Z. Wahrscheinlichkeitstheorie verw. Geb. 24, 309 – 324 (1972)

Linéarité asymptotique d'une statistique de rang

André Antille

Introduction

On considère une suite $X_1, X_2, \ldots$ de variables aléatoires indépendantes de même fonction de répartition $F(x-\theta)$ où $F(x)$ est symétrique, absolument continue de densité $f(x)$ et où θ est un paramètre inconnu à estimer.

Désignons par T l'estimateur de Hodges Lehmann (médiane empirique des moyennes arithmétiques des couples (X_i, X_j), $i<j$).

En supposant

$$\int f^2(x)\,dx<\infty.$$

Lehmann [1] a démontré le résultat suivant:

$$\frac{n^{\frac{1}{2}}(T_2-T_1)}{2\,a_\alpha}=n^{\frac{1}{2}}\,S(T)\xrightarrow{P}\sigma_T.$$

σ_T désigne la variance asymptotique de T.

a_α est tel que $\Phi(-a_\alpha)=\alpha$ avec $\Phi(x)$ égale à la fonction de répartition de la loi normale $N(0,1)$.

T_2 et T_1 sont définis de la manière suivante:

$$n^{-\frac{3}{2}}\sqrt{12}\sum_{i<j}[I(X_i+X_j>2\,T_1)-\tfrac{1}{2}]=a_\alpha$$

$$n^{-\frac{3}{2}}\sqrt{12}\sum_{i<j}[I(X_i+X_j>2\,T_2)-\tfrac{1}{2}]=-a_\alpha$$

($I(A)$ désigne la fonction indicatrice de A).

Ces deux dernières relations ne sont pas à prendre dans le sens d'égalités strictes.

Par exemple la première relation signifie que T_1 est la valeur pour laquelle le membre de gauche passe de $>a_\alpha$ à $\leqq a_\alpha$.

La statistique

$$T_n(0,t)=n^{-\frac{3}{2}}\sum_{i<j}I(X_i+X_j\leqq 2\,t\,n^{-\frac{1}{2}})-n^{-\frac{3}{2}}\sum_{i<j}I(X_i+X_j\leqq 0)$$

joue un rôle important dans la démonstration du résultat obtenu par Lehmann.

En particulier celui-ci démontre d'une manière élémentaire que

$$T_n(0,t)=t\cdot\int f^2(x)\,dx+R_n(t)$$

où $R_n(t)\xrightarrow{P}0$ uniformément pour $0\leqq t\leqq M$. (M est un nombre positif quelconque.)

L'étude de la rapidité avec laquelle $n^{\frac{1}{2}} S(T)$ tend vers sa limite a donné naissance au présent travail.

Un moyen d'en connaître plus sur les propriétés asymptotiques de $n^{\frac{1}{2}} S(T)$ est d'en savoir plus sur la statistique $T_n(0, t)$. Par exemple est-il vrai que

$$\sup_{|t| \leqq M} |T_n(0, t) - t \cdot \int f^2(x)\, dx| \xrightarrow[n\to\infty]{P} 0?$$

Un travail de Jana Jurečková [2] rend plausible une telle conjecture.

En ce qui concerne la rapidité de convergence de $n^{\frac{1}{2}} S(T)$, Huber [3] a émis l'hypothèse suivante:

$$n^{\frac{1}{2}} \left(\frac{n^{\frac{1}{2}} S(T)}{\sigma_T} - 1 \right)$$

est asymptotiquement normale $N(0, d^2)$ avec

$$d^2 = 4\left[\int f^3(x)\, dx - \left(\int f^2(x)\, dx\right)^2\right] \cdot \frac{1}{\left(\int f^2(x)\, dx\right)^2}.$$

La motivation de cette hypothèse conduit naturellement à étudier le processus

$$Y_n(t) = n^{\frac{1}{2}} \left(T_n(0, t) - t \cdot \int f^2(x)\, dx\right).$$

L'étude de ce processus constitue la partie principale de ce travail.

Le résultat obtenu est le suivant:

Le processus $(Y_n(t))_{t \in [01]}$ est, sous certaines hypothèses, asymptotiquement linéaire, ceci dans le sens suivant: $(Y_n(t))_{t \in [01]}$ converge faiblement vers le processus $(tZ)_{t \in [01]}$ où Z est une variable normale $N(0, c^2)$, avec

$$c^2 = 4\left[\int f^3(x)\, dx - \left(\int f^2(x)\, dx\right)^2\right].$$

Les processus étant à valeurs dans l'espace $D[01]$ des fonctions réelles définies sur $[01]$, continues à droite et possédant des limites à gauche, la convergence faible signifie donc que les mesures v_n induites sur la tribu de Borel de $D[01]$ convergent faiblement vers la mesure v induite par tZ.

Le résultat se généralise sans peine au cas où t varie de $-M$ à $+M$. L'espace $D[01]$ est dans ce cas remplacé par l'espace $D[-M, +M]$.

Le travail est divisé en 3 chapitres.

Le premier chapitre expose les théorèmes connus relatifs à la convergence faible de mesures dans $D[01]$ qui seront utilisés dans la suite.

Le deuxième chapitre est consacré à la démonstration de la linéarité asymptotique.

Le troisième établit le résultat pressenti par Peter J. Huber.

Chapitre I

I.1. L'espace $D[01]$

Définition. $D[01]$ est l'espace des fonctions réelles x définies sur $[01]$ et telles que

(i) Pour tout $0 \leqq t < 1$ la limite

$$x(t+) = \lim_{s \downarrow t} x(s) \quad \text{existe et} \quad x(t+) = x(t).$$

(ii) Pour tout $0 < t \leqq 1$ la limite

$$x(t-) = \lim_{s \uparrow t} x(s) \quad \text{existe.}$$

Modules de continuité

Soit $x \in D[01]$, $S \subset [01]$, $\delta > 0$. On définit:

$$w_x(S) = \sup_{s, t \in S} |x(s) - x(t)|$$

$$w_x(\delta) = \sup_{0 \leqq t \leqq 1-\delta} w_x([t, t+\delta])$$

$$v_x^1(\delta) = \sup_{\substack{t-\delta \leqq t_1 \leqq t \leqq t_2 \leqq t+\delta \\ t, t_1, t_2 \in [01]}} \min\left(|x(t_1) - x(t)|, |x(t_2) - x(t)|\right)$$

$$v_x^2(\delta) = \sup_{0 \leqq t \leqq \delta} |x(t) - x(0)|$$

$$v_x^3(\delta) = \sup_{1-\delta \leqq t \leqq 1} |x(t) - x(1)|$$

$$\tilde{w}_x(\delta) = \max\left(v_x^1(\delta), v_x^2(\delta), v_x^3(\delta)\right).$$

Topologie de Skorokhod

Soit Λ la classe des fonctions continues strictement croissantes, appliquant $[01]$ sur $[01]$ et soient $x, y \in D[01]$. Posons

$$\|\lambda\| = \sup_{t \in [01]} |\lambda(t) - t|.$$

La topologie de Skorokhod est la topologie définie par la distance:

$$d(x, y) = \inf\left\{\varepsilon \mid \exists\, \lambda \in \Lambda \text{ avec } \|\lambda\| \leqq \varepsilon, \sup_{t \in [01]} \left|x(t) - y(\lambda(t))\right| \leqq \varepsilon\right\}.$$

Il est aisé de voir que si $x \in D[01]$ est continue $d(x_n, x) \to 0$ est équivalent à $d_0(x_n, x) \to 0$ où

$$d_0(x, y) = \max_{t \in [01]} |x(t) - y(t)|.$$

Dans la suite on dénotera par $\mathfrak{S}$ la tribu de Borel de $D[01]$.

