Revue de statistique appliquée

Laurent Leprêtre

Lois limites du Jackknife de statistiques associées à des fonctionnelles dérivables au sens de Von Mises

Revue de statistique appliquée, tome 27, nº 1 (1979), p. 55-79

<http://www.numdam.org/item?id=RSA_1979__27_1_55_0>

LOIS LIMITES DU JACKKNIFE DE STATISTIQUES ASSOCIÉES A DES FONCTIONNELLES DÉRIVABLES AU SENS DE VON MISES

Laurent LEPRÊTRE

La technique de JACKKNIFE est une technique statistique qui est utilisée soit pour réduire le biais des estimateurs soit pour construire des tests ou intervalles de confiance asymptotiques robustes. Considérant ce dernier aspect de la théorie du JACKKNIFE, nous nous proposons de montrer que l'étude des lois limites du JACKKNIFE d'une statistique peut être abordée par l'examen des propriétés des dérivées au sens de VON MISES, des fonctionnelles associées à cette statistique.

Alors qu'une étude directe de la distribution du JACKKNIFE d'une statistique particulière conduit assez souvent à des calculs complexes masquant la nature de la transformation du JACKKNIFE, la théorie des fonctionnelles de VON MISES permet de traiter de manière synthétique l'étude des distributions asymptotiques du JACKKNIFE et apparaît comme un outil particulièrement efficace du point de vue des applications.

1. SOMMAIRE

Soit $(X_1, \ldots, X_N)$ un échantillon aléatoire de taille $N = nm$ tiré d'une v.a.r. dont la loi de probabilité dépend d'un paramètre réel θ et soit T un estimateur de θ; Considérons la partition suivante de $D_n = \{X_1, \ldots, X_N\}$ en n sous-ensembles à m éléments :

$$\{X_1, \ldots, X_m\}, \ldots, \{X_{(i-1)m+1}, \ldots, X_{im}\}, \ldots, \{X_{(n-1)m+1}, \ldots, X_N\}$$

$\{X_1, \ldots, X_{(i-1)m}, X_{im+1}, \ldots, X_N)$ désignant l'échantillon aléatoire de taille $N - m$ obtenu après suppression dans l'échantillon initial des observations appartenant au i sous-ensemble de la partition, posons :

$$T = T[X_1, \ldots, X_N], \quad T^{(1)} = T[X_{m+1}, \ldots, X_N]$$

$$T^{(i)} = T[X_1, \ldots, X_{(i-1)m}, \quad X_{im+1}, \ldots, X_N; \quad i = 2, \ldots, n-1;$$

$$T^{(n)} = T[X_1, \ldots, X_{(n-1)m}];$$

$$J_i[T] = nT - (n-1)T^{(i)}; \quad i = 1, \ldots, n; \tag{1.1}$$

$$J[T] = n^{-1} \sum_{i=1}^{n} J_i[T]. \tag{1.2}$$

Mots-clés : **Jackknife, lois limites, intervalles de confiance, robustesse, fonction d'influence, U-statistiques.**

A toute partition de $\{X_1, \ldots, X_N\}$ du type défini précédemment, il est donc possible d'associer une transformation J désignée communément sous le nom de 'JACKKNIFE'. J (T) est appelé le JACKKNIFE de la statistique T et $J_i(T)$ représente la $i^{\text{ème}}$ pseudo-valeur du JACKKNIFE. Nous nous proposons d'étudier la convergence en loi de la v.a.r. :

$$\frac{n^{1/2}\,[J(T) - \theta]}{\tilde{s}} \tag{1.3}$$

où :

$$\tilde{s}^2 = [n-1]^{-1} \sum_{i=1}^{n} [J_i[T] - J[T]]^2 \tag{1.4}$$

dans les différents cas suivants :

1) La taille m des blocs d'observations reste fixée, alors que le nombre de blocs n croît indéfiniment.

2) Le nombre de blocs d'observations n est fixé, tandis que m la taille de blocs croit indéfiniment.

3) n et m tendent vers l'infini simultanément.

Pour une classe de statistique T associées à des fonctionnelles dérivables au sens de VON MISES, nous démontrons que la distribution de la v.a.r. (1.3) tend respectivement vers la loi de STUDENT à $(n-1)$ degrés de liberté dans le premier cas et vers la loi normale centrée, réduite dans les cas 2 et 3.

La méthode employée consiste à remarquer que chacun des termes du développement de TAYLOR d'une statistique associée à une fonctionnelle dérivable au sens de VON MISES est décomposable en une combinaison linéaire de U-statistiques. En utilisant la linéarité de la transformation du JACKKNIFE, il est alors possible de déterminer la distribution asymptotique de la v.a.r. (1.3) en appliquant les théorèmes relatifs aux distributions limites du JACKKNIFE de statistiques fonctions de U-statistiques. (cf. [1]).

L'utilisation de la théorie des fonctionnelles de MISES fait d'autre part apparaître les points suivants :

1) La notion de pseudo-valeur de JACKKNIFE prend une signification concrete, lorsqu'elle est exprimée par rapport à la fonction d'influence.

2) Parmi les différentes transformations du JACKKNIFE il est possible de déterminer celle qu'il convient d'appliquer à une statistique, dans le cadre d'un modèle probabiliste donné.

3) La stabilité et la robustesse du JACKKNIFE d'une statistique se revèlent être dépendantes de la continuité des dérivées des fonctionnelles de VON MISES associées à cette statistique.

2. INTRODUCTION

Désignons par F_N la fonction de répartition empirique associée à un échantillon $(X_1, \ldots, X_N)$ de taille N tiré d'une v.a.r. X de fonction de répartition F. La valeur d'une statistique quelconque au point $(X_1, \ldots, X_N)$ peut être considérée comme l'image du point F_N par une fonctionnelle à valeurs réelles, définie sur un ensemble $\mathcal{F}$ de fonctions de répartition, Notons que pour un grand nombre de statistiques, cette fonctionnelle associée est connue. (cf. [7]).

Définition 2.1

Soit $\mathcal{C} \subset \mathcal{F}$ un ensemble de fonctions de répartition tel que : $\forall G \in \mathcal{C}$ et $\forall t \in [0, 1]$, on ait : $F + t(G - F) \in \mathcal{C}$. Une fonctionnelle $T : \mathcal{F} \to R$ est dite m fois dérivable au point $F \in \mathcal{F}$ par rapport à l'ensemble $\mathcal{C} \subset \mathcal{F}$ lorsque :

1) $\forall t \in [0, 1]$; $\forall G \in \mathcal{C}$; $\forall p = 1, \dots, m$;

$$\frac{d^p}{dt^p} T[F + t(G - F)] \quad \text{existe}$$

2) Pour $p = 1, \dots, m$, il existe des fonctionnelles $T^{(p)}[F ; y_1, \dots, y_p]$, dépendant de F et telles que :

$$\forall G \in \mathcal{C}, \quad \forall p = 1, \dots, m,$$

$$\frac{d^p}{dt^p} T[F + t(G - F)]_{t=0} = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} T^{(p)}[F ; y_1, \dots, y_p] \prod_{i=1}^{p} d[G(y_i) - F(y_i)] .$$

Remarque 2.2

Notons que parmi les fonctions

$$(y_1, \dots, y_p) \to T^{(p)}[F ; y_1, \dots, y_p]$$

correspondant à la définition précédente, il existe toujours une fonction invariante par rapport aux permutations des arguments : $y_1, \dots, y_p$ entre eux et telle que :

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} T^{(p)}[F ; y_1, \dots, y_p] \prod_{i=1}^{p} dF(y_i) = 0 .$$

Définition 2.3

$\mathcal{F}$ étant un ensemble de fonctions de répartition, on désigne par fonctionnelle de Mises d'ordre m, au point $F \in \mathcal{F}$, une fonctionnelle $T : \mathcal{F} \to R$ telle que :

1) il existe $\mathcal{C}_T \subset \mathcal{F}$ satisfaisant les propriétés suivantes :

a) $F \in \mathcal{C}_T$

b) $\forall G \in \mathcal{C}_T$ et $\forall t \in [0, 1]$, $F + t(G - F) \in \mathcal{C}_T$

c) $\lim\limits_{N \to +\infty} P[F_N \in \mathcal{C}_T] = 1$

2) T est m fois dérivable au point $F \in \mathcal{F}$ par rapport à l'ensemble $\mathcal{C}_T$.

3) $\forall \epsilon > 0$, $\forall \delta > 0$ et $\forall p = 1, \ldots, m$, on a :

$$\lim_{N \to +\infty} P \left[N^{(p/2)-\delta} \sup_{t \in [0,1]} \left| \frac{d^p}{dt^p} T[F_N^t] \right| > \epsilon \right] = 0$$

pour $F_N^t = F + t[F_N - F]$.

Définition 2.4

Notons $\overline{L}_p^{(m)}(F)$ la classe des fonctions $f : R^m \to R$ telles que :

$$\forall r \in \{1, \ldots, m\}$$

et quelle que soit l'application surjective : $k \to i_k$ de $\{1, \ldots, m\}$ dans $\{1, \ldots, r\}$, on ait :

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} |f^p(x_{i_1}, \ldots, x_{i_m})| \prod_{j=1}^{r} dF(x_j) < +\infty$$

On vérifie aisément que l'espace $\overline{L}_2^{(m)}(F)$ muni de la norme :

$$\|f\|_{\overline{L}_2^{(m)}(F)} = \sum_{r=1}^{m} \sum_{\{i_1, \ldots, i_m\} = \{1, \ldots, r\}} \left[\int \ldots \int f^2(x_{i_1}, \ldots, x_{i_m}) \prod_{j=1}^{r} dF(x_j) \right]^{1/2}$$

est un espace de BANACH (cf. [4]).

