ANNALES DE L'I. H. P., SECTION B

DENIS BOSQ
JACQUES BLEUEZ

Étude d'une classe d'estimateurs non-paramétriques de la densité

Annales de l'I. H. P., section B, tome 14, n° 4 (1978), p. 479-498
<http://www.numdam.org/item?id=AIHPB_1978__14_4_479_0>

Ann. Inst. Henri Poincaré,
Vol. XIV, n° 4, 1978, p. 479-498.

Section B :
Calcul des Probabilités et Statistique.

Étude d'une classe d'estimateurs non-paramétriques de la densité (*)

par

Denis BOSQ et Jacques BLEUEZ
Université des Sciences et Techniques de Lille I,
U. E. R. de Mathématiques Pures et Appliquées

SUMMARY. — The purpose of this paper is the construction and study of a class of density estimators which contains the estimators obtained by the kernel method and the estimators obtained by the orthogonal functions method. Necessary and sufficient conditions are given for the pointwise and uniform convergence of estimators.

PLAN

(*) L'essentiel des résultats contenus dans ce mémoire a été exposé au Congrès des Statisticiens Européens de Grenoble (septembre 1976).

INTRODUCTION

Ce travail est consacré à la construction et à l'étude d'une classe d'estimateurs de la densité qui contient de nombreux estimateurs usuels.

Il s'agit d'estimateurs de la forme

$$\hat{f}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{K}_{r(n)}(\mathrm{X}_i, t)\,; \qquad t \in \mathrm{E},\ n \in \mathbb{N}^* \tag{1}$$

où E est l'espace des observations; $\mathrm{X}_1, \ldots, \mathrm{X}_n$ un échantillon de la loi dont on cherche à estimer la densité et (K_r) une famille de fonctions numériques définie sur $\mathrm{E} \times \mathrm{E}$. On suppose que les X_i sont définis sur le même espace probabilisé $(\Omega, \mathscr{A}, \mathrm{Pr})$.

Cette classe contient les estimateurs obtenus par la *méthode du noyau* c'est-à-dire les estimateurs de la forme

$$\hat{f}_n^{(1)}(t) = \frac{r(n)^s}{n}\sum_{i=1}^{n} \mathrm{K}[r(n)(\mathrm{X}_i - t)]\,; \qquad t \in \mathbb{R}^s,\ n \in \mathbb{N}^*$$

où K est défini sur $\mathbb{R}^s$.

Il en est de même pour les estimateurs

$$\hat{f}_n^{(2)}(t) = \sum_{j=0}^{r(n)}\left[\sum_{i=1}^{n} \frac{1}{n} e_j(\mathrm{X}_i)\right] e_j(t)\,; \qquad t \in \mathrm{E},\ n \in \mathbb{N}^*$$

où $(e_j,\ j \in \mathbb{N})$ est un système orthonormal de $\mathrm{L}^2(\mu)$ *(méthode des fonctions orthogonales)* [1].

Dans la première partie nous essayons de justifier la méthode d'estimation utilisée. La démarche est la suivante : on suppose que la densité à estimer est à valeurs dans $\mathrm{L}^2(\mu)$ et on commence par chercher un estimateur sans biais d'opérateur de covariance minimum.

Or l'existence d'un estimateur sans biais d'ordre 1 pour la densité signifie que l'identité de $\mathrm{L}^2(\mu)$ est un opérateur intégral. On en déduit qu'une condition nécessaire et suffisante pour qu'un tel estimateur existe est que μ soit une mesure discrète (proposition 2). Ce résultat est plus précis que ceux que l'on donne généralement (cf. [5] et [*19*]).

---

[1] Les estimateurs splines de la densité font également partie de cette classe d'estimateurs.

On peut alors construire un estimateur de la densité, soit en estimant sans biais une fonction de la densité (qui doit être un opérateur intégral), soit en restreignant la famille $\mathscr{D}$ des densités possibles de façon à obtenir un estimateur sans biais (l'espace engendré par $\mathscr{D}$ doit être à noyau reproduisant). Dans les deux cas on obtient un estimateur de la forme (1).