I.2. Ensembles relativement compacts de $D[01]$

Définition. On dit que A est un ensemble relativement compact si la fermeture de A est un ensemble compact.

Théorème I.2 [5, p. 240]. *A est un ensemble relativement compact si et seulement si les deux conditions suivantes sont satisfaites:*

(i) $\sup_{x \in A} \sup_{t \in [01]} |x(t)| < \infty$,

(ii) $\lim_{\delta \to 0} \sup_{x \in A} \tilde{w}_x(\delta) = 0$.

I.3. Probabilités sur $\mathfrak{S}$

Un cylindre de dimension finie est un ensemble de la forme

$$\left\{x \in D[01] \mid \left(x(t_1), x(t_2), \ldots, x(t_k)\right) \in A\right\}$$

où A est un borélien de R^k.

Pour $t_1, \cdots, t_k \in [01]$ on note par $\pi_{t_1, \ldots, t_k}$ l'application de $D[01]$ dans R^k: $\pi_{t_1, \ldots, t_k}(x) = \left(x(t_1), \ldots, x(t_k)\right)$.

Lemme I.3.1 [4, p. 121]. *L'application* $\pi_{t_1,\dots,t_k}$ *est mesurable.*

Lemme I.3.2 [4, p. 121]. *Soit S un ensemble dénombrable, dense dans* [01] *avec* $1 \in S$.

Soit $J(S)$ *la classe des sous-ensembles de la forme* $\pi^{-1}_{t_1,\dots,t_k}(A)$ *avec k arbitraire, les* t_i *points de S et A borélien de* R^k.

Nous avons alors

1) $J(S)$ *est une algèbre d'ensembles.*

2) $J(S)$ *engendre* $\mathfrak{S}$.

Corollaire. *Soient u et v deux probabilités sur* $\mathfrak{S}$.

Si $u\,\pi^{-1}_{t_1,\dots,t_k} = v\,\pi^{-1}_{t_1,\dots,t_k}$ *pour tout k, pour tous* $t_1, \dots, t_k \in S$ *(S contenant 1, dénombrable et dense dans* [01]*) les deux mesures sont identiques sur* $\mathfrak{S}$. *Ceci est une conséquence immédiate du Lemme* I.3.2.

($u\,\pi^{-1}_{t_1,\dots,t_k}$ désigne la probabilité image de u par l'application $\pi_{t_1,\dots,t_k}$.)

I.4. Convergence faible de mesures

On dit que la suite v_n de probabilités converge faiblement vers v si $\int f\,dv_n \to \int f\,dv$ pour toute fonction continue et bornée.

Théorème I.4 [5, p. 253]. *Soient* $v_1, v_2, \dots$ *et* v *des probabilités sur* $\mathfrak{S}$. *Si on suppose:*

(i) $\lim_{\delta\to 0} \limsup_n v_n\{x \mid \tilde{w}_x(\delta) > \varepsilon\} = 0$, *pour tout* $\varepsilon > 0$,

(ii) *la suite* $v_n\,\pi^{-1}_{t_1,\dots,t_k}$ *converge faiblement vers* $v\,\pi^{-1}_{t_1,\dots,t_k}$ *pour tout k, pour tous* $t_1, \dots, t_k \in [01]$,

la suite v_n *converge faiblement vers* v.

Chapitre II

Soit $X_1, X_2, \dots$ une suite de variables aléatoires indépendantes, de même fonction de répartition $F(x)$ ayant une densité symétrique $f(x)$ $\big(f(x)=f(-x)\big)$.

Posons:

$$T_n(0,t) = n^{-\frac{3}{2}} \sum_{\substack{i<j \\ 1\leqq i,\, j\leqq n}} I(X_i + X_j \leqq 2\,t\,n^{-\frac{1}{2}}) - n^{-\frac{3}{2}} \sum_{\substack{i<j \\ 1\leqq i,\, j\leqq n}} I(X_i + X_j \leqq 0)$$

$$Y_n(t) = n^{\frac{1}{2}}\big(T_n(0,t) - t \cdot \textstyle\int f^2(x)\,dx\big)$$

$$Z_n(t) = Y_n(t) - E\big(Y_n(t)\big).$$

Théorème II.1. *Supposons que* $\int f^3(x)\,dx < \infty$. *La suite des processus* $\big(Z_n(t)\big)_{t\in[01]}$ *converge alors faiblement vers un processus de la forme* $(t\,Z)_{t\in[01]}$ *où Z est une variable normale d'espérance* 0 *et de variance*

$$c^2 = 4\big[\textstyle\int f^3(x)\,dx - (\int f^2(x)\,dx)^2\big].$$

Théorème II.2. *En supposant:*

$$\frac{1}{\Delta^2} \int_{-\infty}^{+\infty} \left\{ \int_0^{\Delta} [f(z+y) - f(y)]^2\,dz \right\} dy \xrightarrow[\Delta\to 0]{} 0$$

nous avons:

$$\sup_{t\in[01]} \big|E\big(Y_n(t)\big)\big| \xrightarrow[n\to\infty]{} 0.$$

Corollaire. *Sous les hypothèses des théorèmes* II.1 *et* II.2 *la suite des processus* $(Y_n(t))_{t\in[0\,1]}$ *converge faiblement vers un processus de la forme* $(tZ)_{t\in[0\,1]}$ *où* Z *est une variable normale* $N(0, c^2)$.

Remarques. 1) Soient v_n et v les mesures induites sur $\mathfrak{S}$ respectivement par les processus $(Y_n(t))_{t\in[0\,1]}$ et $(tZ)_{t\in[0\,1]}$.

Si la suite $Y_n(t)$ converge faiblement vers le processus tZ, alors pour toute fonction continue (ou seulement p.s. continue par rapport à v) f, la suite des mesures $v_n f^{-1}$ converge faiblement vers la mesure $v f^{-1}$. ($v f^{-1}$ signifie la probabilité image de v par f.)

Comme v concentre toute sa masse sur l'ensemble des fonctions continues, cette dernière propriété reste valable pour toute fonction f continue par rapport à la topologie uniforme.

2) L'hypothèse du théorème II.2 est satisfaite si par exemple

a) $f(x)$ remplit une condition de Lipschitz de la forme

$$|f(x+t)-f(x)| \leqq t^\alpha h(x), \quad \text{avec} \quad \alpha > \tfrac{1}{2} \quad \text{et} \quad h(x) \in L_2(-\infty, +\infty),$$

ou

b) $f(c)$ est absolument continue et sa densité $f'(x)$ appartient à $L_2(-\infty, +\infty)$.

