



HAL
open science

Contribution à la théorie de l'estimation fonctionnelle. Chap. IV-VI

D. Bosq

► **To cite this version:**

D. Bosq. Contribution à la théorie de l'estimation fonctionnelle. Chap. IV-VI. Annales de l'ISUP, 1970, XIX (3), pp.97-177. hal-04083378

HAL Id: hal-04083378

<https://hal.science/hal-04083378v1>

Submitted on 27 Apr 2023

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

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



Distributed under a Creative Commons Attribution 4.0 International License

TABLE DES MATIÈRES

	Pages
Chapitre IV - ESTIMATION DE LA DENSITÉ CONDITIONNELLE ET DES FONCTIONS DE RÉGRESSION	99
I - Généralités	99
II - Estimation de la densité conditionnelle	100
III - Estimation des fonctions de régression	113
Chapitre V - ESTIMATION DE LA FONCTION FRONTIÈRE	122
I - Positivité à P et M	122
II - Construction de la fonction frontière	123
III - Cas d'un espace localement compact	128
IV - Cas où la fonction frontière est de signe quelconque	131
V - Application: Définition d'une méthode de Monte Carlo	140
Chapitre VI - ÉLÉMENTS EXTRÊMAUX D'UN ÉCHANTILLON VECTORIEL	142
I - Généralités	142
II - Loi du nombre d'éléments maximaux associés à une rotation de R^n	145
III - Comportement asymptotique du nombre d'éléments maximaux (resp. minimaux)	150
IV - Théorème de M. Gailly et application à l'estimation de la courbe maximale	166
BIBLIOGRAPHIE	173

D. BOSQ



- [12] VAN RYZIN (J.) (1969). On strong consistency of density estimates.
Ann. Math. Stat. 40 p. 1765-1772.
- [19] WOODROOFE (M.) (1967). On the maximal deviation of the sample density.
Ann. Math. Stat. 38 p. 475-481.

CONTRIBUTION A LA THEORIE DE

L'ESTIMATION FONCTIONNELLE

0208 0

TABLE DES MATIÈRES

	Pages
Chapitre IV - ESTIMATION DE LA DENSITÉ CONDITIONNELLE ET DES FONCTIONS DE RÉGRESSION	99
I - Généralités et notations	99
II - Estimation de la densité conditionnelle	100
III - Estimation des fonctions de régression	115
Chapitre V - ESTIMATION DE LA FONCTION FRONTIÈRE	122
I - Position du Problème	122
II - Construction d'un estimateur convergent	123
III - Cas d'un espace localement compact	128
IV - Cas où la fonction frontière est de signe quelconque	131
V - Application: Définition d'une méthode de Monte Carlo	140
Chapitre VI - ÉLÉMENTS EXTRÉMAUX D'UN ÉCHANTILLON VECTORIEL	142
I - Généralités	142
II - Loi du nombre d'éléments maximaux associés à une rotation de \mathbf{R}^2	145
III - Comportement asymptotique du nombre d'éléments maximaux (resp. minimaux)	150
IV - Théorèmes de M. Geffroy - Application à l'estimation de la courbe maximale	166
BIBLIOGRAPHIE	173

TABLE DES MATIÈRES

Pages

173	BIBLIOGRAPHIE
166	IV - Théorèmes de M. Gelfand - Application à l'estimation de la courbe maximale
150	III - Comportement asymptotique du nombre d'éléments maximaux (resp. minimaux)
145	II - Loi du nombre d'éléments maximaux associés à une rotation de \mathbb{R}^n
142	I - Généralités
Chapitre VI - ÉLÉMENTS EXTRÊMAUX D'UN ÉCHANTILLON VECTORIEL	
140	V - Application: Définition d'une méthode de Monte Carlo
131	IV - Cas où la fonction frontière est de signe quelconque
128	III - Cas d'un espace localement compact
123	II - Construction d'un estimateur convergent
122	I - Position du Problème
Chapitre V - ESTIMATION DE LA FONCTION FRONTIÈRE	
115	III - Estimation des fonctions de régression
100	II - Estimation de la densité conditionnelle
99	I - Généralités et notations
90	RÉGRESSION
Chapitre IV - ESTIMATION DE LA DENSITÉ CONDITIONNELLE ET DES FONCTIONS DE	

IV

ESTIMATION DE LA DENSITÉ CONDITIONNELLE
ET DES FONCTIONS DE RÉGRESSION

I - GÉNÉRALITÉS ET NOTATIONS

Dans ce chapitre nous montrons qu'à partir de toute suite d'estimateurs de la densité possédant certaines propriétés asymptotiques, on peut construire une suite d'estimateurs de la régression qui converge uniformément presque sûrement. Pour cela, nous passons par l'intermédiaire d'un estimateur de la densité conditionnelle.

Nous connaissons un seul article sur cette question ⁽¹⁾: celui de BHATTACHARYA et PARTHASARATHY [36] (1961). Ces deux auteurs utilisent la méthode de l'analyse graphique fractile.

Les résultats que nous donnons sont ceux de notre note aux C.R. du 13 Octobre 1969.

Les notations que nous emploierons sont les suivantes :

- $F(x,y)$ est la fonction de répartition du couple de variables aléatoires (X, Y) . Nous supposons que F est absolument continue, que sa densité $f(x,y)$ est bornée et qu'elle est définie partout par la relation

$$\frac{\partial^2 F}{\partial x \partial y} (x,y) = f(x,y).$$

- $g(x)$ désigne la densité marginale de X .

- $h_{x,a}(y)$ est la densité de Y conditionnée par

$$X \in [x-a, x+a[, \quad x \in \mathbf{R}, \quad a > 0.$$

- $h_x(y)$ est la densité de Y conditionnée par $X = x, \quad x \in \mathbf{R}$.

- $B(x,a) = \{(x', y') \in \mathbf{R}^2 \mid x-a \leq x' < x+a, \quad y' \in \mathbf{R}\}$

(1) A part celui de NADARAYA [18] qui traite un cas particulier.

- $p(x,a) = \text{Prob} [(X, Y) \in B(x,a)]$

- $(x_i, y_i)_{1 \leq i \leq n}$ est un échantillon de taille n tiré suivant F .

- $l_n(x,a)$ est le nombre de points de l'échantillon appartenant à $B(x,a)$; $l_n(x,a)$ est la réalisation d'une variable aléatoire $L_n(x,a)$.

Enfin, $y_{i_1}, \dots, y_{i_{l_n}}$ sont les ordonnées des points de l'échantillon appartenant à $B(x,a)$ ⁽¹⁾.

Nous supposons que a est une fonction positive de

$n(a = a(n) > 0, \forall n \in \mathbf{N})$ et telle que

$$\lim_{n \rightarrow \infty} a(n) = 0 \quad (1)$$

II - ESTIMATION DE LA DENSITÉ CONDITIONNELLE

1. Définition d'une classe d'estimateurs.

Pour définir un estimateur de $h_x(y)$, il semble naturel de partir d'un estimateur de la densité univariée.

Comme nous n'aurons besoin, dans la suite, que de certaines propriétés d'un estimateur de la densité, il est inutile de préciser la méthode utilisée pour construire cet estimateur.

Nous désignons donc par $\hat{f} = (\hat{f}_n)_{n=0,1,2,\dots}$ une suite d'estimateurs de la densité univariée ⁽²⁾. Pour des raisons qui apparaîtront ci-dessous, nous posons arbitrairement :

$$\hat{f}_0(y) = 0, \quad \forall y \in \mathbf{R} \quad (2)$$

Soit alors $\mathcal{F}(\hat{f})$ l'ensemble des densités univariées telles que :

$$1^{\circ} f_n^g(y) = E_g [(\hat{f}_n(y))] = \int_{\mathbf{R}^n} \hat{f}_n(y_1, \dots, y_n; y) g(y_1) \dots g(y_n) dy_1 \dots dy_n$$

(1) Pour abrégé, nous écrirons l_n au lieu de $l_n(x,a)$ quand il n'y aura pas ambiguïté.

(2) Nous supposons, une fois pour toutes, que $\hat{f}_n(y_1, \dots, y_n; y)$ est une application mesurable de \mathbf{R}^{n+1} dans \mathbf{R} , $\forall n \in \mathbf{N}$.

existe $\forall y \in \mathbf{R}, \forall n \in \mathbf{N}, \forall g \in \mathcal{F}(\hat{f})$.

$$2^{\circ}) \sup_{y \in \mathbf{R}} |f_n^g(y) - g(y)| \rightarrow 0 \quad \text{avec} \quad \frac{1}{n}, \quad \forall g \in \mathcal{F}(\hat{f}).$$

3^o) Il existe deux constantes positives C_1 et C_2 et une fonction de n , α_n , non décroissante, tendant vers $+\infty$ avec n et vérifiant la condition

$$\lim_{n \rightarrow \infty} \frac{n}{\alpha_n^2} = +\infty, \quad \text{telles que}$$

$$\text{Prob} \left[\sup_{y \in \mathbf{R}} |\hat{f}_n(y) - f_n^g(y)| > \varepsilon \right] \leq C_1 \exp \left[-C_2 \frac{n}{\alpha_n^2} \varepsilon^2 \right]$$

$\forall \varepsilon > 0, \forall n \in \mathbf{N}, \forall g \in \mathcal{F}(\hat{f})$.

REMARQUE :

Les conditions précédentes impliquent que

$$\sup_{y \in \mathbf{R}} | \hat{f}_n(y) - g(y) | \rightarrow 0 \quad \text{quand} \quad n \rightarrow +\infty$$

$\forall g \in \mathcal{F}(\hat{f})$.

\hat{f} étant donnée et x étant fixé, nous prendrons comme estimateur de h_x la fonction aléatoire définie par :

$$\hat{h}_{n,x}(y) = \hat{f}_{1_n}(y_{i_1}, \dots, y_{i_{1_n}}; y) \tag{3}$$

$\forall y \in \mathbf{R}, \forall n \in \mathbf{N}$.

2. Convergence de l'estimateur.

Dans ce paragraphe, x est fixé une fois pour toutes.

LEMME 17 :

Si $g(x) > 0$ et s'il existe $\delta > 0$ tel que, pour $a \leq \delta$, f est

lipchitzienne d'ordre γ_1 dans $B(x,a)$ et g est lipchitzienne d'ordre γ_2 dans $[x-a, x+a]$, alors, il existe $\delta_1 \in]0, \delta]$ tel que $a < \delta_1$ entraîne

$$\sup_{y \in \mathbf{R}} |h_{x,a}(y) - h_x(y)| \leq A(x) \cdot a^\gamma \quad (4)$$

où $A(x)$ est indépendante de y et de a et où l'on a posé

$$\gamma = \max(\gamma_1, \gamma_2).$$

DÉMONSTRATION.

Si $g(x) > 0$, on a $F(x+a, +\infty) - F(x-a, +\infty) > 0$, $\forall a > 0$ et l'on peut écrire :

$$h_{x,a}(y) = \frac{\frac{\partial F}{\partial y}(x+a,y) - \frac{\partial F}{\partial y}(x-a-y)}{F(x+a, +\infty) - F(x-a, +\infty)}$$

Or, pour $a \leq \delta$, f est Lipchitzienne donc continue et, a fortiori, continue par rapport à x ; il en est de même pour g et l'on peut écrire :

$$h_{x,a}(y) = \frac{f(x-a + 2\theta a, y)}{g(x-a + 2\theta' a)}, \quad 0 < \theta < 1, \quad 0 < \theta' < 1$$

D'autre part, $h_x(y) = \frac{f(x,y)}{g(x)}$ et comme, pour a assez petit ($a < \delta_1$) on

a $g(x-a + 2\theta' a) > \frac{g(x)}{2}$, il vient :

$$|h_{x,a}(y) - h_x(y)| \leq \frac{2}{g(x)} |f(x-a + 2\theta a, y) - f(x,y)| + \frac{2f(x,y)}{g(x)^2} |g(x) - g(x-a + 2\theta' a)|$$

Or, il existe deux constantes k_1 et k_2 telles que, $\forall a \leq \delta$

$$|f(x-a + 2\theta a, y) - f(x,y)| \leq k_1 a^{\gamma_1} |1 - 2\theta|^{\gamma_1} \leq k_1 a^{\gamma_1}$$

et $|g(x) - g(x-a + 2\theta' a)| \leq k_2 a^{\gamma_2} |1 - 2\theta'|^{\gamma_2} \leq k_2 a^{\gamma_2}$

d'où :

$$|h_{x,a}(y) - h_x(y)| \leq \frac{2}{g(x)} k_1 a^{\gamma_1} + \frac{2f(x,y)}{g(x)^2} k_2 a^{\gamma_2}$$

Or, on sait que $M = \sup_{(x,y) \in \mathbb{R}^2} f(x,y) < +\infty$, on a donc immédiatement

l'inégalité (4) avec

$$A(x) = \frac{2}{g(x)} k_1 + \frac{2M}{g(x)^2} k_2 \tag{5}$$

Pour simplifier l'énoncé du lemme suivant, nous allons poser

$$h_{n,x}(y) = f_{L_n(x,a)}^{h_{x,a}}(y) \quad \forall n \in \mathbb{N}, \quad \forall y \in \mathbb{R} \tag{6}$$

Autrement dit, $\forall y \in \mathbb{R}$ et $\forall n \in \mathbb{N}$, $h_{n,x}(y)$ est une variable aléatoire liée à $L_n(x,a)$ par la relation fonctionnelle suivante :

$$L_n(x,a) = \ell \implies h_{n,x}(y) = E_{h_{x,a}}[\hat{f}_1(y)] \tag{7}$$

$\forall \ell = 0, 1, \dots, n$.

Nous avons alors le

LEMME 18 :

Sous les hypothèses du lemme 17 et en supposant, de plus, que $h_{x,a} \in \mathcal{F}(\hat{f})$ au moins pour a assez petit et que

$$\lim_{n \rightarrow \infty} \frac{n a_n}{\gamma_n^2} = +\infty \quad \text{on a}$$

$$\sup_{y \in \mathbb{R}} | \hat{h}_{n,x}(y) - h_{n,x}(x) | \xrightarrow{p} 0 \quad \text{quand } n \rightarrow \infty. \tag{8}$$

DÉMONSTRATION.

Désignons par d la distance de la convergence uniforme. En appliquant la formule des probabilités totales, on obtient :

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon] = \sum_{\ell=0}^n C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell}$$

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon / L_n = \ell]$$

Or, d'après (7), on a

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon / L_n = \ell] = \text{Prob} [d(\hat{f}_1; f_{\ell}^{h_{x,a}}) > \varepsilon]$$

cette dernière probabilité étant calculée par rapport à la loi conditionnelle de densité $h_{x,a}$.

Comme on a supposé que, pour a assez petit, $h_{x,a} \in \mathcal{F}(\hat{f})$, on peut, pour n assez grand, appliquer 3°) et écrire :

$$\text{Prob} [d(\hat{h}_{r,x}; h_{n,x}) > \varepsilon] \leq C_\ell \sum_{\ell=0}^n C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell} \exp \left[-C_2 \frac{1}{\alpha_\ell^2} \varepsilon^2 \right]$$

et comme α_1 a été supposée non décroissante :

$$\begin{aligned} \text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon] &\leq C_\ell \sum_{\ell=0}^n C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell} \exp \left[-C_2 \frac{1}{\alpha_n^2} \varepsilon^2 \right] \\ &\leq C_1 [1-p_{x,a}] [1 - \exp(-C_2 \frac{\varepsilon^2}{\alpha_n^2})]^{1^n} \end{aligned}$$

Or, pour a assez petit

$$\begin{aligned}
 p(x, a) &= F(x+a, +\infty) - F(x-a, +\infty) \\
 &= 2a g(x) \quad , \quad \forall x \in]x-a, x+a[\\
 &> 2a \cdot \frac{g(x)}{2} = a \cdot g(x)
 \end{aligned}$$

il vient alors :

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon] \leq C_1 [1 - a_n g(x) [1 - \exp(-C_2 \frac{\varepsilon^2}{\alpha_n^2})]^n \quad (9)$$

Le deuxième membre de cette inégalité tend vers 0 si et seulement si :

$$n \text{ Log} [1 - a_n g(x) [1 - \exp(-C_2 \frac{\varepsilon^2}{\alpha_n^2})]] \rightarrow -\infty$$

donc, compte tenu du fait que $\lim_{n \rightarrow \infty} a_n = 0$ si

$$n a_n [1 - \exp(-C_2 \frac{\varepsilon^2}{\alpha_n^2})] \rightarrow +\infty$$

or $\alpha_n \rightarrow +\infty$ donc $1 - \exp(-C_2 \frac{\varepsilon^2}{\alpha_n^2}) \sim \frac{C_2 \varepsilon^2}{\alpha_n^2}$ et la condi-

tion s'écrit finalement

$$\lim_{n \rightarrow \infty} \frac{n a_n}{\alpha_n^2} = +\infty \quad (10)$$

C.Q.F.D.

LEMME 18 bis :

Sous les hypothèses du lemme 18 et l'hypothèse supplémentaire

$$\sum_{r=0}^{+\infty} \exp[-\lambda \frac{n a_n}{\alpha_n^2}] < +\infty, \quad \forall \lambda > 0 \quad (11)$$

On a $d(\hat{h}_{n,x}; h_{n,x}) \xrightarrow{\text{p.s.}} 0$ avec $\frac{1}{n}$.

DÉMONSTRATION.

On sait, puisque $\text{Log}(1-u) \leq -u, \forall u < 1$, que

$$n \text{Log} \left[1 - a_n g(x) \left[1 - \exp \left(-C_2 \frac{\varepsilon^2}{\alpha_n^2} \right) \right] \right] \leq -n a_n g(x) \left[1 - \exp \left(-C_2 \frac{\varepsilon^2}{\alpha_n^2} \right) \right]$$

pour n assez grand. (12)

D'autre part, on a, pour n assez grand :

$$1 - \exp \left(-\frac{C_2 \varepsilon^2}{\alpha_n^2} \right) < 2 \frac{C_2 \varepsilon^2}{\alpha_n^2} \tag{13}$$

(12) et (13) entraînent alors :

$$n \text{Log} \left[1 - a_n g(x) \left[1 - \exp \left(-\frac{C_2 \varepsilon^2}{\alpha_n^2} \right) \right] \right] \leq -2 C_2 \varepsilon^2 g(x) \cdot \frac{n a_n}{\alpha_n^2}$$

D'où, d'après (9) et toujours pour n assez grand :

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon] \leq C_1 \exp \left[-2 C_2 \varepsilon^2 g(x) \frac{n a_n}{\alpha_n^2} \right]$$

(14)

La condition (11) permet alors de conclure.

