



HAL
open science

Bootstrap de l'estimateur de Hill: théorèmes limites.

Salim Bouzebda

► **To cite this version:**

Salim Bouzebda. Bootstrap de l'estimateur de Hill: théorèmes limites.. Annales de l'ISUP, 2010, LIV (1-2), pp.61-72. hal-01764380

HAL Id: hal-01764380

<https://hal.science/hal-01764380>

Submitted on 8 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
LIV, fasc. 1-2, 2010, 61 à 72

Bootstrap de l'Estimateur de Hill: Théorèmes Limites

Salim BOUZEBDA

L.S.T.A., Université Pierre et Marie Curie-Paris 6, 175, rue du Chevaleret, 8^{ème} étage, bâtiment A,
75013 PARIS FRANCE.

Résumé. Nous développons une utilisation du bootstrap pour l'estimation de l'indice de Pareto d'une distribution extrême. Nous considérons, au départ, une suite X_1, \dots, X_n de variables aléatoires indépendantes et identiquement distribuées de fonction de répartition $F(\cdot)$, vérifiant

$$\lim_{x \rightarrow \infty} \frac{1 - F(xt)}{1 - F(x)} = t^{-\frac{1}{c}} \text{ pour tout } t > 0.$$

Nous établissons des propriétés de normalité asymptotique des sommes d'extrêmes bootstrappées, ainsi que pour, l'estimateur de Hill bootstrappé. Nous obtenons nos résultats, en faisant usage aux approximations pondérées du processus empirique uniforme et quantile par des pont browniens appropriés.

AMS : 62F40, 60F17, 62G32, 62G30.

Keywords: Théorie des valeurs extrêmes; Bootstrap; Processus empirique; L'estimateur de Hill.

1 Introduction

Soit $X = X_1, X_2, \dots$ une suite de variables aléatoires indépendantes et identiquement distribuées [i.i.d.], de fonction de répartition $F(x) = P(X \leq x)$, pour $x \in \mathbb{R}$. Soit $Q(\cdot)$ la fonction de quantiles associée, définie par

$$Q(s) := \inf\{x : F(x) \geq s\} \text{ pour } 0 < s < 1, \quad Q(0) = \lim_{s \downarrow 0} Q(s) \text{ et } Q(1) = \lim_{s \uparrow 1} Q(s).$$

Nous supposons que la fonction de queue, $1 - F(\cdot)$, est de *type Pareto*, d'indice $1/c > 0$, c'est à dire, telle que

$$1 - F(x) = P(X > x) = x^{-\frac{1}{c}} L(x), \text{ pour } x > 0, \quad (1)$$

où $L(\cdot)$ désigne une fonction à *variation lente* à l'infini. En d'autres termes, cette fonction est telle que

$$\lim_{x \rightarrow \infty} L(\lambda x)/L(x) = 1, \text{ pour tout } \lambda > 0. \quad (2)$$

L'estimateur de c le plus classique est l'estimateur de Hill (1975), généralisé par la famille d'estimateurs à noyaux, introduite par Csörgő *et al.* (1985). Nous limitons ici notre étude à l'estimateur de Hill (1975), tout en notant que nos résultats peuvent être généralisés aux estimateurs de Csörgő *et al.* (1985), au prix de développements additionnels. Pour tout choix de $n \geq 1$, désignons par $X_{1,n} \leq \dots \leq X_{n,n}$ la statistique ordonnée de l'échantillon X_1, \dots, X_n . Faisant usage de la notation $\log_+ u = \log(u \vee e)$ pour $u \in \mathbb{R}$, l'estimateur \hat{c}_n de c , introduit par Hill (1975), est défini par

$$\hat{c}_n = \hat{c}_{n;k_n} := k_n^{-1} \sum_{i=1}^{k_n} \log_+ X_{n-i+1,n} - \log_+ X_{n-k_n,n}, \quad (3)$$

où $1 \leq k_n \leq n-1$ est un paramètre de lissage entier, dépendant de n . Pour assurer la convergence asymptotique de \hat{c}_n vers c , les conditions suivantes sont requises.

$$k_n \rightarrow \infty \text{ et } n^{-1}k_n \rightarrow 0. \quad (4)$$

Mason (1982) a montré que ces conditions étaient nécessaires et suffisantes pour que $\hat{c}_n \rightarrow c$ en probabilité lorsque $n \rightarrow \infty$, pour toute loi de probabilité $F(\cdot)$, sous-jacente, vérifiant (1). La caractérisation de la

convergence presque sûre de \hat{c}_n vers c requiert, comme l'ont montré Deheuvels *et al.* (1988), la condition supplémentaire que $k_n / \log \log n \rightarrow \infty$ lorsque $n \rightarrow \infty$. La normalité asymptotique de \hat{c}_n est de nature sensiblement plus délicate. Dans le cas d'une loi de Pareto exacte, obtenue en choisissant $L(t) = 1$ pour tout $t \geq 1$ dans (1), on obtient le résultat simplifié que, indépendamment du choix de k_n vérifiant (4), on a

$$k_n^{1/2}(\hat{c}_n - c) \xrightarrow{d} N(0, c^2), \quad (5)$$

lorsque $n \rightarrow \infty$. Par contre, lorsque $L(\cdot)$ est quelconque, la relation n'est plus satisfaite en général, du fait que l'estimateur \hat{c}_n est alors asymptotiquement biaisé. Sa validité requiert des hypothèses de régularité additionnelles portant sur $F(\cdot)$, ainsi que sur le comportement asymptotique de k_n . On consultera, à ce sujet, Davis et Resnick (1984), Csörgő et Mason (1985), Haeusler et Teugels (1985) et Hall et Welsh (1985). Dans un cadre très général, et sous des conditions peu restrictives, Csörgő *et al.* (1985) ont établi la normalité asymptotique d'estimateurs à noyau, comprenant, comme cas particulier, l'estimateur de Hill \hat{c}_n . L'objet du présent article est d'explorer le comportement du *bootstrap multinomial*, introduit par Efron (1979), appliqué au rééchantillonnage de l'estimateur \hat{c}_n . Les travaux de Deheuvels *et al.* (1993) ont montré que le rééchantillonnage des valeurs extrêmes posaient de multiples problèmes, et il s'agit donc, dans ce qui suit, d'explorer les propriétés correspondantes pour l'estimateur de Hill. Le bootstrap de l'estimateur de Hill, que nous étudions, est également lié au bootstrap de sommes d'extrêmes, lui-même associé au bootstrap de sommes de variables aléatoires indépendantes de même loi. On consultera, entre autres, à ce sujet, Arcones et Giné (1989), Hall (1990) et Deheuvels *et al.* (1993). Ces derniers auteurs ont étudié le bootstrap de sommes globales, comprenant l'ensemble de l'échantillon, ainsi que de sommes, *intermédiaires* et *extrêmes* de statistiques ordonnées. Nous renvoyons à leur article pour des définitions plus explicites de ces notions.