Pour qu'une suite de tels estimateurs converge, il faut que la suite d'opérateurs intégraux associés converge vers l'identité (ou encore que la suite d'espaces autoreproduisants associés recouvre $\mathscr{L}^2(\mu)$). Dans la deuxième partie, des hypothèses étant faites sur (K_r), on cherche des conditions sur $(r(n), n \in \mathbb{N}^*)$ pour que l'estimateur converge : on obtient ainsi des conditions nécessaires et suffisantes de convergence simple et de convergence uniforme suivant divers modes stochastiques. On donne aussi des majorations et des minorations de la vitesse de convergence. Enfin des applications à des estimateurs particuliers sont indiquées : il s'agit d'estimateurs construits à l'aide des fonctions trigonométriques, des fonctions d'Hermite, d'une base de Haar.

Les démonstrations de certaines conditions nécessaires (en particulier de conditions nécessaires de convergence simple et uniforme en probabilité) utilisent des méthodes introduites par J. Geffroy et M. Bertrand-Retali dans [*10*].

D'autre part, dans [*17*], A. Földes et P. Revesz ont étudié une classe tout à fait analogue d'estimateurs de la densité : ils ont obtenu des conditions suffisantes de convergence uniforme dans le cas où $E = \mathbb{R}$.

Il semble que les résultats du présent mémoire soient les plus généraux obtenus jusqu'à présent.

Signalons pour terminer que la plupart des résultats obtenus s'étendent à l'estimation de paramètres tels que la régression (quand les marges sont données) et les dérivées de la densité.

I. CONSTRUCTION D'UNE CLASSE D'ESTIMATEURS DE LA DENSITÉ

1. Préliminaires

Soit $(E, \mathscr{B}, \mu)$ un espace mesuré où μ est une mesure σ-finie. Soit $\mathscr{P}_0$ l'ensemble des probabilités sur $(E, \mathscr{B})$ qui admettent, par rapport à μ, une densité de carré μ-intégrable.

On définit une application φ de $\mathscr{P}_0$ dans $L^2(\mu)$ par la formule

$$\varphi(P) = \frac{dP}{d\mu}, \qquad P \in \mathscr{P}_0$$

Enfin on se donne $\mathscr{P} \subseteq \mathscr{P}_0$ et une application Ψ de $\varphi(\mathscr{P})$ dans $L^2(\mu)$. On cherche alors à construire un estimateur sans biais (e. s. b.) de $\Psi \circ \varphi$ à partir d'un échantillon de taille finie tiré suivant $P \in \mathscr{P}$. Quand un tel estimateur existe on cherche un e. s. b. de $\Psi \circ \varphi$ d'opérateur de covariance minimum, les opérateurs symétriques de $L^2(\mu)$ étant munis de leur relation d'ordre usuelle (cf. [*1*], p. 2). Enfin on étudie le cas particulier où Ψ est l'identité, ce qui permet d'obtenir des conditions nécessaires et suffisantes d'existence d'un e. s. b. de la densité.

Dans toute cette partie on supposera que $L^2(\mu)$ est séparable.

2. Existence d'un estimateur sans biais pour un paramètre à valeurs dans $L^2(\mu)$

DÉFINITION 1. — *Soit* T *un estimateur d'ordre n de* $\Psi \circ \varphi$. *Il sera dit sans biais (pour la version* T^**) si* T^* *est une version régulière de* T *telle que :*

a) $\forall P \in \mathscr{P}$, $T^*(., \ldots, . ; t) \in \mathscr{L}^1(P^{\otimes n})$ *pour* μ *presque tous les* t.

b) $\forall P \in P$, $\int T^*(x_1, \ldots, x_n ; .) dP^{\otimes n}(x_1, \ldots, x_n)$ *est une version de* $(\Psi \circ \varphi)$ (P).

Remarques. — 1. Rappelons que T^*, version de T, est dite régulière si $T^*(., \ldots, . ; .)$ est $\mathscr{B}^{n+1}$-mesurable. Dans la suite nous supposerons que les estimateurs considérés possèdent une telle version.