Soient a et ℓ deux entiers naturels positifs et φ une fonction de $R^{2a+\ell}$ dans R. $\mathcal{F}$ étant un ensemble de fonctions de répartition vérifiant les conditions 1 de la définition (2.3), considérons la fonctionnelle $T : \mathcal{F} \to R$ définie par la relation :

$$T[F] = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \varphi[x_1, \ldots, x_a ; F(x_1), \ldots, F(x_a) ; S_1(F), \ldots, S_\ell(F)] \prod_{i=1}^{b} dG_i(x_i) \prod_{i=b+1}^{a} dF(x_i)$$

$$0 \leqslant b \leqslant a \qquad (2.0)$$

où les G_i ; $i = 1, \ldots, b$; représentent des fonctions monotones, non décroissantes et les S_i, $i = 1, \ldots, \ell$; sont des fonctionnelles de $\mathcal{F}$ dans R.

Posons :

$$x = (x_1, \ldots, x_a) ; \quad F(x) = (F(x_1), \ldots, F(x_a)) ;$$
$$S = (S_1, \ldots, S_\ell) ; \quad u = (u_1, \ldots, u_a) .$$

Théorème 2.5

1) Supposons que $S_1, \ldots, S_\ell$ soient des fonctionnelles de Mises d'ordre m au point F.

2) Pour tout point $(S_1, \ldots, S_\ell)$ appartenant à un voisinage du point $(S_1(F), \ldots, S_\ell(F))$ et quels que soient les u_i ; $i = 1, \ldots, a$; tels que : $0 \leqslant u_i \leqslant 1$, supposons que les dérivées partielles :

$$\frac{\partial^{p+q} \varphi(x, u, S)}{\partial u^\mu \, \partial S^{\sigma+\sigma'}}$$

où : $0 \leqslant \mu_i \leqslant p$; $1 \leqslant i \leqslant a$; $0 \leqslant \sigma_j \leqslant p$; $1 \leqslant j \leqslant \ell$;

$$\sum_{i=1}^{a} \mu_i + \sum_{j=1}^{\ell} \sigma_j = p \; ; \; p = 0, 1, \ldots, m \quad \text{si} \quad m > 1 \; ;$$

$$\sum_{j=1}^{\ell} \sigma'_j = q \; ; \; 0 \leqslant \sigma'_j \leqslant 1 \; ; \; j = 1, \ldots, \ell \; ;$$

existent, soient continues par rapport à l'ensemble des arguments u, S et soient inférieures en module à $f(x)$, f étant une fonction de R^a dans R telle que :

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f(x) \prod_{i=1}^{b} dG_i(x_i) \in \overline{L}_2^{(a-b)}(F) \; .$$

Sous ces hypothèses, la fonctionnelle T définie en (2.0) est une fonctionnelle de Mises d'ordre m au point F.

Remarque 2.5

Le théorème (2.5) est un corollaire du théorème (2.6) (cf. [4]) qui établit une classe de fonctionnelles de Mises T définies par une relation de la forme :

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \psi[T(F) ; x ; F(x) ; S(F)] \prod_{i=1}^{b} dG_i(x_i) \prod_{i=b+1}^{a} dF(x_i) = 0 \tag{2.1}$$

où ψ est une fonction de $R^{2a+\ell+1}$ dans R,

$$x = (x_1, \ldots, x_a), \quad F(x) = (F(x_1), \ldots, F(x_a))$$

et $\quad S(F) = (S_1(F), \ldots, S_\ell(F))$.

Il est à noter que de nombreuses statistiques d'usage courant sont associées à ce type de fonctionnelles (cf. paragraphe 4).

Posons : $F_N^t = F + t(F_N - F)$; $t \in [0, 1]$, ainsi que :

$$H(t, T) = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \psi[T, x, F_N^t(x), S(F_N^t)] \prod_{i=1}^{b} dG_i(x_i) \prod_{i=b+1}^{a} dF_N^t(x_i)$$

où : $\qquad 0 \leqslant b \leqslant a \, . \qquad (2.2)$

D'après la relation (2.1), $T(F_N^t)$ est une valeur de T telle que :

$$H[t, T] = 0 \; ; \quad \forall t \in [0, 1] \, .$$

Théorème 2.6 ([4]).

Supposons les hypothèses suivantes vérifiées :

1) Il existe un réel T_0 tel que : $H[0, T_0] = 0$ et $\dfrac{\partial}{\partial T} H[0, T] \neq 0$.

2) Les fonctionnelles $S_1, \ldots, S_\ell$ sont des fonctionnelles de Mises d'ordre m au point $F \in \mathcal{F}$.

3) Pour tout point $(T, S_1, \ldots, S_\ell)$ appartenant à un voisinage du point $(T_0, S_1(F), \ldots, S_\ell(F))$ et quels que soient les u_i ; $1 \leqslant i \leqslant a$; tels que $0 \leqslant u_i \leqslant 1$, les dérivées partielles :

$$\frac{\partial^p \psi[T, x, u, S]}{\partial T^\tau \partial u^\mu \partial S^\sigma}$$

où : $0 \leqslant \tau \leqslant p$; $0 \leqslant \mu_i \leqslant p$; $1 \leqslant i \leqslant a$; $0 \leqslant \sigma_j \leqslant p$;

$$j = 1, \ldots, \ell ; \quad \tau + \sum_{i=1}^{a} \mu_i + \sum_{j=1}^{\ell} \sigma_j = p ;$$

$$p = 0, 1, \ldots, m \quad \text{si} \quad m > 1 ,$$

existent, sont continues par rapport à l'ensemble des arguments T, u, S et sont inférieures en module à $f(x)$, f étant une fonction de R^a dans R telle que :

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f(x) \prod_{i=1}^{b} dG_i(x_i) \in \overline{L}_1^{(a-b)}(F)$$

4) Pour tout point $(T, S_1, \ldots, S_\ell)$ appartenant à un voisinage du point $(T_0, S_1(F), \ldots, S_\ell(F))$, les dérivées partielles :

$$\frac{\partial^{p+q} \psi(T, x, F(x), S)}{\partial T^{\tau+\tau'} \partial u^\mu \partial S^{\sigma+\sigma'}}$$

existent, sont continues par rapport à l'ensemble des arguments T, S et sont inférieures en module à $g(x)$, g étant une fonction de R^a dans R telle que :

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} g(x) \prod_{i=1}^{b} dG_i(x_i) \in \overline{L}_2^{(a-b)}(F) ;$$

p, τ, μ, σ ont la même signification que dans l'hypothèse 3 et de plus :

$$\tau' + \sum_{i=1}^{\ell} \sigma_i' = q ; \quad 0 \leqslant q \leqslant \ell .$$

Sous ces hypothèses, il existe une solution unique : $T = T(F_N^t)$ pour l'équation $H[t, T] = 0$ et la fonctionnelle $T : \mathcal{F} \to R$ ainsi définie est une fonctionnelle de Mises d'ordre m.

Remarque 2.7

La démonstration du théorème (2.6) est basée sur le théorème des fonctions implicites dont nous déduisons en particulier la relation :

$$\frac{d\mathrm{T}(\mathrm{F}_{\mathrm{N}}^{t})}{dt} = -\frac{\left[\dfrac{\partial}{\partial t}\mathrm{H}[t,\mathrm{T}]\right]_{\mathrm{T}=\mathrm{T}[\mathrm{F}_{\mathrm{N}}^{t}]}}{\left[\dfrac{\partial}{\partial \mathrm{T}}\mathrm{H}[t,\mathrm{T}]\right]_{\mathrm{T}=\mathrm{T}[\mathrm{F}_{\mathrm{N}}^{t}]}}$$

Remarque 2.8

La validité des théorèmes (2.5) et (2.6) subsiste lorsque les conditions 3 et 4 sont satisfaites presque partout relativement aux arguments : $x_1, \ldots, x_b$, $x_{i_1}, \ldots, x_{i_{a-b}}$ et par rapport à la mesure

$$\prod_{i=1}^{b} d\mathrm{G}_i(x_i) \prod_{k=1}^{r} d\mathrm{F}(x_{j_k}) \quad \text{où :} \quad 1 \leqslant r \leqslant a-b$$

pour toute application injective : $k \to j_k$ de $\{1, \ldots, r\}$ dans $\{b+1, \ldots, a\}$ et pour toute application surjective : $k \to i_k$ de $\{1, \ldots, a-b\}$ dans $\{j_k / k = 1, \ldots, r\}$.

Théorème 2.9 ([10])

Soient $\mathrm{T}_1, \ldots, \mathrm{T}_k$ des fonctionnelles de Mises d'ordre m au point $\mathrm{F} \in \mathcal{F}$ et ϕ une fonction de R^k dans R admettant au voisinage du point : $(\mathrm{T}_1(\mathrm{F}), \ldots, \mathrm{T}_k(\mathrm{F}))$ les dérivées partielles continues :

$$\frac{\partial^p \phi[\mathrm{T}_1, \ldots, \mathrm{T}_k]}{\partial \mathrm{T}_1^{\tau_1} \ldots \partial \mathrm{T}_k^{\tau_k}} \quad \text{où :} \quad p = 1, \ldots, m\,;$$

$$\tau_i = 0, 1, \ldots, m\,; \quad i = 1, \ldots, k\,; \quad \sum_{i=1}^{k} \tau_i = p\,.$$

$\phi[\mathrm{T}_1, \ldots, \mathrm{T}_k]$ est alors une fonctionnelle de Mises d'ordre m au point F.