Nous adopterons les notations suivantes. Désignons par $F_n(x) = n^{-1} \#\{X_i \leq x : 1 \leq i \leq n\}$ la fonction de répartition empirique associée à X_1, \dots, X_n , où $\#E$ désigne la cardinalité de E . Nous introduisons ensuite une suite X_1^*, X_2^*, \dots de variables aléatoires bootstrappées, constituant, conditionnellement à F_n , une suite de variables aléatoires i.i.d., de fonction de répartition F_n . Pour tout choix de $m \geq 1$, on désigne par $X_{1,m}^* \leq \dots \leq X_{m,m}^*$ la statistique ordonnée de X_1^*, \dots, X_m^* . Nous choisissons alors un entier $1 \leq \ell_m \leq m - 1$, et définissons l'estimateur de Hill bootstrappé \hat{c}_{m,ℓ_m}^* , correspondant à m, ℓ_m et n , par

$$\hat{c}_{m,\ell_m}^* := \ell_m^{-1} \sum_{i=1}^{\ell_m} \log_+ X_{m-i+1,m}^* - \log_+ X_{m-\ell_m,m}^*. \quad (6)$$

Nous supposons, de plus, que, lorsque $m \rightarrow \infty$,

$$\ell_m \rightarrow \infty \quad \text{et} \quad m^{-1} \ell_m \rightarrow 0. \quad (7)$$

La fonction de répartition empirique bootstrappée et la fonction de quantiles empiriques bootstrappée correspondant aux paramètres m, ℓ_m et n ci-dessus, sont notées

$$F_{m;n}(t) := m^{-1} \#\{X_i^* \leq t : 1 \leq i \leq m\}, \quad (8)$$

et

$$Q_{m;n}(s) := \inf\{t : F_{m;n}(t) \geq s\} \quad \text{pour} \quad 0 < s \leq 1 \quad \text{et} \quad Q_{m;n}(0) = X_{1,m}^*. \quad (9)$$

Il sera commode, par la suite de nous placer dans une asymptotique où $m = m_n$ et $\ell = \ell(n) = \ell_{m_n}$ dépendent de n . Nous adopterons alors les hypothèses suivantes sur le comportement relatif de ces quantités. Nous supposons qu'il existe des constantes A, B et A', B' , telles que $0 < A < B < \infty$ et $0 < A' < B' < \infty$, pour lesquelles on ait, pour tout $n \geq 1$, d'une part, $m = m_n \geq 1$, et, d'autre part,

$$A \leq \frac{\ell(n)}{k_n} = \frac{\ell_{m_n}}{k_n} = \frac{\ell_m}{k_n} \leq B \quad \text{et} \quad A' \leq \frac{m}{n} = \frac{m_n}{n} \leq B'. \quad (10)$$

Ces préliminaires étant posés, nous sommes en mesure de présenter les résultats principaux de cette note. Ceux-ci seront donnés dans le §1.1 ci-dessous, sous la forme de deux corollaires de théorèmes plus généraux, exposés

par la suite dans le §2, ou sont reportées l'ensemble de nos démonstrations. Nous anticipons sur ce paragraphe pour mentionner que les preuves de ces résultats reposent, pour l'essentiel, sur les approximations pondérées du processus empirique uniforme et quantile par des pont browniens appropriés dues à Csörgő *et al.* (1986) et Csörgő et Mason (1989). Dans le corollaire 1.2 ci-dessous, nous établissons la normalité asymptotique du bootstrap de l'estimateur de Hill. Dans le corollaire 1.1 qui suit, nous établissons la normalité asymptotique des sommes d'extrêmes bootstrappées

1.1 Résultats

Selon l'usage, $N(\mu, \sigma^2)$ désigne une variable aléatoire normale de moyenne μ et de variance σ^2 .

Corollaire 1.1 *Supposons que $F(\cdot)$ vérifie la condition (1), et que les suites k_n , $\ell(n) = \ell_{m_n}$ et $m = m_n$ vérifient les conditions (4), (7) et (10). Supposons de plus que $\lim_{n \rightarrow \infty} \ell(n)/k_n \rightarrow \zeta$ avec $0 < \zeta < \infty$. Alors, lorsque $n \rightarrow \infty$,*

$$\sqrt{\ell(n)} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} \log_+ X_{m-i+1, m}^* - \frac{1}{k_n} \sum_{i=1}^{k_n} \log_+ X_{n-i+1, n} - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \xrightarrow{d} N(0, (1+\zeta)2c^2).$$

La démonstration du corollaire 1.1 est reportée au §2.

Corollaire 1.2 *Supposons que $F(\cdot)$ vérifie la condition (1), que les suites k_n , $\ell(n) = \ell_{m_n}$ et $m = m_n$ vérifient les conditions (4), (7) et (10). Supposons de plus que $\lim_{n \rightarrow \infty} \ell(n)/k_n \rightarrow \zeta$ avec $0 < \zeta < \infty$. Alors, lorsque $n \rightarrow \infty$,*

$$\sqrt{\ell(n)} \left\{ \log_+ (X_{m-\ell(n), m}^*) - \log_+ (X_{n-k_n, n}) - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \xrightarrow{d} N(0, (1+\zeta)c^2)$$

et

$$\sqrt{\ell(n)} \left\{ \hat{c}_{m; \ell(n)}^* - \hat{c}_n \right\} \xrightarrow{d} N(0, (1+\zeta)c^2). \quad (11)$$

La démonstration du corollaire 1.2 est reportée au §2.

2 Démonstrations

Nous énonçons, en préalable, quelques résultats techniques qui nous seront utiles dans la suite de nos démonstrations. Nous supposons que les objets aléatoires considérés par la suite sont définis sur l'espace de probabilités (Ω, \mathcal{A}, P) introduit par Csörgő *et al.* (1986). Sur cet espace de probabilités nous construisons, en premier lieu, une suite U_1, U_2, \dots de variables aléatoires indépendantes, de même loi uniforme sur $[0, 1]$. Pour tout $n \geq 1$, nous désignons par $U_{1,n} < \dots < U_{n,n}$, la statistique d'ordre de l'échantillon U_1, \dots, U_n , et posons $U_{0,n} = 0$ et $U_{n+1,n} = 1$. Pour tout $n \geq 1$, le processus empirique uniforme $\alpha_n(\cdot)$, et le processus de quantiles uniforme $\beta_n(\cdot)$, sont définis, respectivement, pour $0 \leq s \leq 1$, par

$$\alpha_n(s) = n^{1/2} \{G_n(s) - s\}, \quad \text{et} \quad \beta_n(s) = n^{1/2} \{s - \mathbf{U}_n(s)\}. \quad (12)$$

Dans (12), on pose, pour tout $s \in [0, 1]$,

$$G_n(s) := n^{-1} \# \{U_i \leq s : 1 \leq i \leq n\}, \quad (13)$$

et

$$\mathbf{U}_n(s) := \inf \{t : G_n(t) \geq s\} \quad \text{pour} \quad 0 < s \leq 1 \quad \text{et} \quad \mathbf{U}_n(0) = U_{1,n}. \quad (14)$$

