métodos bootstrap discutidos na secção 3.2 permitem a construção de intervalos de confiança numa variedade considerável de situações. Nesta secção veremos formas de refinar estes métodos em termos de diminuição do erro cometido devido à utilização do bootstrap. Existem vários processos de refinamento como os sugeridos por Hall (1983, 1986), Abromovitch e Singh (1985). Contudo, aqui, faremos apenas referência à pré-pivotagem (bootstrap iterado) introduzida por Beran (1987), pela sua característica generalista e até unificadora de alguns dos últimos refinamentos referidos.

De forma a construir o intervalo de confiança (3.2.1) é necessário escolher uma raíz. Em particular consideremos a raíz

$$R_n\left(\mathbb{X}_n,\tau\right)=\sqrt{n}\left(\widehat{\tau}_n-\tau\right)$$

e a sua versão na forma normalizada

$$S_n\left(\mathbb{X}_n,\tau\right)=\sqrt{n}\left(\widehat{\tau}_n-\tau\right)/s_n$$

onde $\widehat{\tau}_n$ é um estimador consistente de τ e $s_n/\sqrt{n}$ um estimador consistente do desvio padrão de $\widehat{\tau}_n$.

Em amostras grandes, a distribuição de R_n sob F, $H_n\left(\cdot,F\right)$, é tipicamente normal com média 0 e variância $\sigma^2\left(F\right)$. Por outro lado, a f.d. de S_n sob F, $K_n\left(\cdot,F\right)$, converge fracamente (quando estamos perante v.a. tipicamente Gaussianas) para a $N\left(0,1\right)$. Isto sugere que a distribuição de S_n depende "menos" de F que a distribuição de R_n. De facto, segundo Beran (1982), em casos regulares, aplicar o bootstrap com base numa raíz na forma normalizada resulta numa diminuição do erro cometido.

Em particular,

$$H_n\left(x,F\right)-H_n\left(x,\widehat{F}_n\right)=O_P\left(n^{-1/2}\right)$$

e

$$K_n\left(x,F\right)-K_n\left(x,\widehat{F}_n\right)=O_P\left(n^{-1}\right).$$

Dever-se-á notar, no entanto, que o erro $O_P\left(n^{-1/2}\right)$ baseado na raíz R_n, não é um defeito do bootstrap, já que Beran (1982) mostrou que não há estimativa de $H_n\left(\cdot,F\right)$ que melhore esta ordem de convergência. O problema parece residir numa má escolha da raíz.

4.2. **Pré-pivotagem.** Como na secção 2, seja $H_n\left(\cdot,F\right)$ a f.d. da raíz $R_n\left(\mathbb{X}_n,\tau\right)$. A região de confiança resultante C_B é dada por (3.2.1). Definamos uma nova raíz

$$R_{n1}\left(\mathbb{X}_n,\tau\right)=H_n\left(R_n\left(\mathbb{X}_n,\tau\right),\widehat{F}_n\right)=P\left[R_n^*\leq R_n\mid\mathbb{X}_n\right],$$

com f.d. $H_{n1}\left(\cdot,F\right)$ sob F. Como anteriormente, para construirmos o intervalo de confiança com base na raíz $R_{n1}\left(\mathbb{X}_n,\tau\right)$, podemos aproximar $H_{n1}\left(\cdot,F\right)$ de duas formas:

- Usando a sua distribuição assimptótica, que é a distribuição uniforme (como pode ser verificado na prova da proposição 3.2), o que resulta, novamente, no intervalo de confiança C_B, já que:

  $$\left\{\tau:R_{n1}\left(\mathbb{X}_n,\tau\right)\leq 1-\alpha\right\}=\left\{\tau:H_n\left(R_n,\widehat{F}_n\right)\leq 1-\alpha\right\}=C_B$$

  voltando-se, desta feita, ao ponto de partida.
- Usando a distribuição bootstrap $H_{n1}\left(\cdot,\widehat{F}_n\right)$ para obter o intervalo de confiança:

  $$C_{B1}=\left\{\tau:R_{n1}\left(\mathbb{X}_n,\tau\right)\leq H_{n1}^{-1}\left(1-\alpha,\widehat{F}_n\right)\right\}\tag{4.2.1}$$

  descrevendo-se, desta forma, em que consiste o método de **pré-pivotagem** de Beran (1987).

Na prática, o intervalo de confiança (4.2.1) é calculado de forma equivalente usando a fórmula:

$$C_{B1}=\left\{\tau:R_n\left(\mathbb{X}_n,\tau\right)\leq H_n^{-1}\left(H_{n1}^{-1}\left(1-\alpha,\widehat{F}_n\right),\widehat{F}_n\right)\right\}.$$

Assim, o valor crítico de C_{B1}, nesta forma, é calculado em dois passos:

(1) calcula-se o quantil de ordem $(1-\alpha)$ de $H_{n1}\left(\cdot,\widehat{F}_n\right)$, que designamos por $c_{n1}=H_{n1}^{-1}\left(1-\alpha,\widehat{F}_n\right)$;

(2) calcula-se o quantil de ordem c_{n1} de $H_n\left(\cdot,\widehat{F}_n\right)$.

Beran (1987) mostrou que geralmente o erro cometido na construção do intervalo de confiança C_{B1} é inferior aos erros inerentes aos intervalos C_A e C_B.

Beran argumentou, também, que a distribuição $H_{n1}(\cdot,F)$ é "menos" dependente de F que a distribuição $H_n(\cdot,F)$. Isto pode ser verificado no caso de uma raíz assimptoticamente normal, cuja variância depende de F. Neste contexto, a distribuição assimptótica de $R_{n1}(\mathbb{X}_n,\tau)$ é a distribuição uniforme que não depende de F. Ainda neste caso, começando com uma raíz da forma:

$$R_n(\mathbb{X}_n,\tau)=\sqrt{n}\left(\widehat{\tau}_n-\tau\right),$$

Beran mostrou que a operação de pré-pivotagem é assimptoticamente equivalente a tomar a raíz correspondente na forma normalizada, S_n. Em suma, os erros cometidos ao usar o intervalo de confiança definido como em (4.2.1) para a raíz R_n, e os erros inerentes ao intervalo de confiança obtido como em (3.2.1) baseado na raíz S_n, são assimptoticamente da mesma ordem de grandeza.

O método da pré-pivotagem pode ser generalizado a um método iterativo da forma que se segue. Dada uma raíz $R_{nj}(\mathbb{X}_n,\tau)$, seja $H_{nj}(\cdot,F)$ a sua f.d. sob F. Forme-se uma nova raíz $R_{nj+1}(\mathbb{X}_n,\tau)$ dada por:

$$R_{nj+1}(\mathbb{X}_n,\tau)=H_{nj}\left(R_{nj}(\mathbb{X}_n,\tau),\widehat{F}_n\right)$$

e definamos

$$C_{Bj}=\left\{\tau:R_{nj}(\mathbb{X}_n,\tau)\le H_{nj}^{-1}\left(1-\alpha,\widehat{F}_n\right)\right\}.$$

Em certos casos, o erro cometido diminui quando j aumenta. (Beran (1987))

4.2.1. *Aproximações de Monte Carlo para o Cálculo de* $H_{n1}\left(\cdot,\widehat{F}_n\right)$. Considere-se $\widehat{\tau}^*=T\left(\widehat{F}_n^*\right)$. Na notação sugerida por Beran (1987), temos que:

$$\begin{aligned}H_n(x,F)&=P\left[R_n(\mathbb{X}_n,\tau)\le x\mid F\right],\\ H_{n1}(x,F)&=P\left[P\left[R_n(\mathbb{X}_n^*,\widehat{\tau})\le R_n(\mathbb{X}_n,\tau)\mid\widehat{F}_n\right]\le x\mid F\right].\end{aligned}$$

As distribuições bootstrap correspondentes a estas f.d. podem ser expressas da seguinte forma:

$$\widehat{H}_B(x)=H_n\left(x,\widehat{F}_n\right)=P\left[R_n(\mathbb{X}_n^*,\widehat{\tau})\le x\mid\widehat{F}_n\right],\tag{4.2.2}$$

$$\widehat{H}_{B1}(x)=H_{n1}\left(x,\widehat{F}_n\right)=P\left[P\left[R_n(\mathbb{X}_n^{**},\widehat{\tau}^*)\le R_n(\mathbb{X}_n^*,\widehat{\tau})\mid\widehat{F}_n^*\right]\le x\mid\widehat{F}_n\right].\tag{4.2.3}$$

As representações (4.2.2) e (4.2.3) motivam o seguinte algoritmo, introduzido por Beran (1987), para aproximar $\widehat{H}_B$ e $\widehat{H}_{B1}$.