Démonstration. On a

$$\begin{aligned}\frac{1}{\Delta^2} \int_{-\infty}^{+\infty} \int_0^\Delta [f(z+y)-f(y)]^2\, dz\, dy &\leqq \frac{1}{\Delta^2} \int_{-\infty}^{+\infty} \int_0^\Delta z^{2\alpha} h^2(x)\, dz\, dy \\ &= \int_{-\infty}^{+\infty} h^2(x)\, dx \cdot \frac{1}{2\alpha+1} \Delta^{2\alpha-1}.\end{aligned}$$

Le dernier terme tend vers 0 avec Δ, donc a) implique l'hypothèse du théorème II.2.

En outre

$$\begin{aligned}&\frac{1}{\Delta^2} \int_0^\Delta \int_{-\infty}^{+\infty} [f(x+z)-f(x)]^2\, dx\, dz \\ &= \frac{1}{\Delta^2} \int_0^\Delta \int_{-\infty}^{+\infty} \left[\int_0^Z f'(x+u)\, du\right]^2 dx\, dz \\ &\leqq \frac{1}{\Delta^2} \int_0^\Delta \int_{-\infty}^{+\infty} \left[z \int_0^Z f'^2(x+u)\, du\right] dx\, dz\end{aligned}$$

(D'après l'inégalité de Schwarz)

$$= \frac{1}{\Delta^2} \int_0^\Delta \left[z \int_0^Z du\right] dz \cdot \int_{-\infty}^{+\infty} f'^2(x)\, dx = \frac{\Delta}{3} \cdot \int f'^2(x)\, dx.$$

Le dernier terme tend vers 0 avec Δ et on a donc la même conclusion.

Démonstration du théorème II.1. Soient $X_1, X_2, \ldots, X_n$ n variables aléatoires indépendantes. L'ensemble des statistiques $S = s(X_1, \ldots, X_n)$ de carré intégrable forme un espace de Hilbert. Pour étudier une telle statistique S il s'est avéré commode de l'approximer par une statistique $\hat{S}$ appartenant au sous-espace L formé par les statistiques de la forme

$$\sum_{i=1}^n k_i(X_i) \left(E\, k_i^2(X_i) < \infty, \forall i\right) \quad \text{et minimisant } E(S-S')^2, \quad S' \in L.$$

Une telle statistique $\hat{S}$ est naturellement la projection orthogonale de S sur L. Nous avons alors les propriétés suivantes [7]:

$$\hat{S} = \sum_{i=1}^{n} E(S|X_i) - (n-1)\, E(S),$$

$$E(\hat{S}) = E(S), \qquad E(S-\hat{S})^2 = \mathrm{Var}(S) - \mathrm{Var}(\hat{S}).$$

$E(S|X_i)$ désigne ici l'espérance conditionnelle de S par rapport à X_i.

Désignons par $\hat{Z}_n(t)$ le processus projection de $Z_n(t)$ et posons $X_n(t) = Z_n(t) - \hat{Z}_n(t)$.

Lemme II.1. *Soit $t > s$. En posant*

$$f(x,t,s,n) = \frac{n^{\frac{1}{2}}}{2(t-s)} \left(F(-x+2t\,n^{-\frac{1}{2}}) - F(-x+2s\,n^{-\frac{1}{2}})\right)$$

nous avons:

1. $\mathrm{Var}(Z_n(t) - Z_n(s))$

$$= \frac{(n-1)}{n^2}(n-2)\, 4(t-s)^2 \left[\int f^2(x,t,s,n)\, f(x)\, dx - \left(\int f(x,t,s,n)\, f(x)\, dx\right)^2\right]$$

$$+ n^{-\frac{3}{2}}(n-1)(t-s) \cdot \int f(x,t,s,n)\, f(x)\, dx$$

$$- n^{-2}(n-1)\, 2(t-s)^2 \left(\int f(x,t,s,n)\, f(x)\, dx\right)^2.$$

2. $\hat{Z}_n(t) = n^{-1}(n-1) \sum_{i=1}^{n} (R_i - E(R_i))$

avec $R_i = F(-X_i + 2t\,n^{-\frac{1}{2}}) - F(-X_i)$.

3. $\mathrm{Var}(\hat{Z}_n(t) - \hat{Z}_n(s))$

$$= \frac{(n-1)^2}{n^2} 4(t-s)^2 \left[\int f^2(x,t,s,n)\, f(x)\, dx - \left(\int f(x,t,s,n)\, f(x)\, dx\right)^2\right].$$

4. $\mathrm{Var}(X_n(t) - X_n(s))$

$$= \frac{(1-n)}{n^2} 4(t-s)^2 \left[\int f^2(x,t,s,n)\, f(x)\, dx - \left(\int f(x,t,s,n)\, f(x)\, dx\right)^2\right]$$

$$+ n^{-\frac{3}{2}}(n-1)(t-s) \cdot \int f(x,t,s,n)\, f(x)\, dx$$

$$- n^{-2}(n-1)\, 2(t-s)^2 \left(\int f(x,t,s,n)\, f(x)\, dx\right)^2.$$

Démonstration. 1. Posons $Z_{ij} = I(2s\,n^{-\frac{1}{2}} < X_i + X_j \leqq 2t\,n^{-\frac{1}{2}})$. On a:

$$\begin{aligned}
\mathrm{Var}(Z_n(t) - Z_n(s)) &= \mathrm{Var}(Y_n(t) - Y_n(s)) = \mathrm{Var}\Big(n^{-1} \sum_{i<j} Z_{ij}\Big) \\
&= n^{-2} E\Big(\sum_{i<j} [Z_{ij} - E(Z_{ij})]\Big) \\
&= 2^{-1} n^{-1}(n-1)\, \mathrm{Var}(Z_{12}) + n^{-1}(n-1)(n-2)\, \mathrm{Cov}(Z_{12}, Z_{13}).
\end{aligned}$$

En outre:

$$\operatorname{Var}(Z_{12}) = \int [F(-x+2t\,n^{-\frac{1}{2}}) - F(-x+2s\,n^{-\frac{1}{2}})]\, f(x)\,dx$$
$$-\left(\int [F(-x+2t\,n^{-\frac{1}{2}}) - F(-x+2s\,n^{-\frac{1}{2}})]\, f(x)\,dx\right)^2$$
$$= n^{-\frac{1}{2}}\, 2(t-s) \int f(x,t,s,n)\, f(x)\,dx - n^{-1}\, 4(t-s)^2 \left(\int f(x,t,s,n)\, f(x)\,dx\right)^2,$$

$$\operatorname{Cov}(Z_{12}, Z_{13}) = n^{-1}\, 4(t-s)^2 \left[\int f^2(x,t,s,n)\, f(x)\,dx - \left(\int f(x,t,s,n)\, f(x)\,dx\right)^2\right].$$

2. Posons $Z_{ij} = I(0 < X_i + X_j \leqq 2t\,n^{-\frac{1}{2}})$

$$\hat{Z}_n(t) = \sum_{i=1}^{n} E(Z_n(t)|X_i)$$
$$= \sum_{i=1}^{n} E\left(n^{-1} \sum_{k<m} [Z_{km} - E(Z_{km})] | X_i\right)$$
$$= n^{-1}(n-1) \sum_{i=1}^{n} [R_i - E(R_i)] \quad \text{avec } R_i = F(-X_i + 2t\,n^{-\frac{1}{2}}) - F(-X_i).$$