De plus, nous supposons systématiquement, et sans perte de généralité, que ces variables aléatoires sont liées à X_1, \dots, X_n , et à la statistiques d'ordre $X_{1,n}, \dots, X_{n,n}$ correspondante, par les identités

$$X_i = Q(U_i) \quad \text{et} \quad X_{i,n} = Q(U_{i,n}) \quad \text{pour} \quad i = 1, \dots, n. \quad (15)$$

Désignons par $F_n(t) := n^{-1} \# \{X_i \leq t : 1 \leq i \leq n\}$ la fonction de répartition empirique de l'échantillon X_1, \dots, X_n . La fonction de quantile empirique associée à $F_n(\cdot)$ est alors donnée par

$$Q_n(s) := \inf\{t : F_n(t) \geq s\} \quad \text{pour } 0 < s \leq 1 \quad \text{et} \quad Q_n(0) = X_{1,n}.$$

On a pour tout $x \in \mathbb{R}$ et $0 < s < 1$

$$F_n(x) = G_n(F(x)) \quad \text{et} \quad Q_n(s) = Q(\mathbf{U}_n(s)). \quad (16)$$

L'espace de probabilités (Ω, A, P) de Csörgő *et al.* (1986) supporte, en second lieu, une suite $\{B_n(s) : 0 \leq s \leq 1\}$, $n = 1, 2, \dots$, de ponts browniens liés aux variables U_1, U_2, \dots , par les relations suivantes (voir (17) et (18) ci-dessous). Sur cet espace de probabilités, pour tout choix de ν_1 vérifiant $0 \leq \nu_1 < 1/4$, et, pour tout choix de $\gamma > 0$, le processus empirique uniforme $\{\alpha_n(s) : 0 \leq s \leq 1\}$ associé à U_1, \dots, U_n vérifie, lorsque $n \rightarrow \infty$,

$$\sup_{\frac{\gamma}{n} \leq s \leq 1 - \frac{\gamma}{n}} \frac{n^{\nu_1} |\alpha_n(s) - B_n(s)|}{(s(1-s))^{1/2-\nu_1}} = \mathcal{O}_P(1). \quad (17)$$

La version de ce résultat pour le processus uniforme de quantiles, permet d'affirmer que, pour tout choix de ν_2 vérifiant $0 \leq \nu_2 < 1/2$, et, pour tout choix de $\gamma > 0$, lorsque $n \rightarrow \infty$,

$$\sup_{\frac{\gamma}{n} \leq s \leq 1 - \frac{\gamma}{n}} \frac{n^{\nu_2} |\beta_n(s) - B_n(s)|}{(s(1-s))^{1/2-\nu_2}} = \mathcal{O}_P(1). \quad (18)$$

(On consultera, en particulier, les théorème 2.1, p.195, et théorème 3.1, p.213 de Csörgő et Horváth (1993)).

Définissons le *processus empirique uniforme bootstrappé* $\alpha_{m;n}(\cdot)$, et le *processus de quantiles uniforme bootstrappé* $\beta_{m;n}(\cdot)$, d'une façon analogue à (12). Nous donnons, ci-dessous, quelques détails sur leur construction.

Tout d'abord, nous construisons, sur l'espace de probabilités (Ω, A, P) supportant $\{U_n : n \geq 1\}$ (éventuellement élargi par produits), une suite auxiliaire, notée $\{\xi_m : m \geq 1\}$, de variables aléatoires indépendantes de même loi uniforme sur $(0, 1)$. Nous supposons que les deux suites, $\{\xi_m : m \geq 1\}$ et $\{U_n : n \geq 1\}$, sont mutuellement indépendantes. Pour tout $m \geq 1$, nous notons $0 < \xi_{1,m} \leq \dots \leq \xi_{m,m} < 1$ la statistique d'ordre de l'échantillon ξ_1, \dots, ξ_m , et posons de plus $\xi_{0,m} = 0$ et $\xi_{m+1,m} = 1$. Notons les fonctions de répartition et de quantiles empiriques correspondantes, à l'ordre $m \geq 1$, par

$$E_m(s) = m^{-1} \#\{\xi_i \leq s : 1 \leq i \leq m\} \quad \text{pour } 0 \leq s \leq 1, \quad (19)$$

et

$$K_m(s) = \inf\{t : E_m(t) \geq s\} \quad \text{pour } 0 < s \leq 1 \quad \text{et} \quad K_m(0) = \xi_{1,m}. \quad (20)$$

Dans une deuxième étape, nous posons, pour tout $n \geq 1$ et $m \geq 1$,

$$U_j^* = U_n(\xi_j) \quad \text{pour } j = 1, \dots, m. \quad (21)$$

Désignons la statistique d'ordre de l'échantillon U_1^*, \dots, U_m^* par $0 \leq U_{1,m}^* \leq \dots \leq U_{m,m}^* \leq 1$, et posons, également, $U_{0,m}^* = 0$ et $U_{m+1,m}^* = 1$. On notera que, conformément à (21), pour tout $n \geq 1$,

$$U_{j,m}^* = U_n(\xi_{j,m}) \quad \text{pour } j = 1, \dots, m. \quad (22)$$

On supposera, par la suite et sans perte de généralité que pour tout $i \geq 1$ et $1 \leq j \leq m$,

$$X_i^* = Q(U_i^*) = Q(\mathbf{U}_n(\xi_i)) \quad \text{et} \quad X_{j,m}^* = Q(U_{j,m}^*) = Q(\mathbf{U}_n(\xi_{j,m})). \quad (23)$$

On pose ensuite pour tout $t \in [0, 1]$,

$$G_{m;n}(t) := E_m(G_n(t)) = m^{-1} \#\{U_i^* \leq t : 1 \leq i \leq m\}, \quad (24)$$

et, pour tout $s \in [0, 1]$,

$$\mathbf{U}_{m;n}(s) := \mathbf{U}_n(K_m(s)) = \inf\{t : G_{m;n}(t) \geq s\}, \text{ pour } 0 < s \leq 1 \text{ et } \mathbf{U}_{m;n}(0) = U_{1,m}^*. \quad (25)$$

Notons $e_m(t) = m^{1/2}(E_m(t) - t)$, le processus empirique basé sur ξ_1, \dots, ξ_m , et désignons par $\kappa_m(s) = m^{1/2}(s - K_m(s))$, le processus de quantiles correspondant. On pose, pour tout $t \in \mathbb{R}$,

$$\alpha_{m;n}(t) = m^{1/2}\{G_{m;n}(t) - G_n(t)\} = m^{1/2}\{E_n(G_n(t)) - G_n(t)\} = e_m(G_n(t)), \quad (26)$$

et, pour tout $s \in [0, 1]$,

$$\beta_{m;n}(s) = m^{1/2}\{\mathbf{U}_n(s) - \mathbf{U}_{m;n}(s)\} = m^{1/2}\{\mathbf{U}_n(s) - \mathbf{U}_n(K_m(s))\}. \quad (27)$$