(1) Obtenham-se amostras bootstrap $\mathbb{X}_{n,1}^*,\mathbb{X}_{n,2}^*,...,\mathbb{X}_{n,M}^*$ de v.a. i.i.d. com f.d. $\widehat{F}_n$;

(2) Denote-se por $\widehat{F}_{n,j}^*$ um estimador de F calculado com base na j-ésima amostra bootstrap $\mathbb{X}_{n,j}^*$. Para cada j, construam-se N a.a. $\mathbb{X}_{n,j,1}^{**},\mathbb{X}_{n,j,2}^{**},...,\mathbb{X}_{n,j,N}^{**}$ com f.d. $\widehat{F}_{n,j}^*$;

(3) Calcule-se $\widehat{\tau}_j^*=T\left(\widehat{F}_{n,j}^*\right)$. Defina-se Z_j como sendo a proporção de valores do conjunto $\left\{R_n\left(\mathbb{X}_{n,j,k}^{**},\widehat{\tau}_j^*\right):1\le k\le N\right\}$ que são inferiores ou iguais a $R_n\left(\mathbb{X}_{n,j}^*,\widehat{\tau}\right)$;

(4) Construa-se a f.d.e. com base no conjunto de v.a. i.i.d. $\{Z_j:1\le j\le M\}$, que constitui uma aproximação de $\widehat{H}_{B1}$ para N e M suficientemente grandes.

5. Expansões Edgeworth

5.1. **Introdução.** Nesta secção, desenvolveremos séries de potências, introduzidas por Edgeworth, tendo em vista a aproximação da função distribuição de pivots assimptóticos. Basicamente, se considerarmos como exemplo a raíz $R_n(\mathbb{X}_n,\tau)=\sqrt{n}\left(\widehat{\tau}_n-\tau\right)$, que supomos assimptoticamente

Normal de média 0 e variância σ^2, então numa vasta colecção de casos de interesse a f.d. de $R_n(\mathbb{X}_n, \tau)$ pode ser expandida como uma série de potências em $n^{-\frac{1}{2}}$,

$$P\left[\frac{\sqrt{n}\left(\widehat{\tau}_n - \tau\right)}{\sigma} \leq x\right] = \Phi(x) + n^{-\frac{1}{2}} p_1(x)\,\phi(x) + ... + n^{-\frac{j}{2}} p_j(x)\,\phi(x) + ...,$$

que designamos por **Expansão Edgeworth**.

5.2. **Somas de Variáveis Aleatórias Independentes.** Estudaremos, aqui, o caso em que $\widehat{\tau}_n$ é a média de uma amostra que, apesar de se tratar do exemplo mais básico de aplicação das expansões Edgeworth, evidencia as principais características técnicas gerais. Para além disso, no decorrer da exposição, os únicos resultados necessários sobre expansões Edgeworth serão os referentes a este caso.

Considere-se uma a.a. $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ obtida de uma f.d. F com média μ e variância σ^2. Considere-se a raíz:

$$S_n(\mathbb{X}_n, \tau) = \sqrt{n}\left(\widehat{\tau}_n - \tau\right)/\sigma$$

onde $\tau = \mu$ e $\widehat{\tau}_n = \overline{X}_n = 1/n \sum_{i=1}^{n} X_i$. Pelo Teorema do Limite Central, S_n é assimptoticamente Normal de média 0 e variância 1.O próximo Teorema pode ser visto, por exemplo, na referência [13].

Teorema 5.1. *Considere-se o pivot assimptótico $S_n(\mathbb{X}_n, \tau)$ como acima definido. Suponhamos que $E\left(|X_i|^{j+2}\right) < \infty$ e que a função característica de X_i, que denotamos por χ, satisfaz a condição:*

$$\limsup_{t \to \infty} |\chi(t)| < 1, \tag{5.2.1}$$

então para cada $j \geq 1$

$$P\left[\sqrt{n}\left(\widehat{\tau}_n - \tau\right)/\sigma \leq x\right] = \Phi(x) + n^{-\frac{1}{2}} p_1(x)\,\phi(x) + n^{-1} p_2(x)\,\phi(x) + ... \\ ... + n^{-\frac{j}{2}} p_j(x)\,\phi(x) + o\left(n^{-\frac{j}{2}}\right) \tag{5.2.2}$$

uniformemente em x, onde p_j é um polinómio de grau inferior ou igual a $3j-1$, que é par para j ímpar e ímpar para j par e cujos coeficientes dependem dos momentos de X_i até à ordem $j+2$.

Observação 5.1. *A condição (5.2.1) é satisfeita se a distribuição de X_i for absolutamente contínua e, por conseguinte, tiver função densidade de probabilidade. Veja-se, por exemplo, referência* [13]*, secção 2.4.*

5.3. **Expansões Edgeworth e Bootstrap.** Vejamos, agora, a utilidade destes desenvolvimentos na explicação de algumas propriedades dos métodos bootstrap para estimação de distribuições e construção de intervalos de confiança.

Comecemos por enunciar um teorema (ver, por exemplo, referência [13]) que estabelece a existência da versão bootstrap do desenvolvimento de Edgeworth (5.2.2).

Teorema 5.2. *Considere-se o pivot assimptótico $S_n(\mathbb{X}_n, \tau)$ como na secção 5.2. Suponhamos que $E\left(|X_i|^l\right) < \infty$, para algum l suficientemente grande, e que a função característica de X_i, que denotamos por χ, satisfaz a condição:*

$$\limsup_{t \to \infty} |\chi(t)| < 1,$$

então para cada $j \geq 1$

$$P\left[\sqrt{n}\left(\widehat{\tau}_n^* - \widehat{\tau}_n\right)/\sigma \leq x \mid \mathbb{X}_n\right] = \Phi(x) + n^{-\frac{1}{2}} \widehat{p}_1(x)\,\phi(x) + n^{-1} \widehat{p}_2(x)\,\phi(x) + ... \\ ... + n^{-\frac{j}{2}} \widehat{p}_j(x)\,\phi(x) + o_P\left(n^{-\frac{j}{2}}\right) \tag{5.3.1}$$

uniformemente em x, onde $\widehat{p}_j$ é um polinómio que se obtém de p_j substituindo os momentos da população, pelos momentos empíricos calculados com base em $\mathbb{X}_n$.

Com base nas expansões Edgeworth verifiquemos a vantagem da utilização da distribuição bootstrap (sob certas condições), face à distribuição assimptótica. Como já foi visto, $S_n(\mathbb{X}_n, \tau)$, definida como na secção 5.2, tem uma distribuição assimptótica Normal de parâmetros 0 e 1. Recorrendo às expressões (5.2.2) e (5.3.1) decorre que

$$P\left[\sqrt{n}\left(\widehat{\tau}_n^* - \widehat{\tau}_n\right)/\sigma \leq x \mid \mathbb{X}_n\right] - P\left[\sqrt{n}\left(\widehat{\tau}_n - \tau\right)/\sigma \leq x\right] = O_P\left(n^{-1}\right),$$

uma vez que $\widehat{p}_j - p_j = O_P\left(n^{-\frac{1}{2}}\right)$, já que os coeficientes de $\widehat{p}_j$ diferem dos de p_j de termos de ordem $O_P\left(n^{-\frac{1}{2}}\right)$. Assim sendo, constatamos que o erro cometido usando a distribuição bootstrap para aproximar a distribuição de $S_n(\mathbb{X}_n, \tau)$ é da ordem de $O_P\left(n^{-1}\right)$. Porém, o erro cometido usando a distribuição assimptótica é, como a fórmula (5.2.2) indica, da ordem de $n^{-\frac{1}{2}}$.

