STATISTIQUE ET ANALYSE DES DONNÉES

JEAN-FRANÇOIS INGENBLEEK

Normalité asymptotique d'une classe de statistiques de rangs sérielles multivariées

Statistique et analyse des données, tome 6, nº 2 (1981), p. 49-59

<http://www.numdam.org/item?id=SAD_1981__6_2_49_0>

Statistique et Analyse de données
1981 - Vol. 6 n° 2 - pp. 49-59

NORMALITE ASYMPTOTIQUE D'UNE CLASSE DE STATISTIQUES DE RANGS SERIELLES MULTIVARIEES

Jean-François INGENBLEEK

Université Libre de Bruxelles
Institut de Statistique
B - 1050 BRUXELLES

Résumé : *Dans un précédent article ([2]), nous avons introduit une nouvelle classe de statistiques de rangs permettant de construire un test d'hypothèse nulle de bruit blanc multivarié qui soit libre. Pour ces statistiques, nous démontrons un théorème du type central limite d'application sous l'hypothèse nulle.*

Abstract : *In a recent paper ([2]), we have introduced a new class of rank statistics. With these statistics, we have constructed a distribution-free rank test for the null hypothesis of white noise. We establish here a permutational central limit theorem for these rank statistics.*

1 - INTRODUCTION.

Soit $X_\alpha = (X_{1\alpha}, X_{2\alpha}, \ldots, X_{m\alpha})'$ ($\alpha = 1,2,\ldots$) une suite de vecteurs aléatoires m-variés et identiquement distribués suivant une fonction de répartition F absolument continue à marginales F_i ($i = 1,2,\ldots,m$). Soient également $R^n_{i1}, R^n_{i2}, \ldots, R^n_{in}$ les rangs des variables $X_{i1}, X_{i2}, \ldots, X_{in}$ ($i = 1,2,\ldots,m$).

Introduisons une classe de statistiques de rangs dont les éléments sont définis de la manière suivante. Soient $2 \leq q \in \mathbb{N}$ et $J(u_1, u_2, \ldots, u_q)$ une fonction définie pour $u_\ell \in (0,1)^m$ ($\ell = 1,2, \ldots, q$); posons

$$S_n = \frac{1}{n} \sum_{\alpha=1}^{n} J\left(\frac{R_\alpha^n}{n+1}, \frac{R_{\alpha+1}^n}{n+1}, \ldots, \frac{R_{\alpha+q-1}^n}{n+1}\right), \tag{1}$$

où $R_\alpha^n = (R_{1\alpha}^n, R_{2\alpha}^n, \ldots, R_{m\alpha}^n)'$ $(\alpha = 1,2,\ldots,n)$ et où $R_{n+k}^n = R_k^n$ $(k = 1,2,\ldots,n)$. Les statistiques ainsi définies sont appelées statistiques de rangs *sérielles*. Dans un précédent article ([2]), nous avons montré comment, sur la base de telles statistiques, construire un test de rangs ϕ pour tester l'hypothèse nulle de bruit blanc multivarié. Remarquons que c'est par application du principe de permutation introduit par Puri et Sen ([3]) qu'il est possible d'obtenir un tel test ϕ qui soit libre sous l'hypothèse nulle. Ce test est conditionnel. D'une manière plus précise, soient $R^n = (R_1^n, R_2^n, \ldots, R_n^n)$ la matrice (aléatoire) des rangs et r_n une réalisation de cette matrice. Notons $\mathcal{A}(r_n)$ l'ensemble des n! matrices obtenues en permutant les colonnes de r_n. La donnée de $\mathcal{A}(r_n)$ est équivalente à la donnée de r_n^*, matrice appartenant à $\mathcal{A}(r_n)$ mais dont la première ligne est 1,2,...,n. Appelons R_n^* la matrice aléatoire dont les réalisations sont les matrices r_n^*. Le test de rangs ϕ est conditionnel étant donné R_n^*.

Pour étudier les propriétés asymtotiques de ϕ, il est nécessaire de connaître la distribution asymptotique conditionnelle de S_n étant donné R_n^*. Notre but est de démontrer un théorème du type central limite pour S_n et donc d'étudier le comportement asymptotique de la fonction de répartition conditionnelle :

$$F_n(x;\, R_n^*) = \mathbb{P}[S_n \leq x \mid R_n^*] .$$

2 - LEMMES PRELIMINAIRES.