Les résultats de Csörgő et Mason (1989) (voir les théorème 2.1, p.195, et théorème 3.1, p.213 de Csörgő et Horváth (1993)), montrent qu'il est possible de choisir initialement l'espace de probabilités (Ω, A, P) , de telle sorte qu'il supporte une deuxième suite $\{B_m^*(s) : 0 \leq s \leq 1\}$, $m = 1, 2, \dots$, de ponts browniens, vérifiant les propriétés énoncées aux (28) et (29) ci-dessous. Il est également possible (voir (2.2), p.1453 dans Csörgő et Mason (1989)) de construire une suite $\{B_{m;n}^*(\cdot) : m \geq 1\}$ de ponts browniens standard tels que les propriétés suivantes soient satisfaites:

(A) Les suites $(\{U_n\}_{n \geq 1}, \{B_n\}_{n \geq 1})$ et $(\{\xi_m\}_{m \geq 1}, \{B_m^*\}_{m \geq 1})$ sont indépendantes;

(B) De même qu'en (17), pour tout choix de ν_1 vérifiant $0 \leq \nu_1 < 1/4$, et, pour tout choix de $\gamma > 0$, on a, lorsque $m \rightarrow \infty$,

$$\sup_{\frac{\gamma}{m} \leq s \leq 1 - \frac{\gamma}{m}} \frac{m^{\nu_1} |e_m(s) - B_m^*(s)|}{(s(1-s))^{1/2-\nu_1}} = \mathcal{O}_P(1). \quad (28)$$

De même qu'en (18), pour tout choix de ν_2 vérifiant $0 \leq \nu_2 < 1/2$, et, pour tout choix de $\gamma > 0$, on a, lorsque $m \rightarrow \infty$,

$$\sup_{\frac{\gamma}{m} \leq s \leq 1 - \frac{\gamma}{m}} \frac{m^{\nu_2} |\kappa_m(s) - B_m^*(s)|}{(s(1-s))^{1/2-\nu_2}} = \mathcal{O}_P(1). \quad (29)$$

Compte tenu de ces résultats, sous l'hypothèse que $m = m_n \rightarrow \infty$, on a alors le résultat suivant (voir le théorème 2.1, p.1454 de Csörgő et Mason (1989)). Pour tout $0 \leq \nu < 1/4$ et tout $\gamma > 0$, lorsque $n \rightarrow \infty$,

$$\sup_{U_{1,n} \leq s \leq U_{n,n}} \frac{|\alpha_{m,n}(s) - B_{m;n}^*(s)|}{(s(1-s))^{1/2-\nu}} = \mathcal{O}_P((n \wedge m)^{-\nu}). \quad (30)$$

Si de plus la suite $m = m_n$ vérifie (10), alors, pour tout $0 \leq \nu < 1/4$ et tout $\gamma > 0$, on a, lorsque $n \rightarrow \infty$,

$$\sup_{\frac{\gamma}{m} \leq s \leq 1 - \frac{\gamma}{m}} \frac{|\beta_{m,n}(s) - B_{m;n}^*(s)|}{(s(1-s))^{1/2-\nu}} = \mathcal{O}_P(n^{-\nu}). \quad (31)$$

Nous aurons besoin des notations additionnelles suivantes

$$\mu_n(k_n) := n \int_{1-\frac{k_n}{n}}^1 Q(u) du, \quad \sigma^2(s) := \int_{1-s}^1 \int_{1-s}^1 (\min(u, v) - uv) dQ(u)dQ(v) \text{ pour } 0 < s < 1,$$

$$m_n(k_n) := \int_{1-\frac{k_n}{n}}^1 Q_n(u) du, \quad S_n^2(k_n) := \int_{1-\frac{k_n}{n}}^1 Q_n(u)^2 du - m_n(k_n)^2.$$

Introduisons en plus, pour $0 < \gamma < \infty$, la famille de fonctions de répartition [fdr] suivante.

$$\mathcal{E}_\gamma := \left\{ G \text{ fdr} : \lim_{x \rightarrow \infty} (1 - G(x))^{-1} \int_x^\infty (1 - G(y)) dy = \gamma \right\}. \quad (32)$$

Théorème 2.1 *Supposons que $F(\cdot)$ soit dans la classe \mathcal{E}_c . Supposons de plus que k_n vérifie (4), que $\ell(n)$ vérifie (7) et que la condition (10) soit satisfaite. Alors il existe deux suites de ponts browniens, $\{B_{m;n}^* : m \geq 1, n \geq 1\}$ et $\{B_n : m \geq 1, n \geq 1\}$ indépendantes, telles que, lorsque $n \rightarrow \infty$,*

$$\begin{aligned} & \frac{\sqrt{\ell(n)}}{S_n(k_n)} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} X_{m-i+1,m}^* - \frac{1}{k_n} \sum_{i=1}^{k_n} X_{n-i+1,n} - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \\ &= -\frac{\sqrt{\frac{m}{\ell(n)}}}{\sigma \left(\frac{k_n}{n} \right)} \int_{1-\ell(n)/m}^1 B_{m;n}^*(u) dQ(u) + \frac{\sqrt{\frac{\ell(n)}{k_n}}}{\sigma \left(\frac{k_n}{n} \right)} \sqrt{\frac{n}{k_n}} \int_{1-k_n/n}^1 B_n(s) dQ(s) + o_P(1). \quad (33) \end{aligned}$$

Démonstration. Rappelons tout d'abord que, sous les hypothèses de ce théorème, la fonction de quantiles $Q(\cdot)$ de X_1, X_2, \dots , vérifie

$$Q(1-t) = -c \log t + \log L^*(1/t), \quad (34)$$

où $L^*(\cdot)$ est à variation lente à l'infini. Compte tenu de la définition (9) de $Q_{m;n}(\cdot)$, nous pouvons écrire

$$\begin{aligned} & \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} X_{m-i+1,m}^* - \frac{1}{k_n} \sum_{i=1}^{k_n} X_{n-i+1,n} = \frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 Q_{m;n}(s) ds - \frac{n}{k_n} \int_{1-k_n/n}^1 Q_n(s) ds \\ &= \frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 Q(s) ds - \frac{n}{k_n} \int_{1-k_n/n}^1 Q(s) ds + \frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 \{Q_{m;n}(s) - Q(s)\} ds \\ & \quad - \frac{n}{k_n} \int_{1-k_n/n}^1 \{Q_n(s) - Q(s)\} ds \\ &= \frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 (1-s) dQ(s) - \frac{n}{k_n} \int_{1-k_n/n}^1 (1-s) dQ(s) + Q\left(1 - \frac{k_n}{n}\right) - Q\left(1 - \frac{\ell(n)}{m}\right) + o_P(1). \end{aligned}$$

En utilisant le lemme 3.1, p.552 de Csörgő et Mason (1985), nous obtenons, sous les hypothèses du théorème 2.1, que lorsque $n \rightarrow \infty$

$$\frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 (1-s) dQ(s) - \frac{n}{k_n} \int_{1-k_n/n}^1 (1-s) dQ(s) \rightarrow 0.$$