Pour le lemme suivant, nous aurons besoin d'une nouvelle propriété de la classe d'estimateurs utilisée: nous dirons que \hat{f} est *uniformément asymptotiquement sans biais* sur $\mathcal{F}'(\hat{f}) \subset \mathcal{F}(\hat{f})$ si la condition 2°) est valable uniformément pour $g \in \mathcal{F}'(\hat{f})$, autrement dit si

$$\lim_{n \rightarrow \infty} \sup_{g \in \mathcal{F}'(\hat{f})} d(f_n^g; g) = 0 \tag{15}$$

Nous dirons, pour abrégé, que \hat{f} est u.a.b. sur $\mathcal{F}'(\hat{f})$.

LEMME 19 :

Si $g(x) > 0$, si g est Lipchitzienne dans un voisinage de x , si $h_{x,a} \in \mathcal{F}(\hat{f}), \forall n \in \mathbb{N}$, si \hat{f} est u.a.b. sur $\{h_{x,a_n}\}_{n \in \mathbb{N}}$

et si $n a_n \rightarrow +\infty$

alors

$$d(h_{n,x}; h_{x,a}) \xrightarrow{p} 0.$$

DÉMONSTRATION.

D'après (6), on peut écrire :

$$\begin{aligned} \text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] &= \sum_{\ell=0}^n C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell} \\ &\quad \text{Prob} [d(f_\ell^{h_{x,a}}; h_{x,a}) > \varepsilon / L_n = \ell] \end{aligned} \tag{16}$$

Les probabilités qui figurent au 2ème membre de (16) valant 0 ou 1 puisque $d(f_\ell^{h_{x,a}}; h_{x,a})$ est une variable certaine quand $L_n = \ell$. Comme \hat{f} est u.a.b. sur $\{h_{x,a_n}\}_{n \in \mathbf{N}}$, il existe un entier ℓ_0 indépendant de a_n (donc de n) tel que $\ell > \ell_0$ entraîne $d(f_\ell^{h_{x,a}}; h_{x,a}) < \varepsilon, \forall a_n$ donc $\forall n \in \mathbf{N}$.

Dans ces conditions, pour $n > \ell_0$, on peut écrire :

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] \leq \sum_{\ell=0}^{\ell_0} C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell} \tag{17}$$

Or, pour n assez grand, on aura

$$\ell_0 < n p(x,a) - [1-p(x,a)] \tag{18}$$

En effet, on a vu dans la démonstration du lemme 18 que, à partir d'un certain rang, on a l'inégalité $p(x,a) > a g(x)$; et comme, par hypothèse, $n a_n \rightarrow +\infty$, on aura toujours $\ell_0 < n a_n g(x)$, à partir d'un certain rang, ce qui entraîne (18).

Donc, pour $n \geq \nu$, $C_n^\ell p(x,a)^\ell [1-p(x,a)]^{n-\ell}$ sera fonction non

décroissante de ℓ pour $\ell \leq \ell_0$ (propriété connue de la loi binominale).

Dans ces conditions, (17) implique :

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] \leq \ell_0 C_n^{\ell_0} p(x,a)^{\ell_0} [1-p(x,a)]^{n-\ell_0} \quad (19)$$

et en choisissant n assez grand pour que

$$a g(x) < p(x,a) < 3a g(x) \quad (19 \text{ bis})$$

on obtient facilement :

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] \leq \ell_0 \cdot \frac{n^{\ell_0}}{\ell_0!} [3a g(x)]^{\ell_0} \exp[-(n-\ell_0) a g(x)] \quad (20)$$

A partir de (20), on obtient donc une inégalité de la forme :

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] \leq A (n a_n)^{\ell_0} e^{-B n a_n} \quad (21)$$

et, à partir d'un certain rang

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon] \leq A e^{-B n a_n} \quad (22)$$

où A et B sont indépendants de n .

L'hypothèse $n a_n \rightarrow +\infty$ permet de conclure.

LEMME 19 bis :

Si les conditions du lemme 19 sont vérifiées et si, de plus,

$$\sum_{n=0}^{+\infty} \exp(-\gamma n a_n) < +\infty, \quad \forall \gamma > 0 \quad (23)$$

alors : $d(h_{n,x}; h_{x,a}) \xrightarrow{\text{p.s.}} 0$.

DÉMONSTRATION.

C'est une conséquence immédiate de l'inégalité (22).

THÉORÈME 19 :

Sous les conditions suivantes :

1°) Pour a assez petit, f est Lipchitzienne dans $B(x, a)$ et g est Lipchitzienne dans $[x-a, x+a]$.

2°) $h_{x, a_n} \in \mathcal{F}(\hat{f})$, $\forall n \in \mathbf{N}$ et \hat{f} est u.a.b. sur $\{h_{x, a_n}\}_n \in \mathbf{N}$.

$$3°) \sum_{n=0}^{+\infty} \exp \left[-\lambda \frac{n a_n}{\alpha_n^2} \right] < +\infty, \quad \forall \lambda > 0$$

on a

$$d(\hat{h}_{n, x}; h_x) \xrightarrow{\text{p.s.}} 0.$$

DÉMONSTRATION.

On écrit :

$$d(\hat{h}_{n, x}; h_x) \leq d(\hat{h}_{n, x}; h_{n, x}) + d(h_{n, x}; h_{x, a}) + d(h_{x, a}; h_x) \quad (24)$$

et il suffit d'appliquer les lemmes 17, 18 bis et 19 bis pour conclure.

3. Etude de la vitesse de convergence.

Les inégalités (4), (14) et (22) montrent que, pour n assez grand, on a l'inégalité :

$$\text{Prob} \left[d(\hat{h}_{n, x}; h_x) > \varepsilon \right] \leq K_1 \exp \left[-\frac{n a_n}{\alpha_n^2} K_2 \varepsilon^2 \right] \quad (25)$$

En examinant les démonstrations des lemmes précédents, on constate que (25) reste valable, uniformément sur un intervalle compact, si les hypothèses faites sont uniformes sur cet intervalle (cf. les hypothèses du lemme 17, la formule (5), les inégalités (19 bis), la formule (15), l'inégalité (14) et les valeurs de A et B dans l'inégalité (22). On peut donc énoncer le

THÉORÈME 20 :

Soit $[u, v]$ un intervalle compact tel que :

1°) Il existe un voisinage V de $[u, v]$ dans \mathbf{R} tel que f soit Lipchitzienne sur $V \times \mathbf{R}$ et g Lipchitzienne sur V .

2°) $g(x) > 0, \quad \forall x \in [u, v]$.

3°) $h_{x, a_n} \in \mathcal{F}(\hat{f}), \quad \forall x \in [u, v], \quad \forall n \in \mathbf{N}$.

4°) \hat{f} est u.a.b. sur $\{h_{x, a_n}\}_{u \leq x \leq v; n \geq v} \quad (v \in \mathbf{N})$

Alors si $\frac{n a_n}{\alpha_n^2} \rightarrow \infty$ et si $\varepsilon > 0$ est donné, il existe deux nombres positifs $K_1(\varepsilon)$ et $K_2(\varepsilon)$ et un entier $N(\varepsilon)$ tels que

$$\sup_{u \leq x \leq v} \text{Prob} [d(\hat{h}_{n, x}; h_x) > \varepsilon] \leq K_1(\varepsilon) \exp \left[-K_2(\varepsilon) \frac{n a_n}{\alpha_n^2} \right] \quad (26)$$

$$\forall n \geq N(\varepsilon).$$

Pour préciser un peu plus la vitesse de convergence, supposons que l'on ait :

$$\beta_n = \sup_{\substack{u \leq x \leq v \\ k \geq v}} d(f_n^{h_{x, a_k}}; h_{x, a_k}) \leq C'_1 n^{-b}, \quad C'_1 > 0, \quad b > 0 \quad (27)$$

$$\alpha_n = C'_2 \cdot n^c, \quad 0 < c < \frac{1}{2} \quad (28)$$

$$a_n = C'_3 \cdot n^{-d}, \quad 0 < d < 1 - 2c \quad (29)$$

On peut énoncer alors le

COROLLAIRE 15 :

Si $\gamma_1 = \gamma_2 = 1$ et si $0 < \delta < \inf(d, b(1-d), \frac{1-d}{2} - c)$

Alors, sous les hypothèses du théorème 20 :

$$n^\delta \cdot d(\hat{h}_{n,x}; h_x) \xrightarrow{\text{p.s.}} 0, \quad \forall x \in [u, v].$$

DÉMONSTRATION.

Tout d'abord, compte tenu de (4), (29) et du fait que $\delta < d$, on a $\forall x \in [u, v]$:

$$n^\delta \cdot d(h_{x,a}; h_x) \rightarrow 0 \tag{30}$$

D'autre part, si l'on reprend les démonstrations des lemmes 18 et 18 bis, on arrive à l'inégalité :

$$\text{Prob} [d(\hat{h}_{n,x}; h_{n,x}) > \varepsilon \cdot n^{-\delta}] \leq C_1 \exp [-C_2 \varepsilon^2 n^{-2\delta} g(x) \frac{n a_n}{\alpha_n^2}] \tag{14 bis}$$

et le deuxième membre de l'inégalité (14 bis) est le terme général d'une série convergente puisque

$$n^{-2\delta} \cdot \frac{n a_n}{\alpha_n^2} = \frac{C'_3}{C'^2_2} \cdot n^{1-d-2\delta-2c}$$

et que $\delta < \frac{1-d}{2} - c$. On a donc, $\forall x \in [u, v]$:

$$n^\delta \cdot d(\hat{h}_{n,x}; h_{n,x}) \xrightarrow{\text{p.s.}} 0 \tag{31}$$

Enfin, si l'on reprend la démonstration du lemme 19, on constate que l'inégalité :

$$d(f_{\ell}^{h_{x,a}}; h_{x,a}) < \varepsilon \cdot n^{-\delta}$$

est vérifiée pour $\ell > \ell_0$ avec ℓ_0 tel que

$$C'_\ell \cdot \ell_0^{-b} < \varepsilon \cdot n^{-\delta} \tag{32}$$

En prenant $\ell_0 = \ell_0(n, \varepsilon) = K n^{\frac{\delta}{b}}$ avec $K > (\frac{C'_\ell}{\varepsilon})^{\frac{b}{\delta}}$, on a (32) et comme

$\delta < b$, on a aussi $l_0 < n$, au moins pour n assez grand: les inégalités (18), (19) et (20) restent donc valables si l'on remplace ε par $\varepsilon n^{-\delta}$. Il en est de même pour (21) puisque :

$$\exp [a_n l_0 g(x)] \rightarrow l \quad \text{quand } n \rightarrow +\infty.$$

D'autre part, comme

$$\text{Log} (n a_n)^{l_0} = K n^{\frac{\delta}{b}} (1-d) \text{Log} C'_3 \frac{l}{l^{-d}} \cdot n$$

On a, $\forall u > 0$

$$(n a_n)^{l_0} < e^{\frac{\delta}{b} + u}$$

au moins pour n assez grand. En choisissant $u < 1-d\frac{\delta}{b}$, il vient :

$$\text{Prob} [d(h_{n,x}; h_{x,a}) > \varepsilon n^{-\delta}] < A' e^{-B'n^{\alpha_n}}$$

où A' et B' sont positifs et indépendants de n . Donc :

$$n^{\delta} d(h_{n,x}; h_{x,a}) \xrightarrow{\text{p.s.}} 0 \tag{33}$$

ce qui, joint à (30) et (31), permet de conclure.

4. Exemples.

Examinons comment on peut appliquer les résultats obtenus aux deux classes usuelles d'estimateurs de la densité.

a) *Méthode du noyau.*

L'estimateur utilisé est de la forme :

$$\hat{f}_n(y) = \frac{\alpha_n}{n} \sum_{i=1}^n K[\alpha_n(y-y_i)] \tag{34}$$

où les y_i sont les points de l'échantillon et où K est une densité à variation bornée. Alors d'après un théorème de NADARAYA [18] on sait que

si g est une densité uniformément continue, $g \in \mathcal{F}(\hat{f})$ ⁽¹⁾.

De plus, $\forall g \in \mathcal{F}(\hat{f})$, on a l'inégalité : ⁽²⁾

$$d(E_g(\hat{f}_n), g) \leq \sup_{y \in \mathbf{R}} \cdot \sup_{|z| < \delta} |g(y-z) - g(y)| \\ + 2 \sup_{y \in \mathbf{R}} g(y) \cdot \int_{|u| \geq \delta} K(u) du$$

$\forall \delta > 0, \quad \forall n \in \mathbf{N}.$

On en déduit immédiatement que si $(g_i)_{i \in I}$ est une famille de densités uniformément continues, uniformément bornées et uniformément équicontinues, alors \hat{f} est u.a.b. sur $(g_i)_{i \in I}$.

CONSÉQUENCE.

Pour \hat{f}_n , si les hypothèses 1°) et 2°) du théorème 20 sont vérifiées, les hypothèses 3°) et 4°) le sont aussi.

Car la relation $h_{x,a}(y) = \frac{f(x-a + 2\theta a, y)}{g(x-a + 2\theta'a)} \tag{35}$

$0 < \theta < 1, \quad 0 < \theta' < 1$

montre que la famille $\{h_{x,a_n}\}_{n \in \mathbf{N}}$ vérifie les conditions indiquées ci-dessus.

Enfin, d'après les résultats de SCHUSTER ⁽³⁾, on voit que, sous les hypothèses 1°) et 2°) du théorème 20, et en supposant, de plus.

$\sup_{n \in \mathbf{N}} |h'_{x,a_n}| = D < +\infty, \quad \text{si l'on choisit } \delta \in \left] 0, \frac{1}{4} \right[\quad \text{on aura}$

$n^\delta \cdot d(\hat{h}_{n,x}; h_x) \xrightarrow{\text{p.s.}} 0, \quad \forall x \in [u,v].$

(1) La réciproque est vraie moyennant des hypothèses supplémentaires sur K (cf SCHUSTER [14]).

(2) cf [18] p.

(3) Lemme 2.3 et Corollaire 2.6.

b) Méthode des fonctions orthogonales.

On utilise l'estimateur défini au chapitre II en prenant, par exemple, les fonctions d'Hermite comme système orthonormal et $q(n)$ tel que

$$\sum_{n=1}^{\infty} \exp \left[-\gamma \frac{n}{q(n)^4} \right] < +\infty, \quad \forall \gamma > 0.$$

D'après l'inégalité (23 bis) (Chapitre II) et les lemmes 11 et 15 (Chapitre III), on voit que $\alpha_n = q(n)^2$ et que si g est une densité telle que

$$e^{y^2/2} \frac{d^3}{dy^3} [e^{-y^2/2} g(y)]$$

existe et soit de carré intégrable, alors $g \in \mathcal{F}(\hat{f})$.

D'autre part, si $(g_i)_{i \in I}$ est une famille de densités telles que

$$A = \sup_{i \in I} \left\| e^{y^2/2} \frac{d^3}{dy^3} [e^{-y^2/2} g_i(y)] \right\|_{L^2} < +\infty$$

alors \hat{f} est u.a.b. sur $(g_i)_{i \in I}$. On aura, en effet, une majoration de la forme ⁽¹⁾.

$$\sup_{i \in I} d(E_{g_i}(\hat{f}_n); g_i) \leq \frac{A \cdot c}{2^{3/2}} \cdot \sum_{j > q(n)} \frac{1}{j^{3/2}}$$

CONSEQUENCE.

Si $f(x,y)$ admet une dérivée partielle troisième $f'''_{y^3}(x,y)$ et s'il existe $\eta > 0$ tel que

$$\sup_{|x-x_0| \leq \eta} \left\| e^{y^2/2} \frac{d^3}{dy^3} [e^{-y^2/2} f(x,y)] \right\|_{L^2} < +\infty$$

On en déduit, d'après (35), que \hat{f} est u.a.b. sur $\{h_{x_0, a_n}\}_n \nu$
 ν étant convenablement choisi.

(1) cf S.C. SCHWARTZ [27] lemme 3 p. 1262.

III - ESTIMATION DES FONCTIONS DE RÉGRESSION.

1. Cas où Y est bornée.

Soit $[u, v]$ un intervalle compact pour lequel les hypothèses du théorème 20 sont vérifiées. On suppose que $|Y| \leq M$ où M est un nombre positif connu.

Choisissons a_n tel que $k_n = \frac{v-u}{2a_n}$ soit entier, $\forall n \in \mathbb{N}$ et posons :

$$u_1 = u + a_n, \quad u_2 = u + 3a_n, \quad \dots, \quad u_{k_n} = u + (2k_n - 1)a_n.$$

Soit alors $E(x) = \int_{-M}^{+M} y^\alpha h_x(y) dy$, $\alpha > 0$, une fonction de régression de Y en X . On définira un estimateur de $E_\alpha(x)$ en posant :

$$\hat{E}_{n,\alpha}(x) = \int_{-M}^{+M} y^\alpha \hat{h}_{n,u_i}(y) dy \quad \text{pour } x \in [u + (2i-2)a_n, u + 2ia_n[$$

$$i = 1, \dots, k_n^{-1}$$

et

$$\hat{E}_{n,\alpha}(x) = \int_{-M}^{+M} y^\alpha \hat{h}_{n,u_{k_n}}(y) dy \quad \text{pour } x \in [u + (2k_n - 2)a_n, u + 2k_n a_n]$$

THÉORÈME 21 :

Sous les hypothèses du théorème 20 et l'hypothèse supplémentaire

$$\sum_{n=0}^{\infty} a_n^{-1} \exp \left[-\lambda \frac{n a_n}{\alpha^2} \right] < +\infty, \quad \forall \lambda > 0 \quad (36)$$

$$\text{on a } \lim_{n \rightarrow \infty} \text{p.s.} \sup_{x \in [u, v]} |\hat{E}_{n,\alpha}(x) - E_\alpha(x)| = 0, \quad \forall \alpha > 0.$$

DÉMONSTRATION.

Tout d'abord, compte tenu des hypothèses 1°) et 2°) du théorème

20, on voit immédiatement que $h_x(y)$ est Lipchitzienne sur $[u, v] \times \mathbf{R}$, d'où l'on déduit que $E_\alpha(x)$ est Lipchitzienne sur $[u, v]$, $\forall \alpha > 0$. Il existe donc deux constantes positives δ et γ telles que

$$|E_\alpha(x) - E_\alpha(u_i)| \leq \delta |x - u_i|^\gamma, \quad \forall x \in [u, v], \quad \forall i \in \{1, \dots, k_n\}$$

d'où, pour $x \in [u_i - a_n, u_i + a_n]$

$$|E_\alpha(x) - E_\alpha(u_i)| \leq \delta \cdot a_n^\gamma, \quad \forall i \in \{1, \dots, k_n\}$$

Compte tenu du fait que $\lim_{n \rightarrow \infty} a_n = 0$ et de l'inégalité

$$|\hat{E}_{n,\alpha}(x) - E_\alpha(x)| \leq |\hat{E}_{n,\alpha}(u_i) - E_\alpha(u_i)| + |E_\alpha(u_i) - E_\alpha(x)|$$

$\forall x \in [u_i - a_n, u_i + a_n]$, $\forall i \in \{1, \dots, k_n\}$, il vient pour n assez grand

$$\text{Prob} \left[\sup_{x \in [u, v]} |\hat{E}_{n,\alpha}(x) - E_\alpha(x)| > \varepsilon \right] \leq \text{Prob} \left[\sup_{1 \leq i \leq k_n} |\hat{E}_{n,\alpha}(u_i) - E_\alpha(u_i)| > \frac{\varepsilon}{2} \right] \quad (37)$$

où ε est un nombre positif donné.