6. Bootstrap de Cauda para a Estimação do Índice de Cauda de Pareto

6.1. **Introdução.** Sejam $X_1, X_2, ...$ v.a. independentes, não negativas e com f.d. F comum. Assuma-se que F é de variação regular na cauda superior, ou seja, existe um $0 < c < \infty$, tal que, para todo $t > 0$

$$\lim_{x\to\infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-c} \tag{6.1.1}$$

ou, de forma equivalente,

$$1 - F(x) = x^{-c} L(x), \text{ para } x > 0,$$

onde L é uma função de variação lenta em infinito, isto é,

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1.$$

O problema consiste em estimar o expoente c com base numa a.a. (finita) de v.a. i.i.d., $\mathbb{X}_n = (X_1, X_2, ..., X_n)$. Uma vez que o objecto de estudo incide sobre a cauda de F, parece intuitivamente claro que os estimadores de c deverão fazer uso das maiores estatísticas de ordem da amostra, com o intuito de extrair informação acerca do comportamento da cauda superior de F. Diversos estimadores de c foram introduzidos. O estimador de c habitualmente mais usado é o estimador proposto por Hill (1975), que passamos a apresentar. Sejam $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$ as estatísticas de ordem da amostra e considere-se uma sucessão de inteiros positivos $(k_n)_{n\in\mathbb{N}}$ que satisfaz as seguintes condições,

$$1 \leq k_n < n, \qquad k_n \underset{n\to\infty}{\longrightarrow} \infty \quad e \quad \frac{k_n}{n} \underset{n\to\infty}{\longrightarrow} 0. \tag{6.1.2}$$

O estimador de Hill de c^{-1}, que denotamos por EH_n, é definido por:

$$EH_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k_n,n}\right).$$

Haeusler e Teugels (1985) mostraram a normalidade assimptótica de EH_n.

Nos últimos tempos, tem havido um interesse acrescido em usar os métodos bootstrap aplicados a estatísticas de ordem. J. Bacro e M. Brito (1998) introduziram um procedimento **bootstrap de cauda** aplicado a extremos. A ideia consiste no seguinte: em vez de aplicar o naïve bootstrap à amostra original $\mathbb{X}_n = (X_1, X_2, ..., X_n)$, propuseram aplicar o naïve bootstrap à amostra $\mathbb{W}_{l_n} = (W_1, W_2, ..., W_{l_n})$, onde $W_i = \log\left(X_{n-l_n+i,n}\right) - \log\left(X_{n-l_n,n}\right)$, $1 \leq i \leq l_n$, e $(l_n)_{n\in\mathbb{N}}$ é uma sequência de inteiros positivos satisfazendo algumas condições de regularidade. Uma motivação heurística para este procedimento é o facto de a informação acerca da cauda de F estar concentrada nas maiores estatísticas de ordem da amostra, e daí que o natural seja a reamostragem incidir sobre estas "estatísticas informativas".

6.2. **Procedimento Bootstrap e Resultados Principais.** O bootstrap de cauda pode ser descrito da forma que se segue. Seja l_n uma sequência de inteiros positivos e considerem-se as l_n excedências do nível aleatório $\log(X_{n-l_n,n})$ definidas por:

$$W_i = \log(X_{n-l_n+i,n}) - \log(X_{n-l_n,n}), \qquad 1 \leq i \leq l_n. \tag{6.2.1}$$

Note-se que $W_1 \leq W_2 \leq ... \leq W_{l_n}$. Com uma escolha adequada da sequência l_n o estimador de Hill pode ser facilmente escrito à custa destas v.a. Seja k_n uma sequência de inteiros positivos satisfazendo a condição (6.1.2). Se para cada $n \in \mathbb{N}$ tomarmos $l_n \equiv k_n$, então

$$EH_n = \frac{1}{k_n}\sum_{i=1}^{k_n} W_i. \tag{6.2.2}$$

Introduzamos agora o bootstrap. Considere-se a amostra $\mathbb{W}_{k_n}^* = \left(W_1^*, W_2^*, ..., W_{k_n}^*\right)$ obtida por naïve bootstrap da amostra $\mathbb{W}_{k_n} = (W_1, W_2, ..., W_{k_n})$. Usando a representação acima para EH_n, a sua versão bootstrap é definida como:

$$EH_n^* = \frac{1}{k_n}\sum_{i=1}^{k_n} W_i^*. \tag{6.2.3}$$

O modelo caracterizado pela equação (6.1.1) pode também ser escrito com base na função quantil da f.d. F, obtendo-se:

$$F^{-1}(1-y) = y^{-\frac{1}{c}}\widetilde{L}(y), \qquad 0 < y < 1, \tag{6.2.4}$$

onde $\widetilde{L}$ é uma função de variação lenta em zero, isto é,

$$\widetilde{L}(y) = L'\left(\frac{1}{y}\right),$$

onde L' é uma função de variação lenta em infinito.

Vejamos, agora, os resultados principais relativos à distribuição assimptótica de EH_n e EH_n^*.

Teorema 6.1. *(Heusler e Teugels (1985))*

Assumamos que F satisfaz a condição (6.1.1) e que k_n é uma sucessão de inteiros positivos satisfazendo (6.1.2). Se

$$\sqrt{k_n}\int_1^\infty \left\{ y^{-c} - \frac{1 - F\left[F^{-1}\left(1 - \frac{k_n}{n}\left(1 + \frac{x}{\sqrt{k_n}}\right)\right)y\right]}{\frac{k_n}{n}\left(1 + \frac{x}{\sqrt{k_n}}\right)} \right\}\frac{dy}{y} \underset{n\to\infty}{\longrightarrow} 0,$$

uniformemente em x num conjunto compacto, então

$$c\sqrt{k_n}\left(EH_n - \frac{1}{c}\right) \xrightarrow{D} N(0,1), \qquad \textit{quando } n \to \infty.$$

Teorema 6.2. *(J. Bacro e M. Brito (1998))*

Assumamos que F satisfaz a condição (6.1.1) e que k_n é uma sucessão de inteiros positivos satisfazendo (6.1.2). Se

$$\sqrt{k_n}\sup_{\frac{1}{k_n}\leq y\leq 1}\left|\log\left(\frac{\widetilde{L}\left(yt\frac{k_n}{n}\right)}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right)\right| \underset{n\to\infty}{\longrightarrow} 0 \tag{6.2.5}$$

uniformemente em t num conjunto compacto contido em $]0,\infty[$, então, em probabilidade, para todo o y

$$P\left[c\sqrt{k_n}\left(EH_n^* - EH_n\right) \leq y \mid X_{n-k_n,n}, ..., X_{n,n}\right] \underset{n\to\infty}{\longrightarrow} \Phi(y). \tag{6.2.6}$$

Observação 6.1. *Tendo em conta os resultados de Bickel e Freedman (1981) acerca do bootstrap da média, é fácil verificar que o Teorema 6.2 ainda prevalece se o tamanho da reamostra for diferente do número k_n de excedências usado. É, pois, suficiente substituir $\sqrt{k_n}$ por $\sqrt{m_n}$ nas equações (6.2.5) e (6.2.6), onde m_n denota o tamanho da reamostra.*

Observação 6.2. *Usando o facto de EH_n ser um estimador consistente de c^{-1}, conclui-se imediatamente pelo Teorema de Slutsky que o Teorema 6.2 permanece verdadeiro se c for substituído por EH_n^{-1}, em (6.2.6).*

Posto isto, vejamos como utilizar estes resultados para construir intervalos de confiança para o parâmetro c.

Considere-se a raíz

$$R_n = R_n\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right), c\right) = c\sqrt{k_n}\left(EH_n - \frac{1}{c}\right) \tag{6.2.7}$$

e a sua versão bootstrap de cauda

$$R_n^* = R_n\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n}^*\right), EH_n\right) = EH_n\sqrt{k_n}\left(EH_n^* - EH_n\right). \tag{6.2.8}$$

Denotemos, por $H_n(\cdot, F)$ a f.d. da raíz R_n definida em (6.2.7), por $H_A(\cdot, F)$ a respectiva f.d. assimptótica e por $\widehat{H}_{BC}(\cdot)$ a devida f.d. bootstrap de cauda que se define através da relação:

$$\widehat{H}_{BC}(x) = P\left[R_n^* \leq x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right)\right].$$

Pelo Teorema 6.1 concluímos que $H_A(\cdot, F) = \Phi(\cdot)$. Pelo Teorema 6.2 e notas subsequentes temos que $\forall\varepsilon > 0 \lim_{n\to\infty} P\left[\left\|\widehat{H}_{BC}(\cdot) - H_A(\cdot, F)\right\|_\infty > \varepsilon\right] = 0$, donde deduzimos que $\widehat{H}_{BC}(R_n) \longrightarrow U \sim \mathcal{U}_{[0,1]}$. Assim sendo, em analogia à secção 3.2, considere-se o intervalo de confiança unilateral

$$\begin{aligned} C_{BC} &= \left\{c : R_n \leq \widehat{H}_{BC}^{-1}(1-\alpha)\right\} \\ &\simeq \left\{c : \widehat{H}_{BC}(R_n) \leq 1-\alpha\right\}. \end{aligned}$$

Pelo facto de $\forall\varepsilon > 0 \lim_{n\to\infty} P\left[\left\|\widehat{H}_{BC}(\cdot) - H_A(\cdot, F)\right\|_\infty > \varepsilon\right] = 0$, é possível mostrar, como foi feito na prova da Proposição 3.2, que o intervalo de confiança C_{BC} é assimptoticamente consistente.