Compte tenu de (34), on obtient $Q\left(1 - \frac{k_n}{n}\right) - Q\left(1 - \frac{\ell(n)}{m}\right) = c \log\left(\frac{n\ell(n)}{mk_n}\right) + o(1)$. Après intégration par parties, nous obtenons les relations

$$\begin{aligned} & \sqrt{\ell(n)} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} X_{m-i+1,m}^* - \frac{1}{k_n} \sum_{i=1}^{k_n} X_{n-i+1,n} - c \log\left(\frac{n\ell(n)}{mk_n}\right) \right\} \\ &= -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 \alpha_{m;n}(s) dQ(s) + \frac{m}{\sqrt{\ell(n)}} \int_{\xi_{m-\ell(n),m}}^{1-\ell(n)/m} (1-E_m(s) - (1-s)) dQ(s) \\ & \quad + \frac{m}{\sqrt{\ell(n)}} \int_{\xi_{m-\ell(n),m}}^{1-\ell(n)/m} (1-s - \ell(n)/m) dQ(s) \\ & \quad + \sqrt{\frac{\ell(n)}{k_n}} \sqrt{\frac{n}{k_n}} \int_{1-k_n/n}^1 \alpha_n(s) dQ(s) - \sqrt{\frac{\ell(n)}{k_n}} \frac{n}{\sqrt{k_n}} \int_{U_{n-k_n,n}}^{1-k_n/n} (1-G_n(s) - (1-s)) dQ(s) \\ & \quad - \sqrt{\frac{\ell(n)}{k_n}} \frac{n}{\sqrt{k_n}} \int_{U_{n-k_n,n}}^{1-\frac{k_n}{n}} \left(1-s - \frac{k_n}{n}\right) dQ(s) := \Delta_{1,n}^* + \Delta_{2,n}^* + \Delta_{3,n}^* + \Delta_{1,n} + \Delta_{2,n} + \Delta_{3,n}. \end{aligned}$$

Nous cherchons, maintenant, à montrer que, lorsque $n \rightarrow \infty$,

$$\Delta_{1,n}^* + \Delta_{2,n}^* + \Delta_{3,n}^* = -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 B_{m;n}^*(u) dQ(u) + o_P(1).$$

Pour $0 < d < \infty$, définissons $T_{m;n}^{(1)}(d)$ et $T_{m;n}^{(2)}(d)$, respectivement, par

$$\begin{aligned} T_{m;n}^{(1)}(d) &:= \frac{m}{\sqrt{\ell(n)}} \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} |1-E_m(s) - (1-s)| dQ(s) \\ &= \frac{m}{\sqrt{\ell(n)}} \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} |s-E_m(s)| dQ(s), \\ T_{m;n}^{(2)}(d) &:= \frac{m}{\sqrt{\ell(n)}} \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} \left|1-s - \frac{\ell}{m}\right| dQ(s). \end{aligned}$$

Il est bien connu que, si k_n vérifie la condition (4), étant telle que $k_n \rightarrow \infty$ et $n^{-1}k_n \rightarrow 0$, alors (voir, par exemple, Balkema et de Haan (1975)), on a la loi limite

$$\frac{n}{\sqrt{k_n}} \left\{ U_{n-k_n,n} - \left(1 - \frac{k_n}{n}\right) \right\} \xrightarrow{d} N(0,1) \quad \text{lorsque } n \rightarrow \infty, \quad (35)$$

Notons au passage qu'on a l'inégalité

$$\begin{aligned} \lim_{d \rightarrow \infty} \liminf_{n \rightarrow \infty} P \left(|\Delta_{2,n}^*| + |\Delta_{3,n}^*| \leq T_{m;n}^{(1)}(d) + T_{m;n}^{(2)}(d) \right) \\ \geq \lim_{d \rightarrow \infty} \liminf_{n \rightarrow \infty} P \left(\frac{m}{\sqrt{\ell(n)}} \left| \xi_{m-\ell(n),m} - (1-\ell(n)/m) \right| \leq d \right), \end{aligned} \quad (36)$$

où la borne inférieure est égale à 1, comme conséquence directe de la relation (35). Pour montrer que

$$\Delta_{i,n}^* \xrightarrow{P} 0, \text{ pour } i = 2, 3, \quad (37)$$

il suffit de montrer que

$$T_{m;n}^{(i)}(d) \xrightarrow{P} 0, \text{ pour } i = 1, 2. \quad (38)$$

Traisons le cas $i = 1$. Nous remarquons que

$$\begin{aligned} E(T_{m;n}^{(1)}(d)) &= E \left(\frac{m}{\sqrt{\ell(n)}} \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} |s - E_m(s)| dQ(s) \right) \\ &\leq \sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} (1-s)^{1/2} dQ(s) \leq 2 \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} dQ(s). \end{aligned} \quad (39)$$

Pour n assez grand, et $1 < \lambda < \infty$, nous obtenons que

$$\begin{aligned} E(T_{m;n}^{(1)}(d)) &\leq 2 \left\{ Q \left(1 - \frac{\ell(n)}{\lambda m} \right) - Q \left(1 - \frac{\lambda \ell(n)}{m} \right) \right\} \\ &= 4c \log_+ \lambda + \log \left\{ L^* \left(\frac{\lambda m}{\ell(n)} \right) / L^* \left(\frac{m}{\lambda \ell(n)} \right) \right\}. \end{aligned} \quad (40)$$

Du fait que $L^*(\cdot)$ est une fonction à variation lente à l'infini, il est facile de voir que

$$\lim_{n \rightarrow \infty} \log \left\{ L^* \left(\frac{\lambda m}{\ell(n)} \right) / L^* \left(\frac{m}{\lambda \ell(n)} \right) \right\} = 0.$$

Ainsi, nous obtenons que, pour tout $0 < \lambda < \infty$,

$$\limsup_{n \rightarrow \infty} E(T_{m;n}^{(1)}(d)) \leq 4c \log \lambda. \quad (41)$$

La relation (41) implique, lorsque $\lambda \rightarrow 1$,

$$\lim_{n \rightarrow \infty} E(T_{m;n}^{(1)}(d)) \rightarrow 0. \quad (42)$$

Une simple application de l'inégalité de Markov permet d'obtenir la convergence

$$\Delta_{2,n}^* \xrightarrow{P} 0. \quad (43)$$

Considérons maintenant le cas $i = 2$. On a

$$E(T_{m;n}^{(2)}(d)) \leq 2d \int_{1-\ell(n)/m-d\frac{\ell^{1/2}}{m}}^{1-\ell(n)/m+d\frac{\ell^{1/2}}{m}} dQ(s). \quad (44)$$

Le terme de droite de (44) est égal à d fois la borne supérieure de $E(T_{m;n}^{(1)}(d))$. Nous en déduisons que

$$T_{m;n}^{(2)}(d) \xrightarrow{P} 0, \text{ lorsque } n \rightarrow \infty, \quad (45)$$

ce qui implique que

$$\Delta_{3,n}^* \xrightarrow{P} 0. \quad (46)$$

Traisons à présent le terme $\Delta_{1,n}^*$. Celui-ci admet la décomposition

$$\begin{aligned}\Delta_{1,n}^* &= -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 \alpha_{m;n}(s) dQ(s) = -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 B_{m;n}^*(s) dQ(s) \\ &+ \sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^{1-1/m} (B_{m;n}^*(s) - \alpha_{m;n}(s)) dQ(s) + \sqrt{\frac{m}{\ell(n)}} \int_{1-1/m}^1 (B_{m;n}^*(s) - \alpha_{m;n}(s)) dQ(s) \\ &:= -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 B_{m;n}^*(s) dQ(s) + \nabla_{1,n}^* + \nabla_{2,n}^*.\end{aligned}$$