3. $\operatorname{Var}(\hat{Z}_n(t) - \hat{Z}_n(s)) = n^{-2}(n-1)^2\, n \operatorname{Var}(F(-X_1 + 2t\,n^{-\frac{1}{2}}) - F(-X_1 + 2s\,n^{-\frac{1}{2}}))$

$$= \frac{(n-1)^2}{n^2}\, 4(t-s)^2 \left[\int f^2(x,t,s,n)\, f(x)\,dx - \left(\int f(x,t,s,n)\, f(x)\,dx\right)^2\right].$$

4. Cela découle des propriétés de la projection orthogonale et des points 1. et 3.

Lemme II.2. *En supposant:*

$$\int f^3(x)\,dx < \infty$$

nous avons:

a) $\int f(x,t,s,n)\, f(x)\,dx \leqq \int f^2(x)\,dx, \forall n, \forall t,\ s \in [01],$

b) $\int f^2(x,t,s,n)\, f(x)\,dx \leqq \int f^3(x)\,dx, \forall n, \forall t,\ s \in [01],$

c) $\int (f(x,t,0,n) - f(x))^2 f(x)\,dx \xrightarrow[n\to\infty]{} 0, \quad \forall t \in [01].$

Démonstration. a) Nous avons:

$$f(x,t,s,n) = \frac{1}{\Delta} \int_{\Delta'}^{\Delta''} f(-x+y)\,dy,$$

où $\Delta' = 2s\,n^{-\frac{1}{2}}$, $\Delta'' = 2t\,n^{-\frac{1}{2}}$ et $\Delta = \Delta'' - \Delta'$. Donc

$$\int f(x,t,s,n)\, f(x)\,dx = \frac{1}{\Delta} \int \left(\int_{\Delta'}^{\Delta''} f(-x+y)\, f(x)\,dy\right) dx$$
$$= \frac{1}{\Delta} \int_{\Delta'}^{\Delta''} \left(\int f(-x+y)\, f(x)\,dx\right) dy$$

(d'après le théorème de Fubini, applicable ici puisque les fonctions sont positives).

$$\leqq \frac{1}{\Delta} \int_{\Delta'}^{\Delta''} \left(\int f^2(-x+y)\,dx \cdot \int f^2(x)\,dx\right)^{\frac{1}{2}} dy$$

(inégalité de Schwarz)

$$= \frac{1}{\Delta} \int_{\Delta'}^{\Delta''} \left(\int f^2(x)\,dx\right) dy = \int f^2(x)\,dx.$$

b) Nous avons, en appliquant successivement le théorème de Fubini et les inégalités de Schwarz et de Hölder:

$$\begin{aligned}
\int f^2(x,t,s,n)\,f(x)\,dx &= \frac{1}{\Delta^2}\int\left(\int_{\Delta'}^{\Delta''} f(-x+y)\,dy \int_{\Delta'}^{\Delta''} f(-x+z)\,dz\right) f(x)\,dx\\
&= \frac{1}{\Delta^2}\int_{\Delta'}^{\Delta''}\int_{\Delta'}^{\Delta''}\left(\int f(-x+y)\,f(-x+z)\,f(x)\,dx\right) dy\,dz\\
&\leqq \frac{1}{\Delta^2}\int_{\Delta'}^{\Delta''}\int_{\Delta'}^{\Delta''}\left(\int f^2(-x+y)\,f(x)\,dx \int f^2(-x+z)\,f(x)\,dx\right)^{\frac{1}{2}} dy\,dz\\
&= \left(\frac{1}{\Delta}\int_{\Delta'}^{\Delta''}\left[\int f^2(-x+y)\,f(x)\,dx\right]^{\frac{1}{2}} dy\right)^2\\
&\leqq \frac{1}{\Delta}\int_{\Delta'}^{\Delta''}\left(\int f^2(-x+y)\,f(x)\,dx\right) dy\\
&\leqq \frac{1}{\Delta}\int_{\Delta'}^{\Delta''}\left[\left(\int f^3(-x+y)\,dx\right)^{\frac{2}{3}}\cdot\left(\int f^3(x)\,dx\right)^{\frac{1}{3}}\right] dy\\
&= \frac{1}{\Delta}\int_{\Delta'}^{\Delta''}\left(\int f^3(x)\,dx\right) dy = \int f^3(x)\,dx.
\end{aligned}$$

c) Montrons que

$$\lim_{n\to\infty}\int f^2(x,t,0,n)\,f(x)\,dx = \int f^3(x)\,dx.$$

On a

$$f(x,t,0,n) = \frac{1}{\Delta_n}\int_0^{\Delta_n} f(-x+y)\,dy, \quad \text{où} \quad \Delta_n = 2t\,n^{-\frac{1}{2}},$$

et donc

$$\lim_{n\to\infty}\frac{1}{\Delta_n}\int_0^{\Delta_n} f(-x+y)\,dy = f(-x) = f(x)$$

pour presque tout x (mesure de Lebesgue).

En appliquant le lemme de Fatou il vient:

$$\int f^3(x)\,dx \leqq \liminf_n \int\left(\frac{1}{\Delta_n}\int_0^{\Delta_n} f(-x+y)\,dy\right)^2 f(x)\,dx.$$

En vertu du point b) cela implique

$$\lim_{n\to\infty}\int f^2(x,t,0,n)\,f(x)\,dx = \int f^3(x)\,dx.$$

Le point c) du lemme en résulte alors aisément.

Lemme II.3 [8, p. 103]. *Soient $X_1, X_2, \ldots$ et X des vecteurs aléatoires:*

$$X_n = (X_{n1}, X_{n2}, \ldots, X_{nk}), \qquad X = (X_1, X_2, \ldots, X_k).$$

Supposons que la suite des variables aléatoires $\sum_{j=1}^{k} s_j X_{nj}$ converge faiblement vers la variable $\sum_{j=1}^{k} s_j X_j$ pour n tendant vers l'infini, ceci pour tout $(s_1, s_2, \ldots, s_k)\in R^k$. Alors la suite X_n converge faiblement vers X.