Définition 2.10

Soit $(\mathrm{X}_i\,;\, i \in \mathrm{N})$ une suite de vecteurs aléatoires de R^p, indépendants, équidistribués et soit d'autre part une fonction : $f : (x_1, \ldots, x_m) \to f(x_1, \ldots, x_m)$ de $(\mathrm{R}^p)^m$ dans R, invariante par rapport aux permutations sur les arguments : $x_1, \ldots, x_m$. Nous supposerons l'existence de l'espérance : $\mathrm{E}[f(\mathrm{X}_1, \ldots, \mathrm{X}_m)] = \eta$. La U-statistique de noyau f est alors définie comme suit :

$$\mathrm{U}_{\mathrm{N}} = \mathrm{U}[\mathrm{X}_1, \ldots, \mathrm{X}_{\mathrm{N}}] = \binom{\mathrm{N}}{m}^{-1} \sum_{\mathrm{C}_{\mathrm{N}}} f(\mathrm{X}_{\alpha_1}, \ldots, \mathrm{X}_{\alpha_m}) \tag{2.3}$$

où C_N désigne la sommation sur les $\binom{N}{m}$ combinaisons de m nombres : $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ choisis parmi $1, \ldots, N$. Nous utiliserons les notations suivantes :

$$f_c(x_1, \ldots, x_c) = E[f(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m)/X_1 = x_1, \ldots, X_c = x_c] ,$$

$$\zeta_c = \mathrm{Var}\, f_c(X_1, \ldots, X_c) ; \quad c = 1, \ldots, m \quad \text{et} \quad \zeta_0 = 0 . \tag{2.4}$$

Lemme 2.11 ([2])

Soient $(X_i ; i \in N)$ une suite de vecteurs aléatoires de R^p, indépendants, équidistribués et $f : (R^p)^m \to R$ un noyau symétrique tel que

$$E\,|f(X_1, \ldots, X_m)| < +\infty .$$

Sous ces hypothèses, on a :

$$U_N \underset{p.s}{\rightarrow} \eta \quad \text{et} \quad U_N \underset{L_1}{\rightarrow} \eta$$

lorsque $N \to +\infty$. □

Lemme 2.12 ([5])

$X_1, \ldots, X_N$ désignant N vecteurs aléatoires de R^p indépendants, équidistribués, soit $U_N = (U_N^{(1)}, \ldots, U_N^{(q)})$ un vecteur dont chaque composante $U_N^{(j)}$ est une U-statistique de noyau symétrique $f^{(j)} : (R^p)^{m_j} \to R$ tel que :

$$E[f^{(j)}(X_1, \ldots, X_{m_j})] = \eta^j \quad \text{et} \quad E[f^{(j)}(X_1, \ldots, X_{m_j})]^2 < +\infty ;$$

$$\forall j = 1, \ldots, q .$$

Si g est une fonction de R^q dans R dont les dérivées partielles secondes sont bornées au voisinage de $\eta = (\eta_1, \ldots, \eta^q)$, on obtient la propriété suivante :

$$\mathcal{L}[N^{1/2}[g(U_N) - g(\eta)]] \xrightarrow[N \to +\infty]{} N\left[0, \sum_{i,j=1}^{q} m_i m_j g_i g_j \zeta_1^{(i,j)}\right]$$

où :
$$g_k = \left[\frac{\partial g}{\partial x_k}\right]_{x=\eta} ; \quad k = 1, \ldots, q ;$$

et $\zeta_c^{(i,j)} = \mathrm{Cov}(f_c^{(i)}(X_1, \ldots, X_c),\ f_c^{(j)}(X_1, \ldots, X_c)) ; \quad i, j = 1, \ldots, q .$

Lemme 2.13 ([1])

Considérons la transformation de JACKKNIFE associée à la partition de $\{X_1, \ldots, X_N\}$ en N sous ensembles à un élément. Sous les hypothèses du lemme (2.12), on a :

$$\mathcal{L}[N^{1/2}[J[g(U_N)] - g(\eta)]] \xrightarrow[N \to +\infty]{} N\left[0, \sum_{i,j=1}^{q} m_i m_j g_i g_j \zeta_1^{(i,j)}\right]$$

D'autre part, sous l'hypothèse moins restrictive de l'existence pour la fonction g, de dérivées partielles premières continues au voisinage de η, la statistique :

$$\tilde{s}^2 = (\mathrm{N}-1)^{-1} \cdot \sum_{i=1}^{\mathrm{N}} [\mathrm{J}_i [g(\mathrm{U_N})] - \mathrm{J}[g(\mathrm{U_N})]]^2$$

vérifie la propriété suivante :

$$\tilde{s}^2 \underset{p}{\to} \sum_{i,j=1}^{q} m_i m_j g_i g_j \zeta_1^{(ij)}, \quad \text{lorsque} \quad \mathrm{N} \to +\infty .$$

En conséquence, on obtient :

$$\mathcal{L}\left[\frac{\mathrm{N}^{1/2} [\mathrm{J}[g(\mathrm{U_N})] - g(\eta)]}{\tilde{s}}\right] \xrightarrow[\mathrm{N}\to+\infty]{} \mathrm{N}(0,1) .$$

Lemme 2.14

Soient $(\mathrm{X}_i \,;\, i \in \mathrm{N})$ une suite de vecteurs aléatoires de dimension p quelconque, indépendants, équidistribués et $(\mathrm{X}_1, \dots, \mathrm{X}_m) \to f(\mathrm{X}_1, \dots, \mathrm{X}_m)$ une fonction d'échantillon à valeurs réelles, invariante par rapport à une permutation des arguments $\mathrm{X}_1, \dots, \mathrm{X}_m$ et de moment d'ordre 2 fini ; U désignera la U-statistique de noyau f. Pour un échantillon : $\mathrm{X}_1, \dots, \mathrm{X_N}$ de taille $\mathrm{N} = nk$, considérons la transformation du JACKKNIFE associée à la partition de $\{\mathrm{X}_1, \dots, \mathrm{X_N}\}$ en n sous-ensembles à k éléments :

$$\{\mathrm{X}_i / i \in \mathrm{C}_1\}, \dots, \{\mathrm{X}_i / i \in \mathrm{C}_n\}$$

où : $\quad \mathrm{C}_j = \{(j-1)\cdot k + 1, \dots, jk\} \,;\quad j = 1, \dots, n .$

Posons : $\quad \mathrm{U_N} = \mathrm{U}(\mathrm{X}_1, \dots, \mathrm{X_N})$ et $\mathrm{U}_k^{(*i)} = \mathrm{U}(\mathrm{X}_{(i-1)k+1}, \dots, \mathrm{X}_{ik})$;

$i = 1, \dots, n$; pour $k \geqslant m$.

Quel que soit $\delta > 0$, on a :

$$\mathrm{Var}\,[k^{1-\delta}[\mathrm{J}_i(\mathrm{U_N}) - \mathrm{U}_k^{(*i)}]] \xrightarrow[k\to+\infty]{} 0 , \tag{2.5}$$

uniformément par rapport à $i = 1, \dots, n$, cette relation restant vérifiée lorsque n tend vers l'infini.

De plus, pour un nombre fini, quelconque de U-statistiques $\overset{(s)}{\mathrm{U}}$; $s = 1, \dots, r$; admettant des noyaux de moments d'ordre 2 finis et pour une suite réelle $(a_s \,;\, s = 1, \dots, r)$ quelconque, on a :

$$\mathrm{Var}\left[k^{1-\delta}\left[\mathrm{J}_i\Big(\sum_{s=1}^{r} a_s \overset{(s)}{\mathrm{U}}_{\mathrm{N}}\Big) - \sum_{s=1}^{r} a_s \overset{(s)}{\mathrm{U}}{}_k^{(*i)}\right]\right]_{k\to+\infty} 0 \tag{2.6}$$

Démonstration

Par définition : $J_i[U_N] = nU_N - (n-1)\,U_{N-k}^{-i}$; $i = 1, \dots, n$;

où : $$U_{N-k}^{-i} = U[X_1, \dots, X_{(i-1)k}, X_{ik+1}, \dots, X_N]$$

Calculons l'expression :

$$\begin{aligned}\mathrm{Var}\,[J_i(U_N) - U_k^{(*i)}] &= \mathrm{Var}\,[nU_N - (n-1)\,U_{N-k}^{-i} - U_k^{(*i)}] \\ &= E\,[nU_N - (n-1)\,U_{N-k}^{-i} - U_k^{(*i)}]^2 \\ &= n^2 E\,[U_N^2] + (n-1)^2 E\,[U_{N-k}^{-i}]^2 + E\,[U_k^{(*i)}]^2 \\ &\quad - 2n(n-1)\,E\,[U_N \cdot U_{N-k}^{-i}] - 2n \cdot E\,[U_N \cdot U_k^{(*i)}] \qquad (2.7)\end{aligned}$$