En combinant le théorème 2.1, p.1454 de Csörgő et Mason (1989) (voir (30)) et le fait que $0 \leq \nu_1 < \frac{1}{4}$, on peut en déduire que

$$\begin{aligned}|\nabla_{1,n}^*| &\leq \sup_{0 \leq s \leq 1-1/m} \frac{|\alpha_{m;n}(s) - B_{m;n}^*(s)|}{(1-s)^{1/2-\nu_1}} \sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 (1-s)^{1/2-\nu_1} dQ(s) \\ &= \mathcal{O}_P(1) \left(\frac{m}{\ell(n)}\right)^{1/2-\nu_1} \ell(n)^{-\nu_1} \int_{1-\ell(n)/m}^1 (1-s)^{1/2-\nu_1} dQ(s).\end{aligned}$$

D'après le lemme 3.2, p.552 de Csörgő et Mason (1985), il est clair que

$$\left(\frac{m}{\ell(n)}\right)^{1/2-\nu_1} \int_{1-\ell(n)/m}^1 (1-s)^{1/2-\nu_1} dQ(s) = \frac{c}{1/2-\nu_1}.$$

Il ressort que, lorsque $n \rightarrow \infty$,

$$\nabla_{1,n}^* = \mathcal{O}_P(\ell(n)^{-\nu_1}) = o_P(1). \quad (47)$$

Étudions le terme $\nabla_{2,n}^*$. Nous avons

$$E|\nabla_{2,n}^*| \leq \sqrt{\frac{m}{\ell(n)}} E \int_{1-1/m}^1 (|B_{m;n}^*(s)| + |\alpha_{m;n}(s)|) dQ(s) \leq 2\sqrt{\frac{m}{\ell(n)}} \int_{1-1/m}^1 (1-s)^{1/2} dQ(s).$$

D'après le lemme 3.2, p.552 de Csörgő et Mason (1985), on en déduit que

$$\lim_{n \rightarrow \infty} \sqrt{m} \int_{1-1/m}^1 (1-s)^{1/2} dQ(s) = 2c. \quad (48)$$

Ainsi en tenant compte de (48), nous obtenons $E(\nabla_{2,n}^*) = \mathcal{O}_P(\ell(n)^{-1/2})$, et en appliquant l'inégalité de Markov on obtient immédiatement

$$\nabla_{2,n}^* \xrightarrow{P} 0, \text{ lorsque } n \rightarrow \infty. \quad (49)$$

En combinant (43), (46), (47) et (49), on peut constater que

$$\frac{m}{\ell(n)} \int_{1-\ell(n)/m}^1 \{Q_{m;n}(s) - Q(s)\} ds = -\sqrt{\frac{m}{\ell(n)}} \int_{1-\ell(n)/m}^1 B_{m;n}^*(s) dQ(s) + o_P(1). \quad (50)$$

En utilisant des techniques similaires, nous obtenons également le résultat suivant

$$\sqrt{\frac{\ell(n)}{k_n}} \frac{n}{k_n} \int_{1-k_n/n}^1 \{Q_n(s) - Q(s)\} ds = -\sqrt{\frac{\ell(n)}{k_n}} \sqrt{\frac{n}{k_n}} \int_{1-k_n/n}^1 B_n(s) dQ(s) + o_P(1). \quad (51)$$

Nous allons montrer que la convergence en probabilité suivante a lieu, lorsque n tend vers l'infini

$$S_n(k_n) / \sigma \left(\frac{k_n}{n}\right) \xrightarrow{P} 1. \quad (52)$$

Nous allons procéder comme dans la preuve du lemme 3.3, p.552 de Csörgő et Mason (1985). En combinant les lemmes 3.1 et 3.3, p.552 de Csörgő et Mason (1985), on obtient

$$\sigma^2(s) / (2sC^2(s)) \rightarrow 1 \text{ lorsque } s \downarrow 0. \quad (53)$$

En utilisant (53), il suffit de montrer que

$$\frac{S_n^2(k_n)}{2 \frac{k_n}{n} C^2\left(\frac{k_n}{n}\right)} \xrightarrow{P} 1, \text{ lorsque } n \rightarrow \infty.$$

Remarquons qu'on peut décomposer $S_n^2(k_n)$ de la manière suivante

$$\begin{aligned} S_n^2(k_n) &= \int_{1-k_n/n}^1 v(1-v) \left\{ \frac{1}{1-v} \int_v^1 (1-u) dQ_n(u) \right\} dQ_n(v) \\ &+ \int_{1-k_n/n}^1 u(1-u) \left\{ \frac{1}{1-u} \int_{u+}^1 (1-v) dQ_n(v) \right\} dQ_n(u) := S_{1,n} + S_{2,n}. \end{aligned}$$

On va étudier uniquement $S_{1,n}$. L'étude de $S_{2,n}$ se fait selon le même procédé. Nous remarquons que

$$\begin{aligned} \frac{n}{k_n} \left(1 - \frac{k_n}{n}\right) \int_{1-k_n/n}^1 (1-v) dQ_n(v) &\inf_{1-\frac{k_n}{n} \leq v < 1} \frac{1}{1-v} \int_v^1 (1-u) dQ_n(u) \\ &\leq \frac{n}{k_n} S_{1,n} \leq \frac{n}{k_n} \int_{1-k_n/n}^1 (1-v) dQ_n(v) \sup_{1-\frac{k_n}{n} \leq v < 1} \frac{1}{1-v} \int_v^1 (1-u) dQ_n(u). \end{aligned}$$

Ce dernier terme est équivalent à $\left\{ \frac{n}{k_n} \int_{1-k_n/n}^1 (1-v) dQ_n(v) \right\}^2$. De plus $S_{1,n}$ est majorée par $\frac{k_n}{n} C^2\left(\frac{k_n}{n}\right)$, lorsque $n \rightarrow \infty$. D'autre part on peut minorer $S_{1,n}$ par $\left\{ \frac{k_n}{n} C^2\left(\frac{k_n}{n}\right) \right\}$, lorsque $n \rightarrow \infty$. En combinant (50), (51) et (52), on en déduit le résultat (33). ■