2. Si T est un estimateur d'ordre n dont la norme est de carré $P^{\otimes n}$-intégrable, pour tout $P \in \mathscr{P}$, alors T est sans biais au sens de la définition 1 si et seulement si il l'est au sens usuel de l'intégrale de Bochner (cf. [*2*], lemme 2).

PROPOSITION 1. — 1° *Une condition nécessaire et suffisante pour que* $\Psi \circ \varphi$ *possède un e. s. b. d'ordre 1 est que* Ψ *soit un opérateur intégral dont le noyau* K *soit tel que* $K(X, .) \in \mathscr{L}^2(\mu)$ *pour tout* $x \in E$.

2° *Si la tribu des événements symétriques est complète pour* $\mathscr{P}$ *et si* $K(., .) \in \mathscr{L}^2(P \otimes \mu)$ *pour tout* $P \in \mathscr{P}$, *alors, pour tout entier* $n \geqslant 1$, *la formule*

$$T_n^*(x_1, \ldots, x_n ; t) = \frac{1}{n} \sum_{i=1}^{n} K(x_i, t) ; \qquad (x_1, \ldots, x_n, t) \in E^{n+1}$$

définit une version régulière de l'unique estimateur sans biais d'opérateur de covariance minimum pour $\Psi \circ \varphi$.

Démonstration. — Le 1° est une conséquence directe de la définition 1.

Quant au 2° il se déduit aisément de la version hilbertienne du théorème de Lehmann-Scheffé (cf. [1] proposition 5). ■

Exemple. — Sur $([0, 1], \mathscr{B}_{[0,1]}, \lambda)$ (λ : mesure de Lebesgue) considérons le noyau K défini par

$$K(x, t) = 1_{[0,t]}(x), \qquad (x, t) \in [0, 1]^2$$

alors

$$\int K(x, t) f(x) dx = F(t) ; \qquad t \in [0, 1]$$

où F désigne la f. d. r. associée à f.

La proposition 1 montre alors que la f. d. r. empirique est sans biais et d'opérateur de covariance minimum comme estimateur de la f. d. r. dans la famille $\mathscr{P}_0$ (²).

3. Existence d'un estimateur sans biais de la densité

Dans ce paragraphe, on fait les hypothèses supplémentaires suivantes :

A_1 — La diagonale de E × E, soit Δ, est telle que

$$\Delta^c = \bigcup_{\mathbb{N}} (F_n \times G_n) \quad \text{avec} \quad F_n \cap G_n = \varnothing ; \qquad F_n, G_n \in \mathscr{B} ; \ n \in \mathbb{N}$$

A_2 — $\mathscr{P}$ contient les lois uniformes sur les ensembles de μ-mesure finie et strictement positive.

D'autre part, étant donné T estimateur sans biais (pour la version T*), on dira qu'il est localement μ^2-intégrable si $1_{A \times A} T^* \in \mathscr{L}^2(\mu \otimes \mu)$ chaque fois que $\mu(A) < +\infty$.

Dans ces conditions on a le résultat suivant :

Proposition 2. — 1° *Une condition nécessaire et suffisante pour qu'il existe un e. s. b. de la densité, d'ordre 1 et localement μ^2-intégrable, est que μ soit une mesure discrète.*

2° *Soit n un entier $\geqslant 1$. Si $\mathscr{P}$ est convexe et si μ n'est pas discrète il n'existe aucun estimateur d'ordre n de la densité qui soit sans biais et de norme de carré intégrable.*

Démonstration. — 1° On commence par remarquer que, d'après la proposition 1, l'existence d'un e. s. b. d'ordre 1 de la densité signifie que la restriction de l'identité de $L^2(\mu)$ à $\varphi(\mathscr{P})$ est un opérateur intégral.

(²) Dans cette phrase un élément de $\mathscr{L}^2(\lambda)$ est identifié à sa classe.