$C_{N,m}$ désignant l'ensemble des combinaisons de m nombres choisis parmi $\{1, 2, \dots, N\}$, soient : $\alpha = (\alpha_1, \dots, \alpha_m) \in C_{N,m}$ et $\beta = (\beta_1, \dots, \beta_{m'}) \in C_{N',m'}$. Définissons les ensembles suivants :

$$\mathcal{E}_{N,m;N',m'} = \{(\alpha, \beta)/\alpha \in C_{N,m}\ ,\ \beta \in C_{N',m'},\ \{\alpha_1, \dots, \alpha_m\} \cap \{\beta_1, \dots, \beta_{m'}\} = \phi\}\,.$$

$$\mathcal{E}_{N,m;N',m'}^{j,(\gamma_1,\dots,\gamma_j)} = \{(\alpha, \beta)/\alpha \in C_{N,m},\ \beta \in C_{N',m'}\ ;\ \{\alpha_1, \dots, \alpha_m\} \cap \{\beta_1, \dots, \beta_{m'}\} = \{\gamma_1, \dots, \gamma_j\}\}\ ;$$

$$j = 1, \dots, \mathrm{Inf}\,(m, m')\,.$$

Posons : $\mathcal{E}_{N,m;N,m}^{j,(\gamma_1,\dots,\gamma_j)} = \mathcal{E}_{N,m}^{j,(\gamma_1,\dots,\gamma_j)}$

De la définition de U_N (cf. 2.3), on déduit :

$$[U_N]^2 = \binom{N}{m}^{-2} \sum_{\substack{\alpha \in C_{N,m} \\ \alpha \in C_{N,m}}} f[X_{\alpha_1}, \dots, X_{\alpha_m}]\,f[X_{\beta_1}, \dots, X_{\beta_m}]$$

$$[U_N]^2 = \binom{N}{m}^{-2} \sum_{(\alpha,\beta) \in \mathcal{E}_{N,m}^{0}} f[X_{\alpha_1}, \dots, X_{\alpha_m}]\,f[X_{\beta_1}, \dots, X_{\beta_m}]$$

$$\binom{N}{m}^{-2} \sum_{j=1}^{m} \sum_{(\gamma_1,\dots,\gamma_j) \in C_{N,j}} \sum_{(\alpha,\beta) \in \mathcal{E}_{N,m}^{j,(\gamma_1,\dots,\gamma_j)}} f[X_{\alpha_1}, \dots, X_{\alpha_m}]\,f[X_{\beta_1}, \dots, X_{\beta_m}]$$

Notons que :

$$E\,[f[X_{\alpha_1}, \dots, X_{\alpha_m}]\,f[X_{\beta_1}, \dots, X_{\beta_m}]] = \begin{cases} 0 & \text{si } (\alpha, \beta) \in \mathcal{E}_{N,m}^{0} \\ \zeta_j & \text{si } (\alpha, \beta) \in \mathcal{E}_{N,m}^{j,(\gamma_1,\dots,\gamma_j)} \end{cases}$$

$$j = 1, \dots, m\,.$$

On a donc

$$E\,[U_N]^2 = \binom{N}{m}^{-2} \cdot \sum_{j=1}^{m} \binom{N}{j} \binom{N-j}{m-j} \binom{N-m}{m-j} \zeta_j\,, \quad \text{pour } N \geqslant m\,. \qquad (2.8)$$

En conséquence, pour $N - k \geqslant m$ et $k \geqslant m$, il vient :

$$E[U_{N-k}^{-i}]^2 = \binom{N-k}{m}^{-2} \sum_{j=1}^{m} \binom{N-k}{j}\binom{N-k-j}{m-j}\binom{N-k-m}{m-j} \cdot \zeta_j , \tag{2.9}$$

$$E[U_k^{(*i)}]^2 = \binom{k}{m}^{-2} \cdot \sum_{j=1}^{m} \binom{k}{j}\binom{k-j}{m-j}\binom{k-m}{m-j} \cdot \zeta_j . \tag{2.10}$$

Des définitions de U_N et U_{N-k}^{-i}, il résulte que :

$$E[U_N \cdot U_{N-k}^{-i}] = \binom{N}{m}^{-1} \binom{N-k}{m}^{-1} \sum_{\substack{\alpha \in C_{N,m} \\ \beta \in C_{N-k,m}^{-i}}} E[f(X_{\alpha_1}, \dots, X_{\alpha_m}] \cdot f[X_{\beta_1}, \dots, X_{\beta_m}]]$$

$$= \binom{N}{m}^{-1} \binom{N-k}{m}^{-1} \sum_{j=1}^{m} \sum_{(\gamma_1, \dots, \gamma_j) \in C_{N-k,j}^{-i}} \sum_{(\alpha,\beta) \in \mathscr{E}_{N,m;N-k,m}^{j,(\gamma_1,\dots,\gamma_j)}} \zeta_j$$

$$= \binom{N}{m}^{-1} \binom{N-k}{m}^{-1} \sum_{j=1}^{m} \binom{N-k}{j}\binom{N-k-j}{m-j}\binom{N-m}{m-j} \cdot \zeta_j \tag{2.11}$$

De manière identique, on obtient :

$$E[U_N U_k^{(*i)}] = \binom{N}{m}^{-1} \binom{k}{m}^{-1} \sum_{j=1}^{m} \binom{k}{j}\binom{k-j}{m-j}\binom{N-m}{m-j} \cdot \zeta_j \tag{2.12}$$

Remplaçons chacun des termes de la somme (2.7), respectivement par leurs expressions obtenues en (2.8), (2.9), (2.10), (2.11) et (2.12). Lorsque les termes contenant les mêmes coefficients ζ_j ; $j = 1, \dots, m$; sont regroupés, on observe alors que :

$$\operatorname{Var}[k^{1/2}[J_i[U_N] - U_k^{(*i)}]] = 0\left(\frac{1}{k}\right)$$

d'où : $\quad \forall \delta > 0 , \quad \operatorname{Var}[k^{1-\delta}[J_i[U_N] - U_k^{(*i)}]] \xrightarrow[k \to +\infty]{} 0 ,$

uniformément par rapport à $i = 1, \dots, n$, n étant fixé ou tendant vers l'infini. La relation (2.6) résulte de la propriété précédente et de la relation suivante :

Si $Z_1, \dots, Z_r$ sont des v.a.r. de moments d'ordre 2 finis, on a :

$$\left[\sum_{s=1}^{r} [\operatorname{Var} Z_s]^{1/2}\right]^2 \geqslant \operatorname{Var}\left[\sum_{s=1}^{r} Z_s\right] .$$

Lemme 2.15

Soient (X_i ; $i \in N$) une suite de vecteurs aléatoires de dimension quelconque p, indépendants et équidistribués et soit $(X_1, \dots, X_m) \to f(X_1, \dots, X_m)$ une

5

fonction d'échantillon à valeurs réelles, symétrique, de moment d'ordre 2 fini. U désignant la U-statistique de noyau f, posons :

$$U_N = U(X_1, \ldots, X_N) \quad \text{pour} \quad N = nk$$

et soit U_{N-k}^{-i} la valeur de la U-statistique pour l'échantillon obtenu en supprimant dans l'échantillon $(X_1, \ldots, X_N)$ le $i^{\text{ème}}$ bloc d'observations de taille k.

On obtient alors la relation :

$$\text{Var}\,[N^{1-\delta}(U_N - U_{N-k}^{-i})] \xrightarrow[N\to+\infty]{} 0\,, \quad \forall \delta > 0\,.$$

Il en résulte, lorsque $N \to +\infty$, $N^{1-\delta}\, |U_N - U_{N-k}^{-i}| \underset{p}{\to} 0$, uniformément par rapport à $i = 1, \ldots, n$.

Démonstration

$$\text{Var}\,[U_N - U_{N-k}^{-i}] = \text{Var}\,[U_N] + \text{Var}\,[U_{N-k}^{-i}] - 2\,\text{Cov}\,(U_N, U_{N-k}^{-i})$$

On déduit des relations (2.8), (2.9) et (2.11) que :

$$\text{Var}\,[U_N] = \frac{m^2 . \zeta_1}{N} + 0\left(\frac{1}{N^2}\right), \quad \text{Cov}\,(U_N, U_{N-k}^{-i}) = \frac{m^2 . \zeta_1}{N} + 0\left(\frac{1}{N^2}\right)$$

D'où :

$$\text{Var}\,[U_N - U_{N-k}^{-i}] = \frac{m^2 . \zeta_1}{N} + \frac{m^2 . \zeta_1}{N-k} - 2\,\frac{m^2 . \zeta_1}{N} + 0\left(\frac{1}{N^2}\right) = 0\left(\frac{1}{N^2}\right).$$

3. LOIS LIMITES DU JACKKNIFE

Les résultats énoncés au paragraphe 2, portant sur les fonctionnelles de Mises ainsi que sur certaines propriétés du JACKKNIFE de U-statistiques nous permettent d'étudier la convergence en loi du JACKKNIFE de statistiques associées à ces fonctionnelles.

Théorème 3.1

Soient $(Y_i\,;\, i \in N)$ une suite de v.a.r. indépendantes, équidistribuées, de fonction de répartition F et $T : \mathcal{F} \to R$ une fonctionnelle de Mises d'ordre 4 au point $F \in \mathcal{F}$, telle que pour $j = 1, 2, 3$, la fonction :

$$(Z_1, \ldots, Z_j) \to T^{(j)}[F\,; Z_1, \ldots, Z_j]$$

(cf. définition 2.1) appartienne à l'espace $\overline{L}_2^j(F)$.

$(X_1, \ldots, X_N)$ désignant l'échantillon ordonné associé à l'échantillon aléatoire $(Y_1, \ldots, Y_N)$, considérons la transformation du JACKKNIFE associée à la partition de $\{X_1, \ldots, X_N\}$ en N sous ensembles à un élément. Les relations suivantes sont alors vérifiées :

1) $\mathcal{L}[N^{1/2}[J[T(F_N)] - T(F)]] \xrightarrow[N\to+\infty]{} N(0, \sigma)^2$,

2) $\tilde{s}^2 \underset{P}{\to} \sigma^2$, lorsque $N \to +\infty$,

3) $\mathcal{L}\left[\dfrac{N^{1/2}[J[T(F_N)] - T(F)]}{\tilde{s}}\right] \xrightarrow[N\to+\infty]{} N(0, 1)$.