Or, $\forall i \in \{1, \dots, k_n\}$, on a l'inégalité :

$$|\hat{E}_{n,\alpha}(u_i) - E_\alpha(u_i)| \leq \int_{-M}^{+M} |y|^\alpha dy \cdot \sup_{|y| \leq M} |\hat{h}_{n,u_i}(y) - h_{u_i}(y)|$$

d'où

$$\text{Prob} \left[\sup_{x \in [u, v]} |\hat{E}_{n,\alpha}(x) - E_\alpha(x)| > \varepsilon \right] \leq \sum_{i=1}^{k_n} \text{Prob} \left[\sup_{|y| \leq M} |\hat{h}_{n,u_i}(y) - h_{u_i}(y)| > \frac{\alpha+1}{4} \cdot \frac{\varepsilon}{M^{\alpha+1}} \right]$$

en appliquant (26), on obtient, en posant $\frac{\alpha+1}{4} \cdot \frac{\varepsilon}{M^{\alpha+1}} = \eta$:

$$\text{Prob} \left[\sup_{x \in [u, v]} |\hat{E}_{n, \alpha}(x) - E_{\alpha}(x)| > \varepsilon \right] \leq \frac{v-u}{2a_n} K_1(h) \exp[-K_2(h) \frac{n a_n}{\alpha^2}]$$

$$\forall n > N(h)$$

et (36) permet de conclure.

2. Cas où Y n'est pas nécessairement borné.

Moyennant de légères modifications, la méthode précédente s'applique au cas où Y n'est pas nécessairement borné: on suppose que les hypothèses du théorème 20 et du corollaire 15 sont vérifiées et que :

$$E_{\alpha}(x) = \int_{\mathbf{R}} y^{\alpha} h_x(y) dy \quad \alpha > 0$$

existe, $\forall x \in [u, v]$. On définira alors un estimateur de $E_{\alpha}(x)$ en posant

$$E_{n, \alpha}^*(x) = \int_{-T_n}^{+T_n} y^{\alpha} \hat{h}_{n, u_i}(y) dy \quad \text{pour } x \in [u + (2i-2)a_n, u + 2i a_n]$$

$$i = 1, \dots, k_n - 1$$

et

$$E_{n, \alpha}^*(x) = \int_{-T_n}^{+T_n} y^{\alpha} \hat{h}_{n, u_{k_n}}(y) dy \quad \text{pour } x \in [u + 2k_n - 2)a_n, v]$$

on aura alors

THÉORÈME 21 bis :

Si

1°) Les hypothèses du théorème 20 et du corollaire 15 sont vérifiées.

2°) $\lim_{T \rightarrow \infty} \sup_{x \in [u, v]} \int_{|y| \geq T} |y|^{\alpha} h_x(y) dy = 0$

3°) $T_n = n^{\mu}$ avec $0 < \mu < \frac{1}{\alpha+1} \cdot \inf(d, b(1-d), \frac{1-d}{2} - c)$

alors

$$\lim_{n \rightarrow \infty} \text{p.s.} \sup_{x \in [u, v]} |E_{n, \alpha}^*(x) - E_{\alpha}(x)| = 0.$$

DÉMONSTRATION.

$$\text{Posons } E_{n, \alpha}(x) = \int_{-T_n}^{+T_n} y^{\alpha} h_x(y) dy.$$

D'après le 2°), il suffit de montrer que

$$\lim_{n \rightarrow \infty} \text{p.s.} \sup_{x \in [u, v]} |E_{n, \alpha}^*(x) - E_{n, \alpha}(x)| = 0.$$

La démonstration est alors analogue à celle du théorème 21 : $E_{n, \alpha}(x)$ est Lipchitzienne d'ordre 1 et il existe une constante δ telle que :

$$|E_{n, \alpha}(x) - E_{n, \alpha}(u_i)| \leq \frac{2}{\alpha+1} \cdot \delta \cdot T_n^{\alpha+1} \cdot a_n$$

$$\forall x \in [u_i - a_n, u_i + a_n], \quad \forall i \in \{1, \dots, k_n\}.$$

Comme on a $\mu < \frac{d}{\alpha+1}$, il vient $\lim_{n \rightarrow \infty} T_n^{\alpha+1} \cdot a_n = 0$

On a donc, pour n assez grand et ε positif donné, une inégalité analogue à (37), c'est-à-dire que :

$$\text{Prob} \left[\sup_{x \in [u, v]} |E_{n, \alpha}^*(x) - E_{n, \alpha}(x)| > \varepsilon \right]$$

$$\leq \text{Prob} \left[\sup_{1 \leq i \leq k_n} |E_{n, \alpha}^*(u_i) - E_{n, \alpha}(u_i)| > \frac{\varepsilon}{2} \right]$$

Or, $\forall i \in \{1, \dots, k_n\}$ on a l'inégalité

$$|E_{n, \alpha}^*(u_i) - E_{n, \alpha}(u_i)| \leq \int_{-T_n}^{+T_n} |y|^{\alpha} dy \cdot \sup_{|y| \leq T_n} |\hat{h}_{n, u_i}(y) - h_{u_i}(y)|$$

d'où :

$$\begin{aligned} \text{Prob} \left[\sup_{x \in [u, v]} |E_{n, \alpha}^*(x) - E_{n, \alpha}(x)| > \varepsilon \right] \\ \leq \sum_{i=1}^k \text{Prob} \left[\sup_{y \in \mathbf{R}} |\hat{h}_{n, u_i}(y) - h_{u_i}(y)| > \frac{\varepsilon (\alpha+1)}{4 \cdot T_n^{\alpha+1}} \right] \end{aligned} \quad (38)$$

Or, l'examen de la démonstration du corollaire 15 montre que pour $0 < \delta < \inf(d, b(1-d), \frac{1-d}{2} - c)$, on a une inégalité de la forme

$$\sup_{x \in [u, v]} \text{Prob} [n^\delta \cdot d(\hat{h}_{n, x}; h_x) > \varepsilon] \leq C_1 \exp [-C_2 n^\beta] \quad (39)$$

où C_1, C_2 et β sont positifs.

Alors (38) et (39) entraînent,

$$\text{Prob} \left[\sup_{x \in [u, v]} |E_{n, \alpha}^*(x) - E_{n, \alpha}(x)| > \varepsilon \right] \leq \frac{v-u}{2a_n} C_1 \exp [-C_2 n^\beta] \quad (40)$$

car $0 < \mu (\alpha+1) < \inf(d, b(1-d), \frac{1-d}{2} - c)$.

Or, on a $a_n = C'_3 \cdot n^{-d}$ ($d > 0$), donc le deuxième membre de (40) est le terme général d'une série convergente et la démonstration est terminée.

REMARQUE.

Dans [36], où seul le cas $\alpha = 1$ est traité, les conditions imposées à la fonction de régression sont un peu moins fortes que les nôtres (continuité au lieu de conditions du type Lipchitz); en revanche, les conditions d'existence de moments conditionnels jusqu'à l'ordre 3 qui figurent dans [36] ne sont pas exigées ici.

3. Un cas particulier important : le régressogramme.

TUKEY dans [8] a proposé d'estimer la courbe de régression sur un intervalle compact $[u, v]$ à l'aide du *régressogramme*. Cet estimateur est

défini de la manière suivante: on fait une partition de $[u, v]$ en intervalles de même longueur et on estime la fonction de régression par la fonction en escalier qui, dans chaque intervalle de la partition, est constante et égale à la moyenne des ordonnées des points de l'échantillon dont l'abscisse appartient à l'intervalle considéré ⁽¹⁾. TUKEY pose alors le problème du choix de la partition: le théorème 21 permet de répondre partiellement à cette question; en effet, le régressogramme appartient à la famille d'estimateurs de la régression que nous venons d'étudier.

Pour le voir, il suffit d'estimer $h_x(y)$ à l'aide d'un noyau K possédant les propriétés suivantes: il existe un nombre positif A tel que $K(t) = 0$ pour $|t| \geq A$ et tel que:

$$\int_{-A}^{+A} K(t) dt = 1 \quad (41) \quad \int_{-A}^{+A} t.K(t) dt = 0 \quad (42)$$

L'estimateur de $E_1(x)$ sera alors défini par

$$E_{n,1}^*(x) = 0 \quad , \quad \text{si } l_n = 0$$

$$= \frac{\alpha_{l_n}}{l_n} \sum_{k=1}^{l_n} \int_{-T_n}^{+T_n} y K[\alpha_{l_n} (y - y_{i_k})] dy, \quad \text{si } l_n \geq l.$$

D'où en faisant les changements de variables $t_{i_k} = \alpha_{l_n} (y - y_{i_k})$,
 $k = 1, \dots, l_n$.

$$E_{n,1}^*(x) = 0, \quad \text{si } l_n = 0$$

$$= \frac{1}{l_n} \sum_{k=1}^{l_n} y_{i_k} \int_{\alpha_{l_n} (-T_n - y_{i_k})}^{\alpha_{l_n} (T_n - y_{i_k})} K(t_{i_k}) dt_{i_k} -$$

$$- \frac{1}{l_n \cdot \alpha_{l_n}} \sum_{k=1}^{l_n} \int_{\alpha_{l_n} (-T_n - y_{i_k})}^{\alpha_{l_n} (T_n - y_{i_k})} t_{i_k} K(t_{i_k}) dt_{i_k}$$

si $l_n \geq l$.

(1) Si aucun point de l'échantillon n'est dans ce cas, on prend 0 comme valeur de l'estimateur sur l'intervalle considéré.

Comme $\lim_{n \rightarrow \infty} \alpha_n = +\infty$ on pourra toujours choisir $\alpha_n \geq 1$, $\forall n \in \mathbb{N}$; alors, si l'on suppose $|\mathcal{Y}| \leq M$ et si l'on tient compte du fait que $\lim_{n \rightarrow \infty} T_n = +\infty$, il vient, pour n assez grand :

$$\alpha_n (T_n - y_{i_k}) \geq T_n - M > A$$

et

$$\alpha_n (-T_n - y_{i_k}) \leq -T_n + M < A$$

d'où, en tenant compte de (41) et (42) :

$$E_{n,1}^*(x) = \begin{cases} 0 & \text{si } \ell_n = 0 \\ \frac{1}{\ell_n} \sum_{k=1}^{\ell_n} y_{i_k} & \text{si } \ell_n \geq 1 \end{cases}$$

On obtient donc le régressogramme associé à la partition de $[u, v]$ en intervalles de longueur $2a_n$. Remarquons que l'estimateur obtenu ne dépend pas de la suite α_n . On aura donc le

COROLLAIRE 16 :

Sous les hypothèses du théorème 21 bis, le régressogramme associé aux intervalles de longueur $O(n^{-d})$, $0 < d < 1$, converge presque sûrement vers $E_1(x)$, uniformément pour $x \in [u, v]$.

V

ESTIMATION DE LA FONCTION FRONTIÈRE

I - POSITION DU PROBLÈME.

1. Introduction.

Dans ce chapitre, ainsi que dans le chapitre VI, nous étudierons des questions liées plus ou moins directement à l'estimation du *support* ⁽¹⁾ d'une loi de probabilité.

RENYI et SULANKE [38] ont étudié le cas où le support est un sous-ensemble convexe de \mathbf{R}^2 ; l'enveloppe convexe de l'échantillon est alors un estimateur naturel de la frontière du support.

Le problème, beaucoup plus compliqué dans le cas non-convexe, n'est pas résolu dans toute sa généralité. Cependant, M. GEFFROY [39] a traité un cas particulier important: nous donnons ci-dessous quelques extensions de certains de ses résultats, notamment celles que nous avons indiquées dans [40].

2. Fonction frontière.

Soit K un espace topologique compact, à base dénombrable, muni de sa tribu Borélienne \mathcal{B} et soit λ une mesure de probabilité sur (K, \mathcal{B}) qui charge tout ouvert de K . Soit, d'autre part, s une application de K dans \mathbf{R} , continue et positive. Alors l'ensemble

$$S = \{(x, y) \in K \times \mathbf{R} \mid 0 \leq y \leq s(x)\}$$

est compact et, si K est connu, S est complètement déterminé par s ; nous dirons que s est la *fonction frontière* associée à S .

Soit, enfin, \mathcal{D} l'ensemble des mesures de probabilité sur $(K \times \mathbf{R}, \mathcal{B}_{\mathbf{R}})$ qui ont pour support S et qui chargent tout ouvert de S . On se propose d'estimer s à partir d'un échantillon $(x_i, y_i)_{1 \leq i \leq n}$ tiré suivant $P \in \mathcal{D}$.

(1) Ou de sa frontière, ce qui est équivalent dans les cas usuels.

II - CONSTRUCTION D'UN ESTIMATEUR CONVERGENT.

1. Construction d'un estimateur.

Généralisant la méthode employée dans [39], nous allons construire une suite d'estimateurs de s à partir d'une suite de partitions de K .

LEMME 20 :

Si λ est non-atomique, il existe une suite (P_n) de partitions finies de K formées de n ouverts quarrables Q_1^n, \dots, Q_n^n et d'un ensemble de mesure nulle N_n et telle que :

$$\lim_{n \rightarrow \infty} \text{Max}_{1 \leq \ell \leq n} |1 - n \lambda(Q_\ell^n)| = 0 \quad (1)$$

DÉMONSTRATION.

Tout d'abord, $\forall \varepsilon > 0$, il existe une partition finie de K formée d'ouverts quarrables de λ -mesure $\leq \varepsilon$ et d'un ensemble de λ -mesure nulle ⁽¹⁾.

Soit $n \in \mathbb{N}$; il existe donc une partition de K formée d'ouverts quarrables $Q_1^n, \dots, Q_{m_n}^n$ et d'un ensemble de mesure nulle N_n et telle que

$$\lambda(Q_i^n) \leq \frac{e^{-n}}{n}, \quad \forall i \in \{1, \dots, m_n\}$$

On peut toujours supposer que $\lambda(Q_i^n) > 0, \forall i \in \{1, \dots, m_n\}$.

Soit alors h_1 l'entier tel que

$$\lambda(Q_1^n + \dots + Q_{h_1-1}^n) < \frac{1}{n}$$

et
$$\lambda(Q_1^n + \dots + Q_{h_1}^n) \geq \frac{1}{n}$$

Posons $Q_1^n = Q_1^n + \dots + Q_{h_1}^n$; on aura $\frac{1}{n} \leq \lambda(Q_1^n) \leq \frac{1}{n} + \frac{e^{-n}}{n}$.

(1) cf. BOURBAKI [41], § 5.

Soit ensuite h_2 l'entier tel que

$$\lambda(Q'_{h_1+1} + \dots + Q'_{h_2-1}) < \frac{1}{n}$$

et

$$\lambda(Q'_{h_1+1} + \dots + Q'_{h_2}) \geq \frac{1}{n}$$

Posons de même $Q_2^n = Q'_{h_1+1} + \dots + Q'_{h_2}$; on aura

$$\frac{1}{n} \leq \lambda(Q_2^n) \leq \frac{1}{n} + \frac{e^{-n}}{n}$$

En continuant le processus, on définit des ensembles

$$Q_\ell^n \quad (1 \leq \ell \leq n-1) \quad \text{tels que} \quad \frac{1}{n} \leq \lambda(Q_\ell^n) \leq \frac{1}{n} + \frac{e^{-n}}{n}$$

Enfin on peut poser

$$Q_n^n = \sum_{i=1}^n Q_i^n - \sum_{\ell=1}^{n-1} Q_\ell^n$$

on aura alors

$$\frac{1}{n} - (n-1) \frac{e^{-n}}{n} \leq \lambda(Q_n^n) < \frac{1}{n}$$

La partition $\{Q_1^n, \dots, Q_n^n; N_n\}$ répond à la question. En effet, les Q_ℓ^n sont bien des ouverts quarrables disjoints tels que

$$\sum_{\ell=1}^n \lambda(Q_\ell^n) = 1 \quad \text{et l'inégalité}$$

$$\text{Max}_{1 \leq \ell \leq n} \left| \frac{1}{n} - \lambda(Q_\ell^n) \right| \leq e^{-n}$$

entraîne bien (1) . C.Q.F.D.

DÉFINITION DE L'ESTIMATEUR.

Soit $k = k(n) \in \mathbb{N}^{(1)}$ et soit $\{N_1^k, \dots, N_k^k\}$ une partition de N_k en ensembles mesurables. Posons :

$$Q_{kj} = Q_j^k + N_j^k \quad j = 1, \dots, k$$

De même que M. GEFFROY dans [39] on définira l'estimateur de s par les relations :

$$\hat{s}_n(x) = \sup_{i \in J_{nj}} y_i, \quad \forall x \in Q_{kj}, \quad \forall j \in \{1, \dots, k\} \quad (3)$$

où J_{nj} est l'ensemble des indices $i \in \{1, \dots, n\}$ tels que $x_i \in Q_{kj}$ et où l'on pose conventionnellement

$$\sup_{i \in J_{nj}} y_i = 0 \quad \text{si } J_{nj} = \emptyset.$$

REMARQUE :

Comme, d'après le lemme 20, les ensembles Q_{kj}^0 forment, à un ensemble de mesure nulle près, une partition de K , on en déduit que $\hat{s}_n(x)$ est λ -presque partout continue.