7. Refinamento do Método Bootstrap de Cauda

7.1. **Introdução.** As simulações apresentadas por J. Bacro e M. Brito (1998) sugerem que a velocidade de convergência do bootstrap de cauda é lenta. Tendo isto em conta, com o intuito de acelerar a ordem de convergência, resolvemos adaptar a pré-pivotagem ao bootstrap de cauda. Designaremos esta adaptação por **pré-pivotagem de cauda**.

Numa primeira etapa faremos a introdução ao procedimento de adaptação da pré-pivotagem.

Posto isto, averiguaremos a sua consistência, no sentido de mostrar a aplicabilidade do método.

Numa etapa posterior, sob certas condições, determinaremos a ordem de convergência do método de bootstrap de cauda para depois compararmos com a ordem de convergência da adaptação da pré-pivotagem ao boostrap de cauda.

7.2. **Descrição do Procedimento Pré-Pivotagem de Cauda.** Considerem-se a raíz R_n definida como em (6.2.7), a respectiva versão bootstrap de cauda R_n^* definida pela equação (6.2.8) e as suas distribuições assimptótica e bootstrap de cauda como foram definidas na secção 6.2.

Definamos a nova raíz:

$$R_{n1} = R_{n1}\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right), c\right) = \widehat{H}_{BC}(R_n),$$

com f.d. $H_{n1}(\cdot, F)$. O procedimento consiste em aplicar o bootstrap de cauda introduzido na secção 6.2 à nova raíz R_{n1}, calculando a respectiva distribuição bootstrap de cauda $\widehat{H}_{BC1}(\cdot)$. Deste modo, à semelhança do que foi feito em (4.2.1), é possível construir o intervalo de confiança:

$$C_{BC1} = \left\{c : R_{n1}\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right), c\right) \leq \widehat{H}_{BC1}^{-1}(1-\alpha)\right\}.$$

7.2.1. *Aproximações de Monte Carlo para o Cálculo de* $\widehat{H}_{BC1}(\cdot)$. Na notação habitual temos que,

$$\begin{aligned} H_n(x,F) &= P[R_n \leq x], \\ H_{n1}(x,F) &= P[P[R_n^* \leq R_n \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})] \leq x]. \end{aligned}$$

As distribuições bootstrap correspondentes a estas f.d. podem ser expressas da seguinte forma:

$$\widehat{H}_{BC}(x) = P[R_n^* \leq x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})], \tag{7.2.1}$$

$$\widehat{H}_{BC1}(x) = P\left[P\left[R_n^{**} \leq R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] \leq x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right], \tag{7.2.2}$$

onde

$$R_n^{**} = R_n\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n}^{**}\right), EH_n^*\right) \tag{7.2.3}$$

e $\mathbb{W}_{k_n}^{**}$ é uma amostra obtida por naïve bootstrap da amostra $\mathbb{W}_{k_n}^*$.

As representações (7.2.1) e (7.2.2) sugerem o seguinte algoritmo para aproximar $\widehat{H}_{BC}(\cdot)$ e $\widehat{H}_{BC1}(\cdot)$.

(1) Obtenham-se amostras bootstrap $\mathbb{W}_{k_n,1}^*, \mathbb{W}_{k_n,2}^*, ..., \mathbb{W}_{k_n,M}^*$ cada uma conseguida por naïve boostsrap da amostra $\mathbb{W}_{k_n}$;
(2) Para cada j, construam-se N amostras $\mathbb{W}_{k_n,j,1}^{**}, \mathbb{W}_{k_n,j,2}^{**}, ..., \mathbb{W}_{k_n,j,N}^{**}$ por naïve bootstrap da amostra $\mathbb{W}_{K_n,j}^*$;
(3) Defina-se Z_j como sendo a proporção de valores do conjunto
$$\left\{R_n\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n,j,k}^{**}\right), EH_n^*\right) : 1 \leq k \leq N\right\},$$
que são inferiores ou iguais a $R_n\left(\left(X_{n-k_n,n}, \mathbb{W}_{k_n,j}^*\right), EH_n\right)$;
(4) Construa-se a f.d.e. com base no conjunto de v.a. i.i.d. $\{Z_j : 1 \leq j \leq M\}$, que constitui uma aproximação de $\widehat{H}_{BC1}$ para N e M suficientemente grandes.

7.3. **Notação e algumas considerações.** Antes de continuarmos, façamos uma pausa para introduzirmos alguma notação e considerações prévias que nos serão úteis posteriormente. A notação e definições introduzidas nesta secção serão as utilizadas no resto da sequela.

Seja $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ uma a.a. obtida de uma f.d. F que satisfaz (6.1.1). Considerem-se $(W_i)_{1 \leq i \leq k_n}$ definidos pela expressão (6.2.1). Considerem-se as amostras $\mathbb{W}_{k_n} = (W_1, W_2, ..., W_{k_n})$, $\mathbb{W}_{k_n}^* = \left(W_1^*, W_2^*, ..., W_{k_n}^*\right)$ obtida por naïve bootstrap de $\mathbb{W}_{k_n}$ e $\mathbb{W}_{k_n}^{**} = (W_1, W_2, ..., W_{k_n})$ obtida por naïve bootstrap de $\mathbb{W}_{k_n}^*$. Considerem-se, ainda, o estimador de Hill EH_n definido pela equação (6.2.2), a sua versão bootstrap EH_n^* definida pela equação (6.2.3) e a versão bootstrap de EH_n^* definida por

$$EH_n^{**} = \frac{1}{k_n} \sum_{i=1}^{k_n} W_i^{**}. \tag{7.3.1}$$

Considere-se, por fim, a raíz de Hill definida pela equação (6.2.7), com f.d. que denotamos por $H_n(\cdot)$. Definamos uma versão bootstrap de cauda de R_n da seguinte forma:

$$R_n^* = c\sqrt{k_n}\left(EH_n^* - EH_n\right), \tag{7.3.2}$$

e uma versão boostrap de cauda desta última pela equação:

$$R_n^{**} = c\sqrt{k_n}\left(EH_n^{**} - EH_n^*\right). \tag{7.3.3}$$

Denotemos por $\widehat{H}_{BC}$ a distribuição bootstrap de cauda de R_n definida por

$$\widehat{H}_{BC}(x) = P[R_n^* \leq x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})].$$

Definamos, agora, a raíz

$$R_{n1} = \widehat{H}_{BC}(R_n) = P[R_n^* \leq R_n \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})],$$

e a sua versão bootstrap de cauda

$$R_{n1}^* = P\left[R_n^{**} \leq R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right].$$

Denotamos por $H_{n1}(\cdot)$ a f.d de R_{n1} e por $\widehat{H}_{BC1}$ a sua distribuição bootstrap de cauda dada pela equação:

$$\begin{aligned} \widehat{H}_{BC1}(x) &= P\left[R_{n1}^{*} \leq x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] = \\ &= P\left[P\left[R_n^{**} \leq R_n^{*} \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^{*}\right)\right] \leq x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right]. \end{aligned}$$

Atentemos, agora, a alguns argumentos que se revelarão bastante úteis.