Démonstration

Posons : $g_N(t) = T(F_N^t) = T(F + t(F_N - F))$ pour $t \in [0, 1]$

Considérons le développement de Taylor :

$$g_N(1) = g_N(0) + g_N^{(1)}(0) + \frac{1}{2!} g_N^{(2)}(0) + \frac{1}{3!} g_N^{(3)}(0) + \frac{1}{4!} g_N^{(4)}(\theta)$$

où : $\theta \in [0, 1]$ (3.1)

On en déduit que : $\forall \delta > 0$,

$$N^{2-\delta}\left| T(F_N) - T(F) - \left[\frac{d}{dt} T(F_N^t)\right]_{t=0} - \frac{1}{2!}\left[\frac{d^2}{dt^2} T(F_N^t)\right]_{t=0} - \frac{1}{3!}\left[\frac{d^3}{dt^3} T(F_N^t)\right]_{t=0}\right| \leqslant \frac{N^{2-\delta}}{4!} \sup_{\theta\in[0,1]} \left|\frac{d^4}{dt^4} T(F_N^\theta)\right|$$

Puisque T est une fonctionnelle de Mises d'ordre 4 au point F, on a :

$$T(F_N) = T(F) + \int_{-\infty}^{+\infty} T^{(1)}(F; y)\, d(F_N - F)(y)$$
$$+ \frac{1}{2!} \int^{+\infty} \int_{-\infty} T^{(2)}(F; y_1, y_2)\, d(F_N - F)(y_1)\, d(F_N - F)(y_2)$$
$$+ \frac{1}{3!} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} T^{(3)}(F; y_1, y_2, y_3)\, d(F_N - F)(y_1)\, d(F_N - F)(y_2)$$
$$d(F_N - F)(y_3) + o_P(N^{\delta-2}) \qquad (3.2)$$

La $i^{\text{ème}}$ pseudo-valeur du JACKKNIFE de $T(F_N)$ a pour expression :

$$J_i[T(F_N)] = N\,T(F_N) - (N-1)\,T(F_{N-1}^{-i})\,; \quad i = 1, \ldots, N\,;$$

où : F_{N-1}^{-i} désigne la fonction de répartition empirique associée à l'échantillon ordonné : $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_N$.

De la relation (3.1), il résulte donc :

$$J_i[T(F_N)] = T(F) + J_i\left[\left[\frac{d}{dt} T(F_N^t)\right]_{t=0}\right] + \frac{1}{2!} J\left[\left[\frac{d^2}{dt^2} T(F_N^t)\right]_{t=0}\right]$$
$$+ \frac{1}{3!} J_i\left[\left[\frac{d^3}{dt^3} T(F_N^t)\right]_{t=0}\right] + o_P(N^{\delta-1})\,; \quad i = 1, \ldots, N. \qquad (3.3)$$

D'où :

$$J[T(F_N)] = T(F) + J\left[\left[\frac{d}{dt}T(F_N^t)\right]_{t=0}\right] + \frac{1}{2!}J\left[\left[\frac{d^2}{dt^2}T(F_N^t)\right]_{t=0}\right]$$
$$+ \frac{1}{3!}J\left[\left[\frac{d^3}{dt^3}T(F_N^t)\right]_{t=0}\right] + o_P(N^{\delta-1}) \qquad (3.4)$$

Procédons à la décomposition de chacune des intégrales figurant dans l'expression (3.2), en une combinaison linéaire de U-statistiques.

α) D'après la remarque (2.2), on a :

$$I_1 = \int T^{(1)}[F; y]\, d(F_N - F)(y) = \frac{1}{N}\sum_{i=1}^{N} T^{(1)}[F; y_i]$$
$$= \overset{1}{U}(Y_1, \dots, Y_N) = \overset{1}{U}_N, \qquad (3.5)$$

où $\overset{1}{U}$ représente la U-statistique de noyau : $y \to T^{(1)}(F; y)$.

β) Soit : $I_2 = \dfrac{1}{2!}\displaystyle\iint T^{(2)}[F; y_1, y_2]\, d(F_N - F)(y_1)\, d(F_N - F)(y_2)$

D'après la remarque (2.2), $(y_1, y_2) \to T^{(2)}[F; y_1, y_2]$ est une fonction symétrique par rapport aux arguments : y_1, y_2 ; on a donc :

$$I_2 = (2N^2)^{-1} \cdot \sum_{i,j=1}^{N} T^{(2)}[F; x_i, x_j] - N^{-1} \cdot \sum_{i=1}^{N} \int T^{(2)}[F; x_i, y]\, dF(y)$$
$$= (2N^2)^{-1} \cdot \sum_{i=1}^{N} T^{(2)}[F; x_i, x_i] + N^{-2} \cdot \sum_{i<j}^{N} T^{(2)}[F; x_i, x_j] -$$
$$- N^{-1} \sum_{i=1}^{N} \int T^{(2)}[F; x_i, y]\, dF(y).$$

Soient : $\overset{2,1}{U}, \overset{2,2}{U}, \overset{2,3}{U}$ les U-statistiques de noyaux respectifs :

$$f^{(2,1)} : x \to T^{(2)}[F; x, x]; \quad f^{(2,2)} : (x_1; x_2) \to T^{(2)}[F; x_1, x_2];$$
$$f^{(2,3)} : x \to \int T^{(2)}[F; x, y]\, dF(y).$$

En conséquence :

$$I_2 = (2N)^{-1} \cdot \overset{2,1}{U}_N + (2N)^{-1} \cdot (N-1)\, \overset{2,2}{U}_N - \overset{2,3}{U}_N = (2)^{-1} \cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N +$$
$$+ (2N)^{-1} \cdot (\overset{2,1}{U}_N - \overset{2,2}{U}_N) \qquad (3.6)$$

γ) Soit : $I_3 = \dfrac{1}{3!}\displaystyle\iiint T^{(3)}[F; y_1, y_2, y_3]\, d(F_N - F)(y_1)$

$$d(F_N - F)(y_2)\, d(F_N - F)(y_3)$$

$$\mathrm{I}_3 = 6^{-1} \cdot \mathrm{N}^{-3} \cdot \sum_{i,j,k=1}^{\mathrm{N}} \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, x_k]$$

$$- 2^{-1} \cdot \mathrm{N}^{-2} \cdot \sum_{i,j=1}^{\mathrm{N}} \int \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, y]\, d\mathrm{F}(y)$$

$$+ (2\mathrm{N})^{-1} \cdot \sum_{i=1}^{\mathrm{N}} \iint \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, y_1, y_2]\, d\mathrm{F}(y_1)\; d\mathrm{F}(y_2)$$

Désignons respectivement par I_3^1, I_3^2, I_3^3 chacune des sommes figurant dans l'expression précédente.

$$\mathrm{I}_3^1 = 6^{-1} \cdot \mathrm{N}^{-3} \cdot \sum_{i,j,k=1}^{\mathrm{N}} \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, x_k]$$

$$\mathrm{I}_3^1 = 6^{-1} \cdot \mathrm{N}^{-3} \cdot \sum_{i=1}^{\mathrm{N}} \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_i, x_i] + 3^{-1} \cdot \mathrm{N}^{-3} \cdot \sum_{i<j} \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, x_j]$$

$$+ \mathrm{N}^{-3} \cdot \sum_{i<j<k} \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, x_k]$$

Soient : $\overset{3,1}{\mathrm{U}}$, $\overset{3,2}{\mathrm{U}}$, $\overset{3,3}{\mathrm{U}}$ les U-statistiques de noyaux respectifs :

$$f^{(3,1)} : x \to \mathrm{T}^{(3)}[\mathrm{F}\,; x, x, x]\,; \quad f^{(3,2)} : (x_1, x_2) \to \mathrm{T}^{(3)}[\mathrm{F}\,; x_1, x_2, x_2]$$

$$f^{(3,3)} : (x_1, x_2, x_3) \to \mathrm{T}^{(3)}[\mathrm{F}\,; x_1, x_2, x_3].$$

On a donc :

$$\mathrm{I}_3^1 = 6^{-1} \cdot \mathrm{N}^{-2} \cdot \overset{3,1}{\mathrm{U}}_{\mathrm{N}} + 6^{-1} \cdot \mathrm{N}^{-2} \cdot (\mathrm{N}-1) \cdot \overset{3,2}{\mathrm{U}}_{\mathrm{N}} + 6^{-1} \cdot \mathrm{N}^{-2} \cdot (\mathrm{N}-1)(\mathrm{N}-2)\, \overset{3,3}{\mathrm{U}}_{\mathrm{N}} \tag{3.7}$$

Décomposons de manière identique, l'expression :

$$\mathrm{I}_3^2 = -(2\mathrm{N}^2)^{-1} \cdot \sum_{i,j=1}^{\mathrm{N}} \int \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, y]\, d\mathrm{F}(y)$$

$$= -(2\mathrm{N}^2)^{-1} \cdot \sum_{i=1}^{\mathrm{N}} \int \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_i, y]\, d\mathrm{F}(y)$$

$$- \mathrm{N}^{-2} \cdot \sum_{i<j} \int \mathrm{T}^{(3)}[\mathrm{F}\,; x_i, x_j, y]\, d\mathrm{F}(y)$$