NOTATIONS :

De même que dans le cas traité par M. GEFFROY, on peut poser :

$$M_{kj} = \text{Max}_{x \in Q_{kj}} s(x) = \sup_{x \in Q_{kj}} s(x); \quad j = 1, \dots, k.$$

$$m_{kj} = \text{Min}_{x \in Q_{kj}} s(x) = \inf_{x \in Q_{kj}} s(x), \quad j = 1, \dots, k.$$

$$S_n = \text{Max}_{1 \leq j \leq k} (M_{kj} - U_{nj})$$

(1) On suppose que $\lim_{n \rightarrow \infty} k(n) = +\infty$.

$$\text{avec } U_{nj} = \sup_{i \in J_{nj}} y_i$$

et on peut considérer les ensembles :

$$D_{kj} = \{ (x, y) \in K \times \mathbf{R} \mid x \in Q_{kj}, 0 \leq y \leq s(x) \}$$

$$D'_{kj}(z) = \{ (x, y) \in K \times \mathbf{R} \mid x \in Q_{kj}, s(x) \geq M_{kj} - z \}$$

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

Enfin, si l'on désigne par $E_{nj}(z)$ l'évènement

$$\bigcap_{i=1}^n \{ (x_i, y_i) \in D'_{kj}(z) \}$$

et par $E'_{nj}(z)$ l'évènement contraire on aura encore

$$\text{Prob} [Z_n < z] = \text{Prob} \left[\bigcap_{j=1}^k E'_{nj}(z) \right].$$

2. Convergence de l'estimateur.

Le théorème 2 de [39] se généralise de la façon suivante :

THÉORÈME 22 :

Si $P \in \mathcal{D}$ est absolument continue par rapport à $\lambda \otimes \ell$ (ℓ désignant la mesure de Lebesgue sur \mathbf{R}), et s'il existe une version de sa dérivée de Radon-Nikodym, soit $\varphi(x, y)$ telle que l'on ait :

$$A \leq \varphi(x, y) \leq B, \quad \forall (x, y) \in S \quad (4)$$

où A et B sont deux constantes positives, alors, pour que \hat{s}_n tende vers s uniformément presque sûrement, il est nécessaire que $k(n) = o\left(\frac{n}{\text{Log } n}\right)$.

Si de plus, le maximum des diamètres des ensembles Q_{kj} (relatif à une distance compatible avec la topologie de K) tend vers 0, la condition est aussi suffisante.

DEMONSTRATION.

Soit d la distance de la convergence uniforme.

On a successivement

$$d(\hat{s}_n, s) \xrightarrow{\text{p.s.}} 0 \implies Z_n \xrightarrow{\text{p.s.}} 0 \implies \text{Prob} \left[\bigcup_{j=1}^k E_{nj}(z) \right] \rightarrow 0.$$

D'autre part

$$\begin{aligned} \text{Prob} [D_{kj}(z)] &= \iint_{D_{kj}(z)} \varphi(x, y) d\lambda(x) d\ell(y) \\ &\leq \iint_{Q_{kj} \times [M_{kj}^{-s}, M_{kj}]} \varphi(x, y) d\lambda(x) d\ell(y) \\ &\leq B \int_{Q_{kj}} d\lambda(x) \int_{M_{kj}^{-z}}^{M_{kj}} d\ell(y) \end{aligned}$$

$$\text{d'où } \text{Prob} [D_{kj}(z)] \leq Bz \lambda(Q_{kj}) \leq \frac{Bz}{k} (1 + ke^{-k})$$

la dernière inégalité étant une conséquence de (2).

On aura donc, pour k assez grand

$$\text{Prob} [D_{kj}(z)] < \frac{Bz}{k} (1 + \varepsilon) \quad (\varepsilon > 0, \text{ donné})$$

et l'on termine comme dans [39] (p. 197-98).

RÉCIPROQUE :

Soit $z \leq \frac{1}{2} [\inf_{x \in K} s(x)] = \frac{m}{2}$. On aura alors

$$\text{Prob} [D_{kj}(z)] = \iint_{D_{kj}(z)} \varphi(x, y) d\lambda(x) d\ell(y) \geq A \frac{m}{2} \lambda(Q_{kj})$$

donc, d'après (2), on aura

$$\text{Prob} [D_{kj}(z)] \geq \frac{Am}{2k} (1 - ke^{-k}) \geq \frac{Am}{2k} (1 - \varepsilon) \quad \text{pour } k \text{ assez grand.}$$

Alors, de même que dans [39] (p. 198-99), on en déduit que $Z_n \xrightarrow[p]{0}$ puis que $Z_n \xrightarrow[p.s.]{0}$ ⁽¹⁾.

Enfin, compte tenu de l'inégalité

$$d(\hat{s}_n, s) \leq Z_n + \sup_{1 \leq j \leq k} (M_{kj} - m_{kj})$$

et de l'hypothèse faite sur les diamètres des Q_{kj} , on en déduit que

$$\lim_{n \rightarrow \infty} p.s. \quad d(\hat{s}_n, s) = 0.$$

C.Q.F.D.

REMARQUE:

La convergence est même presque complètement sûre.

III - CAS D'UN ESPACE LOCALEMENT COMPACT.

Les résultats précédents s'étendent facilement au cas où K est localement compact à base dénombrable ⁽²⁾. On reprend les mêmes hypothèses que dans 1.2 et on suppose, en outre, que s admet, suivant le filtre des voisinages de l'infini, une limite σ positive.

Soit alors K' le compactifié d'Alexandroff de K : $K' = K + \{\omega\}$, ω désignant le point à l'infini de K . On peut se ramener au problème précédent en posant :

$$\begin{cases} s'(x) = s(x) & \text{pour } x \in K \\ s'(\omega) = \sigma \end{cases}$$

$$\begin{cases} \lambda'(A) = \lambda(A) & \text{si } A \text{ est un Borélien de } K \\ \lambda'(\{\omega\}) = 0 \end{cases}$$

et, $\forall P \in \mathcal{F}$,

(1) D'après la démonstration du corollaire 2 de [39] (p.199).

(2) K est supposé non compact.

$$\begin{cases} P'(A) = P(A) & \text{si } A \text{ est un Borélien de } K \\ P'(\{\omega\}) = 0 \end{cases}$$

Le lemme 20 s'applique alors à λ' donc, par projection sur K , à λ et, \hat{s}'_n désignant l'estimateur de s' défini par (3), on voit donc que le théorème 22 s'applique à \hat{s}'_n donc à sa restriction à K , \hat{s}_n . Finalement, on peut énoncer le

COROLLAIRE 17 :

Si K est localement compact à base dénombrable, le lemme 20 et le théorème 22 restent valables.

REMARQUE.

Dans l'énoncé du théorème 22 appliqué au cas localement compact, la distance considérée doit être compatible avec la topologie de K' .

APPLICATION :

Le corollaire 17 s'applique au cas particulier important où $K = \mathbf{R}$.

Commençons par exhiber une famille de suites de partitions de \mathbf{R} qui satisfont aux conditions du lemme 20 et du théorème 22. Nous nous appuyerons sur le lemme suivant :

LEMME 21 :

Soit E un espace métrique compact et soit μ une mesure sur E qui charge tout ouvert, alors si à tout $n \in \mathbf{N}$ on associe des Boules $B_1, \dots, B_{h(n)}$ de telle façon que

$$\lim_{n \rightarrow \infty} \sup_{1 \leq i \leq h(n)} \mu(B_i) = 0 \tag{5}$$

on a aussi

$$\lim_{n \rightarrow \infty} \sup_{1 \leq i \leq h(n)} D(B_i) = 0 \tag{6}$$

$D(B_i)$ désignant le diamètre de B_i ; $i = 1, \dots, h(n)$.

DÉMONSTRATION.

On notera $B(a, r)$ la boule ouverte de centre a et de rayon r . Soit $B(y_n, r_n)$ une suite de Boules telle que $\mu [B(y_n, r_n)] \rightarrow 0$. Si $r_n \not\rightarrow 0$, il existe $\varepsilon > 0$ et une suite extraite (r_{n_k}) telle que $r_{n_k} \geq \varepsilon$, $\forall k \in \mathbb{N}$.

On a alors

$$\mu [B(y_{n_k}, r_{n_k})] \geq \mu [B(y_{n_k}, \varepsilon)] , \quad \forall k \in \mathbb{N}.$$

donc

$$\lim_{n \rightarrow \infty} \mu [B(y_{n_k}, \varepsilon)] = 0.$$

Or, de (y_{n_k}) on peut extraire une suite $(y_{n_{k_\ell}})$ qui converge vers un élément de Y de E donc, pour ℓ assez grand on aura

$$B(y, \frac{\varepsilon}{2}) \subset B(y_{n_{k_\ell}}, \varepsilon)$$

mais d'après l'hypothèse faite sur μ , $\mu [B(y, \frac{\varepsilon}{2})] > 0$ et l'on aboutit à une contradiction. On a donc nécessairement $r_n \rightarrow 0$.

Cela dit, la démonstration du lemme 21 s'achève de la façon suivante: Parmi les Boules $B_1, \dots, B_{h(n)}$, soit B_n celle qui a le plus petit indice parmi celles de plus grand diamètre; d'après (5), on a $\mu (B_n) \rightarrow 0$ donc, d'après ce qui précède, le rayon de B_n tend vers 0, ce qui entraîne immédiatement (6). C.Q.F.D.

Soit alors λ une mesure de probabilité sur $\mathbf{R}^{(1)}$ dont la fonction

(1) On posera ici $\bar{\mathbf{R}} = \mathbf{R} + \{-\infty, +\infty\}$: il est clair que le corollaire 17 s'applique encore moyennant les hypothèses

$$\lim_{x \rightarrow -\infty} s(x) = \sigma > 0 ; \quad \lim_{x \rightarrow +\infty} s(x) = \sigma' > 0.$$

de répartition F_λ est continue et strictement croissante; on peut lui associer la suite de partitions :

$$[-\infty, a_1[; [a_1, a_2[; \dots [a_{k-2}, a_{k-1}[; [a_{k-1}, +\infty] \quad (7)$$

$$\text{avec } a_i = F_\lambda^{-1} \left(\frac{i}{k} \right); \quad i = 1, \dots, k-1; \quad k = 1, 2, \dots \quad (8)$$

Il est clair que la condition (5) est vérifiée donc, d'après le lemme 21, la condition (6) est également vérifiée.

REMARQUE:

Dans (6), D désigne le diamètre relatif à une distance d compatible avec la topologie de \mathbf{R} et pour laquelle les ensembles de la partition précédente ont pour intérieur des Boules ouvertes. Il est facile de construire une telle distance. On pourra prendre, par exemple

$$d(x, y) = | \text{Arc tg } x - \text{Arc tg } y |; \quad x, y \in \mathbf{R}.$$

CONCLUSION.

Si P est une mesure de probabilité sur \mathbf{R}^2 dont le support est limité par la droite $y = 0$ et le graphe d'une fonction $s(x)$ continue, strictement positive et admettant des limites positives pour $x \rightarrow +\infty$; si, de plus, P est absolument continue par rapport à $\lambda \otimes \ell$ et si sa densité vérifie (4), alors le corollaire 17 s'applique, c'est-à-dire que, pour

$$k(n) = 0 \quad \left(\frac{n}{\text{Log } n} \right) \quad \text{on a :}$$

$$\lim_{n \rightarrow \infty} \text{p.s. } d(\hat{s}_n, s) = 0$$

\hat{s}_n étant l'estimateur défini par (3) associé à la partition définie par (7) et (8).

IV - CAS OU LA FONCTION FRONTIÈRE EST DE SIGNE QUELCONQUE.

1. Préliminaires.

Dans ce paragraphe, nous nous proposons de montrer que les résul-

tats précédents subsistent quand $s(x)$ peut s'annuler ou changer de signe un nombre fini de fois sur \mathbf{R} .

LEMME 22 :

Soit (X_n) une suite de variables aléatoires telle que :

$$\lim_{n \rightarrow \infty} p. X_n = 0$$

Soit (N_k) une suite de variables aléatoires à valeurs entières telle que :

$$\lim_{n \rightarrow \infty} p. N_k = \infty$$

Alors, si les suites (X_n) et (N_k) sont indépendantes, on a aussi

$$\lim_{n \rightarrow \infty} p. X_{n_k} = 0$$

DÉMONSTRATION.

Soit $x > 0$; on peut écrire $\forall k \in \mathbf{N}$

$$\text{Prob} [|X_{N_k}| > x] = \sum_{n \in A_k} \text{Prob} [N_k = n] \text{Prob} [|X_{N_k}| > x | N_k = n]$$

$$\text{avec } A_k = \{n \in \mathbf{N} | \text{Prob} [N_k = n] > 0\}$$

Compte tenu de l'hypothèse d'indépendance, il vient :

$$\begin{aligned} \text{Prob} [|X_{N_k}| > x] &= \sum_{n \in A_k} \text{Prob} [N_k = n] \cdot \text{Prob} [|X_n| > x] \\ &= \sum_{n=0}^{+\infty} \text{Prob} [N_k = n] \cdot \text{Prob} [|X_n| > x] \end{aligned}$$

$$\text{Alors, } \forall \epsilon > 0, \quad N(\epsilon, x) \text{ tel que } n \geq N(\epsilon, x) \implies \text{Prob} [|X_n| > x] < \frac{\epsilon}{2}$$

d'où :

$$\text{Prob} [|X_{N_k}| > x] < \sum_{n=0}^{N(\epsilon, x)} \text{Prob} [N_k = n] + \frac{\epsilon}{2}$$

mais $\exists [K \quad [N(\epsilon, x)] \quad \text{tel que}$

$$k \geq K [N(\epsilon, x)] \implies \text{Prob} [N_k \leq N(\epsilon, x)] < \frac{\epsilon}{2}$$

d'où finalement

$$\text{Prob} [|X_{N_k}| > x] < \epsilon \quad \text{dès que } k \geq K [N(\epsilon, x)]$$

$\forall x > 0, \quad \forall \epsilon > 0. \quad \text{C.Q.F.D.}$

NOTATIONS

\hat{s}_n désignant l'estimateur défini par (3), nous dirons dorénavant que \hat{s}_n est l'estimateur de Geffroy associé à la partition $\{Q_{kj}\}$.

D'autre part, nous supposons que K est un intervalle de \mathbf{R} . Soit alors J un sous-intervalle de K qui est réunion d'ensembles Q_{kj} .

$$\text{Posons } S_j = \{(x, y) \in J \times \mathbf{R} \mid 0 \leq y \leq s(x)\}$$

et désignons par s_j et $\hat{s}_{n,J}$ les restrictions respectives de s et de \hat{s}_n à J .

Enfin P étant une mesure de probabilité sur \mathbf{R}^2 , appelons P_j la mesure de probabilité induite par P sur S_j .

On a alors le

LEMME 23 :

Si les points de l'échantillon sont tirés suivant la loi de probabilité P , $\hat{s}_{n,J}$ est l'estimateur de Geffroy de s_j associé à la partition induite par $\{Q_{kj}\}$ sur J , à la loi de probabilité P_j et à un échantillon dont la taille est une variable aléatoire indépendante des variables de l'échantillon et qui suit une loi binominale de paramètres $P(S_j)$ et n .

DÉMONSTRATION.

Posons $J = \sum_{j=1}^{k_1} Q_{kj}$ (ce que l'on peut faire au besoin en changeant les notations) et

$$D'_{kj}(z) = \{ (x,y) \in Q_{kj} \times \mathbf{R} \mid z < y < f(x) \} ,$$

$$\forall z \in [0, f(x)] , \quad \forall x \in J , \quad \forall j \in \{ 1, \dots, k_1 \} .$$

On aura successivement et $\forall j \in \{ 1, \dots, k_1 \}$:

$$\begin{aligned} \text{Prob} [U_{nj} < z] &= \text{Prob} [\bigcap_{i=1}^n (x_i, y_i) \notin D'_{kj}(z)] \\ &= [1 - P(D'_{kj}(z))]^n \\ &= [1 - P(S_j) + P(S_j) - P(D'_{kj}(z))]^n \\ &= \sum_{\ell=0}^n C_n^\ell [1 - P(S_j)]^{n-\ell} \left[\frac{P(S_j - D'_{kj}(z))}{P(S_j)} \right]^\ell \\ &= \sum_{\ell=0}^n \text{Prob} [L_n = \ell] P_J^{\otimes \ell} \left[\bigcap_{i=1}^{\ell} (x_i, y_i) \notin D'_{kj}(z) \right] \end{aligned}$$

où L_n désigne une variable aléatoire qui suit une loi binominale de paramètres $P(S_j)$ et n .

Cette dernière expression de $\text{Prob} [U_{nj} < z]$ correspond bien aux conditions de l'énoncé du lemme 23.

2. Cas où s s'annule à l'infini.

COROLLAIRE 18 :

Si s est continue, positive, définie $\forall x \in \mathbf{R}$ et telle que $\lim_{|x| \rightarrow \infty} s(x) = 0$, si, d'autre part, P est absolument continue par rapport à la mesure de Lebesgue sur \mathbf{R}^2 et s'il existe une version de sa densité,

soit $\varphi(x,y)$, telle que, $\forall I$ intervalle borné de \mathbf{R} il existe deux constantes positives $A(I)$ et $B(I)$ telles que

$$A(I) \leq \varphi(x,y) \leq B(I) \quad , \quad \forall (x,y) \in S_I \quad (9)$$

alors, pour que l'estimateur de Geffroy associé aux ensembles

$$\left[p + \frac{j-1}{k} \right] , \quad p + \frac{j}{k} [\quad , \quad p \in \mathbf{Z} \quad ,$$

$j = 1, \dots, k$, converge uniformément en probabilité, il suffit que

$$k(n) = o\left(\frac{n}{\text{Log } n}\right) \quad (1) \quad .$$

DÉMONSTRATION.

Soit $\varepsilon > 0$, il existe un entier A tel que $|x| \geq A$ entraîne $s(x) < \varepsilon$.

Soient s_A et $\hat{s}_{n,A}$ les restrictions de s et de \hat{s}_n à $[-A, +A]$: d'après le lemme 23, $\hat{s}_{n,A}$ est l'estimateur de Geffroy associé à la partition

$$\left[p + \frac{j-1}{k} \right] , \quad p + \frac{j}{k} [\quad , \quad p \in \{-A, -A+1, \dots, A-1\} \quad j \in \{1, \dots, k\} ,$$

à la loi $\frac{P}{P(S_{[-A, +A]})}$, et à un échantillon dont la taille L_n est une variable aléatoire qui suit une loi binominale de paramètres $P(S_{[-A, +A]})$ et n .

On a donc $\lim_{n \rightarrow \infty} p \cdot L_n = +\infty$.

Or, d'après le théorème 22 :

$$k = o\left(\frac{n}{\text{Log } n}\right) \implies 2A k = o\left(\frac{n}{\text{Log } n}\right) \implies \lim_{q \rightarrow \infty} p \cdot d(s_A; \hat{s}'_{q,A}) = 0$$

(1) Remarquons que, dans ce corollaire, λ n'est plus une mesure bornée mais est la mesure de Lebesgue sur \mathbf{R} et que $\{Q_{kj}\}$ est formée d'une infinité d'intervalles.

$\hat{s}'_{q,A}$ désignant ici l'estimateur de Geffroy de s_A associé à la partition précédente et à un échantillon de taille q tiré suivant $P_{[-A, +A]}$.