En raisonnant comme dans [4] p. 288-289 on montre alors que

$$1_{\Delta^c}T^* = 0 \qquad \mu^2\text{-p. p.} \tag{2}$$

or, A_2 permet d'écrire

$$\int_B T^*(x, t)\mu(B)^{-1}d\mu(x) = \mu(B)^{-1} \tag{3}$$

chaque fois que $0 < \mu(B) < +\infty$ et pour μ-presque tous les t de B.

En intégrant (3) sur B par rapport à μ et en utilisant successivement (2), le théorème de Fubini et le théorème de Radon Nikodym, on trouve

$$T^*(t, t)\mu(t) = 1 \qquad \mu\text{-p. p.}$$

μ est donc discrète. De plus, il est facile de voir que T^* est donné par la formule

$$T^*(x, t) = 1_\Delta(x, t)1_{S_\mu}(t)\mu(t)^{-1}\ ; \qquad (x, t) \in E \times E$$

où $S_\mu = \{ t : \mu(t) > 0 \}$.

La réciproque est claire.

2° On vérifie aisément que le théorème de Bickel-Lehmann ([5] p. 1525-1526) reste valable quand l'espace des paramètres est un espace de Hilbert séparable et si l'on remplace « T sans biais » par « T sans biais et de norme de carré $P^{\otimes n}$-intégrable ($P \in \mathscr{P}$) ». Dans ces conditions, l'existence d'un tel estimateur implique celle d'un estimateur d'ordre 1 possédant les mêmes propriétés ; on conclut à l'aide du 1° de la présente proposition. ∎

4. Estimateurs sans biais de la densité et noyaux reproduisants

Dans la suite on supposera que les densités de $\varphi(\mathscr{P})$ sont définies par une version privilégiée et l'on notera $\mathscr{D}$ le sous-ensemble de $\mathscr{L}^2(\mu)$ formé par les versions considérées. D'autre part, $\mathscr{V}$ désignera l'espace vectoriel engendré par $\mathscr{D}$.

Un estimateur d'ordre n à valeurs dans $\mathscr{L}^2(\mu)$ sera toujours défini par une application numérique $\mathscr{B}^{n+1}$-mesurable et un estimateur T_n de la densité sera dit sans biais si

$$E_p[T_n(X_1, \ldots, X_n\ ; t)] = f_p(t)\ ; \qquad P \in \mathscr{P},\ t \in E$$

où f_p est la version privilégiée de $\dfrac{dP}{d\mu}$ et où E_p désigne l'espérance prise par rapport à P.

Enfin on dira que T_n, estimateur d'ordre n de la densité, vérifie la condition $\mathscr{S}$ si $T(., x_2, \ldots, x_n\ ; t) \in \mathscr{V}$ pour tout $(x_2, \ldots, x_n, t) \in E^n$.

On peut alors énoncer :

PROPOSITION 3. — 1° *Une condition nécessaire et suffisante pour qu'il existe un estimateur des densités de $\mathcal{D}$, d'ordre 1, sans biais et vérifiant la condition $\mathcal{S}$, est que $\mathcal{V}$, muni du produit scalaire usuel de $\mathcal{L}^2(\mu)$, soit un espace préhilbertien séparé à noyau reproduisant mesurable (e. p. s. n. r. m.).*

2° *Si $\mathcal{D}$ est convexe, si $1_E \in \mathcal{V}$ et si $\mathcal{V}$ est un e. p. s. n. r. m. de noyau* K *alors la formule*

$$T_n(X_1, \ldots, X_n ; t) = \frac{1}{n}\sum_{i=1}^{n} K(X_i, t) ; \qquad t \in E$$

définit l'unique e. s. b. symétrique vérifiant $\mathcal{S}$ et tel que

$T(., \ldots, . ; t) \in \mathcal{L}^1(P^{\otimes n})$ *pour tout* $P \in \mathcal{P}$ *et tout* $t \in E$.