Como $X_i \stackrel{D}{=} F^{-1}(U_i)$, *onde* $U_i \sim \mathcal{U}_{[0,1]}$, $1 \leq i \leq k_n$, então é possível escrever sem perda de generalidade

$$W_i = \log F^{-1}(U_{n-k_n+i,n}) - \log F^{-1}(U_{n-k_n,n}).$$

Usando a equação (6.2.4) obtém-se a seguinte representação para $1 \leq i \leq k_n$:

$$W_i = -\frac{1}{c}\log Y_i + \log \frac{\widetilde{L}\left(Y_i\left(1-U_{n-k_n,n}\right)\right)}{\widetilde{L}\left(1-U_{n-k_n,n}\right)} = Q\left(Y_i, U_{n-k_n,n}\right),$$

onde

$$Y_i = \frac{1-U_{n-k_n+i,n}}{1-U_{n-k_n,n}}.$$

Seja $\widetilde{G}_{k_n}$ a f.d.e. associada a $Y_1, Y_2, ..., Y_{k_n}$ e suponhamos que $Y_1^*, Y_2^*, ..., Y_{k_n}^*$ são, condicionalmente, v.a. i.i.d. com f.d. $\widetilde{G}_{k_n}$, dada a amostra $Y_1, Y_2, ..., Y_{k_n}$. Note-se que a distribuição condicional de W_{i,k_n}^*, dada a amostra $Y_1, Y_2, ..., Y_{k_n}$ e $U_{n-k_n,n}$, é a mesma de $Q\left(Y_{i,k_n}^*, U_{n-k_n,n}\right)$. Analogamente, assumindo que $\widetilde{G}_{k_n}^*$ é a f.d.e. associada a $Y_1^*, Y_2^*, ..., Y_{k_n}^*$ e que $Y_1^{**}, Y_2^{**}, ..., Y_{k_n}^{**}$ são, condicionalmente, v.a. i.i.d. com f.d. $\widetilde{G}_{k_n}^*$, observa-se que, dada a amostra $Y_1^*, Y_2^*, ..., Y_{k_n}^*$ e $U_{n-k_n,n}$, a distribuição de W_{i,k_n}^{**} é a mesma de $Q\left(Y_{i,k_n}^{**}, U_{n-k_n,n}\right)$. Designemos por $\widetilde{G}_{k_n}^{**}$ a f.d.e. associada a $Y_1^{**}, Y_2^{**}, ..., Y_{k_n}^{**}$.

Note-se que

$$\left(-\log\left(1-U_{1,n}\right), -\log\left(1-U_{2,n}\right), ..., -\log\left(1-U_{n,n}\right)\right) \stackrel{D}{=} \left(Z_{1,n}, Z_{2,n}, ..., Z_{n,n}\right)$$

onde $(Z_{1,n}, Z_{2,n}, ..., Z_{n,n})$ é um vector de estatísticas de ordem de uma a.a., de dimensão n, com uma lei exponencial de parâmetro 1, $Exp(1)$. Conclui-se, assim, que

$$-\log(Y_i) \stackrel{D}{=} Z_{n-k_n+i,n} - Z_{n-k_n,n} \stackrel{D}{=} Z_{i,k_n} \tag{7.3.4}$$

onde Z_{i,k_n} é a i-ésima estatística de ordem de uma a.a., de dimensão k_n, com lei $Exp(1)$. A segunda igualdade da expressão (7.3.4) resulta do Lema 1.4.3. e do Teorema 1.6.1. da referência [15]. A título de nota, realce-se que

$$\left(1-Y_1, 1-Y_2, ..., 1-Y_{k_n}\right) \stackrel{D}{=} \left(U_{1,k_n}, U_{2,k_n}, ..., U_{n,k_n}\right), \tag{7.3.5}$$

onde $(U_{1,k_n}, U_{2,k_n}, ..., U_{n,k_n})$ é um vector de estatísticas de ordem de uma a.a., de dimensão k_n, com uma lei $\mathcal{U}_{[0,1]}$.

Estamos, agora, em condições de escrever

$$\begin{aligned} R_n &= \sqrt{k_n}\left[\frac{1}{k_n}\sum_{i=1}^{k_n} Z_{i,k_n} - 1\right] + c\sqrt{k_n}\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{\widetilde{L}\left(Y_i\left(1-U_{n-k_n,n}\right)\right)}{\widetilde{L}\left(1-U_{n-k_n,n}\right)}\right) = \\ &= I(n) + II(n), \end{aligned}$$

$$\begin{aligned} R_n^* &= \sqrt{k_n}\left[\frac{1}{k_n}\sum_{i=1}^{k_n} Z_{i,k_n}^* - \frac{1}{k_n}\sum_{i=1}^{k_n} Z_{i,k_n}\right] + c\sqrt{k_n}\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{\widetilde{L}\left(Y_{i,k_n}^*\left(1-U_{n-k_n,n}\right)\right)}{\widetilde{L}\left(Y_i\left(1-U_{n-k_n,n}\right)\right)}\right) = \\ &= I'(n) + II'(n), \end{aligned}$$

$$\begin{aligned} R_n^{**} &= \sqrt{k_n}\left[\frac{1}{k_n}\sum_{i=1}^{k_n} Z_{i,k_n}^{**} - \frac{1}{k_n}\sum_{i=1}^{k_n} Z_{i,k_n}^{*}\right] + c\sqrt{k_n}\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{\widetilde{L}\left(Y_{i,k_n}^{**}\left(1-U_{n-k_n,n}\right)\right)}{\widetilde{L}\left(Y_{i,k_n}^{*}\left(1-U_{n-k_n,n}\right)\right)}\right) = \\ &= I''(n) + II''(n). \end{aligned}$$

7.4. **Consistência Assimptótica.** A próxima proposição permitirá concluir a consistência assimptótica do intervalo de confiança C_{BC1}. Faremos a prova da proposição num esquema demonstrativo muito semelhante ao usado na demonstração do Teorema 6.2, que pode ser vista na referência, já citada, J.Bacro e M. Brito (1998).

Teorema A. *Considere-se R_n^{**} definido por (7.3.3). Nas condições do Teorema 6.2 temos que*

$$P\left[R_n^{**} \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^{*}\right)\right] \underset{n\to\infty}{\longrightarrow} \Phi(x),$$

em probabilidade.

Demonstração. (Teorema A)

Pelo Teorema de Malmquist (*vide* referência [15], secção 1.6), é fácil ver que o vector $(Y_1, Y_2, ..., Y_{k_n})$ é independente de $U_{n-k_n,n}$, e por conseguinte, o vector $\left(Y_1^*, Y_2^*, ..., Y_{k_n}^*\right)$ é também independente de $U_{n-k_n,n}$, pelo que:

$$\begin{aligned} &P\left[I''(n) \le x \mid Y_1^*, Y_2^*, ..., Y_{k_n}^*, Y_1, Y_2, ..., Y_{k_n}, U_{n.k_n,n}\right] = \\ &= P\left[I''(n) \le x \mid Y_1^*, Y_2^*, ..., Y_{k_n}^*, Y_1, Y_2, ..., Y_{k_n}\right]. \end{aligned}$$

Uma vez que a ordenação das v.a. de uma amostra não altera o valor da soma global, $k_n^{-\frac{1}{2}} I''(n)$ pode ser considerado como a diferença entre a média de v.a. i.i.d. com f.d. $\widetilde{G}_{k_n}$ e a correspondente média bootstrap. Aplicando o Teorema 3.3 concluímos que:

$$P\left[I''(n) \le x \mid Y_1^*, Y_2^*, ..., Y_{k_n}^*, Y_1, Y_2, ..., Y_{k_n}, U_{n-k_n,n}\right] \underset{n\to\infty}{\longrightarrow} \Phi(x), \quad q.c.$$

Considere-se, agora, o erro $II''(n)$. Tendo em conta que, dado $Y_1^*, Y_2^*, ..., Y_{k_n}^*, Y_1, Y_2, ..., Y_{k_n}, U_{n-k_n,n}$,

$$\frac{1}{c}k_n^{-\frac{1}{2}} II''(n) = \int_{Y_{k_n}}^{1} \log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))} d\widetilde{G}_{k_n}^{**}(y) - \int_{Y_{k_n}}^{1} \log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))} d\widetilde{G}_{k_n}^{*}(y)$$

resulta que

$$\begin{aligned} \left|\tfrac{1}{c}k_n^{-\frac{1}{2}} II''(n)\right| &\le \left|\int_{Y_{k_n}}^{1} \log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))} d\widetilde{G}_{k_n}^{**}(y)\right| + \left|\int_{Y_{k_n}}^{1} \log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))} d\widetilde{G}_{k_n}^{*}(y)\right| \\ &\le \int_{Y_{k_n}}^{1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right| d\widetilde{G}_{k_n}^{**}(y) + \int_{Y_{k_n}}^{1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right| d\widetilde{G}_{k_n}^{*}(y) \\ &\le 2 \sup_{Y_{k_n} < y \le 1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right|. \end{aligned}$$