Lemme II.4. *Supposons que* $\int f^3(x)\,dx < \infty$. *Nous avons alors:*

a) $E(X_n(t))^2 \leqq n^{-1} t^2 (4B + 6A^2) + n^{-\frac{1}{2}} t A$,

où

$$A = \int f^2(x)\,dx \quad et \quad B = \int f^3(x)\,dx,$$

ce qui entraîne en particulier,

$$|X_n(t)| \xrightarrow{P} 0, \qquad \forall t \in [01].$$

b) $E\left(\hat{Z}_n(t) - n^{-\frac{1}{2}} 2t \sum_{i=1}^{n} [f(X_i) - \int f^2(x)\,dx]\right)^2 \to 0, \qquad \forall t \in [01].$

Démonstration.

a)
$$\begin{aligned} E(X_n(t))^2 &= E(Z_n(t))^2 - E(\hat{Z}_n(t))^2 \\ &= \frac{(1-n)}{n^2} 4t^2 \left[\int f^2(x,t,0,n) f(x)\,dx - \left(\int f(x,t,0,n) f(x)\,dx\right)^2\right] \\ &\quad + n^{-\frac{3}{2}} (n-1)\, t \int f(x,t,0,n) f(x)\,dx - \frac{(n-1)}{n^2} 2t^2 \left(\int f(x,t,0,n) f(x)\,dx\right)^2 \end{aligned}$$

(d'après le lemme II.1)

$$\begin{aligned} &= n^{-1} 4t^2 (B + A^2) + n^{-\frac{1}{2}} t A + n^{-1} 2t^2 A^2 \\ &= n^{-1} t^2 (4B + 6A^2) + n^{-\frac{1}{2}} t A \qquad \text{(d'après le lemme II.2).} \end{aligned}$$

b)
$$\begin{aligned} &E\left(\hat{Z}_n(t) - n^{-\frac{1}{2}} 2t \sum_{i=1}^{n} [f(X_i) - \int f^2(x)\,dx]\right)^2 \\ &= \mathrm{Var}\left(\frac{(n-1)}{n} \sum_{i=1}^{n} [F(-X_i + 2t n^{-\frac{1}{2}}) - F(-X_i)] - n^{-\frac{1}{2}} 2t \sum_{i=1}^{n} f(X_i)\right) \\ &= n\, \mathrm{Var}\left(\frac{(n-1)}{n} [F(-X_1 + t n^{-\frac{1}{2}}) - F(-X_1)] - n^{-\frac{1}{2}} 2t f(X_1)\right) \\ &= 4t^2 \int \left[\frac{(n-1)}{n} f(x,t,0,n) - f(x)\right]^2 f(x)\,dx \\ &\quad - 4t^2 \left(\int \left[\frac{(n-1)}{n} f(x,t,0,n) - f(x)\right] f(x)\,dx\right)^2. \end{aligned}$$

Ces deux derniers termes tendent vers 0 d'après le lemme II.2.

Lemme II.5. *Soient* v_n *et* v *les mesures induites sur* $\mathfrak{S}$ *respectivement par les processus* $Z_n(t)$ *et* $t Z$, *où*

$$Z = N\left(0, 4\left[\int f^3(x)\,dx - \left(\int f^2(x)\,dx\right)^2\right]\right).$$

Sous l'hypothèse du théorème II.1, la suite des mesures $v_n \pi^{-1}_{t_1,\ldots,t_k}$ converge faiblement vers la mesure $v \pi^{-1}_{t_1,\ldots,t_k}$, ceci pour tout k et tous $t_1, \ldots, t_k \in [01]$.

Démonstration. Soient k un nombre arbitraire, $(s_1, \ldots, s_k) \in R^k$, $t_1, \ldots, t_k \in [01]$.

La suite des variables aléatoires $\sum_{i=1}^{k} s_i Z_n(t_i)$ converge alors faiblement vers la variable $\sum_{i=1}^{k} s_i t_i Z$ car, d'après le lemme II.4,

$$\sum_{i=1}^{k} (s_i Z_n(t_i) - s_i \hat{Z}_n(t_i) \xrightarrow[n\to\infty]{P} 0$$

et la suite des variables $\sum_{i=1}^{k} s_i \hat{Z}_n(t_i)$ converge faiblement vers la variable $\sum_{i=1}^{k} s_i t_i Z$. En vertu du lemme II.3, cela entraîne la convergence faible de la suite des mesures $v_n \pi_{t_1,\ldots,t_k}^{-1}$ vers la mesure $v \pi_{t_1,\ldots,t_k}^{-1}$.

Lemme II.6. *Supposons que $\int f^3(x)\,dx < \infty$. Nous avons alors:*

a) $E(\hat{Z}_n(t) - \hat{Z}_n(s))^2 \leqq 4(t-s)^2 (B + A^2)$, *où* $A = \int f^2(x)\,dx$ *et* $B = \int f^3(x)\,dx$.

b) $\forall \varepsilon > 0, \quad \forall \alpha \in [0, \frac{1}{2}), \quad \exists M$ *tel que l'on ait, pour tout n,*

$$P\left\{ \sup_{t,s\in[0\,1]} \frac{|\hat{Z}_n(t) - \hat{Z}_n(s)|}{|t-s|^\alpha} > M \right\} \leqq \varepsilon.$$

Démonstration. a) découle des lemmes II.1 et II.2.

b) découle de a) (voir par exemple [6, p. 517]).

Lemme II.7. *Soient $R_1, \ldots, R_m$ m variables aléatoires non nécessairement indépendantes ou équidistribuées.*

Posons:

$$S_0 = 0, \qquad S_k = \sum_{i=1}^{k} R_i \qquad (k = 1, 2, \ldots, m) \qquad M_m = \max_{1 \leqq k \leqq m} |S_k|.$$

Supposons que

$$E(S_j - S_i)^2 \leqq \sum_{i<k\leqq j} u_k \text{ pour tous } i, j, \quad i<j \quad (i = 0, 1, 2, \ldots, m-1; j = 1, 2, \ldots, m),$$

avec $u_1, \ldots, u_m$ réels et non négatifs.

Nous avons alors:

$$E(M_m^2) \leqq (\log_2 4m)^2 (u_1 + \cdots + u_m).$$

Démonstration. Pour $m = 1$ le lemme est vrai.

Soit $m > 1$. Définissons $r \geqq 0$, entier, par l'inégalité:

$$2^r < m \leqq 2^{r+1} = N \quad \text{et} \quad R_j = 0 \text{ pour } m < j \leqq N.$$

Soit S la somme de toutes les sommes partielles de la forme:

$(R_{\alpha+1} + \cdots + R_\beta)^2, \quad \alpha = u\,2^v, \quad \beta = (u+1)\,2^v, \quad v = 0, \ldots, r+1;\ u = 0, \ldots, 2^{r+1-v} - 1.$

La somme E_v des espérances des termes intervenant dans S et correspondant à une valeur fixe de v satisfait à l'inégalité suivante:

$$E_v \leqq (u_1 + u_2 + \cdots + u_m).$$

On déduit que

$$E(S) \leqq (r+2)(u_1 + u_2 + \cdots + u_m).$$

D'autre part, toute somme $R_1 + R_2 + \cdots + R_j$ se laisse décomposer en somme de sommes partielles de la forme $R_{\alpha+1} + \cdots + R_\beta$ (α, β définis comme plus haut).

Il s'ensuit que

$$R_1+\cdots+R_j=q_1+\cdots+q_k \quad \text{où } q_j \text{ contient } 2^{r_j} \text{ termes avec}$$
$$r+1\geqq r_1>r_2>\cdots>r_k\geqq 0.$$

On obtient alors par l'inégalité de Schwarz

$$(R_1+\cdots+R_j)^2\leqq k\sum_{i=1}^{k}(q_i)^2\leqq (r+2)\,S,$$

ce qui implique

$$E\big(\max_{j\leqq m}(R_1+\cdots+R_j)^2\big)\leqq (r+2)\,E(S)\leqq (r+2)^2\,(u_1+\cdots+u_m)$$
$$\leqq (\log_2 4m)^2\,(u_1+\cdots+u_m),$$

ce qui démontre le lemme.