Corollaire 2.2 Supposons que $F(\cdot)$ vérifie la condition (1). Supposons de plus que les suites $k_n, \ell(n)$ et m vérifient les conditions (4), (7), et (10). Alors les deux suites de ponts browniens, $\{B_{m;n}^* : m \geq 1, n \geq 1\}$ et $\{B_n : n \geq 1\}$ de (33), sont telles que, lorsque $n \rightarrow \infty$,

$$\begin{aligned} &\sqrt{\ell(n)} \{S_n(k_n)\}^{-1} \left\{ \frac{1}{\ell(n)} \sum_{i=1}^{\ell(n)} \log_+ X_{m-i+1,m}^*(n) - \frac{1}{k_n} \sum_{i=1}^{k_n} \log_+ X_{n-i+1,n} - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \\ &= -\frac{\sqrt{\frac{m}{\ell(n)}}}{\sigma\left(\frac{k_n}{n}\right)} \int_{1-\ell(n)/m}^1 B_{m;n}^*(u) dQ(u) + \frac{\sqrt{\frac{\ell(n)}{k_n}}}{\sigma\left(\frac{k_n}{n}\right)} \sqrt{\frac{n}{k_n}} \int_{1-k_n/n}^1 B_n(s) dQ(s) + o_P(1) \\ &:= Z_n + o_P(1). \end{aligned} \quad (54)$$

Démonstration. Pour montrer (54), il suffit d'utiliser la *Fact* 1.2, p.547 de Csörgő et Mason (1985), i.e., si une variable aléatoire X a une fonction de répartition $F(\cdot)$ vérifiant la condition (1), alors la fonction de répartition de $\log X$ appartient à la classe \mathcal{E}_c . ■

Théorème 2.3 Supposons que $F(\cdot)$ vérifie la condition (1). Supposons, de plus, que les suites $k_n, \ell(n)$ et m vérifient les conditions (4), (7) et (10). Alors les deux suites de ponts browniens, $\{B_{m;n}^* : m \geq 1, n \geq 1\}$ et $\{B_n : n \geq 1\}$ de (54), sont telles que, lorsque $n \rightarrow \infty$,

$$\begin{aligned} &\sqrt{\ell(n)} \left\{ \log_+ (X_{m-\ell(n),m}^*(n)) - \log_+ (X_{n-k_n,n}) - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \\ &= -c \sqrt{\ell(n)/m} B_{m;n}^*(1 - \ell(n)/m) + c \sqrt{\frac{\ell(n)}{n}} B_n \left(1 - \frac{k_n}{n} \right) + o_P(1) := Y_n + o_P(1) \end{aligned} \quad (55)$$

et

$$\sqrt{\ell(n)} \left\{ \hat{c}_{m;\ell(n)}^* - \hat{c}_n \right\} = Z_n - Y_n + o_P(1). \quad (56)$$

Démonstration. Pour montrer (55) on s'inspire de la démonstration du théorème 2.4 de Csörgő et Mason (1985). Il est bien connu (voir de Haan (1970)), que $F(\cdot)$ vérifie (1), si et seulement s'il existe une fonction $L^*(\cdot)$, à variation lente (à l'infini), telle que

$$Q(1-x^{-1}) = x^c L^*(x) \text{ pour } x > 1. \quad (57)$$

Par le théorème de Karamata (voir Seneta (1976)), ceci équivaut à l'existence de fonctions mesurables $a(x)$ et $b(x)$, vérifiant $a(x) \rightarrow a_0 \in \mathbb{R}_+^*$ et $b(x) \rightarrow 0$ quand $x \rightarrow \infty$, et telles que

$$L^*(x) = a(x) \exp \left(\int_1^x b(u) u^{-1} du \right), \quad \text{pour } x \geq 1.$$

Nous avons la décomposition suivante

$$\begin{aligned} & \sqrt{\ell(n)} \left\{ \log_+ (X_{m-\ell(n),m}^*) - \log_+ (X_{n-k_n,n}) - c \log \left(\frac{n\ell(n)}{mk_n} \right) \right\} \\ &= \sqrt{\ell(n)} \left\{ \log_+ Q(\xi_{m-\ell(n),m}) - \log_+ Q(1 - \ell(n)/m) \right\} \\ & - \sqrt{\frac{\ell(n)}{k_n}} \sqrt{k_n} \left\{ \log_+ Q(U_{n-k_n,n}) - \log_+ Q \left(1 - \frac{k_n}{n} \right) \right\} + o(1). \end{aligned}$$

Commençons par étudier le terme suivant

$$\begin{aligned} & \sqrt{\ell(n)} \left\{ \log_+ Q(\xi_{m-\ell(n),m}) - \log_+ Q(1 - \ell(n)/m) \right\} \\ &= -c\sqrt{\ell(n)} \left\{ \log_+ (1 - \xi_{m-\ell(n),m}) - \log_+ (1 - (1 - \ell(n)/m)) \right\} \\ & + \sqrt{\ell(n)} \int_{(1-\xi_{m-\ell(n),m})^{-1}}^{m/\ell(n)} b(u) u^{-1} du + o_P(1) := \Theta_{1,n}^m + \Theta_{2,n}^m. \end{aligned}$$

Concernant $\Theta_{1,n}^m$, en utilisant un développement limité de $\log_+(1-s)$, on obtient que

$$\Theta_{1,n}^m = -c \frac{m}{\ell(n)} \{1 - \ell(n)/m - \xi_{m-\ell(n),m}\} + c \frac{\{\ell(n)\}^{1/2}}{2(1-\rho_n)^2} \{1 - \ell(n)/m - \xi_{m-\ell(n),m}\}^2,$$

où ρ_n est compris entre $1 - \ell(n)/m$ et $\xi_{m-\ell(n),m}$. D'une part, d'après la relation (35), on obtient

$$\xi_{m-\ell(n),m} = 1 - \ell(n)/m + \mathcal{O}_P \left(\frac{\sqrt{\ell(n)}}{m} \right), \quad (58)$$

par conséquent, on a

$$\frac{c\{\ell(n)\}^{1/2}}{2(1-\rho_n)^2} \{1 - \ell(n)/m - \xi_{m-\ell(n),m}\}^2 = \mathcal{O}_P(\{\ell(n)\}^{-1/2}) = o_P(1).$$

Et d'autre part, en utilisant le théorème 2.1, p.1454 de Csörgő et Mason (1989) (voir (31)), et, en remplaçant s par $(1 - \ell(n)/m)$, pour $0 \leq \nu < 1/4$, on obtient que

$$\beta_{m,n}(1 - \ell(n)/m) = B_{m;n}^*(1 - \ell(n)/m) + (\ell(n)/m)^{\frac{1}{2}-\nu} \mathcal{O}_P(m^{-\nu}), \quad (59)$$

ce qui implique que

$$\frac{m}{\sqrt{\ell(n)}} \{1 - \ell(n)/m - \xi_{m-\ell(n),m}\} = \left(\frac{m}{\ell(n)} \right)^{1/2} B_{m;n}^*(1 - \ell(n)/m) + \mathcal{O}_P(\ell(n)^{-\nu}). \quad (60)$$

Concernant $\Theta_{2,n}^m$, il est facile de voir que

$$|\Theta_{2,n}^m| \leq \{\ell(n)\}^{1/2} |\log_+(1 - \xi_{m-\ell(n),m}) - \log_+(1 - \ell(n)/m)| \mathfrak{R}_n,$$

où

$$\mathfrak{R}_n = \sup \left\{ b(u) : u \leq \left(\frac{m}{\ell(n)} \right) \vee (1 - \xi_{m-\ell(n),m})^{-1} \right\}. \quad (61)$$