A última desigualdade advém do facto de $1 \ge Y_1 \ge Y_2 \ge ... \ge Y_{k_n}$, $\int_{Y_{k_n}}^{1} d\widetilde{G}_{k_n}^{**}(y) \le 1$ e $\int_{Y_{k_n}}^{1} d\widetilde{G}_{k_n}^{*}(y) \le 1$. Notando que

$$\begin{aligned} \sup_{Y_{k_n} < y \le 1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right| &\le \sup_{Y_{k_n} \le y \le 1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right| + \left|\log \frac{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right| \\ &\le 2 \sup_{Y_{k_n} \le y \le 1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right| \end{aligned}$$

resulta que

$$|II''(n)| \le 4ck_n^{\frac{1}{2}} \sup_{Y_{k_n} \le y \le 1} \left|\log \frac{\widetilde{L}\left(y\left(1-U_{n-k_n,n}\right)\right)}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right|.$$

Seguidamente, escolha-se $(\lambda_1, \lambda_2) \in]1, +\infty[\times]1, +\infty[$ e considere-se o acontecimento $A_n(\lambda_1, \lambda_2) = \left\{\lambda_1^{-1} \le \frac{n}{k_n}(1 - U_{n-k_n,n}) \le \lambda_1, \lambda_2^{-1} \le k_n Y_{k_n} \le \lambda_2\right\}$. Em A_n temos que

$$k_n^{\frac{1}{2}} \sup_{Y_{k_n} \le y \le 1} \left| \log \frac{\widetilde{L}(y(1 - U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)} \right| \le k_n^{\frac{1}{2}} \sup_{\frac{1}{k_n} \le y \le 1} \sup_{\frac{1}{\lambda_1 \lambda_2} \le t \le \lambda_1} \left| \log \frac{\widetilde{L}\left(yt\frac{k_n}{n}\right)}{\widetilde{L}\left(\frac{k_n}{n}\right)} \right|$$

que converge para zero quando $n \to \infty$, pela condição (6.2.5). Uma vez que $P[A_n(\lambda_1, \lambda_2)] \underset{n\to\infty}{\longrightarrow} 1$, obtém-se que, dado $Y_1^*, Y_2^*, ..., Y_{k_n}^*, Y_1, Y_2, ..., Y_{k_n}, U_{n-k_n,n}$, $II''(n) \underset{n\to\infty}{\longrightarrow} 0$, em probabilidade. Deduzimos então que

$$P\left[R_n^{**} \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] \underset{n\to\infty}{\longrightarrow} \Phi(x), \tag{7.4.1}$$

em probabilidade. □

Observação 7.1. *Porque EH_n^* é um estimador consistente de $\frac{1}{c}$, resulta que é possível (pelo Teorema de Slutsky) substituir c em R_n^{**} (observar a sua definição na equação (7.3.3)) por $(EH_n^*)^{-1}$ sem que o teorema anterior fique desprovido de validade.*

Tendo em conta as observações 6.2 e 7.1, a consistência assimptótica do intervalo de confiança C_{BC1} fica estabelecida (usando para R_{n1} a mesma técnica usada para R_n na demostração da proposição 3.2) se mostrarmos que

$$P\left[P\left[R_n^{**} \le R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] \underset{n\to\infty}{\longrightarrow} x, \tag{7.4.2}$$

em probabilidade. Para o efeito, observemos que

$$P\left[R_n^{**} \le R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] = Q + S,$$

onde

$$Q = P\left[R_n^{**} \le R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] - \Phi(R_n^*)$$

e

$$S = \Phi(R_n^*).$$

Usando o Teorema de Polya e a proposição A é possível mostrar que

$$\forall \varepsilon > 0 : P[Q > \varepsilon \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})] \to 0,$$

em probabilidade. Do teorema 6.2 resulta que

$$P[S \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})] \to x,$$

em probabilidade. Por fim, aplicando o teorema de Slutsky, verifica-se a veracidade da equação (7.4.2).

7.5. **Ordem de Convergência do Bootstrap de Cauda e Pré-Pivotagem de Cauda.** Nesta secção, pretende-se determinar o erro cometido na construção dos intervalos de confiança C_{BC} e C_{BC1}. Faremos isso analisando o erro com que as distribuições bootstrap de cauda de R_n e R_{n1} aproximam as suas distribuições exactas.

Teorema B. *Assumamos que a condição (6.1.1) é válida, que k_n é uma sucessão de inteiros positivos satisfazendo (6.1.2). Suponhamos que*

$$P\left[V(n) > \frac{1}{4}k_n^{-j/2}\right] = O\left(k_n^{-j/2}\right), \quad j \in \{1, 2, 3\}, \tag{7.5.1}$$

onde

$$V(n) = \sqrt{k_n} \sup_{Y_{k_n} \le y \le 1} \left| \log \frac{\widetilde{L}(y(1 - U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)} \right|.$$

Então, uniformemente em x, tem-se que:

i): *se $j \in \{1, 2, 3\}$, $|\Phi(x) - H_n(x)| = O\left(k_n^{-1/2}\right)$;*

ii): *se $j \in \{2, 3\}$, $\left|\widehat{H}_{BC}(x) - H_n(x)\right| = O\left(k_n^{-1}\right)$ em probabilidade;*

iii): *se* $j=3$, $\left|\widehat{H}_{BC1}(x)-H_{n1}(x,F)\right|=O\left(k_n^{-3/2}\right)$ *em probabilidade.*

Demonstração. Dividiremos a prova em várias etapas de forma a simplificar a leitura.

1- $P\left[|II(n)|>k_n^{-j/2}\right]=O\left(k_n^{-j/2}\right)$.

Notemos que:

$$II(n)=c\sqrt{k_n}\int_{Y_{k_n}}^{1}\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}d\widetilde{G}_{k_n}(y)+c\sqrt{k_n}\log\frac{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}$$

Resulta assim que:

$$\begin{aligned}|II(n)| &\leq c\sqrt{k_n}\left|\int_{Y_{k_n}}^{1}\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}d\widetilde{G}_{k_n}(y)\right|+c\sqrt{k_n}\left|\log\frac{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}\right| \\ &\leq c\sqrt{k_n}\int_{Y_{k_n}}^{1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}\right|d\widetilde{G}_{k_n}(y)+c\sqrt{k_n}\left|\log\frac{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}\right| \\ &\leq 2c\sqrt{k_n}\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}\right|.\end{aligned}$$

Mas,

$$\begin{aligned}\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(1-U_{n-k_n,n})}\right| &\leq \sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(\frac{k_n}{n})}\right|+\left|\log\frac{\widetilde{L}(1-U_{n-k_n,n})}{\widetilde{L}(\frac{k_n}{n})}\right| \\ &\leq 2\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(\frac{k_n}{n})}\right|.\end{aligned}$$

Assim sendo, temos que $|II(n)|\leq 4V(n)$, logo, usando (7.5.1), $P\left[|II(n)|>k_n^{-j/2}\right]\leq P\left[|V(n)|>\frac{k_n^{-j/2}}{4}\right]=O\left(k_n^{-j/2}\right)$, c.q.d.

2- $P\left[|II'(n)|>k_n^{-j/2}\mid(X_{n-k_n,n},\mathbb{W}_{k_n})\right]=0$**, em probabilidade**.