L'ensemble de ces lemmes va nous permettre maintenant de montrer que la suite v_n de mesures converge vers la mesure v.

Compte tenu du lemme II.5, il suffira pour cela de vérifier que la suite v_n satisfait à l'hypothèse (i) du théorème I.4.

Puisque pour tout $x\in D[01]$, $\tilde{w}_x(\delta)\leqq w_x(\delta)$, il suffit donc de démontrer que

$$\lim_{\delta\to 0}\limsup_{n} P\{w_{Z_n(t)}(\delta)>\varepsilon\}=0, \qquad \forall\varepsilon>0. \tag{$*$}$$

Mais

$$|Z_n(t)-Z_n(s)|\leqq |X_n(t)-X_n(s)|+|\hat{Z}_n(t)-\hat{Z}_n(s)|.$$

La relation $(*)$ est donc une conséquence des deux suivantes:

1. $\lim\limits_{\delta\to 0}\limsup\limits_{n} P\{w_{\hat{Z}_n(t)}(\delta)>\varepsilon\}=0, \qquad \forall\varepsilon>0.$

2. $\lim\limits_{\delta\to 0}\limsup\limits_{n} P\{w_{X_n(t)}(\delta)>\varepsilon\}=0, \qquad \forall\varepsilon>0.$

Démonstration de 1. Soient $\varepsilon>0$, $0<\alpha<\frac{1}{2}$, $\eta>0$.

D'après le lemme II.6 il existe un M tel que l'on ait, pour tout n,

$$P\left\{\sup_{s,t\in[01]}\frac{|\hat{Z}_n(t)-\hat{Z}_n(s)|}{|t-s|^\alpha}>M\right\}\leqq\eta.$$

Pour tout $\delta\leqq\delta_0=\left(\dfrac{\varepsilon}{M}\right)^{1/\alpha}$ et tout n on a alors:

$$P\{\sup_{|t-s|\leqq\delta}|\hat{Z}_n(t)-\hat{Z}_n(s)|>\varepsilon\}$$
$$\leqq P\{\sup_{|t-s|\leqq\delta_0}|\hat{Z}_n(t)-\hat{Z}_n(s)|>M\,\delta_0^\alpha\}$$
$$\leqq P\left\{\sup_{|t-s|\leqq\delta_0}\frac{|\hat{Z}_n(t)-\hat{Z}_n(s)|}{|t-s|^\alpha}>M\right\}\leqq\eta, \qquad \text{ce qui démontre 1.}$$

Démonstration de 2. Soit $\delta=m^{-1}$, m entier positif.

On a:

$$\left\{w_{X_n(t)}\left(\frac{1}{m}\right)>\varepsilon\right\}\subset\left\{\max_{0\leqq j\leqq m-1}\sup_{t\in\left[\frac{j}{m},\frac{j+1}{m}\right]}|X_n(t)-X_n(jm^{-1})|>\frac{\varepsilon}{2}\right\}$$
$$\cup\left\{\max_{1\leqq j\leqq m}\sup_{t\in\left[\frac{j-1}{m},\frac{j}{m}\right]}|X_n(t)-X_n(jm^{-1})|>\frac{\varepsilon}{2}\right\}.$$

Il suffit donc de montrer que

$$\lim_{m\to\infty} \limsup_{n} P\Big\{\max_{0\leqq j\leqq m-1} \sup_{t\in\left[\frac{j}{m},\frac{j+1}{m}\right]} |X_n(t)-X_n(jm^{-1})|>\varepsilon\Big\}=0, \qquad (**)$$

pour tout $\varepsilon>0$.

Voici d'abord un lemme:

Lemme II.8. *Soit* $t_i^j=m^{-1}(j+\lceil n^{\frac{1}{2}}\rceil^{-1} i)$, *j fixé,* $i=0,1,\ldots,\lceil n^{\frac{1}{2}}\rceil$. $\lceil a\rceil$ *désigne ici le plus petit entier supérieur ou égal à a. Nous avons alors:*

$$\begin{aligned}\sup_{t\in\left[\frac{j}{m},\frac{j+1}{m}\right]} &|X_n(t)-X_n(jm^{-1})|\\ &\leqq 2\max_{1\leqq i\leqq\lceil n^{\frac{1}{2}}\rceil}|X_n(t_i^j)-X_n(jm^{-1})|+\sup_{|t-s|\leqq\frac{1}{m}}|\hat Z_n(t)-\hat Z_n(s)|\\ &\quad+m^{-1}\frac{n^{\frac{1}{2}}}{\lceil n^{\frac{1}{2}}\rceil}\int f^2(x)\,dx+\max_{0\leqq i\leqq\lceil n^{\frac{1}{2}}\rceil-1}\sup_{t,s\in[t_i^j,t_{i+1}^j]}|E(Y_n(t)-Y_n(s))|.\end{aligned}$$

Démonstration. Soit $t\in[t_i^j,t_{i+1}^j]$.

Nous avons:

$$\begin{aligned}X_n(t)-X_n(jm^{-1})=n^{-1}\sum_{\substack{k<h\\1\leqq k,h\leqq n}} &I(0<X_k+X_h\leqq 2t\,n^{-\frac{1}{2}})-t\,n^{\frac{1}{2}}\int f^2(x)\,dx\\ &-E(Y_n(t))-\hat Z_n(t)-Z_n(jm^{-1})+\hat Z_n(jm^{-1}).\end{aligned}$$

Supposons $X_n(t)-X_n(jm^{-1})\geqq 0$ (le cas $X_n(t)-X_n(jm^{-1})<0$ se traite de manière analogue).

Nous avons alors:

$$\begin{aligned}|X_n(t)-X_n(jm^{-1})|\leqq|X_n(t_{i+1}^j)-X_n(jm^{-1})|+\frac{n^{\frac{1}{2}}}{\lceil n^{\frac{1}{2}}\rceil}m^{-1}\int f^2(x)\,dx\\ +|\hat Z_n(t_{i+1}^j)-\hat Z_n(t)|+|E(Y_n(t_{i+1}^j)-Y_n(t))|,\end{aligned}$$

ce qui démontre le lemme.

Démonstration de la relation (**). On a, en vertu du lemme II.8,

$$\begin{aligned}P\Big\{&\max_{0\leqq j\leqq m-1}\sup_{t\in\left[\frac{j}{m},\frac{j+1}{m}\right]}|X_n(t)-X_n(jm^{-1})|>\varepsilon\Big\}\\ &\leqq P\left\{\max_{0\leqq j\leqq m-1}2\max_{1\leqq i\leqq\lceil n^{\frac{1}{2}}\rceil}|X_n(t_i^j)-X_n(jm^{-1})|>\frac{\varepsilon}{4}\right\}\\ &\quad+P\left\{\sup_{|t-s|\leqq\frac{1}{m}}|\hat Z_n(t)-\hat Z_n(s)|>\frac{\varepsilon}{4}\right\}+P\left\{\frac{n^{\frac{1}{2}}}{\lceil n^{\frac{1}{2}}\rceil}m^{-1}\int f^2(x)\,dx>\frac{\varepsilon}{4}\right\}\\ &\quad+P\left\{\max_{0\leqq j\leqq m-1}\max_{0\leqq i\leqq\lceil n^{\frac{1}{2}}\rceil-1}\sup_{t,s\in[t_i^j,t_{i+1}^j]}|E(Y_n(t)-Y_n(s))|>\frac{\varepsilon}{4}\right\}\\ &=G_1+G_2+G_3+G_4.\end{aligned}$$

Pour m suffisamment grand G_3 est nul pour tout n et G_2 est arbitrairement petit uniformément en n (voir démonstration de 1.).