Pour alléger l'écriture, nous noterons $\mathbb{E}^*X$, $\mathbb{D}^{*2}X$, $\overset{*}{\mathrm{Cov}}(X,Y)$ les espérance variance et covariance de v.a. prises dans la distribution conditionnelle étant donné R_n^* :

$$\mathbb{E}^*X = \mathbb{E}(X|R_n^*), \text{ etc.}$$

Convenons également de dire (cfr. [3]) que S_n *est, sous les permutations des colonnes de* R_n, *asymptotiquement normale* (m_n, σ_n) si, pour tout x, $F_n\left(\frac{x-m_n}{\sigma_n} ;\, R_n^*\right)$ converge en probabilité, vers $\Phi(x)$, où $\Phi(x)$ est la fonction de répartition d'une v.a. normale réduite.

On dispose du critère de convergence suivant ([1])

Lemme 1. Si tous les moments conditionnels $u_r(\overset{*}{R}_n) = \mathbb{E}^*\left(\frac{S_n - m_n}{\sigma_n}\right)^r$ $(r = 1,2,\ldots)$ *de* $\frac{S_n - m_n}{\sigma_n}$ *convergent en probabilité vers ceux d'une v.a. normale,* S_n *est, sous les permutations des colonnes de* R_n*, asymptotiquement normale* $(m_n,\ \sigma_n)$.

Nous omettons la démonstration immédiate du lemme suivant.

Lemme 2.

$$\mathbb{D}^2 S_n =$$

$$\frac{1}{n}\{\overset{*}{\mathbb{D}}{}^2 J\left(\frac{R_1^n}{n+1},\frac{R_2^n}{n+1},\ldots,\frac{R_q^n}{n+1}\right)+2\sum_{s=1}^{q-1}\overset{*}{\mathrm{C}}\mathrm{ov}\left[J\left(\frac{R_1^n}{n+1},\frac{R_2^n}{n+1},\ldots,\frac{R_q^n}{n+1}\right),J\left(\frac{R_{1+s}^n}{n+1},\frac{R_{2+s}^n}{n+1},\ldots,\frac{R_{q+s}^n}{n+1}\right)\right]\}$$

$$+\ \frac{n-2q+1}{n}\,\overset{*}{\mathrm{C}}\mathrm{ov}\left[J\left(\frac{R_1^n}{n+1},\frac{R_2^n}{n+1},\ldots,\frac{R_q^n}{n+1}\right),J\left(\frac{R_{1+q}^n}{n+1},\frac{R_{2+q}^n}{n+1},\ldots,\frac{R_{2q}^n}{n+1}\right)\right]\ .$$

Si on désire connaître le comportement asymptotique de $n\,\overset{*}{\mathbb{D}}{}^2(S_n)$, il faut étudier le comportement des variances et covariances du premier terme de $\overset{*}{\mathbb{D}}{}^2(S_n)$, d'une part, mais surtout, d'autre part, celui de

$$\overset{*}{\mathrm{C}}\mathrm{ov}\ \left[J\left(\frac{R_1^n}{n+1},\ \frac{R_2^n}{n+1},\ \ldots\right),\ J\left(\frac{R_{1+q}^n}{n+1},\ \frac{R_{2+q}^n}{n+1},\ \ldots\right)\right]\ . \tag{2}$$

Dégageons de la covariance (2) les termes d'ordre $1/n^2$ et ceux qui sont négligeables (en probabilité) face à $1/n^2$.

Mais auparavant, introduisons la notation suivante :

$$\overset{*}{J}(R_{i_1}^n\ ,\ R_{i_2}^n,\ldots,R_{i_k}^n\ ;\ j_1,j_2,\ldots,j_k) \qquad \begin{array}{l}(1 \leq k \leq q,\\ \ 1 \leq i_1,i_2,\ldots,i_k \leq n\ ,\\ \ 1 \leq j_1,j_2,\ldots,j_k \leq q)\end{array}$$

sera l'espérance conditionnelle de $J\left(\frac{R_1^n}{n+1},\ \frac{R_2^n}{n+1},\ \ldots,\ \frac{R_q^n}{n+1}\right)$ étant donné $\overset{*}{R}_n$, $R_{i_1}^n$, $R_{i_2}^n$, ..., $R_{i_k}^n$, où, de plus, les rangs $R_{j_1}^n$, $R_{j_2}^n$, ..., $R_{j_k}^n$ apparaissant dans J ont été remplacés respectivement par $R_{i_1}^n$, $R_{i_2}^n$, ..., et $R_{i_k}^n$.