D'où, en appliquant le lemme 22 aux suites

$d(s_A; \hat{s}'_{q,A})$ et L_n :

$$k = 0 \left(\frac{n}{\text{Log } n} \right) \implies \lim_{n \rightarrow \infty} \text{p. } d(s_A; \hat{s}'_{n,A}) = 0$$

Enfin, l'équivalence $d(s_A; \hat{s}'_{n,A}) < \varepsilon \iff d(s, \hat{s}_n) < \varepsilon$

$\forall n \in \mathbb{N}$ permet de conclure.

3. Cas où s s'annule a distance finie.

Dans cette partie, nous supposons que K est un intervalle compact de \mathbb{R} ; on pourra prendre $K = [0, 1]$ sans restreindre la généralité.

LEMME 24 :

Soit $\{Q_{kj}^r\}_{j=1, \dots, h}$ la partition induite par

$\left[\frac{j-1}{k}, \frac{j}{k} \right[$, $j = 1, \dots, k-1$ sur l'intervalle $[0, r_0]$ où r_0 est un rationnel appartenant à $]0, 1[$. Alors les longueurs des intervalles Q_{kj}^r , sont des infiniment petits de même ordre que $\frac{1}{k}$.

DÉMONSTRATION.

Posons $r_0 = \frac{p}{q_0}$ en supposant cette fraction irréductible. h désignant le nombre d'intervalles de la partition $\{Q_{kj}^r\}$, il est clair que $r_0 k \leq h \leq r_0 k + 1$; h est donc un infiniment grand du même ordre que k .

Si $r_0 k$ est entier, donc si $k = \lambda q_0$, $\lambda \in \mathbb{N}$, on a $h = r_0 k$ et la partition est formée de h intervalles de longueur $\frac{1}{k}$.

Si $r_0 k$ n'est pas entier, on peut écrire :

$$k = \lambda q_0 + \mu, \quad \lambda \in \mathbb{N}, \quad \mu \in \mathbb{N}, \quad \text{avec } 1 \leq \mu \leq q_0 - 1$$

et l'on a alors $h = [r_0 k + 1] = [r_0 k] + 1$: dans ce cas, les $h-1$ premiers intervalles de la partition $\{Q_{kj}^r\}$ sont de longueur $\frac{1}{k}$; montrons que $[\frac{h-1}{k}, r_0]$ a une longueur de l'ordre de $\frac{1}{k}$. On peut écrire

$$r_0 - \frac{h-1}{k} = \frac{1}{k} [r_0 k - [r_0 k]]$$

$$\begin{aligned} \text{Or } r_0 k - [r_0 k] &= \frac{P_0}{q_0} (\lambda q_0 + \mu) - \left[\frac{P_0}{q_0} (\lambda q_0 + \mu) \right] \\ &= r_0 \mu - [r_0 \mu] \end{aligned}$$

Or $r_0 \mu$ ne peut être entier puisque $1 \leq \mu \leq q_0 - 1$, donc $r_0 \mu - [r_0 \mu] \neq 0$ et comme cette quantité ne peut prendre qu'un nombre fini de valeurs, la démonstration est terminée.

COROLLAIRE 19 :

Si s est continue sur $[0,1]$ et positive sauf au point $c \in]0,1[$ où elle s'annule, l'estimateur de Geffroy associé à la partition

$$G_k = \left\{ \left[0, \frac{1}{k}[, \left[\frac{1}{k}, \frac{2}{k}[, \dots, \left[\frac{k-1}{k}, 1\right] \right\}$$

converge uniformément en probabilité si $k(n) = O\left(\frac{n}{\text{Log } n}\right)$ et si la condition (9) est vérifiée pour $I = [0,1]$.

DÉMONSTRATION.

Soit $\varepsilon > 0$; on peut trouver deux rationnels r_0 et $r_1 \in]0,1[$ tels que $c \in]r_0, r_1[$ et $0 \leq s(x) \leq \frac{\varepsilon}{4}, \forall x \in [r_0, r_1]$.

Soit G_k^r la partition induite par G_k sur $[0, r_0]$ et soit \hat{t}_{n,r_0}

l'estimateur de Geffroy associé à l'échantillon induit sur $S_{[0, r_0]}$. Un raisonnement tout à fait analogue à celui qui a été fait dans la démonstration du corollaire 18 montre, compte tenu du lemme 24, que \hat{t}_{n, r_0} converge uniformément en probabilité vers $s_{[0, r_0]}$.

Or, $\hat{t}_{n, r_0} = \hat{s}_{n, [0, r_0]}$ sauf peut-être sur le $h^{\text{ème}}$ intervalle de $G_{k_0}^r$ où l'on a $\hat{t}_{n, r_0} \leq \hat{s}_{n, [0, r_0]}$. Mais on peut choisir n donc k et donc h assez grands pour que, sur cet intervalle, on ait $s(x) < \frac{\varepsilon}{2}$, ce qui entraîne alors $d(\hat{t}_{n, r_0}; \hat{s}_{n, [0, r_0]}) < \frac{\varepsilon}{2}$.

On en déduit donc que $\hat{s}_{n, [0, r_0]}$ converge uniformément en probabilité vers $s_{[0, r_0]}$.

On montre de même que $\hat{s}_{n, [r_1, 1]}$ converge uniformément en probabilité vers $s_{[r_1, 1]}$.

Enfin, il est clair que $d(s_{[r_0, r_1]}, \hat{s}_{n, [r_0, r_1]}) < \frac{\varepsilon}{4}$.

$$\text{Or, } d(\hat{s}_n, s) = \sup [d(s_{[0, r_0]}, \hat{s}_{n, [0, r_0]}); d(s_{[r_0, r_1]}, \hat{s}_{n, [r_0, r_1]}); \\ d(s_{[r_1, 1]}, \hat{s}_{n, [r_1, 1]})]$$

donc, pour n assez grand

$$\text{Prob } [d(\hat{s}_n, s) > \varepsilon] < \varepsilon$$

C.Q.F.D.

4. Cas où s change de signe.

Si s n'est plus supposée positive, il faut généraliser l'estimateur défini par (3). On posera :

$$\hat{v}_n(x) = \sup_{i \in J_{nj}^+} y_i + \inf_{i \in J_{nj}^-} y_i, \quad \forall x \in Q_{kj}, \quad \forall j$$

avec $J_{nj}^+ = \{i \in \{1, \dots, n\} \mid x_i \in Q_{kj}, y_i \geq 0\}$

$J_{nj}^- = \{i \in \{1, \dots, n\} \mid x_i \in Q_{kj}, y_i < 0\}$

et les conventions $\sup_{i \in J_{nj}^+} y_i = 0$ pour $J_{nj}^+ = \emptyset$

et $\inf_{i \in J_{nj}^-} y_i = 0$ pour $J_{nj}^- = \emptyset$.

On a alors le

COROLLAIRE 20 :

Sous les conditions du corollaire 18, excepté le fait que s peut s'annuler ou changer de signe un nombre fini de fois sur \mathbf{R} , \hat{v}_n converge uniformément en probabilité vers s .

DÉMONSTRATION.

Immédiate en appliquant les corollaires 18 et 19 à $|s|$.

5. Convergence presque sûre.

Nous avons énoncé les résultats de ce paragraphe "en convergence en probabilité", mais ils restent valables "en convergence presque sûre"; c'est une conséquence du

LEMME 22 bis :

Soit (X_n) une suite de variables aléatoires telle que

$$\lim_{n \rightarrow \infty} \text{p.s. } X_n = 0$$

Soit (N_k) une suite de variables aléatoires à valeurs entières telle que

$$\lim_{k \rightarrow \infty} \text{p.s. } N_k = \infty$$

Alors, on a aussi

$$\lim_{k \rightarrow \infty} \text{p.s. } X_{N_k} = 0$$

DÉMONSTRATION.

D'après les hypothèses, il existe dans $\Omega^{\mathbb{N}}$ un ensemble E de probabilité nulle et tel que $X_n(\omega) \rightarrow 0$ et $N_k(\omega) \rightarrow +\infty$, $\forall \omega \in \Omega^{\mathbb{N}} - E$. Il s'agit de montrer que l'on a aussi $X_{N_k}(\omega) \rightarrow 0$, $\forall \omega \in \Omega^{\mathbb{N}} - E$.

Or, $\forall \omega \in \Omega^{\mathbb{N}} - E$, $\forall \varepsilon > 0$, il existe un entier $\lambda(\varepsilon, \omega)$ tel que $|X_n(\omega)| < \varepsilon$; il existe alors un entier $K(\varepsilon, \omega)$ tel que $k > K(\varepsilon, \omega)$ entraîne $N_k(\omega) > \lambda(\varepsilon, \omega)$, ce qui entraîne finalement $|X_{N_k}(\omega)| < \varepsilon$.
C.Q.F.D.

V - APPLICATION : DÉFINITION D'UNE MÉTHODE DE MONTE CARLO

Revenons au cas où s est une fonction continue et positive définie sur $[0,1]$. Sous les hypothèses du théorème 2 de [39], il est clair que :

$$\int_0^1 \hat{s}_n(x) dx \xrightarrow{\text{p.s.}} \int_0^1 s(x) dx \tag{10}$$

On peut utiliser (10) pour définir une Méthode de Monte Carlo :

Soit à calculer $\int_0^1 s(x) dx$ où s est une fonction continue et positive connue: on tire $2n$ nombres $x_1, \dots, x_n; z_1, \dots, z_n$ suivant une loi uniforme sur $[0, 1]$ et indépendamment les uns des autres et on estime

$$\int_0^1 s(x) .dx$$

par

$$\int_0^1 \hat{s}_n(x) dx$$

\hat{s}_n étant construit à partir de l'échantillon

$$(x_i, z_i s(x_i))_{1 \leq i \leq n}$$

La précision de cette Méthode dépend évidemment du choix de $k(n)$ et on peut montrer le résultat suivant :

Si s est Lipchitzienne et si $k(n) = O(n^\alpha)$, $\frac{1}{3} < \alpha < \frac{1}{2}$,

alors $\sigma^2 \left[\int_0^1 \hat{s}_n(x) dx \right] = O\left(\frac{1}{n^{1-\alpha}}\right)$

VI

ÉLÉMENTS EXTRÊMAUX D'UN ÉCHANTILLON VECTORIEL

I - GÉNÉRALITES

1. Introduction.

Le problème que nous étudions dans ce chapitre nous a été proposé par M. GEFFROY: il s'agit d'étudier les éléments extrémaux d'un échantillon de variables aléatoires à valeurs dans un espace partiellement ordonné.

Cette étude semble intéressante car ces éléments sont analogues aux valeurs extrêmes d'un échantillon de variables aléatoires réelles. De plus, la répartition des éléments extrémaux donne des indications sur l'aspect du support de la loi des variables de l'échantillon.

Dans la suite, nous étudions surtout le comportement asymptotique du nombre d'éléments extrémaux: nous donnons les résultats qui figurent dans [42] ainsi que certains résultats inédits que M. GEFFROY a eu l'obligeance de nous communiquer.

2. Éléments extrémaux d'une partie d'un espace ordonné.

a) Soit E un espace ordonné. Nous noterons \leq la relation d'ordre de E , \geq la relation d'ordre inverse; enfin si x et y ne sont pas comparables, nous écrivons $x \wedge y$.

DÉFINITION 12 :

Soit A une partie de E . On dit que $y \in A$ est un élément maximal de A si, $\forall x \in A$, on a $x \leq y$ ou $x \wedge y$. On dit que $y \in A$ est un élément minimal de A si, $\forall x \in A$, on a $x \geq y$ ou $x \wedge y$. On dit que $y \in A$ est un élément extrémal de A si y est un élément maximal ou minimal de A .

L'existence d'éléments extrémaux est assurée par le

LEMME 25 :

Si A est une partie finie de E , alors A admet toujours au moins un élément maximal (resp. minimal).

DÉMONSTRATION.

Immédiate par récurrence sur le nombre d'éléments de A .

b)

LEMME 26 :

Soit f une application de E dans E . La relation R_f définie par $x R_f y \iff f(x) \leq f(y)$ est une relation d'ordre si et seulement si f est injective.

DÉMONSTRATION.

Elémentaire.

NOTATION :

$x R_f y$ sera notée $x \leq y (f)$ et appelée relation d'ordre induite par f sur E .

DÉFINITION 12 bis :

Soit A une partie de E . On dira que $y \in A$ est un élément extrémal de A , modulo f , si y est un élément extrémal de A pour la relation d'ordre R_f .

Propriétés élémentaires.

On montre très facilement les propriétés suivantes :

1°) R_f coïncide avec \leq si et seulement si f est un automorphisme de E .

2°) Soient f et g deux bijections de E . f et g induisent le même ordre sur E si et seulement si $g \circ f^{-1}$ est un automorphisme de E .

3°) Si E est un treillis et si f est une bijection de E , alors E est un treillis modulo f et l'on a les relations :

$$\sup_f \{x, y\} = f^{-1} [\sup \{f(x), f(y)\}] \quad (1)$$

$$\inf_f \{ x, y \} = f^{-1} [\inf \{ f(x), f(y) \}] \quad (2)$$

4°) Si E est un groupe ordonné les applications f et $-f$ induisent des relations d'ordre inverses l'une de l'autre.

3. Éléments extrémaux d'un échantillon de variables aléatoires à valeurs dans un espace mesurable ordonné.

a) DÉFINITION 13 :

On dira que E est un espace mesurable ordonné (E.M.O.) si E est un espace mesurable, un espace ordonné et si les intervalles de E sont mesurables.

Si E est un E.M.O., si f est une application injective et mesurable de E dans E et si P est une mesure de probabilité sur E l'application F_f^P définie par

$$F_f^P(x) = P(] \leftarrow, x []_f), \quad \forall x \in E \quad (3)$$

est la fonction de répartition de P , modulo f .⁽¹⁾

La définition de R_f montre que

$$F_f^P(x) = F_u^{fP} [f(x)], \quad \forall x \in E \quad (4)$$

fP désignant la mesure de probabilité image de P par f et u désignant l'application identique de E .

b) Soit x_1, \dots, x_n un échantillon de taille n à valeurs dans E , on dira que x_i est un élément maximal de l'échantillon si x_i est un élément maximal de l'ensemble $\{ x_1, \dots, x_n \}$. On définit de même les éléments minimaux, extrémaux, extrémaux modulo f etc... de l'échantillon.

Remarquons qu'on ne peut pas définir les éléments maximaux comme des variables aléatoires à valeurs dans E , leur nombre n'étant pas fixe.

(1) $] \leftarrow, x []_f$ désigne l'ensemble des $y \in E$ tels que $y \leq x(f)$ et $y \neq x$.

Par contre, le nombre d'éléments maximaux de l'échantillon est une variable aléatoire réelle parfaitement définie.

II - LOI DU NOMBRE D'ÉLÉMENTS MAXIMAUX ASSOCIÉS A UNE ROTATION DE \mathbb{R}^2 .

Nous commençons maintenant l'étude des propriétés du nombre d'éléments maximaux d'un échantillon. Nous étudierons essentiellement le cas où $E = \mathbb{R}^2$ muni de sa tribu Borélienne et de sa relation d'ordre habituelle.

NOTATIONS :

- X_1, \dots, X_n désignent n vecteurs aléatoires à valeurs dans \mathbb{R}^2 , indépendants et équidistribués, de loi commune P .

- θ désigne la rotation de centre 0 et d'angle θ .

- $Y_n(\theta)$ désigne le nombre d'éléments maximaux modulo θ de l'échantillon X_1, \dots, X_n .

- On pose :

$$\varphi_{\theta}^{P,k}(x_1, \dots, x_k) = P \left[\bigcup_{i=1}^k \leftarrow, x_i \mid \theta \right] \quad (5)$$

$\forall x_1, \dots, x_k \in \mathbb{R}^2 ; \forall k \in \mathbb{N}$,

et

$$A_k(\theta) = \{ (x_1, \dots, x_k) \in \mathbb{R}^{2k} \mid (i_1, \dots, i_k) \text{ permutation de } (1, \dots, k) \text{ telle que } x_{i_1} \leq \dots \leq x_{i_k} \left(\theta + \frac{\pi}{2} \right) \} \quad (6)$$

LEMME 27 :

On a les relations suivantes :

$$1^0) \varphi_{\theta}^{P,k}(x_1, \dots, x_k) = \varphi_0^{\theta P,k}(\theta x_1, \dots, \theta x_k) \quad (7)$$

$\forall x_1, \dots, x_k \in \mathbb{R}^2, \forall k \in \mathbb{N}$.

$$2^{\circ}) A_k(\theta) = \theta^{-1} [A_k(0)], \quad \forall k \in \mathbb{N}, \quad \forall \theta \in [0, 2\pi] \quad (8)$$

où θ_k désigne l'application $(x_1, \dots, x_k) \rightarrow (\theta x_1, \dots, \theta x_k)$

$$\forall x_1, \dots, x_k \in \mathbb{R}^2.$$

DÉMONSTRATION.

1^o) D'après (5), on a

$$\varphi_{\theta}^{P, k}(x_1, \dots, x_k) = P \left[\bigcup_{i=1}^k \theta^{-1} \right] \leftarrow, \quad \theta x_i []$$

$$= P \left[\theta^{-1} \bigcup_{i=1}^k \right] \leftarrow, \quad \theta x_i []$$

$$= (\theta P) \left[\bigcup_{i=1}^k \right] \leftarrow, \quad \theta x_i []$$

$$\forall x_1, \dots, x_k \in \mathbb{R}^2; \quad \forall k \in \mathbb{N}. \quad \text{C.Q.F.D.}$$

$$2^{\circ}) (x_1, \dots, x_k) \in \theta^{-1} [A_k(0)] \iff (\theta x_1, \dots, \theta x_k) \in A_k(0)$$

Cette dernière propriété équivaut à l'existence d'une permutation (i_1, \dots, i_k) de $(1, \dots, k)$ telle que

$$\theta x_{i_1} \leq \dots \leq \theta x_{i_k} \left(\frac{\pi}{2} \right) \quad \text{donc telle que} \quad x_{i_1} \leq \dots \leq x_{i_k} \left(\theta + \frac{\pi}{2} \right).$$

THÉORÈME 23 :

Avec les notations précédentes, on a :

$$\text{Prob} [Y_n(\theta) = k] = C_n^k \int_{\theta^{-1}[A_k(0)]} \varphi_{\theta}^{\theta P, k} [\theta x_1, \dots, \theta x_k]^{n-k} dP^{\otimes k}(x_1, \dots, x_k) \quad (9)$$

$$\forall k = 1, \dots, n.$$

DÉMONSTRATION.