Il reste donc à vérifier que

a) $\lim_{m\to\infty} \limsup_n G_1 = 0$,

b) $\lim_{m\to\infty} \limsup_n G_4 = 0$.

Démonstration. On a:

$$G_1 \leqq \sum_{j=0}^{m-1} P\left\{\max_{1\leqq i\leqq \lceil n^{\frac{1}{2}}\rceil} |X_n(t_i^j)-X_n(jm^{-1})| > \frac{\varepsilon}{8}\right\} = \sum_{j=0}^{m-1} K_j.$$

D'autre part, en vertu des lemmes II.1, II.2 et II.4, on a, pour $t>s$,

$$E\big(X_n(t)-X_n(s)\big)^2 \leqq A_1 n^{-\frac{1}{2}}(t-s)+B_1 n^{-1}(t-s)^2, \qquad \text{où} \qquad A_1=A, B_1=4B+6A^2.$$

Soit j fixé. On définit:

$$\begin{aligned}
R_i &= X_n(t_i^j)-X_n(t_{i-1}^j),\\
r_i &= A_1 n^{-\frac{1}{2}} i m^{-1} \lceil n^{\frac{1}{2}}\rceil^{-1} + B_1 n^{-1}(i m^{-1}\lceil n^{\frac{1}{2}}\rceil^{-1})^2,\\
u_i &= r_i - r_{i-1},\\
S_k &= \sum_{i=1}^{k} R_i.
\end{aligned}$$

Nous avons alors pour tous (i,j), $i<j$ $(i=0,1,\ldots,\lceil n^{\frac{1}{2}}\rceil-1, j=1,2,\ldots,\lceil n^{\frac{1}{2}}\rceil)$,

$$\begin{aligned}
E(S_k-S_i)^2 &= E\left(\sum_{h=i+1}^{k} X_h\right)^2 = E\big(X_n(t_k^j)-X_n(t_i^j)\big)^2\\
&\leqq A_1 n^{-\frac{1}{2}}(k-i)\, m^{-1}\lceil n^{\frac{1}{2}}\rceil^{-1} + B_1 n^{-1}\big(m^{-1}\lceil n^{\frac{1}{2}}\rceil^{-1}(k-i)\big)^2\\
&\leqq r_k - r_i = \sum_{i<h\leqq k} u_h.
\end{aligned}$$

En appliquant le lemme II.7 aux variables $R_1, \ldots, R_{\lceil n^{\frac{1}{2}}\rceil}$ définies plus haut, on obtient

$$\begin{aligned}
K_j &\leqq P\left\{\max_{1\leqq i\leqq \lceil n^{\frac{1}{2}}\rceil} |S_i| > \frac{\varepsilon}{8}\right\} \leqq \frac{64}{\varepsilon^2} E\big(\max_{1\leqq i\leqq \lceil n^{\frac{1}{2}}\rceil} |S_i|\big)^2\\
&\leqq \frac{64}{\varepsilon^2} (\log_2 4\lceil n^{\frac{1}{2}}\rceil)^2 \sum_{i=1}^{\lceil n^{\frac{1}{2}}\rceil} u_i\\
&= \frac{64}{\varepsilon^2} (\log_2 4\lceil n^{\frac{1}{2}}\rceil)^2 (A_1 n^{-\frac{1}{2}} m^{-1} + B_1 n^{-1} m^{-2}),
\end{aligned}$$

ce qui entraîne,

$$\sum_{j=0}^{m-1} K_j \leqq \frac{64}{\varepsilon^2} (\log_2 4\lceil n^{\frac{1}{2}}\rceil)^2 (A_1 n^{-\frac{1}{2}} + B_1 n^{-1} m^{-1}).$$

D'où a), puisque le membre de droite tend vers 0 pour n tendant vers l'infini.

Pour démontrer b) considérons $t, s \in [t_i^j, t_{i+1}^j]$, $t > s$.
On a:

$$E\big(Y_n(t) - Y_n(s)\big)$$
$$= E\Big(n^{-1} \sum_{\substack{i<j \\ 1 \leqq i, j \leqq n}} I(2 s n^{-\frac{1}{2}} < X_i + X_j \leqq 2 t n^{-\frac{1}{2}})\Big) - n^{\frac{1}{2}}(t-s) \int f^2(x)\, dx$$
$$= \frac{(n-1)}{2} \int [F(-x + 2 t n^{-\frac{1}{2}}) - F(-x + 2 s n^{-\frac{1}{2}})] f(x)\, dx - n^{\frac{1}{2}}(t-s) \int f^2(x)\, dx$$
$$= M_1 - M_2,$$

où
$$M_1 = (n-1)(t-s)\, n^{-\frac{1}{2}} \int f(x, t, s, n) f(x)\, dx \leqq A m^{-1} \quad \text{(lemme II.2)}$$
et
$$M_2 \leqq A m^{-1}.$$
Cela entraîne
$$\sup_{t, s \in [t_i^j, t_{i+1}^j]} \big|E\big(Y_n(t) - Y_n(s)\big)\big| \leqq 2 A m^{-1}$$
et donc
$$\lim_{m \to \infty} \limsup_n G_4 = 0.$$

La démonstration du théorème II.1 est ainsi terminée.

Démonstration du théorème II.2. On a:

$$E\big(Y_n(t)\big) = 2^{-1} n \int [F(-x + 2 t n^{-\frac{1}{2}}) - F(-x)] f(x)\, dx - n^{\frac{1}{2}} t \int f^2(x)\, dx$$
$$- 2^{-1} \int [F(-x + 2 t n^{-\frac{1}{2}}) - F(-x)] f(x)\, dx.$$
Or,
$$2^{-1} \int [F(-x + 2 t n^{-\frac{1}{2}}) - F(-x)] f(x)\, dx = n^{-\frac{1}{2}} t \int f(x, t, 0, n) f(x)\, dx \leqq n^{-\frac{1}{2}} A.$$

Le dernier terme tend donc vers 0 uniformément en t.

D'autre part,

$$2^{-1} n \int [F(-x + 2 t n^{-\frac{1}{2}}) - F(-x)]\, f(x)\, dx - n^{\frac{1}{2}} t \int f^2(x)\, dx$$
$$= 2^{-1} n \int_{-\infty}^{+\infty} \Bigg[\int_0^{2 t n^{-\frac{1}{2}}} \big(f(z-x) - f(x)\big)\, dz\Bigg] f(x)\, dx$$
$$= -4^{-1} n \int_{-\infty}^{+\infty} \int_0^{2 t n^{-\frac{1}{2}}} [f(z-x) - f(x)]^2\, dz\, dx.$$

Le dernier terme est, en valeur absolue, plus petit que

$$4^{-1} n \int_{-\infty}^{+\infty} \int_0^{2 n^{-\frac{1}{2}}} [f(z+x) - f(x)]^2\, dz\, dx.$$

Le théorème est donc démontré puisque le dernier terme tend vers 0 par hypothèse.