Posons également

$$\hat{J}'_n\left(\frac{R_1^n}{n+1},\frac{R_2^n}{n+1},\ldots,\frac{R_q^n}{n+1}\right) =$$

$$J\left(\frac{R_1^n}{n+1},\frac{R_2^n}{n+1},\ldots,\frac{R_q^n}{n+1}\right) - \frac{(n-1)(n-q+1)}{n(n-q)} \sum_{\ell=1}^{q} \overset{*}{J}(R_\ell^n;\ell) - \frac{n-1}{n(n-q)} \sum_{\ell_1,\ell_2=1}^{q} \overset{*}{J}(R_{\ell_1};\ell_2). \quad (3)$$

L'introduction de la fonction $\hat{J}'_n$ est cruciale : ce sont les termes rajoutés à J qui permettent d'obtenir une covariance (2) d'ordre $1/n^2$; et de plus, ces termes supplémentaires ne changent pas la valeur de S_n. Ces deux résultats sont repris dans les lemmes 3 et 5 suivants.

Lemme 3.

(i) $S_n - \overset{*}{\mathbb{E}}\, S_n = \frac{1}{n} \sum_{\alpha=1}^{n} \hat{J}'_n \left(\frac{R_\alpha^n}{n+1}, \frac{R_{\alpha+1}^n}{n+1}, \ldots, \frac{R_{\alpha+q-1}^n}{n+1}\right)$

(ii) $\overset{*}{\widehat{\hat{J}'_n(R_{\ell_1};\ell_2)}} \equiv 0 \; (\ell_1,\ \ell_2 = 1,2,\ldots,q)$

(iii) $\overset{*}{\mathbb{D}}{}^2(S_n) = \frac{1}{n} \{ \overset{*}{\mathbb{D}}{}^2\ \hat{J}'_n\left(\frac{R_1^n}{n+1}, \frac{R_2^n}{n+1}, \ldots\right)$

$$+ 2 \sum_{s=1}^{q-1} \overset{*}{\mathrm{Cov}} \left[\hat{J}'_n\left(\frac{R_1^n}{n+1}, \frac{R_2^n}{n+1}, \ldots\right), \hat{J}'_n\left(\frac{R_{1+s}^n}{n+1}, \frac{R_{2+s}^n}{n+1}, \ldots\right)\right] \} + B_n ,$$

où

$$B_n = \sum_{\ell=2}^{q} \sum_{1\le i_1,i_2,\ldots,i_\ell\le q} \sum_{\substack{1\le j_1,j_2,\ldots,j_\ell\le q \\ j_k\ne j_i\ (k\ne i)}} (-1)^\ell \frac{\binom{n-\ell}{q-\ell}(q-\ell)!}{\binom{n-q}{q}q!} \frac{n-2q+1}{n} \times$$

$$\overset{*}{\mathbb{E}} \left[\hat{J}'\left(\frac{R_1^n}{n+1},\ldots\right), \overset{*}{\widehat{\hat{J}'_n(R_{i_1}^n, R_{i_2}^n, \ldots, R_{i_\ell}^n\ ;\ j_1,j_2,\ldots,j_\ell)}}\right]$$

Démonstration : Les points (i) et (ii) sont établis par un calcul direct. Le point (iii) est une conséquence immédiate du résultat suivant :

Lemme 4.

$$\mathbb{E}^{*}\left[J\left(\frac{R_1^n}{n+1},\dots\right) \Big| R^n_{q+1},\ R^n_{q+2},\dots,\ R^n_{q+k}\right] =$$

$$\binom{n}{q}\,\mathbb{E}^{*}\, J\left(\frac{R_1^n}{n+1},\dots\right)+\sum_{\ell=1}^{k}\ \sum_{q+1\le i_1<i_2<\dots<i_\ell\le q+k}\ \sum_{\substack{1\le j_1,j_2,\dots,j_q\le q\\ j_r\neq j_i(r\neq i)}}$$

$$(-1)^{\ell}\frac{\binom{n-\ell}{q-\ell}(q-\ell)!}{\binom{n-k}{q}q!}\ \mathfrak{J}^{*}(R^n_{i_1},R^n_{i_2},\dots,R^n_{i_\ell};\ j_1,j_2,\dots,j_\ell)\ .$$

Démonstration.