Soient : $\overset{3,4}{\mathrm{U}}$, $\overset{3,5}{\mathrm{U}}$ les U-statistiques de noyaux respectifs :

$$f^{(3,4)} : x \to \int \mathrm{T}^{(3)}[\mathrm{F}\,; x, x, y]\, d\mathrm{F}(y),$$

$$f^{(3,5)} : (x_1, x_2) \to \int \mathrm{T}^{(3)}[\mathrm{F}\,; x_1, x_2, y]\, d\mathrm{F}(y)$$

D'où : $\quad I_3^2 = -(2N)^{-1} \cdot \overset{3,4}{U}_N - (2N)^{-1}(N-1)\overset{3,5}{U}_N \qquad (3.8)$

D'autre part, l'expression :

$$I_3^3 = (2N)^{-1} \cdot \sum_{i=1}^{N} \iint T^{(3)}[F ; x_i, y_1, y_2] \, dF(y_1)\, dF(y_2)$$

peut s'écrire en fonction de la U-statistique $\overset{3,6}{U}$ de noyau :

$$f^{(3,6)} : x \to \iint T^{(3)}[F ; x, y_1, y_2] \, dF(y_1)\, dF(y_2) \, ,$$

$$I_3^3 = 2^{-1} \cdot \overset{3,6}{U}_N \, . \qquad (3.9)$$

En utilisant les relations (3.7), (3.8) et (3.9), il vient :

$$I_3 = I_3^1 + I_3^2 + I_3^3$$

$$= 6^{-1}(\overset{3,3}{U}_N - 3\,\overset{3,5}{U}_N + 3\,\overset{3,6}{U}_N) + (2N)^{-1} \cdot (3^{-1} \cdot \overset{3,2}{U}_N - \overset{3,3}{U}_N - \overset{3,4}{U}_N + \overset{3,5}{U}_N)$$

$$+ (6N^2)^{-1} \cdot (\overset{3,1}{U}_N - \overset{3,2}{U}_N + 2\,\overset{3,3}{U}_N) \qquad (3.10)$$

Notons d'après le lemme (2.15) que pour chacune des U-statistiques définies précédemment, nous avons :

$$J_i(N^{-1} \cdot U_N) = U_N - U_{N-1}^{-i} = o_P(N^{\delta-1}) \, ,$$

$$J_i(N^{-2} \cdot U_N) = N^{-1} \cdot U_N - (N-1)^{-1} \cdot U_{N-1}^{-i} = o_P(N^{\delta-2}) \, ; \quad i = 1, \ldots, N.$$

Les relations (3.6) et (3.10) entrainent donc :

$$J_i(I_2) = J_i(2^{-1} \cdot \overset{2,2}{U}_N - \overset{3,2}{U}_N) + o_P(N^{\delta-1}) \, , \qquad (3.11)$$

$$J_i(I_3) = 6^{-1} \cdot J_i(\overset{3,3}{U}_N - 3\,\overset{3,5}{U}_N + 3\,\overset{3,6}{U}_N) + o_P(N^{\delta-1}) \, ; \quad i = 1, \ldots, N \, ; \qquad (3.12)$$

En tenant compte des relations (3.1), (3.11), (3.12), l'expression (3.3) devient :

$$J_i(T(F_N)) = T(F)$$

$$+ J_i(\overset{1}{U}_N + 2^{-1} \cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N + 6^{-1} \cdot \overset{3,3}{U}_N - 2^{-1} \cdot \overset{3,5}{U}_N + 2^{-1} \cdot \overset{3,6}{U}_N)$$

$$+ o_p(N^{\delta-1}) \, ; \quad i = 1, \ldots N \, . \qquad (3.13)$$

Par définition : $J(T(F_N)) = N^{-1} \sum_{i=1}^{N} J_i(T(F_N))$.

$$N^{1/2}[J(T(F_N)) - T(F)]$$

$$= N^{1/2} \cdot J(\overset{1,1}{U}_N + 2^{-1} \cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N + 6^{-1} \cdot \overset{3,3}{U}_N - 2^{-1} \cdot \overset{3,5}{U}_N + 2^{-1} \cdot \overset{3,6}{U}_N)$$

$$+ o_P(N^{\delta-(1/2)}) \, .$$

Remarquons que les U-statistiques $\overset{1}{U}, \overset{2,2}{U}, \overset{2,3}{U}, \overset{3,3}{U}, \overset{3,5}{U}$, admettent des noyaux d'espérance nulle et de moment d'ordre 2 fini puisque la fonction :

$$(y_1, \ldots, y_j) \to T^{(j)}[F ; y_1, \ldots, y_j] \in \overline{L}_2^{(j)}(F),$$

quel que soit $j = 1, 2, 3$. Il suffit donc d'appliquer le lemme (2.13) pour obtenir les relations :

$$\mathcal{L}[N^{1/2}[J(T(F_N)) - T(F)]] \underset{N\to\infty}{\longrightarrow} N(0, \sigma^2),$$

$$\tilde{s}^2 \underset{p}{\to} \sigma^2, \text{ lorsque } N \to \infty$$

Théorème 3.2

Soient $(y_i ; i \in N)$ une suite de v.a.r. indépendantes, équidistribuées, de fonction de répartition F et soit $T : \mathfrak{F} \to R$ une fonctionnelle de Mises d'ordre 2 au point $F \in \mathfrak{F}$, telle que la fonction : $z \to T^{(1)}[F ; z]$ appartienne à l'espace $\overline{L}_2^{(1)}(F)$.

Considérons pour $N = nk$ et $k > 1$, la transformation du JACKKNIFE associée à la partition de $\{Y_1, \ldots, Y_N\}$ en n sous-ensembles à k éléments :

$$\{Y_i / i \in C_1\}, \ldots, \{Y_i / i \in C_n\} \quad \text{où :} \quad C_j = \{(j-1)k + 1, \ldots, jk\} ;$$
$$j = 1, \ldots, n.$$

Posons :

$$\tilde{s}^2 = (n-1)^{-1} \sum_{i=1}^{n} [J_i[T(F_N)] - J[T(F_N)]]^2 .$$

Lorsque k tend vers l'infini, tandis que n reste fixé, la distribution de la v.a.r. :

$$\frac{n^{1/2}[J[T(F_N)] - T(F)]}{\tilde{s}} \tag{3.13}$$

tend vers la loi de STUDENT à $(n-1)$ degrés de liberté.

Démonstration

En utilisant un développement limité à l'ordre 2, analogue à celui obtenu en (3.1), on obtient :

$$\forall \delta > 0, \quad T(F_N) = T(F) + \int T^{(1)}[F ; y]\, dF_N(y) + o_p(N^{\delta - 1}),$$

soit :

$$T(F_N) = T(F) + N^{-1} \cdot \sum_{i=1}^{N} T^{(1)}[F ; Y_i] + o_p(N^{\delta-1}).$$

Par définition :

$$J_i[T(F_N)] = n \cdot T(F_N) - (n-1)\, T(F_{N-k}^{-i}) ; \; i = 1, \ldots, n ;$$

où F_{N-k}^{-i} représente la fonction de répartition empirique associée à l'échantillon de taille $(N-k)$ obtenu en supprimant dans l'échantillon : $(Y_1, \ldots, Y_N)$ le i° bloc d'observations : $\{Y_j / j \in C_i\}$.

$$T(F_{N-k}^{-i}) = T(F) + (N-k)^{-1} \cdot \sum_{j \notin C_i} T^{(1)}[F ; Y_j] + o_P[(N-k)^{\delta-1}]$$

D'où :

$$J_i[T(F_N)] = T(F) + k^{-1} \cdot \sum_{j \in C_i} T^{(1)}[F ; Y_j] + o_P(k^{\delta-1}) \tag{3.14}$$

et :

$$k^{1/2}[J_i[T(F_N)] - T(F)] = k^{-1/2} \sum_{j \in C_i} T^{(1)}[F ; Y_j] + o_P(k^{\delta-1/2}) ; \quad i = 1, \ldots, n ; \tag{3.15}$$

De l'hypothèse : $\sigma^2 = \int [T^{(1)}[F ; y]]^2 \, dF(y) < +\infty$, il résulte que :

$$\mathcal{L}[k^{-1/2} \sum_{j \in C_i} T^{(1)}[F ; Y_j]] \xrightarrow[K \to +\infty]{} N(0, \sigma^2) .$$

Les v.a.r. : $k^{1/2} \cdot [J_i[T(F_N)] - T(F)]$, $k^{-1/2} \cdot \sum_{j \in C_i} T^{(1)}[F ; Y_j]$ étant asymptotiquement équidistribuées, quel que soit $i = 1, \ldots, n$; on en conclut que les v.a.r. :

$$V_{i,n,k} = k^{1/2}[J_i[T(F_N)] - T(F)] ; \quad i = 1, \ldots, n ;$$

sont asymptotiquement indépendantes et de même distribution : $N(0, \sigma^2)$.