Chapitre III

On considère une suite $X_1, X_2, \ldots$ de variables aléatoires indépendantes, de même fonction de répartition $F(x-\theta)$, où $F(x)$ est absolument continue et de densité $f(x)$ avec $f(x)=f(-x)$.

Désignons par T l'estimateur de Hodges Lehmann. Considérons à nouveau les quantités $T_1, T_2, a_\alpha, \sigma_T, S(T)$ définies dans l'introduction.

Théorème III. *Sous les hypothèses des théorèmes* II.1 *et* II.2 *nous avons: La statistique* $n^{\frac{1}{2}}(n^{\frac{1}{2}} S(T)-\sigma_T)$ *est asymptotiquement normale* $N(0, e^2)$, *où*

$$e^2=c^2\left(\int f^2(x)\,dx\right)^{-2}\cdot\sigma_T^2.$$

Démonstration. Pour la démonstration on peut admettre sans restreindre la généralité que la vraie valeur du paramètre θ est nulle. Lehmann [1] a alors démontré que

$$n^{\frac{1}{2}} S(T) \xrightarrow{P} \sigma_T \quad \text{et que les quantités } T_1 \text{ et } T_2$$

sont des $O(n^{-\frac{1}{2}})$.

Posons $b=2a_\alpha\sigma_T$. On a alors:

$$\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}\cap\{|n^{\frac{1}{2}} T_1|\leqq M\}$$
$$\subset\Big\{\min_{t\in[-M,+M]}\big(Y_n(t+b)-Y_n(t)\big)<x\Big\}$$

ce qui entraîne:

$$P\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x,\ |n^{\frac{1}{2}} T_1|\leqq M\}$$
$$\leqq P\Big\{\min_{t\in[-M,+M]}\big(Y_n(t+b)-Y_n(t)\big)<x\Big\}, \qquad \text{pour tout } M\geqq 0.$$

Soit $\varepsilon>0$ arbitraire. Il existe alors un M tel que

$$P\{|n^{\frac{1}{2}} T_1|\leqq M\}\geqq 1-\varepsilon, \qquad \text{pour tout } n\geqq n_0.$$

D'autre part $P(A\cap B)\geqq P(A)+P(B)-1$.

On a donc, compte tenu du corollaire suivant le théorème II.2,

$$\limsup_n P\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}-\varepsilon\leqq P\{b\,Z<x\}$$

où Z est une variable normale $N(0, c^2)$.

Comme ε est arbitraire, il s'ensuit que

$$\limsup_n P\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}\leqq P\{b\,Z<x\}.$$

D'autre part

$$\Big\{\max_{t\in[-M,+M]}\big(Y_n(t+b)-Y_n(t)\big)<x\Big\}$$
$$\subset\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}\cap\{|n^{\frac{1}{2}} T_1|\leqq M\}.$$

De manière analogue on montre que

$$P\{b\,Z<x\}\leqq\liminf_n P\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}.$$

Il résulte donc que

$$\lim_{n\to\infty} P\{Y_n(n^{\frac{1}{2}} T_1+b)-Y_n(n^{\frac{1}{2}} T_1)<x\}=P\{b\,Z<x\}.$$

Utilisons maintenant le fait que $n^{\frac{1}{2}}(T_2-T_1) \xrightarrow{P} b$.

On a:

$$\begin{aligned}&\{|Y_n(n^{\frac{1}{2}}T_1+b)-Y_n(n^{\frac{1}{2}}T_2)|>x\}\\ &\qquad\cap\{|n^{\frac{1}{2}}T_1|\leqq M,\ |n^{\frac{1}{2}}T_2|\leqq M,\ |n^{\frac{1}{2}}(T_2-T_1)-b|\leqq\varepsilon\}\\ &\qquad\subset\Big\{\max_{\substack{t,s\in[-M,+M]\\ |t+b-s|\leqq\varepsilon}}|Y_n(t+b)-Y_n(s)|>x\Big\}.\end{aligned}$$

D'où

$$\limsup_n P\{|Y_n(n^{\frac{1}{2}}T_1+b)-Y_n(n^{\frac{1}{2}}T_2)|>x\}\leqq P\{\varepsilon|Z|>x\}.$$

Comme ε est arbitraire on conclut que

$$|Y_n(n^{\frac{1}{2}}T_1+b)-Y_n(n^{\frac{1}{2}}T_2)| \xrightarrow{P} 0.$$

Il résulte alors

$$|Y_n(n^{\frac{1}{2}}T_1+b)-Y_n(n^{\frac{1}{2}}T_1)-Y_n(n^{\frac{1}{2}}T_2)+Y_n(n^{\frac{1}{2}}T_1)| \xrightarrow{P} 0.$$

D'où la convergence faible de la suite des variables aléatoires $Y_n(n^{\frac{1}{2}}T_2)-Y_n(n^{\frac{1}{2}}T_1)$ vers la variable bZ.

En tenant compte de la définition des variables T_1 et T_2 il découle de la dernière relation que la suite

$$n^{\frac{1}{2}}(2a_\alpha(12)^{-\frac{1}{2}}-n^{\frac{1}{2}}(T_2-T_1)\int f^2(x)\,dx)$$

converge faiblement vers la variable bZ.

Le théorème III est ainsi démontré.

Remerciements. Je remercie vivement Monsieur le Professeur P.J. Huber qui m'a donné l'idée de ce travail. Je voudrais également le remercier pour les nombreux conseils qu'il m'a donnés.

Je tiens à adresser aussi mes remerciements à Monsieur le Professeur R. Cairoli pour le soin avec lequel il a lu le manuscrit. Ses remarques judicieuses m'ont beaucoup aidé dans la rédaction du présent travail.

Bibliographie

1. Lehmann, E. L.: Nonparametric Confidence Intervals for a Shift Parameter. Ann. math. Statist. **34**, 1507–1512 (1963).
2. Jurečková, J.: Asymptotic Linearity of a Rank Statistic in Regression Parameter. Ann. math. Statist. **40**, 1889–1900 (1969).
3. Huber, P. J.: Studentizing Robust Estimates. In: Nonparametric Techniques in Statistical Inferences. ed. by M.L. Puri, Cambridge University Press 1970, 453–463.
4. Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968.
5. Parthasarathy, K. R.: Probability Measures on Metric Spaces. New York: Academic Press 1967.
6. Loève, M.: Probability Theory. New York: Van Nostrand 1963.
7. Hájek, J.: Asymptotic Normality of Simple Linear Rank Statistics under Alternatives. Ann. math. Statist. **39**, 325–346 (1968).
8. Rao, C. R.: Linear Statistical Inferences and Its Applications. New York: Wiley 1965.

André Antille, Mathematische Statistik ETH,
Clausiusstraße 55, CH-8006 Zürich, Suisse

(Reçu le 16 mai 1972)