On peut d'abord remarquer que si k points de \mathbf{R}^2 sont maximaux modulo θ ils forment une chaîne modulo $(\theta + \frac{\pi}{2})$, ce qui permet d'écrire:

$$\text{Prob} [\gamma_n(\theta) = k] = \text{Prob} [\exists i_1, \dots, i_k \mid X_{i_1}, \dots, X_{i_k} \text{ forment une chaîne modulo } \frac{\pi}{2} + \theta ;$$

$$\forall j \notin \{i_1, \dots, i_k\} , \exists i \in \{i_1, \dots, i_k\} \mid X_j < X_i]$$

On peut toujours supposer que X_1, \dots, X_k sont les k éléments maximaux et écrire :

$$\begin{aligned} \text{Prob} [\gamma_n(\theta) = k] &= C_n^k \text{ Prob} [X_1, \dots, X_k \text{ forment une chaîne modulo} \\ &\quad (\theta + \frac{\pi}{2}) ; \forall j > k, \exists i \leq k \mid X_j < X_i(\theta)] \\ &= C_n^k \int_{A_k(\theta)} \text{Prob} [\forall j > k, \exists i \leq k \mid X_j < x_i(\theta)] dP^{\otimes k}(x_1, \dots, x_k) \end{aligned}$$

d'où, en tenant compte du fait que les X_i sont indépendants :

$$\text{Prob} [\gamma_n(\theta) = k] = C_n^k \int_{A_k(\theta)} P[\bigcup_{i=1}^k \leftarrow, x_i]^{n-k} \cdot dP^{\otimes k}(x_1, \dots, x_k)$$

et le lemme 27 permet d'écrire le résultat sous la forme annoncée.

LEMME 28 :

1°) $\forall x_1, \dots, x_k \in \mathbf{R}^2$, on peut écrire :

$$\Phi_{\theta}^{P,k}(x_1, \dots, x_k) = \sum_{l=1}^n \sum_{i_1, \dots, i_l} (-1)^{l+1} F_{\theta}^P[\inf_{\theta}(x_{i_1}, \dots, x_{i_l}) + 0] \quad (10)$$

2°) Si $x_1 < \dots < x_k$ ($\theta + \frac{\pi}{2}$), on a :

$$\Phi_{\theta}^{P,k}(x_1, \dots, x_k) = \sum_{i=1}^k F_{\theta}^P(x_i) - \sum_{i=1}^{k-1} F_{\theta}^P[\inf_{\theta}(x_i, x_{i+1}) + \theta] \quad (11)$$

DÉMONSTRATION.

1°) La première partie du lemme se déduit immédiatement du théorème de Poincaré relatif à la probabilité d'une réunion.

2°) Faisons $\theta = 0$. Il faut montrer que

$$x_1 < \dots < x_k \left(\frac{\pi}{2}\right) \implies \Phi_0^{P,k}(x_1, \dots, x_k) = \sum_{i=1}^k F_0^P(x_i) - \sum_{i=1}^{k-1} F_0^P(z_i + 0)$$

avec $z_i = \inf(x_i, x_{i+1})$, $\forall i \in \{1, \dots, k-1\}$

La démonstration se fait immédiatement par récurrence en remarquant que, si l'on pose $x_i = (x_{i1}, x_{i2})$, $\forall i = 1, \dots, k$, on a :

$$x_1 < \dots < x_k \left(\frac{\pi}{2}\right) \implies \begin{cases} x_{11} \leq x_{21} \leq \dots \leq x_{k1} \\ x_{12} \geq x_{22} \geq \dots \geq x_{k2} \end{cases}$$

Pour θ quelconque, on écrit :

$$x_1 < \dots < x_k \left(\theta + \frac{\pi}{2}\right) \iff \theta x_1 < \dots < \theta x_k \left(\frac{\pi}{2}\right)$$

et, compte tenu de (7), on est ramené au cas précédent en remplaçant $\theta x_1, \dots, \theta x_k$ par x_1, \dots, x_k , donc :

$$\Phi_{\theta}^{P,k}(x_1, \dots, x_k) = \sum_{i=1}^k F_{\theta}^{\theta P}(x_i) - \sum_{i=1}^{k-1} F_{\theta}^{\theta P}(\inf(\theta x_i, \theta x_{i+1}) + \theta)$$

et les relations (2) et (4) permettent de conclure.

COROLLAIRE 21 :

Si F_{θ}^P est continue, on a, $\forall k \in \{1, \dots, n\}$:

$$\text{Prob}[Y_n(\theta) = k] = A_n^k \int_{\theta_k^{-1}(B_k)} \left[\sum_{i=1}^k F_\theta^P(x_i) - \sum_{i=1}^{k-1} F_\theta^P[z_i(\theta)] \right]^{n-k} \cdot dP^{\otimes k} (x_1, \dots, x_k) \quad (12)$$

avec $A_n^k =$ nombre d'arrangements de n objets k à k .

$$z_i(\theta) = \inf_{\theta} \{x_i, x_{i+1}\}, \quad i = 1, \dots, k-1.$$

$$B_k = \{(x_1, \dots, x_k) \in \mathbb{R}^{2k} \mid x_1 < \dots < x_k \quad \left(\frac{\pi}{2}\right)\}$$

DÉMONSTRATION.

On peut écrire

$$\theta_k^{-1}(A_k(0)) = \sum_{i_1, \dots, i_k} \theta^{-1} B_{i_1 \dots i_k} + \theta^{-1} A_0$$

où le signe \sum porte sur toutes les permutations de $\{1, \dots, k\}$ et où l'on a posé :

$$B_{i_1, \dots, i_k} = \{(x_1, \dots, x_k) \in A_k(0) \mid x_{i_1} < \dots < x_{i_k} \quad \left(\frac{\pi}{2}\right)\}$$

et $A_0 = \{(x_1, \dots, x_k) \in A_k(0) \mid i \neq j : x_i = x_j\}$

Or, si F_θ^P est continue, F_0^P également donc $\int_{A_0} dF_0^P = 0$.

D'autre part, les événements B_{i_1, \dots, i_k} sont équiprobables, d'où :

$$\int_{\theta_k^{-1}[A_k(0)]} [\varphi_\theta^{P,k}]^{n-k} dP^{\otimes k} = k! \int_{\theta_k^{-1}(B_k)} [\varphi_\theta^{P,k}]^{n-k} \cdot dP^{\otimes k}$$

Le lemme 28 et le théorème 23 permettent alors de conclure.

REMARQUE :

Si on fait $\theta = 0$ (resp. $\theta = \pi$) dans (12), on obtient la loi du nombre d'éléments maximaux (resp. minimaux) pour la relation d'ordre usuelle de \mathbf{R}^2 .

III - COMPORTEMENT ASYMPTOTIQUE DU NOMBRE D'ÉLÉMENTS MAXIMAUX (RESP. MINIMAUX).

Dans ce paragraphe, nous étudierons le comportement asymptotique du nombre d'éléments extrémaux d'un échantillon de variables aléatoires à valeurs dans \mathbf{R}^2 muni de sa relation d'ordre usuelle.

NOTATIONS :

On désignera par Y_n le nombre d'éléments maximaux, par Z_n le nombre d'éléments minimaux et par W_n le nombre d'éléments extrémaux.

1. Espérance Mathématique de Y_n , Z_n , W_n .

THÉORÈME 24 :

Si la loi des variables de l'échantillon est une loi uniforme sur le $[0,1]^2$, on a, quand $n \rightarrow \infty$:

$$1^\circ) \quad E(Y_n) = E(Z_n) \sim \text{Log } n \quad (13)$$

$$2^\circ) \quad E(W_n) \sim 2 \text{ Log } n. \quad (14)$$

DÉMONSTRATION.

1°) On peut écrire :

$$\begin{aligned} P_n &= \text{Prob} [X_i \text{ est maximal}] = \text{Prob} [X_j < X_i \text{ ou } X_j \wedge X_i, \forall j \neq i] \\ &= \iint_{[0,1]^2} \text{Prob} [X_j < (x,y) \text{ ou } X_j \wedge (x,y), \forall j \neq i] \, dx \, dy \end{aligned}$$

$$= \iint_{[0,1]^2} (\text{Prob} [X_j < (x,y) \text{ ou } X_j \wedge (x,y)])^{n-1} dx dy$$

$$= \iint_{[0,1]^2} [1 - (1-x)(1-y)]^{n-1} dx dy$$

en intégrant, on trouve :

$$P_n = \frac{1}{n} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} \right) \quad (15)$$

Soit alors T_i la variable indicatrice de l'évènement " X_i est maximal" ($i = 1, \dots, n$). Il est clair que

$$Y_n = T_1 + \dots + T_n$$

d'où, d'après (15)

$$E(Y_n) = 1 + \frac{1}{2} + \dots + \frac{1}{n} \quad (16)$$

Par raison de symétrie, on a aussi

$$E(Z_n) = 1 + \frac{1}{2} + \dots + \frac{1}{n} \quad (16 \text{ bis})$$

Donc, quand $n \rightarrow \infty$, on a bien

$$E(Y_n) = E(Z_n) \sim \text{Log } n.$$

2°) Commençons par calculer

$$Q_n = \text{Prob} [X_i \text{ est maximal et minimal}]$$

On a :

$$Q_n = \text{Prob} [X_j \wedge X_i, \forall j \neq i]$$

$$= \iint_{[0,1]^2} [x(1-y) + y(1-x)]^{n-1} dx dy$$

en développant suivant la formule du binôme et en utilisant la fonction B, on trouve :

$$Q_n = \frac{1}{n \cdot n!} \sum_{k=0}^{n-1} k! (n-1-k)! \quad (17)$$

Soit alors V_n le nombre d'éléments à la fois maximaux et minimaux, on aura :

$$E(V_n) = \frac{1}{n!} \sum_{k=0}^{n-1} k! (n-1-k)!$$

ou encore

$$E(V_n) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{C_{n-1}^k} \quad (18)$$

En tenant compte du fait que, $\forall k \in \{1, \dots, n-2\}$, on a $C_{n-1}^k \geq n-1$, il vient :

$$E(V_n) \leq \frac{1}{n} \left[2 + \frac{n-2}{n-1} \right] < \frac{3}{n} \quad (19)$$

Or, $W_n = Y_n + Z_n - V_n$, d'où d'après (16), (16 bis) et (18) :

$$E(W_n) = 2 \left(1 + \dots + \frac{1}{n} \right) - \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{C_{n-1}^k} \quad (20)$$

et, d'après (13) et (19), on a bien, quand $n \rightarrow \infty$:

$$E(W_n) \sim 2 \text{Log } n$$

C.Q.F.D.

2. Comportement asymptotique des lois de Y_n et de Z_n .

Posons :

$$P_{n,k} = \text{Prob} [Y_n = k], \quad \forall n \in \mathbb{N}, \quad \forall k \in \{1, \dots, n\} \quad (21)$$

On a le :

LEMME 29 :

Si la loi des variables de l'échantillon est une loi uniforme sur le carré $[0,1]^2$, on a la relation de récurrence :

$$P_{n,k} = \frac{1}{n} \sum_{h=k-1}^{n-1} P_{h,k-1}, \quad k > 1 \quad (22)$$

DÉMONSTRATION.

Nous allons utiliser la *formule de Crofton* :

Soit D un domaine \mathbf{R}^p ($p \in \mathbf{N}$) et soit D' un domaine contenant D , la formule de Crofton s'écrit :

$$V \frac{dP}{dV} = n(P_1 - P) \quad (23)$$

où V est le volume de D , P la probabilité d'un certain évènement associé à la figure formée par n points tirés indépendamment et uniformément, dans D et P_1 la probabilité du même évènement quand on tire indépendamment et uniformément $(n-1)$ points dans D et 1 point dans $D'-D$ suivant une loi uniforme dans D' ⁽¹⁾.

Prenons $D = [0, 1]^2$ et $D' = [0, 1 + \varepsilon] \times [0, 1]$, $\varepsilon > 0$. Il est clair que la probabilité de tirer k maximaux en n expériences suivant une loi uniforme est la même dans D et dans D' car l'évènement "il y a k éléments maximaux" ne fait intervenir que l'"ordre" ⁽²⁾.

On a donc $\frac{dP}{dV} = 0$ et (23) s'écrit $P = P_1$. Dans ces conditions, on a successivement :

$$\begin{aligned} P_{n,k} &= \text{Prob} [k \text{ maximaux} \mid (n-1) \text{ points} \in D, 1 \text{ point} \in D'-D] \\ &= \frac{(1+\varepsilon)^n}{n\varepsilon} \text{Prob} [k \text{ maximaux}; (n-1) \text{ points} D, 1 \text{ point} \in D'-D] \end{aligned}$$

(1) Pour la démonstration de (23), voir R. DELTHEIL [43] p. 35-37.

(2) Nous verrons plus loin une justification plus précise.

$$= \frac{(1+\varepsilon)^n}{n\varepsilon} \cdot n \iint_{\substack{1 \leq x \leq 1+\varepsilon \\ 0 \leq y \leq 1}} \text{Prob} [k \text{ maximaux; } (n-1) \text{ points } \in D|x,y] \frac{dx}{1+\varepsilon} \cdot dy$$

Posons $\Delta y = [0,1] \times [y,1]$ et $\Delta'y = [0,1] \times [0,y]$ et désignons par $E_{h,k-1}$ l'évènement "k-1 points sont maximaux parmi les h qui sont tombés sur Δy ".

La loi induite sur Δy par la loi uniforme sur D' étant encore uniforme, on peut écrire :

$$\begin{aligned} \text{Prob}_{\Delta y} [E_{h,k-1}] &= P_{h,k-1} \quad \text{pour } h \geq k-1 \\ &= 0 \quad \text{pour } h < k-1 \end{aligned}$$

On en tire alors :

$$\begin{aligned} P_{n,k} &= \frac{(1+\varepsilon)^{n-1}}{\varepsilon} \iint_{\substack{h=k-1 \\ 1 \leq x \leq 1+\varepsilon \\ 0 \leq y \leq 1}} \sum_{h=k-1}^{n-1} \text{Prob} [h \text{ points } \in \Delta y; (n-1-h) \text{ points } \in \Delta'y; \\ &\quad E_{h,k-1}] dx dy \\ &= \frac{(1+\varepsilon)^{n-1}}{\varepsilon} \int_1^{1+\varepsilon} dx \sum_{h=k-1}^{n-1} \int_0^1 \frac{C_{n-1}^h (1-y)^h y^{n-1-h}}{(1+\varepsilon)^{n-1}} \cdot P_{h,k-1} dy \end{aligned}$$

d'où

$$P_{n,k} = \sum_{h=k-1}^{n-1} C_{n-1}^h P_{h,k-1} B(h+1, n-h)$$

ce qui entraîne

$$P_{n,k} = \frac{1}{n} \sum_{h=k-1}^{n-1} P_{h,k-1}, \quad k > 1$$

C.Q.F.D.

APPLICATIONS :

Si l'on connaît $P_{n,1}$, la relation (22) permet de calculer $P_{n,k}$ pour $k > 1$. Or on a :

$$P_{n,1} = n \text{Prob} [X_j < X_1, \forall j > 1] = n \iint_{[0,1]^2} x^{n-1} y^{n-1} dx dy$$

d'où $P_{n,1} = \frac{1}{n}$ (24)

Des calculs faciles montrent alors que :

$$P_{n,n} = \frac{1}{n!}, \quad \forall n \in \mathbb{N} \quad (25)$$

$$P_{n,n-1} = \frac{1}{2(n-2)!}, \quad \forall n \geq 2 \quad (26)$$

$$P_{n,2} = \frac{1}{n} \left(1 + \frac{1}{2} + \dots + \frac{1}{n-1} \right), \quad \forall n \geq 2 \quad (27)$$

Sous les mêmes hypothèses que celles du lemme 29, on a le

THÉORÈME 25 :

Quand $n \rightarrow \infty$, $P_{n,k}$ est équivalent à

$$\frac{(\text{Log } n)^{k-1}}{n \cdot (k-1)!}, \quad \forall k \geq 1 \text{ fixe.}$$

DÉMONSTRATION.

Le résultat est vrai pour $k = 1$ car $P_{n,1} = \frac{1}{n} = \frac{(\text{Log } n)^0}{n \cdot 0!}$; supposons-le vrai pour k et montrons qu'il l'est pour $k+1$.

La relation (22) s'écrira :

$$P_{n,k+1} = \frac{1}{n} \sum_{i=k}^{n-1} P_{i,k}$$

Or, $P_{i,k} = \frac{(\text{Log } i)^{k-1}}{i \cdot (k-1)!} \varepsilon(i,k)$ avec $\lim_{i \rightarrow \infty} \varepsilon(i,k) = 1$.

Donc, $\forall \varepsilon > 0$, il existe $N = N(\varepsilon)$ tel que $i \geq N$ entraîne

$$|1 - \varepsilon(i,k)| < \frac{\varepsilon}{2}$$

On peut toujours choisir $N > e^{k-1}$ et assez grand pour que $N - \frac{(\text{Log } N)^k}{k!}$ soit positif. Ecrivons alors, pour $n > N+2$:

$$P_{n,k+1} = \frac{1}{n} \left[\sum_{i=k}^N P_{i,k} + \sum_{i=N+1}^{n-1} P_{i,k} \right]$$

On en tire

$$P_{n,k+1} < \frac{N}{n} + (1 + \frac{\varepsilon}{2}) \frac{1}{n} \sum_{i=N+1}^{n-1} \frac{(\text{Log } i)^{k-1}}{i \cdot (k-1)!}$$

la fonction $\frac{(\text{Log } x)^{k-1}}{x}$ étant décroissante pour $x > e^{k-1}$, il vient :

$$P_{n,k+1} < \frac{N}{n} + \frac{1 + \frac{\varepsilon}{2}}{n \cdot (k-1)!} \left[\frac{(\text{Log } (n-1))^k}{k} - \frac{(\text{Log } N)^k}{k} \right]$$

d'où

$$P_{n,k+1} < \frac{(1 + \frac{\varepsilon}{2}) (\text{Log } (n-1))^k}{n \cdot k!} + \frac{1}{n} \left[N - \frac{(\text{Log } N)^k}{k!} \right]$$

soit

$$P_{n,k+1} < \frac{(1 + \frac{\varepsilon}{2}) (\text{Log } (n-1))^k}{n \cdot k!} + \frac{\beta_1}{n} \tag{28}$$

où β_1 est, d'après l'hypothèse faite sur N , une quantité positive.