Comecemos por notar que, dado $(X_{n-k_n,n},\mathbb{W}_{k_n})$:

$$\begin{aligned}|II'(n)| &= c\sqrt{k_n}\left|\int_{Y_{k_n}}^{1}\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}d\widetilde{G}^*_{k_n}(y)-\int_{Y_{k_n}}^{1}\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}d\widetilde{G}_{k_n}(y)\right| \\ &\leq c\sqrt{k_n}\int_{Y_{k_n}}^{1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right|d\widetilde{G}^*_{k_n}(y) \\ &\quad +c\sqrt{k_n}\int_{Y_{k_n}}^{1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right|d\widetilde{G}_{k_n}(y) \\ &\leq 2c\sqrt{k_n}\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right|.\end{aligned}$$

Como

$$\begin{aligned}\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}\right| &\leq \sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(\frac{k_n}{n})}\right|+\left|\log\frac{\widetilde{L}(Y_{k_n}(1-U_{n-k_n,n}))}{\widetilde{L}(\frac{k_n}{n})}\right| \\ &\leq 2\sup_{Y_{k_n}\leq y\leq 1}\left|\log\frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}(\frac{k_n}{n})}\right|,\end{aligned}$$

então $|II'(n)|\leq 4V(n)$, dado $(X_{n-k_n,n},\mathbb{W}_{k_n})$. Resulta, assim, que

$$P\left[|II'(n)|>k_n^{-j/2}\mid(X_{n-k_n,n},\mathbb{W}_{k_n})\right]\leq P\left[V(n)>\frac{k_n^{-j/2}}{4}\mid(X_{n-k_n,n},\mathbb{W}_{k_n})\right].$$

Mas,

$$\begin{aligned} P\left[V(n) > \frac{k_n^{-j/2}}{4} \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] &= E\left(\mathbb{I}\left\{V(n) > \frac{k_n^{-j/2}}{4}\right\} \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right) = \\ &= \mathbb{I}\left\{V(n) > \frac{k_n^{-j/2}}{4}\right\}. \end{aligned}$$

Como temos por hipótese (7.5.1), então $P\left[V(n) > \frac{k_n^{-j/2}}{4} \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] = 0$ em probabilidade. Em conclusão, obtemos que $P\left[|II'(n)| > k_n^{-j/2} \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] = 0$, em probabilidade.

3- $P\left[|II''(n)| > k_n^{-j/2} \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}^*_{k_n}\right)\right] = 0$**, em probabilidade**.

Em virtude da demonstração deste facto ser inteiramente semelhante à produzida na etapa anterior, omiti-la-emos.

4- Cálculo da expansão Edgeworth de R_n

Pela teorema 5.1 a seguinte expansão Edgeworth é válida:

$$\begin{aligned} P[I(n) \le x] &= \Phi(x) + k_n^{-1/2} p_1(x)\phi(x) + k_n^{-1} p_2(x)\phi(x) + k_n^{-3/2} p_3(x)\phi(x) + o\left(k_n^{-3/2}\right) \\ &= \Phi(x) + k_n^{-1/2} s_1(x) + k_n^{-1} s_2(x) + k_n^{-3/2} s_3(x) + o\left(k_n^{-3/2}\right). \end{aligned}$$

Usando o método delta (secção 2.7 da referência [13]) e tendo em conta o tópico 1 temos que:

$$\begin{aligned} P[R_n \le x] &= P[I(n) + II(n) \le x] \le P\left[I(n) \le x + k_n^{-j/2}\right] + P\left[|II(n)| > k_n^{-j/2}\right] \\ &\le P\left[I(n) \le x + k_n^{-j/2}\right] + O\left(k_n^{-j/2}\right) = P[I(n) \le x] + O\left(k_n^{-j/2}\right). \end{aligned} \tag{7.5.2}$$

A última igualdade, em (7.5.2), pode ser verificada da seguinte forma:

$$\begin{aligned} P\left[I(n) \le x + k_n^{-j/2}\right] &= \Phi\left(x + k_n^{-j/2}\right) + k_n^{-1/2} s_1\left(x + k_n^{-j/2}\right) + k_n^{-1} s_2\left(x + k_n^{-j/2}\right) \\ &\quad + k_n^{-3/2} s_3\left(x + k_n^{-j/2}\right) + o\left(k_n^{-3/2}\right) \\ &= \Phi(x) + k_n^{-1/2} s_1(x) + k_n^{-1} s_2(x) + k_n^{-3/2} s_3(x) + o\left(k_n^{-3/2}\right) \\ &\quad + k_n^{-j/2}\left(\phi(\xi_1) + k_n^{-1/2} s_1'(\xi_2) + k_n^{-1} s_2'(\xi_3) + k_n^{-3/2} s_3'(\xi_4)\right) \\ &= P[I(n) \le x] + O\left(k_n^{-j/2}\right), \end{aligned}$$

onde $\xi_1, \xi_2, \xi_3, \xi_4 \in \left]x, x + k_n^{-j/2}\right[$.

Por outro lado, temos que

$$\begin{aligned} P[R_n \le x] &= P[I(n) + II(n) \le x] \ge P\left[I(n) \le x - k_n^{-j/2}\right] - P\left[|II(n)| > k_n^{-j/2}\right] \\ &\ge P\left[I(n) \le x - k_n^{-j/2}\right] + O\left(k_n^{-j/2}\right) = P[I(n) \le x] + O\left(k_n^{-j/2}\right). \end{aligned} \tag{7.5.3}$$

Como consequência de (7.5.2) e (7.5.3), obtém-se que

$$\begin{aligned} P[R_n \le x] &= P[I(n) \le x] + O\left(k_n^{-j/2}\right) \\ &= \Phi(x) + k_n^{-1/2} s_1(x) + k_n^{-1} s_2(x) + k_n^{-3/2} s_3(x) + O\left(k_n^{-3/2}\right). \end{aligned}$$

Deste desenvolvimento, conclui-se facilmente que $|\Phi(x) - H_n(x)| = O\left(k_n^{-1/2}\right)$, uniformemente em x. (Note-se que todos os desenvolvimentos produzidos são uniformes em x, razão pela qual deixaremos de fazer referência a esse facto).

5- Cálculo da expansão Edgeworth de R_n^*

Pela teorema 5.2, tem-se que é válida a seguinte expansão Edgeworth:

$$P\left[I'(n) \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] = \Phi(x) + k_n^{-1/2}\widehat{s}_1(x) + k_n^{-1}\widehat{s}_2(x) + k_n^{-3/2}\widehat{s}_3(x) + o_P\left(k_n^{-3/2}\right),$$

onde $\widehat{s}_1, \widehat{s}_2$ e $\widehat{s}_3$ advêm de s_1, s_2 e s_3, respectivamente, substituindo os cumulantes (que dependem dos momentos da distribuição $Exp(1)$) pelas respectivas versões bootstrap.

Usando, novamente, o método delta e tendo em conta o tópico 2, facilmente se verifica que

$$\begin{aligned} P\left[R_n^* \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] &= P\left[I'(n) + II'(n) \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] = \\ &= P\left[I'(n) \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] + O\left(k_n^{-j/2}\right) \\ &= \Phi(x) + k_n^{-1/2}\widehat{s}_1(x) + k_n^{-1}\widehat{s}_2(x) + k_n^{-3/2}\widehat{s}_3(x) + O\left(k_n^{-j/2}\right) \end{aligned}$$

em probabilidade.

6- Comparação entre $H_n(\cdot)$ e $\widehat{H}_{BC}(\cdot)$

Uma vez que

$$\begin{aligned} \left|P\left[R_n^* \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] - P\left[R_n \le x\right]\right| &= k_n^{-1/2}\left|\widehat{s}_1(x) - s_1(x)\right| \\ &\quad + k_n^{-1}\left|\widehat{s}_2(x) - s_2(x)\right| + O\left(k_n^{-j/2}\right), \end{aligned}$$

em probabilidade, e que os coeficientes de $\widehat{s}_i$ diferem dos de s_i de $O_P\left(k_n^{-1/2}\right)$, então $\left|\widehat{H}_{BC}(x) - H_n(x)\right| = O_P\left(k_n^{-1}\right)$, desde que $j \in \{2, 3\}$, c.q.d.

7- Comparação entre $H_{n1}(\cdot)$ **e** $\widehat{H}_{BC1}(\cdot)$

Inicialmente, notemos que porque F_n^* é um estimador consistente de F, existe uma expansão Edgeworth para R_n^{**} semelhante à desenvolvida para R_n^* no tópico 4. Assim sendo, de forma semelhante ao que já foi feito, resulta que

$$P\left[R_n^{**} \le x + k_n^{-j/2} \middle| \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] = P\left[R_n^{**} \le x \middle| \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] + O\left(k_n^{-j/2}\right).$$

Posto isto, atendendo ao tópico 3 e cálculos efectuados nos tópicos 4 e 5, vem facilmente que

$$P\left[R_n^{**} \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] = P\left[I''(n) \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] + O\left(k_n^{-j/2}\right),$$

em probabilidade. Considerações prévias permitem ainda escrever:

$$P\left[R_n^{**} \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] =$$

$$= P\left[I''(n) + II'(n) \le x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] + O\left(k_n^{-j/2}\right),$$

em probabilidade.