Démonstration. — 1° Si T est sans biais d'ordre 1 et s'il vérifie $\mathcal{S}$, alors

$$\langle T(., t), g(.) \rangle = g(t) ; \qquad g \in \mathcal{V},\ t \in E$$

et

$$T(., t) \in \mathcal{V}, \qquad t \in E$$

$T(., .)$ est donc bien le noyau de $\mathcal{V}$ (cf. [*6*]). On en déduit en particulier que T est un estimateur à valeurs dans $\mathcal{V}$.

Inversement, si $\mathcal{V}$ est un e. p. s. n. r. m. de noyau K, l'estimateur $K(X_1, t)$ possède les propriétés demandées.

2° Il suffit de montrer que si U est une fonction numérique $\mathcal{B}^n$-mesurable $P^{\otimes n}$-intégrable pour tout $P \in \mathcal{P}$ et telle que $U(., x_2, \ldots, x_n) \in \mathcal{V}$ pour tout $(x_2, \ldots, x_n) \in E^{n-1}$, alors

$$\forall P \in \mathcal{P}, \qquad E_p[U(X_1, \ldots, X_n)] = 0 \Rightarrow U \equiv 0 .$$

Comme $\mathcal{D}$ est convexe, on peut montrer comme dans [*3*] p. 38 que pour une telle fonction U

$$\int U(x_1, \ldots, x_n) f_1(x_1) \ldots f_n(x_n) d\mu(x_1) \ldots d\mu(x_n) = 0 ; \qquad f_1 \in \mathcal{D}, \ldots, f_n \in \mathcal{D}$$

par linéarité on en déduit que

$$\int U(x_1, \ldots, x_n) K(u_1, x_1) \ldots K(u_n, x_n) d\mu(x_1) \ldots d\mu(x_n) = 0 ; \qquad (u_1, \ldots, u_n) \in E^n$$

mais $U(x_1, \ldots, x_{n-1} ; .) \in \mathcal{V}$, donc

$$\int U(x_1, \ldots, x_{n-1}, u_n) K(u_1, x_1) \ldots K(u_{n-1}, x_{n-1}) d\mu(x_1) \ldots d\mu(x_{n-1}) = 0 ; \qquad (u_1, \ldots, u_n) \in E^n$$

en utilisant $(n-1)$ fois le même argument, on obtient finalement

$$U(u_1, \ldots, u_n) = 0 \; ; \qquad (u_1, \ldots, u_n) \in E^n . \qquad \blacksquare$$

Remarques. — 1. Si $\mathscr{V}$ est un espace de Hilbert et si

$$\int K(x, x) dP(x) < \infty \qquad (P \in \mathscr{P})$$

alors T_n est l'unique estimateur d'opérateur de covariance minimum parmi les estimateurs sans biais qui vérifient $\mathscr{S}$ et qui sont à valeurs dans $\mathscr{V}$ (cf. [2] corollaire 1, p. 13).

2. Si $\mathscr{D}$ est convexe, s'il existe $f_0 \in \mathscr{D}$ tel que $\inf_{t \in E} f_0(t) > 0$ et si T est un estimateur sans biais dont la semi-norme (associée au noyau de $\mathscr{V}$) est de carré P-intégrable, alors $\mathscr{V}$ est de dimension finie (cf. [2] corollaire 4, p. 15).

5. Conclusion

Les deux points de vue envisagés dans l'introduction et développés dans les paragraphes précédents pour construire un estimateur d'ordre n de la densité mènent à des estimateurs de la forme $\frac{1}{n} \sum_{i=1}^{n} K(X_i, t)$.

Notons que les propositions 1 et 3 peuvent servir à donner des indications à l'utilisateur sur le choix de K en fonction des informations *a priori* qu'il possède sur la densité qu'il cherche à estimer. Nous comptons revenir ultérieurement sur ce problème.

Maintenant pour construire un estimateur convergent de la densité on peut utiliser le fait que l'identité de $\mathscr{L}^2(\mu)$ est limite d'une suite d'opérateurs intégraux (resp. le fait que $\mathscr{L}^2(\mu)$ est réunion dénombrable d'espaces à noyau reproduisant).