De manière analogue, on peut minorer $P_{n,k+1}$ en écrivant successivement :

$$P_{n,k+1} = \frac{1}{n} \left[\sum_{i=k}^N P_{i,k} + \sum_{i=N+1}^n P_{i,k} \right]$$

$$> \frac{(1 - \frac{\epsilon}{2})}{n} \sum_{i=k}^n P_{i,k} + \frac{1}{n} (1 - \frac{\epsilon}{2}) \frac{1}{k!} [(\text{Log}(n-1))^k - (\text{Log}(N+1))^k]$$

d'où,

$$P_{n,k+1} > \frac{(\text{Log}(n-1))^k}{n \cdot k!} \left(1 - \frac{\epsilon}{2} \right) + \frac{\alpha}{n} \tag{29}$$

avec $\alpha = - (1 - \frac{\epsilon}{2}) \frac{1}{k!} (\text{Log}(N+1))^k = -\beta_2$

Finalement, (28) et (29) entraînent, $\forall n \geq N$:

$$1 - \frac{\epsilon}{2} - \frac{\beta \cdot k!}{(\text{Log}(n-1))^k} < P_{n,k+1} \cdot \frac{n \cdot k!}{(\text{Log}(n-1))^k} < 1 + \frac{\epsilon}{2} + \frac{\beta \cdot k!}{(\text{Log}(n-1))^k}$$

où l'on a posé $\beta = \sup(\beta_1, \beta_2)$.

On en tire, pour n assez grand,

$$1 - \epsilon < P_{n,k+1} \cdot \frac{n \cdot k!}{(\text{Log}(n-1))^k} < 1 + \epsilon$$

ce qui permet de conclure facilement.

COROLLAIRE 22 :

$$\lim_{n \rightarrow \infty} p \cdot Y_n = + \infty$$

DEMONSTRATION.

DÉMONSTRATION.

Il suffit de montrer que, $\forall k \in \mathbf{N}$,

$$\text{Prob} [Y_n \leq k] \rightarrow 0$$

$$\text{Or } \text{Prob} [Y_n \leq k] = P_{n,1} + \dots + P_{n,k}$$

donc, d'après le théorème 25 :

$$\text{Prob} [Y_n \leq k] \sim \frac{(\text{Log } n)^{k-1}}{n^{(k-1)!}} \quad (30)$$

quand $n \rightarrow \infty$. C.Q.F.D.

Remarquons que (30) montre que la convergence de Y_n vers l'infini n'est pas presque complètement sûre. On peut se demander si cette convergence est presque sûre.

M. Geffroy a montré qu'il en est bien ainsi :

THÉORÈME 26 : (M. GEFFROY) :

$$\lim_{n \rightarrow \infty} \text{p.s. } Y_n = +\infty .$$

DÉMONSTRATION.

Posons $E_j = \{ Y_j \geq k \}$, $\forall k \in \mathbf{N}$, $\forall j \in \mathbf{N}$.

Il s'agit de montrer que

$$\lim_{n \rightarrow \infty} \text{Prob} [\bigcup_{j=n}^{\infty} \bar{E}_j] = 0, \quad \forall k \in \mathbf{N}.$$

$$\text{Or } \bigcup_{j=n}^{\infty} \bar{E}_j = \bar{E}_n + \overline{E_{n+1} \cdot \bar{E}_n} + \overline{E_{n+2} \cdot \bar{E}_n \cdot \bar{E}_{n+1}} + \dots$$

$$\text{Donc } \bigcup_{j=n}^{\infty} \bar{E}_j \Rightarrow \bar{E}_n + \overline{E_{n+1} \cdot \bar{E}_n} + \overline{E_{n+2} \cdot \bar{E}_n \cdot \bar{E}_{n+1}} + \dots$$

Or, d'après le corollaire 22, on sait que $\lim_{n \rightarrow \infty} \text{Prob}(\overline{E}_n) = 0$. Il suffit donc de montrer que la série de terme général $\text{Prob}[E_n \cdot \overline{E}_{n+1}]$ est convergente.

Soit alors F_{n+1} l'évènement " X_{n+1} est maximal", on a l'application : $E_n \cdot \overline{E}_{n+1} \implies \overline{E}_{n+1} \cdot F_{n+1}$. D'où :

$$\begin{aligned} \text{Prob}[E_n \cdot \overline{E}_{n+1}] &\leq \text{Prob}[Y_{n+1} < k, X_{n+1} \text{ maximal}] \\ &\leq \sum_{j=1}^{k-1} \text{Prob}[Y_{n+1} = j] \text{Prob}[X_{n+1} \text{ maximal} | Y_{n+1} = j] \\ &\leq \sum_{j=1}^{k-1} \text{Prob}[Y_{n+1} = j] \cdot \frac{j}{n+1} \\ &< \frac{k}{n} \text{Prob}[Y_{n+1} \leq k] \end{aligned}$$

Or (30) implique que la série de terme général $\frac{\text{Prob}(Y_{n+1} \leq k)}{n}$ est convergente, ce qui achève la démonstration.

REMARQUE :

Tous les résultats précédents restent valables quand on remplace Y_n par Z_n .

En ce qui concerne W_n , on a le

COROLLAIRE 23 :

$$\lim_{n \rightarrow \infty} \text{p.s. } W_n = +\infty$$

$$Y_n = Z_n - V_n$$

DÉMONSTRATION.

$W_n = Y_n + Z_n - V_n = Y_n + (Z_n - V_n) \geq Y_n$ et l'on conclut à l'aide du théorème 26.

3. Éléments maximaux et éléments saillants,

Résultats complémentaires.

Les résultats qui précèdent montrent qu'il y a une certaine analogie entre les éléments maximaux d'un échantillon à deux dimensions et les éléments saillants d'une suite d'observations ⁽¹⁾. Cette analogie est mise en évidence par le

LEMME 30 :

Dans les conditions du théorème 24, la loi de Y_n est celle du nombre d'éléments saillants d'une suite d'observations.

DÉMONSTRATION.

Soit $f_n(t) = \sum_{k=1}^n e^{itk} P_{n,k}$ la fonction caractéristique de Y_n .

D'après (22), il vient :

$$f_n(t) = \frac{e^{it}}{n} \sum_{k=0}^{n-1} f_k(t)$$

On en tire

$$n f_n(t) - (n-1) f_{n-1}(t) = e^{it} \left[\sum_{k=0}^{n-1} f_k(t) - \sum_{k=0}^{n-2} f_k(t) \right]$$

$$\text{soit } f_n(t) = f_{n-1}(t) \cdot \left[\frac{1}{n} e^{it} + \left(1 - \frac{1}{n}\right) \right]$$

$$\text{d'où } f_n(t) = \prod_{k=1}^n \left[\frac{1}{k} e^{it} + \left(1 - \frac{1}{k}\right) \right]$$

Ce qui signifie que l'on peut écrire

$$Y_n = B_1 + \dots + B_n \tag{31}$$

(1) Pour la définition et les propriétés des éléments saillants d'une suite d'observations, cf RENEYI [44].

où B_k ($1 \leq k \leq n$) désigne une variable de Bernoulli d'espérance $\frac{1}{k}$ et où l'on suppose les variables B_1, \dots, B_n indépendantes.

On voit donc que (31) est analogue à la relation (9) de [44] (p. 107), ce qui achève la démonstration.

Interprétation (1) :

Si l'on numérote les points de l'échantillon par abscisses décroissantes, on constate qu'un point est maximal si et seulement si son ordonnée est un élément saillant de la suite des ordonnées: il y a donc le même nombre d'éléments maximaux dans l'échantillon que d'éléments saillants dans la suite des ordonnées des points de l'échantillon.

Remarquons cependant que le processus est différent: le nombre d'éléments saillants d'une suite d'observations est presque sûrement croissant par rapport à la taille de l'échantillon alors qu'il n'en est pas de même pour le nombre d'éléments maximaux.

COROLLAIRE 24 :

$(Y_n - \text{Log } n) / \sqrt{\text{Log } n}$ tend, en loi, vers une variable aléatoire qui suit une loi normale de moyenne nulle et d'écart type 1.

DÉMONSTRATION.

Analogue à celle du théorème 2' de [44].

4. Extensions.

a) **Cas où les variables de l'échantillon sont à marges indépendantes.**

On a alors le

COROLLAIRE 25 :

Si les variables de l'échantillon sont à marges indépendantes et si les lois marginales sont continues, les résultats précédents restent valables.

[1] Cette interprétation nous a été signalée par M. Geffroy.

DÉMONSTRATION.

Posons $X_i = (X_{1j}^1; X_1^2)$; $i = 1, \dots, n$; et désignons par F_j ($j = 1, 2$) la loi de X_i .

On sait que $T_i = [F_1(X_1^1), F_2(X_1^2)]$; $i = 1, \dots, n$; suit une loi uniforme sur $[0,1]^2$: la transformation $(X_1, \dots, X_n) \rightarrow (T_1, \dots, T_n)$ conservant l'ordre, conserve aussi les éléments maximaux, ce qui achève la démonstration.

b) Cas où les marges ne sont pas nécessairement indépendantes.

Si l'on suppose les lois marginales continues on voit, en considérant la transformation définie ci-dessus, que l'on peut toujours se ramener au cas où les variables de l'échantillon sont à valeurs dans $[0,1]^2$ et sont à lois marginales uniformes sur $[0,1]$.

D'après un résultat de M. FRECHET [46], on sait alors que la loi H des variables de l'échantillon est telle que :

$$H^0(x,y) \leq H(x,y) \leq H^1(x,y) \quad \forall x,y \in [0,1] \quad (31)$$

avec $H^0(x,y) = \text{Max}[x+y-1, 0] \quad \forall x,y \in [0,1]$

et $H^1(x,y) = \text{Min}[x,y] \quad \forall x,y \in [0,1]$

Pour $H = H^0$, on a $Y_n = n$ p.s., pour $H = H^1$, on a $Y_n = 1$ p.s.

Lorsque H vérifie (31) et qu'elle est absolument continue et à densité continue et positive, on pourrait penser que la loi de Y_n est la même que dans le cas uniforme puisque le nombre d'éléments maximaux ne dépend que de "l'ordre". Le contre-exemple suivant montre qu'il n'en est rien :

Considérons la famille de lois définie par GUMBEL [45] :

$$H_\alpha(x,y) = xy [1 + \alpha(1-x)(1-y)] , \quad \forall x,y \in [0,1] \quad \alpha \in]-1, +1[.$$

Supposons que les variables de l'échantillon ont pour loi H_α et calculons $P_{2,1}(\alpha) = \text{Prob}[Y_2 = 1]$. On aura :

$$P_{2,1}(\alpha) = 2 \iint_{[0,1]^2} xy [1 + \alpha(1-x)(1-y)] [1 + \alpha(1-2x)(1-2y)] dx dy$$

d'où l'on tire $P_{2,1}(\alpha) = \frac{1}{2} - \frac{\alpha}{6} (1 + \frac{\alpha}{3}) \neq P_{2,1} = \frac{1}{2}$ pour $\alpha \neq 0$.

Dans ces conditions, il semble que l'on ne puisse donner que des résultats partiels; on aura, par exemple le :

COROLLAIRE 26 :

Si la loi des variables de l'échantillon s'écrit

$$H(x,y) = xy + \varphi(x,y) ; \forall x,y \in [0,1] \text{ et si, de plus}$$

1°) $\varphi(x,y) \leq 0 ; \forall x,y \in [0,1]$

2°) φ est deux fois dérivable et il existe $A \in]0,1[$ tel que

$$\left| \frac{\partial^2 \varphi}{\partial x \partial y} (x,y) \right| \leq A ; \quad \forall x,y \in]0,1[$$

alors $E(Y_n)$ est un infiniment grand d'ordre supérieur ou égal à $\text{Log } n$.

DÉMONSTRATION.

On peut écrire $\forall i \in \{1, \dots, n\}$:

$$\text{Prob } [X_1 \text{ maximal}] = P_n(H) = \iint_{[0,1]^2} [1 - (1-x)(1-y) - \varphi(x,y)]^{n-1} \cdot [1 + \frac{\partial^2 \varphi}{\partial x \partial y} (x,y)] dx dy$$

d'où $P_n(H) \geq (1-A) P_n$, ce qui permet de conclure en utilisant la relation $E(Y_n) = n P_n(H)$ et le théorème 24.

EXEMPLE :

Les lois H_α vérifient les hypothèses du corollaire 26, $\forall \alpha \in [0,1[$.

Enfin, on peut énoncer le

COROLLAIRE 27 :

$$\text{Si } \lim_{n \rightarrow \infty} P_{n,k} = 0 \quad \text{et} \quad \sum_{n=1}^{\infty} \frac{P_{n,k}}{n} < +\infty,$$

$$\forall k \in \mathbb{N} \quad \text{alors} \quad \lim_{n \rightarrow \infty} \text{p.s. } Y_n = +\infty.$$

DÉMONSTRATION.

Identique à celle du théorème 26.

c) Cas d'un échantillon de variables aléatoires à valeurs dans \mathbb{R}^m ($m > 2$).

THÉORÈME 27 :

Si la loi des variables de l'échantillon est la loi uniforme sur $[0,1]^m$, on peut écrire

$$E(Y_n) = E(Z_n) \sim \frac{(\text{Log } n)^{m-1}}{(m-1)!} \quad (32)$$

DÉMONSTRATION.

Posons $P_n(m) = \text{Prob}[X_1 \text{ maximal}] = \text{Prob}[X_1 \text{ minimal}]$

$\forall i \in \{1, \dots, n\}$. On aura

$$\begin{aligned} P_n(m) &= \int_{[0,1]^m} (1 - x_1 \dots x_m)^{n-1} dx_1 \dots dx_m \\ &= \frac{1}{n} \int_{[0,1]^{m-1}} \left[(1 - x_1 \dots x_m)^n \Big|_{x_m=0}^{x_m=1} \cdot \frac{1}{x_1 \dots x_{m-1}} \right] dx_1 \dots dx_{m-1} \\ &= \frac{1}{n} \int_{[0,1]^{m-1}} \frac{1^n - (1 - x_1 \dots x_{m-1})^n}{x_1 \dots x_{m-1}} dx_1 \dots dx_{m-1} \end{aligned}$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \int_{[0,1]^{m-1}} (1-x_1 \dots x_{m-1})^k dx_1 \dots dx_{m-1}$$

d'où

$$P_n(m) = \frac{P_1(m-1) + \dots + P_n(m-1)}{n} \tag{33}$$

Or on sait que $P_n(1) = \frac{1}{n}$; on en tire donc

$$P_n(m) \sim \frac{(\text{Log } n)^{m-1}}{n(m-1)!}, \quad \forall m \geq 1 \tag{34}$$

La démonstration s'effectuant de manière tout à fait analogue à celle du théorème 25.

(34) entraîne immédiatement (32), C.Q.F.D.

COROLLAIRE 28 :

Sous les hypothèses du théorème 27, on a

$$\lim_{n \rightarrow \infty} \text{p.s. } Y_n = +\infty.$$

DÉMONSTRATION.

On pose $X_i = (X_i^1, \dots, X_i^m)$, $\forall i \in \{1, \dots, n\}$. Soit alors Y'_n le nombre de points de l'échantillon tels que (X_i^1, X_i^2) soit maximal pour l'échantillon $(X_1^1, X_1^2) \dots (X_n^1, X_n^2)$. On a évidemment $Y'_n < Y_n$, $\forall n \in \mathbb{N}$, or, d'après le théorème 26, $\lim_{n \rightarrow \infty} \text{p.d. } Y'_n = +\infty$.

On a donc aussi $\lim_{n \rightarrow \infty} \text{p.s. } Y_n = +\infty$, C.Q.F.D.

REMARQUE :

Le corollaire 27 reste valable dans \mathbb{R}^n et même dans un espace ordonné quelconque puisque sa démonstration ne nécessite aucune hypothèse

sur la dimension de l'espace dans lequel les variables de l'échantillon prennent leurs valeurs.

IV - THÉORÈME DE M. GEFFROY - APPLICATION A L'ESTIMATION DE LA COURBE MAXIMALE.

1. Théorème de M. Geffroy.

Outre le théorème 26, M. Geffroy a obtenu deux Résultats importants sur le comportement asymptotique du nombre d'éléments maximaux.

Nous donnons seulement les énoncés de ces Résultats; les démonstrations paraîtront prochainement dans une note aux C.R. de M. Geffroy.

Hypothèses et Notations.

On suppose que les variables de l'échantillon sont à valeurs dans $[0,1]^2$ et que leur loi commune est continue et à marges uniformes. On désigne par S le support de la loi des variables de l'échantillon, par S_m l'ensemble des éléments maximaux de S et par M_n l'ensemble des éléments maximaux d'un échantillon de taille n .

THÉORÈME 28 (M. GEFFROY):

$\forall s \in S_m$, $\forall V$ voisinage de s , on a

$$\lim_{n \rightarrow \infty} \text{Prob} \left[\bigcap_{r=n}^{\infty} \{M_r \cap V\} \neq \emptyset \right] = 1.$$

COROLLAIRE 29 (M. GEFFROY):

Si S_m est infini, on a :

$$\lim_{n \rightarrow \infty} \text{p.s. } Y_n = +\infty.$$

DÉFINITION 14 :

Si $s \in S$, on dit que s est *non dominant* si

1°) $s \in S_m$.

2°)
$$\lim_{\substack{h \rightarrow 0 \\ k \rightarrow 0}} \frac{\text{Prob}([s_1-h, s_1] \times [s_2-k, s_2])}{\text{Prob}([s_1-h, s_1] \times [0, s_2-k]) + \text{Prob}([0, s_1-h] \times [s_2-k, s_2])} = 0$$

où l'on a posé $s = (s_1, s_2)$.

THÉORÈME 29 (M. GEFFROY) :

Si S_m est fini, si $s \in S_m$, si V désigne un voisinage de s tel que $V \cap S_m = \{s\}$ et si $Y_n(V)$ désigne le nombre d'éléments maximaux de l'échantillon appartenant à V , alors $\lim_{n \rightarrow \infty} p \cdot Y_n(V) = +\infty$, si et seulement si s est non dominant.

2. Estimation de la courbe maximale.

Nous allons voir comment on peut construire un estimateur convergent de l'ensemble des éléments maximaux de S en utilisant le théorème 28.

Les hypothèses et les notations sont les mêmes que dans le paragraphe précédent. Nous supposons, en outre, que S_m est une courbe continue d'équation $y = g_1(x)$ où g_1 est définie par un intervalle de la forme $[a, 1]$, $0 \leq a \leq 1$, telle que $g_1(a) = 1$.

Posons

$$\begin{cases} g(x) = 1 & \text{pour } x \in [0, a] \\ g(x) = g_1(x) & \text{pour } x \in [a, 1] \end{cases}$$

et soit C_m le graphe de g : nous dirons que C_m est la *courbe maximale* de S (1).

(1) Remarquons que g est la fonction frontière de S .

LEMME 31 :

La courbe maximale de S est au-dessus de la droite $y + x - 1 = 0$ et la fonction g est strictement décroissante sur $[a, 1]$.

DÉMONSTRATION.

Soit $F(x, y)$ la f.d.r. de la loi des variables de l'échantillon. On a $(x, y) \in S_m \Rightarrow 1 - x - y + F(x, y) = 0$ d'où $x + y - 1 \geq 0$. D'autre part $(x, y) \in S_m$ et $(x', y') \in S_m$ avec $x' > x$ entraîne bien $y' < y$ puisque (x, y) et (x', y') sont non comparables.