Lembrando que $R_{n1}^* = P\left[R_n^{**} \le R_n^* \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right]$, resulta que

$$\begin{aligned} R_{n1}^* &= P\left[I''(n) + II'(n) \le I'(n) + II'(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] + O\left(k_n^{-j/2}\right) \\ &= P\left[I''(n) \le I'(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}_{k_n}^*\right)\right] + O\left(k_n^{-j/2}\right), \end{aligned}$$

em probabilidade.

Analogamente se constata que

$$R_{n1} = P\left[I'(n) \le I(n) \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] + O\left(k_n^{-j/2}\right),$$

em probabilidade.

Verifica-se, por fim, de acordo com o estudo feito por Beran (1987), que

$$\left|P\left[R_{n1}^* \le x \mid (X_{n-k_n,n}, \mathbb{W}_{k_n})\right] - P\left[R'_{n1} \le x\right]\right| =$$

$$\begin{aligned}
&= P\left[P\left[I''(n) \leq I'(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}^*_{k_n}\right)\right] + O\left(k_n^{-j/2}\right) \leq x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right)\right] \\
&\quad -P\left[P\left[I'(n) \leq I(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right)\right] + O\left(k_n^{-j/2}\right) \leq x\right] \\
&= P\left[P\left[I''(n) \leq I'(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}, \mathbb{W}^*_{k_n}\right)\right] \leq x \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right)\right] \\
&\quad -P\left[P\left[I'(n) \leq I(n) \mid \left(X_{n-k_n,n}, \mathbb{W}_{k_n}\right)\right] \leq x\right] + O\left(k_n^{-j/2}\right) \\
&= O\left(k_n^{-3/2}\right),
\end{aligned}$$

em probabilidade, se $j = 3$. □

7.6. **Exemplo de Aplicação.** Consideremos a família de funções de distribuição cuja cauda tem o seguinte comportamento limite (*vide* Haeusler e Teugels (1985) e referências citadas):

$$1 - F(x) = Kx^{-c}\left[1 + O\left(x^{-\beta}\right)\right].$$

Escolhamos um caso particular desta família tomando $K = 2$, $c = \beta = 1$. Então $F^{-1}(y) = (1-y)^{-1} + 2$, quando $y \to 1$, donde se conclui que $\widetilde{L}(y) = 1 + 2y$. O objectivo é mostrar que para uma escolha adequada da sequência k_n (6.1.2), a condição (7.5.1) é satisfeita. Para o efeito atente-se aos cálculos seguintes.

$$\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)} = \log \frac{1+2y(1-U_{n-k_n,n})}{1+2\frac{k_n}{n}} \Rightarrow \sup_{Y_{k_n} \leq y \leq 1} \left|\log \frac{\widetilde{L}(y(1-U_{n-k_n,n}))}{\widetilde{L}\left(\frac{k_n}{n}\right)}\right| = \left|\log \frac{1+2(1-U_{n-k_n,n})}{1+2\frac{k_n}{n}}\right|$$

$$\left|\log \frac{1+2(1-U_{n-k_n,n})}{1+2\frac{k_n}{n}}\right| \stackrel{D}{=} \left|\log\left(1+2U_{k_n+1,n}\right) - \log\left(1+2\frac{k_n}{n}\right)\right|$$

$$\left|\log\left(1+2U_{k_n+1,n}\right) - \log\left(1+2\frac{k_n}{n}\right)\right| = 2\left|U_{k_n+1,n} - \frac{k_n}{n}\right|(1+o(1));$$

$$\begin{aligned}
P\left[2\sqrt{k_n}\left|U_{k_n+1,n} - \tfrac{k_n}{n}\right| > \tfrac{1}{4}k_n^{-3/2}\right] &= P\left[\left|U_{k_n+1,n} - \tfrac{k_n}{n}\right| > \tfrac{1}{8}k_n^{-2}\right] = \\
P\left[\left(U_{k_n+1,n} - \tfrac{k_n}{n}\right)^2 > \tfrac{1}{64}k_n^{-4}\right] &\leq \frac{E\left(\left(U_{k_n+1,n} - \frac{k_n}{n}\right)^2\right)}{\frac{1}{64}k_n^{-4}},
\end{aligned}$$

onde a última desigualdade é obtida por aplicação da desigualdade de Markov.

Como

$$E\left(\left(U_{k_n+1,n} - \frac{k_n}{n}\right)^2\right) = E\left(\left(U_{k_n+1,n} - \frac{k_n+1}{n+1}\right)^2\right)(1+o(1))$$

e

$$E\left(\left(U_{k_n+1,n} - \frac{k_n+1}{n+1}\right)^2\right) = V\left(U_{k_n+1,n}\right) = \frac{\frac{k_n+1}{n+1} - \left(\frac{k_n+1}{n+1}\right)^2}{n+2},$$

resulta que se escolhermos $k_n \leq n^{4/13}$, então, para n suficientemente grande, a condição (7.5.1) é satisfeita.

Apêndice A. Alguns resultados

Teorema A.1. *(Glivenko- Cantelli)*

Suponhamos que $X_1, X_2, \ldots$ é uma sucessão de v.a independentes com f.d. F. Designemos por F_n a f.d.e. associada às n primeiras variáveis, $X_1, X_2, \ldots, X_n$. Então

$$\|F_n - F\|_\infty \underset{n\to\infty}{\longrightarrow} 0, q.c.$$

Teorema A.2. *(Polya)*

Sejam F, G_1, G_2 ,... f.d. Se $G_n \xrightarrow{w} F$ e F é contínua então

$$\lim_{n\to\infty} \|G_n - F\|_\infty = 0$$

Corolário A.3. *(do Teorema de Slutsky)*
Seja $c \in R$. Se

$$X_n \xrightarrow{D} X \ \ e \ Y_n \xrightarrow{P} c$$

então

$$X_n + Y_n \xrightarrow{D} X + c$$

e

$$Y_n X_n \xrightarrow{D} cX.$$

Os resultados deste apêndice podem ser vistos, por exemplo, na referência [8].

Referências

[1] Abramovitch, L. and Singh, K. (1985), Edgeworth corrected pivotal statistics and the bootstrap, *Annals of Statistics*, Vol. 13, 116-132.
[2] Bacro, J.N. and Brito, M. (1998), A tail bootstrap procedure for estimating the tail Pareto- index, *Journal of Statistical Planning and Inference*, Vol. 71, 245-260.
[3] Beran, Rudolf (1982), Estimated sampling distributions: the bootstrap and competitors, *Annals of Statistics*, Vol. 10, 212-225.
[4] Beran, Rudolf (1987), Prepivoting to reduce level error of confidence sets, *Biometrika*, Vol. 74, No.3, 457-468.
[5] Beran, R. and Ducharme, G.R. (1991), Asymptotic theory for bootstrap methods in statistics, *Publication C.R.M. Universités de Montreal.*
[6] Bickel, P. J. and Freedman, D.A. (1981), Some asymptotic theory for the bootstrap, *Annals of Statistics,* Vol. 9, No. 6, 1196-1217.
[7] Diciccio, T.H. and Romano, J.P. (1988), A review of bootstrap confidence intervals, *Journal of Royal Statistical Society*-B, Vol. 50, No. 3, 338-354.
[8] Dudewicz Edward J. and Mishra, Satya N. (1988), Modern Mathematical Statistics, *Wiley Series in Probability and Mathematical Statistics.*
[9] Efron, B. (1979), Bootstrap Methods: another look at Jacknife, *Annals of Statistics*, Vol. 7, 1-26.
[10] Haeusler, E. and Teugels, J.L. (1985), On asymptotic normality of Hill's estimator for the exponent of regular vatiation, *Annals of Statistics*, Vol. 13, No. 2, 743-756.
[11] Hall, Peter (1983), Inverting an Edgeworth expansion, *Annals of Statistics*, Vol. 11, 569-576.
[12] Hall, Peter (1986), On the bootstrap and confidence intervals, *Annals of Statistics*, Vol. 14, 1431-1452.
[13] Hall, Peter (1992), The bootstrap and Edgeworth expansion, *Springer- Verlag*, New York.
[14] Hill, B.M. (1975), A simple general aproach to inference about the tail of a distribution, *Annals of Statistics*, Vol. 3, 1663-1674.
[15] Reiss, Rolf-Dieter (1989), Aproximate distributions of order statistics, *Springer- Verlag*, New York.