Posons :

$$\overline{V}_{n,k} = n^{-1} \sum_{i=1}^{n} V_{i,n,k}$$

D'après la remarque précédente, la loi de la v.a.r.

$$\frac{n^{1/2} \overline{V}_{n,k}}{\left[(n-1)^{-1} \cdot \sum_{i=1}^{n} [V_{i,n,k} - \overline{V}_{n,k}]^2 \right]^{1/2}} \tag{3.16}$$

tend vers la loi de STUDENT de degré de liberté : $(n-1)$, lorsque k tend vers l'infini. La démonstration est achevée en remarquant que la v.a.r. (3.15) n'est autre que :

$$\frac{n^{1/2}[J[T(F_N)] - T(F)]}{\tilde{s}} \tag{3.17}$$

Remarque 3.2

Les pseudo-valeurs du JACKKNIFE sont liées aux fonctions d'influence par la relation (3.14) :

$$J_i(T(F_N) = T(F) + k^{-1} \cdot \sum_{j \in C_i} T^{(1)}(F ; y_j) + o_P(k^{\delta-1}) ; \quad i = 1, \ldots, n .$$

Pour limiter l'effet d'observations étrangères éventuelles, il apparaît donc nécessaire de n'appliquer la transformation du JACKKNIFE qu'à des statistiques pour lesquelles :

1) la fonction : $x \to T^{(1)}(F\,;x)$ est bornée

2) la fonctionnelle : $G \to T^{(1)}(G\,;x)$ est uniformément continue par rapport à x, au point F.

Théorème 3.3

Supposons les hypothèses du théorème (3.2) vérifiées et supposons de plus que $T : \mathcal{F} \to R$ soit une fonctionnelle de Mises d'ordre 3, au point $F \in \mathcal{F}$, telle que la fonction : $(z_1, \ldots, z_j) \to T^{(j)}(F\,; z_1, \ldots, z_j)$ appartienne à l'espace $\overline{L}_2^{(j)}(F)$; $\forall j = 1, 2$. La loi de la v.a.r.

$$\frac{n^{1/2}\,[J(T(F_N)) - T(F)]}{\tilde{s}}$$

converge alors vers la loi normale centrée, réduite, lorsque n et k tendent simultanément vers l'infini.

Démonstration

En considérant le développement de TAYLOR (3.1) limité à l'ordre 3, il résulte des propriétés d'une fonctionnelle de Mises d'ordre 3, que :

$$\forall \delta > 0\,, \quad T(F_N) = T(F) + \int T^{(1)}(F\,;y)\, dF_N(y)$$

$$+ \frac{1}{2} \iint T^{(2)}(F\,; y_1, y_2)\, d(F_N - F)(y_1)\, d(F_N - F)(y_2) + o_P(N^{\delta - 3/2})$$

On en déduit en utilisant les notations de la démonstration du théorème (3.2) :

$$J_i(T(F_N)) = T(F) + J_i(I_1) + 2^{-1} \cdot J_i(I_2) + o_P(n^{\delta - (1/2)} \cdot k^{\delta - (3/2)})$$

Soient : $\overset{1}{U}, \overset{2,2}{U}, \overset{2,3}{U}$ les U-statistiques de noyaux respectifs :

$$f^{(1)} : x \to T^{(1)}(F\,;x)\,; \quad f^{(2,2)} : (x_1, x_2) \to T^{(2)}(F\,; x_1, x_2)\,;$$

$$f^{(2,3)} : x \to \int T^{(2)}(F\,; x, y)\, dF(y)\,,$$

d'espérances nulles et de moments d'ordre 2 finis. D'après la relation (3.11), on a :

$$J_i(I_2) = J_i(2^{-1} \cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N) + o_P(N^{\delta - 1})$$

On en déduit :

$$J_i(T(F_N)) = T(F) + J_i(\overset{1}{U}_N + 2^{-1} \cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N) + o_P(n^{\delta - (1/2)} \cdot k^{\delta - 1})\,.$$

D'autre part, en vertu du lemme (2.14), on obtient :

$$k^{1/2}[J_i(\overset{1}{U}_N + 2^{-1}\cdot \overset{2,2}{U}_N - \overset{2,3}{U}_N) - (\overset{1}{U}_k^{(*i)} + 2^{-1}\cdot \overset{2,2}{U}_k^{(*i)} - \overset{2,3}{U}_k^{(*i)})] = o_P(k^{\delta-1/2})$$

$$\forall i = 1, \ldots, n ;$$

Il s'en suit :

$$k^{1/2}[J_i(T(F_N)) - T(F)] - k^{1/2}[\overset{1}{U}_k^{(*i)} + 2^{-1}\cdot \overset{2,2}{U}_k^{(*i)} - \overset{2,3}{U}_k^{(*i)}] = o_P(k^{\delta-1/2}) .$$

Il résulte du lemme (2.12) que :

$$\mathcal{L}[k^{1/2}[\overset{1}{U}_k^{(*i)} + 2^{-1}\cdot \overset{2,2}{U}_k^{(*i)} - \overset{2,3}{U}_k^{(*i)}]] \xrightarrow[K\to+\infty]{} N(0, \sigma^2)$$

Les variables aléatoires : $W_{i,n,k} = k^{1/2}[J_i(T(F_N)) - T(F)]$; $i = 1, \ldots, n$; sont donc asymptotiquement indépendantes et même distribution limite $N(0, \sigma^2)$. Soit :

$$\overline{W}_{n,k} = n^{-1}\cdot \sum_{i=1}^{n} W_{i,n,k}$$

D'après la remarque précédente, la loi de la v.a.

$$\frac{n^{1/2}\overline{W}_{n,k}}{\left[(n-1)^{-1}\cdot \sum_{i=1}^{n} (W_{i,n,k} - \overline{W}_{n,k})^2\right]^{1/2}}$$

tend vers la loi de STUDENT à $(n-1)$ degrés de liberté, lorsque k tend vers l'infini, n étant fixé. La distribution limite est donc la loi $N(0, 1)$, pour n tendant vers l'infini.

4. APPLICATIONS

4.1. Lois limites du JACKKNIFE d'estimateurs du maximum de vraisemblance

1) Soit : $(Y_1, \ldots, Y_N)$ un échantillon tiré d'une v.a. dont la fonction de répartition F appartient à une famille de fonctions de répartition $F(\cdot, \alpha)$ de densités $f(\cdot, \alpha)$; $\alpha \in R$. Désignons par $\hat{\alpha}_N$ la valeur de l'estimateur du maximum de vraisemblance, pour l'échantillon considéré ; $\hat{\alpha}_N$ est donc une solution de l'équation en α :

$$\int \varphi(\alpha, y)\, dF_N(y) = 0 \quad \text{où} \quad \varphi(\alpha, y) = \frac{\partial}{\partial \alpha} \operatorname{Log} f(\alpha, y) .$$

La fonctionnelle T : $\mathfrak{F} \to R$ associée à cet estimateur est définie par la relation :

$$\int \varphi(T(G), y)\, dG(y) = 0 ; \quad \forall G \in \mathfrak{F} .$$

Rappelons les conditions requises pour l'application du théorème (3.2) :

a) T est une fonctionnelle de Mises d'ordre 2, au point F.

b) $T^{(1)}(F\,;\cdot) \in \overline{L}_2^{(1)}(F)$

D'après le théorème (2.6), la condition a est vérifiée sous les propriétés A et B suivantes :

A) Il existe un réel α_0 tel que :

$$\int \varphi(\alpha_0, y)\, dF(y) = 0 \quad \text{et} \quad \int \frac{\partial}{\partial \alpha} \varphi(\alpha_0, y)\, dF(y) \neq 0 \qquad (4.1)$$

B) Les fonctions : $\alpha \to \dfrac{\partial^m}{\partial \alpha^m} \varphi(\alpha\,, y)$; $m = 0, 1, 2$; existent, et sont continues en un voisinage $V(\alpha_0)$ de α_0. De plus, il existe une fonction $g \in \overline{L}_2^{(1)}(F)$ telle que :

$$\left| \frac{\partial^m}{\partial \alpha^m} \varphi(\alpha\,, y) \right| \leqslant g(y)\,; \quad \forall \alpha \in V(\alpha_0)\,; \quad \forall m = 0, 1, 2\,; \qquad (4.2)$$

Rappelons qu'il suffit en fait que ces propriétés soient vérifiées presque partout relativement à la variable y et par rapport à la mesure F. D'autre part, en utilisant le résultat de la remarque (2.7), on obtient :

$$T^{(1)}(F\,; y) = - \frac{\varphi(\alpha_0, y)}{\displaystyle\int \frac{\partial}{\partial \alpha} \varphi(\alpha_0, x)\, dF(x)}$$

Finalement, sous les hypothèses A et B, la loi de la v.a.

$$\frac{n^{1/2}\,[J(\hat{\alpha}_N) - \alpha_0]}{\tilde{s}} \qquad (4.3)$$

tend vers la loi de STUDENT à $(n - 1)$ degrés de liberté, lorsque k tend vers l'infini.

2) Cas de distributions discrètes

Soient $[a_0, a_r]$ un intervalle contenant l'ensemble des valeurs observées : $y_j\,; j = 1, \ldots, N$; et $(a_i)_{i=0,\ldots,r}$ une suite strictement croissante.

Posons :

$$n_i = \text{Card}\,\{j / a_{i-1} \leqslant y_j < a_i\}$$

et

$$P_i(\alpha) = F(a_i, \alpha) - F(a_{i-1}, \alpha)\,; \quad i = 1, \ldots, r.$$

La valeur de α maximisant la probabilité :

$$\prod_{i=1}^{r} [F(a_i, \alpha) - F(a_{i-1}, \alpha)]^{n_i}$$

est solution de l'équation :

$$\int \varphi(\alpha\,, y)\, dF_N(y) = 0$$

où : $\varphi(\alpha , y) = P_i'(\alpha) \cdot (P_i(\alpha))^{-1}$ pour $y \in [a_{i-1}, a_i[$; $i = 1 , \ldots , r$. La propriété (4.3) est donc vérifiée sous les hypothèses A et B.