On a

$$\mathbb{E}^{*}\left[J \,|\, R^n_{q+1},\dots,R^n_{q+k}\right]=$$

$$\frac{1}{(n-k)(n-k-1)\dots(n-k-q+1)}\ \sum_{r_1\dots r_q} J(R^n_{r_1}/n+1,\dots,R^n_{r_q}/n+1)\ ,$$

où les indices $r_1,\ r_2,\ \dots,r_q$ parcourent l'ensemble $\{1,\dots,n\}\setminus\{q+1,\dots,q+k\}$ en restant distincts entre eux. Cette sommation se décompose en

$$\Sigma = \Sigma^{(o)} - \Sigma^{(1)}\ ,$$

où $\Sigma^{(o)}$ est une somme étendue à tous les indices distincts entre eux et $\Sigma^{(1)}$ une somme étendue à tous les q-uples d'indices $r_1,\dots,r_q$ ayant une composante au moins dans $\{q+1,\dots,q+k\}$. $\Sigma^{(1)}$ se décompose à son tour en une double somme

$$\Sigma^{(1)} = \Sigma^{1.1}\ \Sigma^{1.2}\ ,$$

où $\Sigma^{1.1}$ s'étend à tous les choix (combinaisons) d'un ensemble de q colonnes pour les q indices $r_1,\ \dots,\ r_q$ comprenant au moins une des colonnes $R^n_{q+1},\ \dots,\ R^n_{q+k}$ et où $\Sigma^{1.2}$ s'étend à tous les arrangements de ces q colonnes entre les q indices $r_1,\ r_2,\ \dots,\ r_q$. Mais par application du principe d'inclusion-exclusion, $\Sigma^{1.1}$ se décompose à son tour en

$$\Sigma^{1.1}=\{\Sigma_1^{1.1}+\Sigma_2^{1.1}+\dots+\Sigma_q^{1.1}\} = \{\Sigma_{12}^{1.1}+\Sigma_{13}^{1.1}+\dots+\Sigma_{1q}^{1.1}+\Sigma_{23}^{1.1}+\Sigma_{24}^{1.1}+\dots+\Sigma_{q-1q}^{1.1}\} +$$

$$\{\Sigma_{123}^{1.1} + \dots\} - \{\Sigma_{1234}^{1.1} + \dots\} + \dots\ ,$$

où $\Sigma_1^{1.1}$ est une somme étendue à tous les choix d'un ensemble de q colonnes comprenant au moins R_{q+1}^n, $\Sigma_2^{1.1}$ s'étend à tous les choix d'un ensemble de q colonnes comprenant au moins R_{q+2}^n ; semblablement $\Sigma_{12}^{1.1}$ s'étend à tous les choix d'un ensemble de q colonnes comprenant R_{q+1}^n et R_{q+2}^n et ainsi de suite pour les autres termes.
Regardons à présent $\Sigma_1^{1.1}\ \Sigma^{1.2} + \Sigma_2^{1.1}\ \Sigma^{1.2} + \ldots + \Sigma_q^{1.1}\ \Sigma^{1.2}$. Cette somme est équivalente à la triple somme