Or, lorsque $n \rightarrow \infty$, $(1 - \xi_{m-\ell(n),m})^{-1} \vee \left(\frac{m}{\ell(n)} \right) \xrightarrow{p.s.} \infty$, ceci nous donne que $b(u) \rightarrow 0$, $u \rightarrow \infty$. On obtient donc, lorsque $n \rightarrow \infty$

$$\mathfrak{R}_n \rightarrow 0 \text{ p.s.} \quad (62)$$

En utilisant le résultat (59), on constate bien que

$$\sqrt{\ell(n)} |\log_+(1 - U_{m-\ell(n),m}^*) - \log_+(1 - U_{m-\ell(n),m})| = \mathcal{O}_P(1). \quad (63)$$

En combinant (62) et (63), il ressort que

$$\Theta_{2,n}^m = o_P(1), \text{ lorsque } n \rightarrow \infty. \quad (64)$$

En utilisant (59) et (64), il découle que

$$\sqrt{\ell(n)} \{ \log_+ Q(\xi_{m-\ell(n),m}) - Q(1 - \ell(n)/m) \} = -c\sqrt{\ell(n)/m} B_{m;n}^* (1 - \ell(n)/m) + o_P(1). \quad (65)$$

Traitions à présent le dernier terme de (58). Notons ici que nous utilisons les mêmes techniques que précédemment. En effet, en tenant compte de la condition (10) et le Theorem 2.1, p.195 de Csörgő et Horváth (1993), nous obtenons un résultat analogue à (65), donné par la relation suivante

$$\begin{aligned} \sqrt{\frac{\ell(n)}{k_n}} \sqrt{k_n} \left\{ \log_+ Q(U_{n-k_n,n}) - \log_+ Q\left(1 - \frac{k_n}{n}\right) \right\} &= -c\sqrt{\frac{\ell(n)}{k_n}} \sqrt{\frac{k_n}{n}} B_n \left(1 - \frac{k_n}{n}\right) + o_P(1) \\ &= -c\sqrt{\frac{\ell(n)}{n}} B_n \left(1 - \frac{k_n}{n}\right) + o_P(1). \end{aligned} \quad (66)$$

En regroupant (65) et (66), on déduit (55).

Il suffit de remarquer que (56), n'est que la différence entre (54) et (55), ce qui achève la démonstration du théorème 2.3. ■

Démonstration du corollaire 1.1. Nous nous inspirons de la démonstration de Csörgő et Mason (1985), p.555. Notons que pour tout $m \geq 1$, la variable $-\sqrt{\frac{m}{\ell(n)}} \int_{1-(\ell(n)/m)}^1 B_{m;n}^*(s) dQ(s)$ suit une loi normale centrée et de variance égale à $\frac{m}{\ell(n)} \sigma^2 (\ell(n)/m)$, qui converge vers $2c^2$ par le lemme 3.3 de Csörgő et Mason (1985). Il en découle que, lorsque $n \rightarrow \infty$,

$$-\sqrt{\frac{m}{\ell(n)}} \int_{1-(\ell(n)/m)}^1 B_{m;n}^*(s) dQ(s) \xrightarrow{d} N(0, 2c^2). \quad (67)$$

Nous avons, lorsque $n \rightarrow \infty$

$$-\sqrt{\frac{n}{k_n}} \int_{1-(k_n/n)}^1 B_n(s) dQ(s) \xrightarrow{d} N(0, 2c^2). \quad (68)$$

En combinant (67) et (68), on achève la démonstration. ■

Démonstration du corollaire 1.2. Nous nous inspirons de la démonstration de Csörgő et Mason (1985), p.557. La preuve du corollaire 1.2 suit exactement les mêmes lignes que la précédente. Notons que la variable

$$c \left(\frac{n}{k_n}\right)^{1/2} B_n \left(1 - \frac{k_n}{n}\right) - \sqrt{\frac{n}{k_n}} \int_{1-(k_n/n)}^1 B_n(s) dQ(s) \quad (69)$$

suit une loi normale centrée et de variance qui converge vers c^2 par le lemme 3.1 et 3.3, p.552 de Csörgő et Mason (1985). Observons aussi que

$$c \left(\frac{m}{\ell(n)}\right)^{1/2} B_{m;n} (1 - \ell(n)/m) - \sqrt{\frac{m}{\ell(n)}} \int_{1-(\ell(n)/m)}^1 B_{m;n}^*(s) dQ(s) \quad (70)$$

est une variable gaussienne de moyenne zéro et de variance qui converge vers c^2 par le lemme 3.1 et 3.3, p.552 de Csörgő et Mason (1985). En combinant (69) et (70) on achève la démonstration. ■

L'auteur remercie sincèrement le Professeur Paul Deheuvels d'avoir contribué par ses suggestions constructives à la qualité de cet article.

References

- Arcones, M. A. et Giné, E. (1989). The bootstrap of the mean with arbitrary bootstrap sample size. *Ann. Inst. H. Poincaré Probab. Statist.*, **25**(4), 457–481.
- Balkema, A. A. et de Haan, L. (1975). Limit laws for order statistics. In *Limit theorems of probability theory (Colloq., Keszthely, 1974)*, pages 17–22. Colloq. Math. Soc. János Bolyai, Vol. 11. North-Holland, Amsterdam.
- Csörgő, M. et Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. John Wiley & Sons, New York.
- Csörgő, M., Csörgő, S., Horváth, L., et Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.*, **14**(1), 31–85.
- Csörgő, S. et Mason, D. M. (1985). Central limit theorems for sums of extreme values. *Math. Proc. Cambridge Philos. Soc.*, **98**(3), 547–558.
- Csörgő, S. et Mason, D. M. (1989). Bootstrapping empirical functions. *Ann. Statist.*, **17**(4), 1447–1471.
- Csörgő, S., Deheuvels, P., et Mason, D. (1985). Kernel estimates of the tail index of a distribution. *Ann. Statist.*, **13**(3), 1050–1077.
- Davis, R. et Resnick, S. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.*, **12**(4), 1467–1487.
- de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*, volume 32 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam.
- Deheuvels, P., Haeusler, E., et Mason, D. M. (1988). Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.*, **104**(2), 371–381.
- Deheuvels, P., Mason, D. M., et Shorack, G. R. (1993). Some results on the influence of extremes on the bootstrap. *Ann. Inst. H. Poincaré Probab. Statist.*, **29**(1), 83–103.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.*, **7**(1), 1–26.
- Haeusler, E. et Teugels, J. L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.*, **13**(2), 743–756.
- Hall, P. (1990). Asymptotic properties of the bootstrap for heavy-tailed distributions. *Ann. Probab.*, **18**(3), 1342–1360.
- Hall, P. et Welsh, A. H. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.*, **13**(1), 331–341.
- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, **3**(5), 1163–1174.
- Mason, D. M. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.*, **10**(3), 754–764.
- Seneta, E. (1976). *Regularly varying functions*. Springer-Verlag, Berlin. Lecture Notes in Mathematics, Vol. 508.