4.2. Estimateurs issus de la méthode du maximum de vraisemblance modifié

1) Etant donné un échantillon $(Y_1, \ldots , Y_N)$ tiré d'une v.a. de fonction de répartition F appartenant à la classe des fonctions : $x \to G(x - \alpha)$; $\alpha \in R$; G étant supposée connue. Définissons l'estimateur $\hat{\alpha}_N$ du paramètre de position α par la relation :

$$\sum_{i=1}^{N} \rho(Y_i - \hat{\alpha}_N) = \operatorname*{Inf}_{\alpha \in R} \sum_{i=1}^{N} \rho(Y_i - \alpha) ,$$

où ρ est une fonction à valeurs réelles dont HUBER (1964) a déterminé les critères de choix. Soient : $\rho' = \psi$ et $\varphi(\alpha , y) = \psi(y - \alpha)$: $\hat{\alpha}_N$ est alors solution de l'équation :

$$\int \varphi(\alpha , y) \, dF_N(y) = 0 .$$

Sous les hypothèses A et B, la loi limite du JACKKNIFE de l'estimateur $\hat{\alpha}_N$ vérifie la propriété (4.3).

2) Supposons que la fonction de répartition F de la v.a. observée Y appartienne à la classe des fonctions : $x \to G\left(\dfrac{x - \alpha}{\sigma}\right)$; $\alpha \in R$; $\sigma \in R^{*+}$. Soit : $S(F_N)$ un estimateur du paramètre d'échelle σ, associé à une fonctionnelle de Mises d'ordre 2 au point F, telle que $S^{(1)}[F ; \cdot] \in \overline{L}_2^{(1)}(F)$. L'estimateur $T(F_N)$ du paramètre de position α, obtenu par la méthode du maximum de vraisemblance modifié est solution de l'équation en α :

$$\int \varphi(\alpha , y , S(F_N)) \, dF_N(y) = 0 .$$

où :
$$\varphi(\alpha , y , S) = \psi\left(\frac{y - \alpha}{S}\right)$$

Supposons les hypothèses suivantes vérifiées :

1) $\exists \alpha_0 \in R$ tel que : $\displaystyle\int \psi\left(\frac{y - \alpha_0}{S(F)}\right) dF(y) = 0$

et : $\displaystyle\int \psi'\left(\frac{y - \alpha_0}{S(F)}\right) dF(y) \neq 0$

2) Pour tout point (α , S) appartenant à un voisinage de $(\alpha_0, S(F))$, les dérivées partielles :

$$\frac{\partial^{p+q} \psi\left(\dfrac{y - \alpha}{S}\right)}{\partial \alpha^{\tau+\tau'} \partial S^{\sigma+\sigma'}} \quad \text{où} \quad \begin{array}{l} 0 \leqslant \tau + \sigma = p \ ; \ p = 0, 1, 2 \ ; \\ 0 \leqslant \tau' + \sigma' = q \ ; \ q = 0, 1 \ ; \end{array}$$

existent $\forall y$ F *p.p.*, sont continues par rapport à l'ensemble des arguments α, S et sont inférieures *p.p.* en module à $g(y)$, g étant une fonction appartenant à l'espace $\overline{L}_2^{(1)}(F)$.

Sous ces hypothèses, la distribution du JACKKNIFE de $T(F_N)$ associé à une partition de l'échantillon de taille $N = nk$ en n sous-ensembles à k éléments vérifie la propriété suivante :

$$\frac{n^{1/2}[J(T(F_N)) - \alpha_0]}{\widetilde{s}}$$

tend en loi vers la distribution de STUDENT à $(n-1)$ degrés de liberté lorsque k tend vers l'infini, n restant fixé.

La démonstration de cette proposition résulte de l'application du théorème (3.2). En effet, les conditions 1 et 2 impliquent en vertu du théorème (2.6) que T est une fonctionnelle de Mises d'ordre 2 au point F. D'autre part :

$$T^{(1)}[F\,;x] = \left[\int \psi'\left(\frac{x-\alpha_0}{S(F)}\right) dF(x)\right]^{-1} \cdot$$

$$\cdot \left[S(F)\cdot\psi\left(\frac{x-\alpha_0}{S(F)}\right) - \frac{S^{(1)}[F\,;x]}{S(F)}\int (x-\alpha_0)\,\psi'\left(\frac{x-\alpha_0}{S(F)}\right) dF(x)\right]$$

Il suffit donc de remarquer que l'hypothèse 2 et $S^{(1)}[F\,;\cdot\,] \in \overline{L}_2^{(1)}(F)$ entraînent : $T^{(1)}[F\,;\cdot\,] \in \overline{L}_2^{(1)}(F)$.

4.3. Combinaisons linéaires de statistiques d'ordre

Soit $(X_1, \ldots, X_N)$ l'échantillon ordonné associé à un échantillon aléatoire $(Y_1, \ldots, Y_N)$ tiré d'une v.a. de fonction de répartition F et de moment d'ordre 2 fini ; désignons par $H : [0\,,1] \to R$ une fonction admettant une dérivée seconde continue. La distribution du JACKKNIFE de la statistique :

$$T(F_N) = N^{-1}\cdot\sum_{i=1}^{N} H((i-1)/N)\,X_i$$

à laquelle est associée la fonctionnelle $T : \mathfrak{F} \to R$ définie par la relation :

$$T(F) = \int y\cdot H(F(y))\,dF(y)$$

vérifie alors la propriété (3.13).

En effet, puisque :

$$T(F) = \int \varphi(y\,, F(y))\,dF(y) \quad \text{où} \quad \varphi(y\,,u) = y\cdot H(u) \;,$$

il suffit d'après le théorème (3.2) que :

a) T soit une fonctionnelle de Mises d'ordre 2, au point F.

b) $T^{(1)}(F\,;\cdot) \in \overline{L}_2^{(1)}(F)$.

On vérifie que la proposition a est satisfaite, en utilisant le théorème (2.5) ; d'autre part, l'existence du moment d'ordre 2 de la v.a. observée assure la validité de la proposition b.

4.4. Estimateurs déduits de la méthode du χ^2 minimum

$[a_0, a_r]$ désignant un intervalle contenant l'ensemble des observations y_i ; $j = 1, \ldots, N$; soient $(a_i)_{i=0,\ldots,r}$ une suite strictement croissante et

$$P_i(\alpha) = F(a_i, \alpha) - F(a_{i-1}, \alpha) ; \quad i = 1, \ldots, r .$$

Puisque :

$$X^2 = N \left[\sum_{i=1}^{r} \int_{a_{i-1}}^{a_i} \int_{a_{i-1}}^{a_i} (P_i(\alpha))^{-1} \, dF_N(y_1) \, dF_N(y_2) - 1 \right] ,$$

l'estimateur $\hat{\alpha}_N$ considéré est donc solution de l'équation :

$$\iint \varphi(\alpha, y_1, y_2) \, dF_N(y_1) \, dF_N(y_2) = 0$$

pour

$$\varphi(\alpha, y_1, y_2) = \sum_{i=1}^{r} (P_i'(\alpha))^{-2} \, P_i'(\alpha) \cdot 1_{[a_{i-1}, a_i[\times [a_{i-1}, a_i[}(y_1, y_2)$$

Pour que la condition a soit vérifiée, il suffit en vertu du théorème (2.6) que les propriétés suivantes soient satisfaites :

A) $\exists \alpha_0 \in R$ tel que $\sum_{i=1}^{r} P_i'(\alpha_0) \neq 0$ et

$$\iint \frac{\partial}{\partial \alpha} \varphi(\alpha_0, y_1, y_2) \, dF(y_1) \, dF(y_2) \neq 0$$

B) Les dérivées : $\alpha \rightarrow P_i^{(m)}(\alpha)$; $m = 0, 1, 2, 3$; $i = 1, \ldots, r$; sont continues au voisinage de α_0 ; on suppose d'autre part que pour une constante $c > 0$: $P_i(\alpha) > c$; $i = 1, \ldots, r$; au voisinage de α_0.

La propriété b se trouvant alors réalisée, la loi de $J(\hat{\alpha}_N)$ vérifie la propriété (4.3), sous les hypothèses A et B.

BIBLIOGRAPHIE

[1] ARVESEN J.N. – Jackknifing U-statistics. *Ann. Math. Statist. 40,* 2076-2100, 1969.

[2] BERK R. – Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist. 37,* 51-58, 1966.

[3] BRILLINGER D.R. – The asymptotic behavior of Tukey's general method of setting approximate confidence limits (the Jackknife) when applied to maximum likelihood estimates. *Rev. Inst. Internat. Statist. 32,* 202-6, 1964.

[4] FILIPPOVA A.A. – Mises' theorem of the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theor. Probability. Appl. 7,* 24-57, 1962.

[5] HOEFFDING W. – A class of statistics with asymptotically normal distribution. *Ann. Math. Statist. 19,* 293-325. 1948.

[6] HUBER P.J. – Robust estimation of a location parameter. *Ann. Math. Statist. 35,* 73-101, 1964.

[7] HUBER P.J. – Robust statistics : A review. *Ann. Math. Statist. 43,* 1041-67, 1972.

[8] LEPRETRE L.F. – *Propriétés asymptotiques du Jackknife.* Thèse de 3ᵉ cycle. (Paris VI), 1976.

[9] MANN H.B. et WALD A. – On stochastic limit and order relationships. *Ann. Math. Statist. 14,* 217-226, 1943.

[10] VON MISES R. – On the asymptotic distributions of differentiable statistical functions. *Ann. Math. Statist. 18,* 309-348, 1947.