$$\Sigma^1_{(1)}\ \Sigma^2_{(1)}\ \Sigma^3_{(1)}\ ,$$

où $\Sigma^1_{(1)}$ s'étend à tous les choix d'une seule colonne priseparmi les colonnes $R_{q+1}^n, \ldots, R_{q+k}^n$, où $\Sigma^2_{(1)}$ s'étend à toutes les façons d'attribuer à l'un des indices $r_1, r_2, \ldots, r_q$ cette colonne précédemment choisie et où $\Sigma_1^{(3)}$ s'étend à toutes les façons de choisir q-1 colonnes distinctes (entre elles et de la précédente) pour les indices restants $r_1, \ldots, r_q$.
Semblablement $\Sigma_{12}^{1.1}\ \Sigma^{1.2} + \ldots + \Sigma_{q-1q}^{1.1}\ \Sigma^{1.2}$ est équivalente à une triple somme $\Sigma^1_{(2)}\ \Sigma^2_{(2)}\ \Sigma^3_{(2)}$ où $\Sigma^1_{(2)}$ s'étend à tous les choix d'un ensemble de 2 colonnes parmi $R_{q+1}^n, \ldots, R_{q+k}^n$.
Analysons chacune de ces sommes $\Sigma_\ell^{(1)}\ \Sigma_\ell^{(2)}\ \Sigma_\ell^{(3)}$ ($\ell = 1,2, \ldots k$). On choisit ℓ colonnes $R_{i_1}^n, \ldots, R_{i_\ell}^n$; on "attribue" à ℓ indices $r_{j_1}, r_{j_2}, \ldots, r_{j_\ell}$ ces colonnes; on effectue ensuite une sommation étendue à toutes les manières d'attribuer q-ℓ colonnes distinctes (entre elles et distinctesde $R_{i_1}^n, \ldots, R_{i_\ell}^n$) aux indices restants. Cette dernière sommation est, à une constante multiplicative près, l'espérance conditionnelle $\mathfrak{J}(R_{i_1}^n, R_{i_2}^n, \ldots, R_{i_\ell}^n ; j_1, j_2, \ldots, j_\ell)$.
La constante multiplicative est un rapport de probabilitésconditionnelles. Ce qui termine la démonstration.

Lemme 5. Si $J(\frac{R_{i_1}^n}{n+1}, \frac{R_{i_2}^n}{n+1}, \ldots, \frac{R_{i_q}^N}{n+1})$ *converge en moyenne quadratique vers* $\mathbb{E}\, J(V_{i_1}, V_{i_2}, \ldots, V_{i_q})$ *où* $V_i = (F_1(X_{1i}), \ldots, F_q(X_{qi}))$ *et* $1 \leq i_\ell \leq q$ ($\ell=1,2,..,q$), *alors le terme* B_n *du lemme 3 (iii) est un* $o_{\mathbb{P}}(1/n)$.

Démonstration. B_n est formé d'une somme dont le nombre de termes est indépendant de n. Etudions chacun de ces termes.

Pour $\ell \geqslant 2$, on a :

$$\frac{\binom{n-\ell}{q-\ell}(q-\ell)!}{\binom{n-q}{q}q!} \quad \frac{n-2q+1}{n} = \mathbf{O}\left(\frac{1}{n}\right) .$$

Il reste à démontrer que :

$$\overset{*}{\mathbb{E}} \left[\hat{J}'_n\left(\frac{R_1}{n+1}, \dots\right) \overbrace{\hat{J}'_n (R_{i_1}, \dots ; j_1, \dots)}^{*}\right] = 0_{\mathbb{P}}(1) .$$

Cette égalité se vérifie aisément en appliquant l'inégalité de Schwartz, la C_r inégalité $(\mathbb{E}(X+Y+Z)^2 \leqslant 8(\mathbb{E}\, X^2 + \mathbb{E}\, Y^2 + \mathbb{E}\, Z^2)$ et l'inégalité suivante :

$$\mathbb{E}\ \overset{*}{J}{}^2(R_{i_1}, R_{i_2}, \dots ; j_1, j_2, \dots) \leqslant \mathbb{E}\ J^2(R_1^n, \dots, R_{i_1}^n, \dots) = O(1) \qquad (5)$$

4 - CONVERGENCE EN PROBABILITE DE $\overset{*}{\mathbb{E}}\, S_n$ et de $\sqrt{n}\, \overset{*}{\mathbb{D}}(S_n)$

Etablissons la convergence en probabilité de $\overset{*}{\mathbb{E}}\, S_n$ et de $\sqrt{n}\, \overset{*}{\mathbb{D}}\, S_n$.

Théorème 1. Si $J\left(\frac{R_{i_1}^n}{n+1}, \frac{R_{i_2}^n}{n+1}, \dots, \frac{R_{i_q}^n}{n+1}\right)$ converge en moyenne quadratique vers $J(V_{i_1}, V_{i_2}, \dots, V_{i_q})$ $(1 \leqslant i_\ell \leqslant q,\ \ell = 1,2,\dots q)$ alors

$$\text{(i) } \overset{*}{\mathbb{E}}\, S_n \xrightarrow{\mathbb{P}} \mathbb{E}\ J(V_1, V_2, \dots, V_q)$$

$$\text{(ii) } \sqrt{n}\, \overset{*}{\mathbb{D}}(S_n) \xrightarrow{\mathbb{P}} \sigma ,$$

où $\sigma^2 = \mathbb{D}^2\ J'(V_1,V_2,\dots,V_q) + 2 \sum_{s=1}^{q-1} \mathrm{Cov}\ [J'(V_1,\dots), J'(V_{1+s}, \dots)]$,

et $J'(V_1,V_2,\dots,V_q) = J(V_1,V_2,\dots,V_q) - \sum_{\ell=1}^{q} \mathbb{E}[J(V_1,V_2,\dots,V_q) | V_\ell]$.