Dans la suite, pour tout $\varepsilon > 0$ et tout entier p , nous poserons:

$$V_\varepsilon = \{ (x, y) \mid x \in [0, 1], \quad g(x) - \varepsilon < y < g(x) + \varepsilon \}$$

et

$$A_{\varepsilon, p, j} = \{ (x, y) \in V_\varepsilon \mid a + j \cdot \frac{1-a}{p} \leq x < a + (j+1) \cdot \frac{1-a}{p} \}$$

$$j = 0, 1, \dots, p-1;$$

LEMME 32 :

$$\forall \varepsilon > 0, \quad \forall \varepsilon' \in]0, \varepsilon[, \quad p_0 \in \mathbb{N} \quad \text{tel que} \quad \forall p \geq p_0 :$$

1°) $\forall M = (x, y), \quad \forall M' = (x', y') \in A_{\varepsilon', p, j}, \quad \forall j \in \{0, \dots, p-1\}$
 le segment MM' est inclus dans V_ε .

$$2°) \forall M = (x, y) \in A_{\varepsilon', p, j}, \quad \forall M' = (x', y') \in A_{\varepsilon', p, j+1}.$$

$\forall j \in \{0, \dots, p-2\}$ le segment MM' est inclus dans V_ε .

DÉMONSTRATION.

Pour p assez grand, on aura, d'après la continuité uniforme de g sur $[a, 1]$:

$$\left| g\left(x + 2 \cdot \frac{1-a}{p}\right) - g(x) \right| < \varepsilon - \varepsilon', \quad \forall x \in \left[a, 1 - 2 \cdot \frac{1-a}{p} \right] \quad (35)$$

Soient alors $M_1 = (x_1, y_1)$ et $M_2 = (x_2, y_2) \in A_{\varepsilon', p, j} \cup A_{\varepsilon', p, j+1}$ et tels que $y_2 < y_1$. $\forall M = (x, y)$ et appartenant au segment $M_1 M_2$, on aura $y_2 < y < y_1$, or :

$$y_2 > \inf_{x \in A_{\varepsilon', p, j} \cup A_{\varepsilon', p, j+1}} g(x) - \varepsilon' = g\left(a + (j+1) \cdot \frac{1-a}{p}\right) - \varepsilon'$$

mais $g\left(a + (j-1) \cdot \frac{1-a}{p}\right) - g\left(a + (j+1) \cdot \frac{1-a}{p}\right) < \varepsilon - \varepsilon'$

d'où $y_2 > g\left(a + (j-1) \cdot \frac{1-a}{p}\right) - \varepsilon + \varepsilon' - \varepsilon'$

et $y_2 > \sup_{x \in A_{\varepsilon', p, h} \cup A_{\varepsilon', p, j+1}} g(x) - \varepsilon$

On montrerait de même que $y_1 < \inf_{x \in A_{\varepsilon', p, j} \cup A_{\varepsilon', p, j+1}} g(x) + \varepsilon$

on a donc bien $M_1 M_2 \subset V_\varepsilon$, C.Q.F.D.

LEMME 33 :

$\forall \varepsilon'' \in]0, \varepsilon'[$ on peut choisir p assez grand pour que $\forall P \in \mathring{S} - V_{\varepsilon'}$, $[P, \rightarrow [$ contienne au moins un ensemble $A_{\varepsilon'', p, j}$.

DÉMONSTRATION.

Pour p assez grand, on a $\left|g\left(x + \frac{1-a}{p}\right) - g(x)\right| < \frac{\varepsilon' - \varepsilon''}{2}$, $\forall x \in [a, 1 - \frac{1-a}{p}]$. Or, si $P = (x_0, y_0) \in \mathring{S} - V_{\varepsilon'}$, on a

$y_0 < g(x_0) - \varepsilon'$; alors la droite $y = y_0$ coupe le graphe de $g(x) - \varepsilon''$ (1) en un tel point que $g(x) = y_0 + \varepsilon''$ donc tel que $g(x) < g(x_0) - \varepsilon' + \varepsilon''$;

(1) Si la droite $y = y_0$ ne coupe pas le graphe de $g(x) - \varepsilon''$, la propriété cherchée est évidente.

ce qui implique $g(x_0) - g(x) > \varepsilon' - \varepsilon''$. Donc, d'après (35) et le fait que g est décroissante, on a :

$$x > x_0 + 2 \cdot \frac{1-a}{p}$$

Ce qui entraîne immédiatement la propriété cherchée.

LEMME 34 :

Soit $M_n(a,p)$ l'ensemble des éléments maximaux de l'échantillon dont l'abscisse appartient à l'intervalle $[a, a - \frac{1-a}{p}]$. Alors, les conditions du lemme 33 étant supposées remplies, on a :

$$\text{Prob} \left[\bigcap_{r=n}^{\infty} M_r(a,p) \subset V_{\varepsilon'} \right] \rightarrow 1 \quad (36)$$

quand $n \rightarrow \infty$.

DÉMONSTRATION.

D'après le théorème 28, on peut écrire :

$$\text{Prob} \left[\bigcap_{r=n}^{\infty} \left[\bigcap_{j=0}^{p-1} \{ M_r \cap A_{\varepsilon''_{r,p,j}} \} \neq \emptyset \right] \right] \rightarrow 1$$

mais d'après le lemme 33, l'évènement

$$\left[\bigcap_{j=0}^{p-1} \{ M_r \cap A_{\varepsilon''_{r,p,j}} \} \neq \emptyset \right]$$

entraîne qu'un point de l'échantillon appartenant à $\hat{S} - V_{\varepsilon'}$, ne peut être maximal, donc, avec une probabilité 1 que $M_r(a,p) \subset V_{\varepsilon'}$, d'où (36), C.Q.F.D.

Soient maintenant μ_1, \dots, μ_k les éléments maximaux d'un échantillon de taille n rangés par abscisses croissantes. Désignons par L_n la ligne brisée $\mu_1 \mu_2 \dots \mu_k$ complétée par des parallèles à $x'ox$.

Soit S_n^{\wedge} la fonction affine par morceaux dont le graphe est L_n . On a le résultat suivant :

THÉORÈME 30 :

\hat{g}_n converge uniformément presque sûrement vers g sur tout intervalle $[\alpha, \beta]$ tel que $0 < \alpha < \beta < 1$.

DÉMONSTRATION.

Commençons par montrer la convergence sur un intervalle $[\alpha, \beta] \subset]a, 1[$: Soit $\varepsilon > 0$ et soit $\varepsilon' \in]0, \varepsilon]$; choisissons p assez grand pour que les conditions du lemme 32 soient remplies et tel que l'on ait $\frac{1-a}{p} < \frac{\alpha-a}{2}$ et $\frac{1-a}{p} < \frac{1-\beta}{2}$.

Alors, si $M_n(\alpha, \beta)$ désigne l'ensemble des éléments maximaux dont l'abscisse appartient à $[\alpha, \beta]$,

l'évènement $\bigcap_{r=n}^{\infty} M_r(\alpha, \beta) \subset V_\varepsilon$, implique que

$$E_n \cap \{(x, y) \mid \alpha \leq x \leq \beta, \quad y \in \mathbf{R}\} \subset V_\varepsilon$$

et le résultat découle immédiatement du lemme 34.

Soit maintenant $h > 0$ donné et soit $\alpha \in]a, 1[$ tel que $g(\alpha) = 1 - h/2$: l'évènement

$$E_n = \bigcap_{r=n}^{\infty} \left\{ \sup_{\alpha \leq x \leq \beta} | \hat{g}_r(x) - g(x) | < \frac{h}{2} \right\}$$

entraîne

$$\bigcap_{r=n}^{\infty} \{ \hat{g}_r(\alpha) > 1 - h \}$$

ce qui, compte tenu du fait que g et \hat{g}_r sont non croissantes, entraîne l'évènement

$$\bigcap_{r=n}^{\infty} \left\{ \sup_{0 \leq x \leq \alpha} | \hat{g}_r(x) - g(x) | < h \right\}$$

Finalement E_n entraîne l'évènement

$$\bigcap_{r=n}^{\infty} \left\{ \sup_{0 \leq x \leq \beta} |g_r^{\wedge}(x) - g(x)| < h \right\}$$

et comme $\text{Prob}(E_n) \rightarrow 1$, la démonstration est terminée.

CONCLUSION.

Le théorème 30 donne un exemple d'utilisation des éléments maximaux pour estimer une partie de la frontière du support. Nous pensons que, dans des cas plus généraux, il doit être possible d'estimer la frontière du support en utilisant les éléments maximaux associés à une famille de relations d'ordre convenablement choisie.

BIBLIOGRAPHIE

- [33] BHATTACHARYA (P.K.) (1967). Estimation of a Probability density function and its derivatives.
Sankya, Ser. A 29 p. 373-382.
- [36] BHATTACHARYA (P.K.). PARTHASARATHY (K.R.) (1961).
Some limit theorems in regression theory.
Sankya, Ser. A 23 p. 91-102.
- [6] BICKEL (P.J.). LEHMANN (E.L.) (1969). Unbiased estimation in convex families.
Ann. Math. Stat., Vol. 40 n° 5 p. 1523-1535.
- [32] BOSQ (D.) (1969). Sur l'estimation d'une densité multivariée par une série de fonctions orthogonales.
Comptes rendus, t. 268 p. 555-557.
- [34] BOSQ (D.) (1969). Estimation non paramétrique de la densité et de ses dérivées.
Comptes rendus, t. 269 p. 1010-1012.
- [37] BOSQ (D.) (1969). Estimation de la densité conditionnelle et de la régression.
Comptes rendus, t. 269 p. 661-664.
- [40] BOSQ (D.) (1968). Sur l'estimation d'une fonction continue.
Comptes rendus, t. 266 p. 837-838.
- [48] BOSQ (D.) (1970). Complément à deux notes sur l'estimation de la densité et de ses dérivées.
Comptes rendus, t. 271 p. 45.
- [42] BOSQ (D.) (1968). Sur les éléments extrémaux d'un échantillon vectoriel.
Comptes rendus, t. 266 p. 570-572.
- [15] BOURBAKI (1955). Espaces vectoriels topologiques (Chapitre IV).
Hermann, Paris.
- [41] BOURBAKI (1952). Intégration (Chapitre IV).
Hermann, Paris.

- [20] CACOULLOS (T.) (1966). Estimation of a multivariate density. Ann. Inst. Stat. Math. 18 p. 179-190.
- [25] CENCOV (N.N.) (1962). Evaluation of an unknown distribution density from observations. Soviet Math. 3 p. 1559-1562.
- [2] COURREGE (P.) (1965). Théorie de la Mesure. CDU, Paris.
- [22] CRASWELL (K.J.) (1965). Density estimation in a topological group. Ann. Math. Stat. 36 p. 1047-1048.
- [43] DELTHEIL (R.) (1926). Probabilités géométriques. Gauthier-Villars, Paris.
- [9] DUGUE (D.) (1958). Traité de Statistique théorique et appliquée. Masson, Paris.
- [4] DUNFORD (D.N.), SCHWARTZ (J.T.) (1958). Linear operators (Tome I). Wiley, New York.
- [13] DVORETZKY (A.), KIEFER (J.), WOLFOWITZ (J.) (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. Ann. Math. Stat. 27 p. 642-669.
- [16] FRASER (D.A.S.) (1959). Non parametric Methods in Statistics. Wiley, New York.
- [46] FRECHET (M.) (1958). Remarques au sujet de la note précédente ([45]). Comptes rendus, t. 246 p. 2719-2720.
- [39] GEFFROY (J.) (1964). Sur un problème d'estimation géométrique. Publ. Inst. Stat. Univ. Paris 13 p. 191-210.
- [47] GEFFROY (J.). Sur le nombre d'éléments maximaux d'un échantillon bi-dimensionnel. Comptes rendus (à paraître).
- [45] GUMBEL (E.J.) (1958). Distributions à plusieurs variables dont les marges sont données. Comptes rendus, t. 246 p. 2717-2719.

- [5] HALMOS (P.R.) (1946). The theory of unbiased estimation.
Ann. Math. Stat. 17 p. 34-43.
- [10] JILOVEC (S.) (1967). On the existence of superconsistent estimates.
4th. conf. Inf. th. Stat. Dec. funct. Rand. Proc. Prague 1965 Trans.
Prague, Academia, 24 p. 343-380.
- [28] KRONMAL (R.) et TARTER (M.) (1968). The estimation of probability
densities and cumulatives by Fourier series methods.
J. Amer. Stat. Assoc., 63 p. 925-952.
- [17] LANNUZEL (N.) (1969). Thèse 3ème cycle.
Paris.
- [23] LEADBETTER (M.R.), WATSON (G.S.) (1963). On the estimation of the
probability density I.
Ann. Math. Stat. 34 p. 480-491.
- [49] LE CAM (L.) (1953). On some asymptotic properties of maximum likeli-
hood and related Bayes estimates.
Univ. Calif. Publ. Stat. Vol. 1 p. 277-300.
- [12] LE CAM (L.) (1956). On the asymptotic theory of estimation and testing
hypotheses.
Proc. of 3^d Berkeley symposium p. 129-156
- [11] LE CAM (L.), SCHWARTZ LORRAINE (1960). A necessary and suffi-
cient condition for the existence of consistent estimates.
Ann. Math. Stat. 31 p. 140-150.
- [3] LEHMANN (E.L.) (1951). A general concept of unbiasedness.
Ann. Math. Stat. 22 p. 587-597.
- [31] LOEVE (M.) (1955). Probability Theory.
Van Nostrand, New York.
- [18] NADARAYA (E.A.) (1965). On non-parametric estimates of density
functions and regression curves.
Theor. Prob. Appl. 10 p. 186-190.
- [1] PARZEN (E.) (1958). On asymptotically efficient consistent estimates
of the spectral density of a stationary time series.
J. Roy. Stat. Soc. Ser B 20 p. 303-322.

- [29] PARZEN (E.) (1962). On estimation of a probability density function and mode.
Ann. Math. Stat. 33 p. 1065-1076.
- [24] PICKANDS (J.) III (1969). Efficient estimation of a probability density function.
Ann. Math. Stat. 40 p. 854-864.
- [30] RENYI (A.) (1966). Calcul des Probabilités
Dunod, Paris.
- [44] RENYI (A.) (1962). Théorie des éléments saillants d'une suite d'observations.
Colloquium, Aarhus p. 104-117.
- [38] RENYI (A.).SULANKE (R.) (1963,1964).Über die convexe hulle von is zufällig gewählten punkten, I et II.
Z. Wahr. 2 p. 75-83 et 3 p. 138-148.
- [7] ROSENBLATT (M.) (1956). Remark on some non parametric estimates of a density function.
Ann. Math. Stat. 27 p. 832-837.
- [35] SANSONE (G.) (1959). Orthogonal functions.
Intersciences, New York.
- [14] SCHUSTER (E.F.) (1969). Estimation of a probability density function and its derivatives.
Ann. Math. Stat. 40 p. 1187-1195.
- [27] SCHWARTZ (S.C.) (1967). Estimation of probability density by an orthogonal series.
Ann. Math. Stat. 38 p. 1261-1265.
- [26] TIAGO DE OLIVEIRA (J.) (1961). Estatística de densidades; Resultados assintoticos.
Rev. Fac. Ci. Univ. Lisboa A 9 1961 p. 165-171.
- [8] TUKEY (J.W.) (1961). Curves as parameters, and touch estimation.
Proc. 4 th. Symp. Math. Stat. Prob.,
Berkeley, p. 681-694.

- [21] VAN RYZIN (J.) (1969). On strong consistency of density estimates. Ann. Math. Stat. 40 p. 1765-1772.
- [19] WOODROOFE (M.) (1967). On the maximal deviation of the sample density. Ann. Math. Stat. 38 p. 475-481.

Reçu en Octobre 1970
 104. Avenue de Verdun
 Esc. 4
 94 - IVRY

1. The Model.

The model with which we shall deal in this paper is the following. Given are three variables, ξ , η , and x , which are considered to be linearly related by the expression

$$\eta = \alpha + \beta \xi + \gamma x$$

in which α , β and γ are unknown parameters.

The variable x can be determined exactly, but the variables ξ and η are observed with errors. This implies that instead of ξ and η we observe the random variables

$$z = \xi + u$$

$$y = \eta + v,$$

in which it is assumed that the random variable u has a normal distribution, with expectation 0 and variance σ^2 , and the random variable v has a normal distribution, with expectation 0 and variance τ^2 . It is furthermore assumed that the random variables u and v do not depend on each other.

Now the problem is to estimate the parameters α , β and γ on the basis of n observations:

$(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)$, on the assumption that both $y_i = \eta_i + v_i$ ($i = 1, 2, \dots, n$) and $x_i = \xi_i + u_i$ ($i = 1, 2, \dots, n$) are mutually independent.

The purpose is to determine the functional relation $\eta = \alpha + \beta \xi + \gamma x$.

*PUBLISHED BY THE SYSTEMS AND ESTIMATION RESEARCH GROUP, Eindhoven University of Technology.

- [19] PARZEN (E.) (1962). On estimation of a probability density function and its derivative. *Ann. Math. Stat.* 33 p. 1065-1076.
- [20] PARZEN (E.) (1967). On estimation of a probability density function and its derivative. *Ann. Math. Stat.* 38 p. 473-481.
- [21] RENYI (A.) (1966). *Calcul des Probabilités*. Révisé par A. L. F. (1970). *Colloques, Ankara* p. 104-117.
- [22] RENYI (A.) (1962). *Théorie des Éléments aléatoires*. *Colloques, Ankara* p. 104-117.
- [23] RENYI (A.), SU LANEE (R.) (1963, 1964). Über die konvexe hülle von n zufällig gewählten punkten. I et II. *Z. Natur.* 2 p. 75-83 et 3 p. 139-142.
- [24] ROSENBLATT (M.) (1956). Remark on some non parametric estimates of a density function. *Ann. Math. Stat.* 27 p. 832-837.
- [25] SANSONE (G.) (1959). *Orthogonal functions*. Interscience, New York.
- [26] SCHUSTER (E. F.) (1969). Estimation of a probability density function and its derivatives. *Ann. Math. Stat.* 40 p. 1187-1195.
- [27] SCHARWY (S. C.) (1967). Estimation of probability density by an orthogonal series. *Ann. Math. Stat.* 38 p. 1261-1265.
- [28] TIAGO DE OLIVEIRA (J.) (1961). *Estadística de densidades*. *Revista dos estatísticos*. *Rev. Fac. Ci. Univ. Lisboa A* 9 1961 p. 365-371.
- [29] TUKEY (J. W.) (1961). Curves as parameters, and their estimation. *Proc. 4th Symp. Math. Stat. Prob.*, Berkeley, p. 671-694.