Démonstration. Le premier point est une conséquence de la propriété générale de convergence en probabilité des statistiques $\mathfrak{U}$([4]) :

$$T_n = \frac{1}{\binom{n}{q}q!} \Sigma_{i_1 \neq i_2 \neq \dots \neq i_q} J(V_{i_1}, V_{i_2}, \dots, V_{i_q}) \xrightarrow{\mathbb{P}} \mathbb{E}\ J(V_1,V_2,\dots,V_q) .$$

En effet, on a :

$$\mathbb{E}[\overset{*}{\mathbb{E}} S_n - T_n]^2 = \frac{1}{\binom{n}{q}^2 (q!)^2} \Sigma_{1\leq j_1\neq j_2\neq..\neq j_q\leq n} \Sigma_{1\leq i_1\neq i_2\neq..\neq i_q\leq n}$$

$$\mathbb{E}[J(\frac{R^n_{i_1}}{n+1},\frac{R^n_{i_2}}{n+1},\ldots,\frac{R^n_{i_q}}{n+1})-J(V_{i_1},\ldots,V_{i_q})][J(\frac{R^n_{j_1}}{n+1},\ldots,\frac{R^n_{j_q}}{n+1})-J(V_{j_1},\ldots,V_{j_q})] .$$

Mais quels que soient les indices, on a

$$\mathbb{E}[J(\frac{R^n_{i_1}}{n+1},\ldots,\frac{R^n_{i_q}}{n+1})-J(V_{i_1},\ldots,V_{i_q})]^2 = \mathbb{E}[J(\frac{R^n_1}{n+1},\frac{R^n_2}{n+1},\ldots,\frac{R^n_q}{n+1})-J(V_1,V_2,\ldots,V_q)]^2 ,$$

et par conséquent

$$\mathbb{E}|\overset{*}{\mathbb{E}} S_n - T_n|^2 \leq \mathbb{E}[J(\frac{R^n_1}{n+1},\frac{R^n_2}{n+1},\ldots,\frac{R^n_q}{n+1}) - J(V_1,V_2,\ldots,V_q)]^2 \to 0 .$$

Pour établir le point (iii), on remarque que

$$\dot{J}'_n(\frac{R_1}{n+1},\frac{R_2}{n+1},\ldots,\frac{R_q}{n+1}) = J(\frac{R_1}{n+1},\frac{R_2}{n+1},\ldots) - \sum_{\ell=1}^{q} \overset{*}{J}(R_\ell;\ell) + o_{\mathbb{P}}(1) .$$

Il suffit alors de s'appuyer sur un raisonnement semblable à celui qui a été développé pour la démonstration du lemme 5 et sur les égalités (3) et (4).

5 - NORMALITE ASYMPTOTIQUE DE S_n.

Notre but est d'obtenir un théorème du type central limite pour la distribution conditionnelle de S_n.

Théorème 2. Si pour tout $r=1,2,\ldots$ $J^r(\frac{R^n_{i_1}}{n+1},\frac{R^n_{i_2}}{n+1},\ldots,\frac{R^n_{i_q}}{n+1})$ $(1 \leq i_\ell \leq q,$ $\ell = 1,2,\ldots,q)$ *converge en moyenne quadratique vers* $J^r(V_{i_1}, V_{i_2}, \ldots, V_{i_q})$, *alors sous les permutations des colonnes de* R_n, S_n *est asymptotiquement normale* $(m, \sigma/\sqrt{n})$ *où* m *et* σ *sont donnés par le théorème 1.*

Démonstration. En vue d'utiliser le lemme 1, nous allons calculer les moments conditionnels

$$\overset{*}{\mathbb{E}}[\sqrt{n}(S_n - \overset{*}{\mathbb{E}} S_n)]^r \qquad (r = 1,2,\ldots).$$

Pour simplifier l'écriture posons $a_\alpha = \hat{J}'_n\left(\frac{R^n_\alpha}{n+1}, \ldots, \frac{R^n_{\alpha+p}}{n+1}\right)$. On a

$$\overset{*}{\mathbb{E}}\, n^r (S_n - \overset{*}{\mathbb{E}}\, S_n) = \overset{*}{\mathbb{E}}\Big(\sum_{\alpha=1}^{n} a_\alpha\Big)^r = \sum_{\alpha_1,\alpha_2 .. \alpha_r = 1}^{n} \overset{*}{\mathbb{E}}\, a_{\alpha_1} a_{\alpha_2} \ldots a_{\alpha_r}$$

Décomposons cette somme dont le nombre de termes varie avec n en une somme de termes convergeant en probabilité,et en nombre indépendant de n.
Pour chaque $\alpha_1, \alpha_2, \ldots, \alpha_r$ posons :

$$\beta_1 = \alpha_{j_1^1} = \alpha_{j_2^1} = \ldots = \alpha_{j_{i_1}^1} < \beta_i = \alpha_{j_1^2} = \alpha_{j_2^2} = \ldots = \alpha_{j_{i_2}^2} < \ldots <$$

$$\beta_p = \alpha_{j_1^p} = \alpha_{j_2^p} = \ldots = \alpha_{j_{i_p}^p} \leqslant n \, .$$

Pour chaque $1 \leqslant \beta_1 < \beta_2 < \ldots < \beta_p \leqslant n$ on distingue

1) le nombre g $(1 \leqslant g \leqslant r)$ de groupes constitués par des β_i successifs tels que

$$1 \leqslant \beta_{i+1} - \beta_i \leqslant q-1 \, .$$

2) le nombre $\ell_1, \ell_2, \ldots, \ell_g$ d'indices α de chaque groupe.

3) pour chaque groupe les différences premières Δ_i^j entre indices successifs d'un même groupe ordonnés par ordre croissant.

On a alors

$$n^r \overset{*}{\mathbb{E}}\, (S_n - \overset{*}{\mathbb{E}}\, S_n)^r =$$

$$\sum_{g=1}^{r} \sum_{\substack{\ell_1+\ell_2+..+\ell_g=r \\ \ell_i \geqslant 0}} \sum_{\begin{cases} 0 \leqslant \Delta_1^1, \Delta_2^1, \ldots, \Delta_{\ell_1-1}^1 \leqslant q-1 \\ \vdots \\ 0 \leqslant \Delta_1^g, \Delta_2^g, \ldots\, \Delta_{\ell_g-1}^g \leqslant q-1 \end{cases}} \sum_{1<\beta_1<\beta_2 ..<\beta_p} \frac{r!}{i_1! .. i_p!} \overset{*}{\mathbb{E}}\, a_{\beta_1}^{i_1} \text{-} a_{\beta_p}^{i_p} \, ,$$

où la dernière somme est étendue aux β compatibles avec les sommations précédentes. On remarque que la distribution de $a_{\beta_1} \ldots a_{\beta_p}$ ne dépend que de la configuration d'un groupe (ℓ_k et Δ_j^i) et non des variables particulières. Or s'il y a g groupes comprenant au total L_g indices distincts, il y a $\binom{n-L_g+g}{g}$ choix de β compatibles avec les sommations (pour ℓ_k et Δ_j^i fixés). On a alors

$$n^{r/2}\,\mathbb{E}^*(S_n - \mathbb{E}^* S_n)^r = \frac{1}{n^{r/2}} \sum_{g=1}^{r} \underbrace{\sum}_{\substack{\ell_1+\ell_2+..+\ell_q=r \\ \ell_1 \geq 0}} \sum_{\begin{cases} 0\leq \Delta_1^1 \Delta_2^1 .., \Delta_{\ell_1-1}^1 \leq q-1 \\ \vdots \\ 0\leq \Delta_1^g, \Delta_2^g, .., \Delta_{q-1}^g \leq q-1 \end{cases}}$$

$$\frac{r!}{i_1!\dots i_p!} \binom{n-L_g+g}{g} \mathbb{E}^* A_1 A_2 \dots A_g$$

où les $A_1 \dots A_g$ sont des expressions du type $a_{\beta_j}^{i_j}\, a_{\beta_{j+1}}^{i_{j+1}}$, ... $a_{\beta_k}^{i_k}$ avec pour $\beta_j\ \beta_{j+1} \dots \beta_k$ un choix quelconque satisfaisant aux conditions du groupe. On remarque que les sommes subsistant sont indépendantes de n.

Si $g < r/2$, $\binom{n-L_g+g}{g} \frac{1}{n^{r/2}}\, \mathbb{E}^* A_1 \dots A_g = \frac{O(n^g)}{n^{r/2}} . \mathbb{E}^* A_1 \dots A_g \xrightarrow{\mathbb{P}} 0$

($A_1 \dots A_g$ est borné en probabilité par hypothèse).

Si $g > r/2$, il y a au moins 2g-r facteurs A_i formés d'un et un seul a_{β_i} et en vertu du lemme 3

$$\binom{N-L_g+g}{g} \frac{1}{n^{r/2}} \mathbb{E}^* A_1 \dots A_g = \frac{O(n^g)}{n^{r/2}} \quad o_{\mathbb{P}}\left(\frac{1}{n^{2g-r}}\right) \xrightarrow{\mathbb{P}} 0$$

Dès à présent, si r est impair ,

$$\mathbb{E}^* [\sqrt{n}\,(S_n - \mathbb{E}^* S_n)]^r \xrightarrow{\mathbb{P}} 0 .$$

Si g est égal à r/2, soit il y a r/2 facteurs A_i composés de deux β, soit il y a au moins un facteur A_i composé d'un et un seul β, et dans ce cas, en vertu du lemme 3

$$\binom{n-L_{r/2}+r/2}{r/2} \frac{1}{n^{r/2}} \mathbb{E}^* A_1 \dots A_{r/2} = \frac{O(n^{r/2})}{n^{r/2}} \; o_{\mathbb{P}}\left(\frac{1}{n}\right) \xrightarrow{\mathbb{P}} 0 .$$

S'il y a r/2 facteurs A_i composés de deux a_β, on a par hypothèse :

$$\binom{n-L_{r/2}+r/2}{r/2} \overset{*}{\mathbb{E}}\, A_1 \ldots A_{r/2} \xrightarrow{\mathbb{P}} \mathbb{E}a_1 a_{1+\Delta^1} \cdot \mathbb{E}a_1 a_{1+\Delta^2} \cdots \cdot \mathbb{E}a_1 a_{1+\Delta^{r/2}} \cdot$$

Dès lors, pour r pair , on a

$$\overset{*}{\mathbb{E}}\, [\sqrt{n}(S_n - \overset{*}{\mathbb{E}}\, S_n)]^r \xrightarrow{\mathbb{P}} \sum_{0 \leq \Delta^1, \Delta^2 .. \Delta^{r/2} \leq q-1} \frac{r!}{(r/2)!\, i_1! .. i_m!}$$

$$\mathbb{E}a_1 a_{1+\Delta^1} \ldots \mathbb{E}a_1 a_{1+\Delta^{r/2}}$$

$$= \frac{r!}{(r/2)!} [\frac{\mathbb{E}\, a_1^2}{2} + \mathbb{E}\, a_1 a_2 + \ldots + \mathbb{E}a_1 a_q]^{r/2}$$

$$= \frac{r!}{(r/2)!\, 2^{r/2}} [\mathbb{E}a_1^2 + 2\, \mathbb{E}a_1 a_2 + \ldots + 2\, \mathbb{E}a_1 a_q]^{r/2}$$

$$= \frac{r!}{(r/2)!\, 2^{r/2}} \sigma^r ,$$

ce qui termine la démonstration puisque les moments d'ordre pair d'une variable normale réduite sont donnés par $\frac{r!}{(r/2)!\ 2^{r/2}}$.

6 - BIBLIOGRAPHIE

[1] FRASER, D.A.S. (1957), *Nonparametric Methods in Statistics,* Wiley.

[2] INGENBLEEK, J-F. (1980), "Un test de rang bruit-blanc multivarié", *C.E.R.O.* Vol.22, n°3-4, pp.369-373.

[3] PURI, M.L., SEN, P.K. (1971), *Nonparametric Methods in Multivariate Analysis,* Wiley.

[4] SEN, P.K. (1977), "Almost Convergence of Generalized U-Statistics", *Ann. of Prob.,* Vol.5, n°2, 287-290.