On se donne alors une partie non bornée de $\mathbb{R}^+$, soit I, et une famille $(K_r, r \in I)$ de fonctions réelles $\mathscr{B}^2$-mesurables telle que la famille d'opérateurs intégraux associée tende vers l'identité lorsque r tend vers l'infini dans I et on considère la suite d'estimateurs de la densité définie par :

$$\hat{f}_n(t) = \frac{1}{n} \sum_{i=1}^{n} K_{r(n)}(X_i, t) \; ; \qquad t \in E$$

où $r(n) \in I$ et $r(n) \to +\infty$.

L'objet de la deuxième partie sera d'étudier la convergence de cette suite d'estimateurs.

II. CONDITIONS NÉCESSAIRES ET SUFFISANTES DE CONVERGENCE

1. Notations et hypothèses

1° Dans la suite, on supposera que la mesure μ est non atomique et on désignera par α un nombre strictement positif attaché à la famille (K_r). On supposera en outre que les densités de $\mathscr{D}$ sont *bornées.*

D'autre part, on posera

$$\mathscr{D}_1 = \left\{ f \,;\, f \in \mathscr{D},\ \lim_{r\to\infty} \int_E K_r(x, t) f(x) d\mu(x) = \bar{f}(t),\ t \in E \right\}$$

et pour $F \subset E$

$$\mathscr{D}_F = \left\{ f \,;\, f \in \mathscr{D},\ \lim_{r\to\infty} \sup_{t\in F} \left| \int_E K_r(x, t) f(x) d\mu(x) - f(t) \right| = 0 \right\}$$

Enfin on définit $D_n(.)$, $f_n(.)$, $\hat{S}_n(.)$ et Δ_n par les formules :

$$D_n(t) = | \hat{f}_n(t) - \bar{f}(t) | \,; \qquad t \in E,\ n \geqslant 1$$

$$f_n(t) = \int_E K_{r(n)}(x, t) f(x) d\mu(x) \,; \qquad t \in E,\ n \geqslant 1$$

$$\hat{S}_n(t) = \hat{f}_n(t) - f_n(t) \,; \qquad t \in E,\ n \geqslant 1$$

$$\Delta_n = \sup_{t\in F} | \hat{f}_n(t) - f(t) | \,; \qquad n \geqslant 1$$

2° On utilisera notamment les hypothèses ci-dessous (les hypothèses qui font intervenir r sont valables pour r assez grand ($r \geqslant r_0$) ; r_0 dépend de t dans les hypothèses I, II, III) :

I (resp. I'). $\exists A_1(.) > 0$ définie sur E (resp. définie et bornée sur F) telle que :

$$\sup_{x\in E} | K_r(x, t) | \leqslant A_1(t) r^\alpha$$

II (resp. II'). $\exists B(.) > 0$ définie sur E (resp. définie et bornée sur F) telle que :

$$\int_E [K_r(x, t)]^2 d\mu(x) \leqslant B(t) r^\alpha$$

III (resp. III'). $\exists g_1 \in \mathscr{D}_1$ (resp. $\in \mathscr{D}_F$), $\exists t_1 \in E$ (resp. $\in F$), $\exists A_2 > 0$ tels que

$$\int_E [K_r(x, t_1)]^2 g_1(x) d\mu(x) \geqslant A_2 . r^\alpha$$

IV (resp. IV'). $\exists f_0 \in \mathcal{D}_1$ (resp. $\in \mathcal{D}_F$), $\exists t_0 \in E$ (resp. $\in F$) et une famille $(B_r, r \in I)$ d'éléments de $\mathcal{B}$ tels que

a) $\bar{f}_0(t_0) > 0$ (resp. $f_0(t_0) > 0$)

b) $\mu(B_r) = lr^{-\alpha}(l > 0)$ et

$$\int_{B_r} K_r(x, t_0) f_0(x) d\mu(x) \geqslant \delta \qquad (\delta > 0)$$

IV'' $\exists A_3 > 0$ tel que $\inf_{x \in E} K_r(x, x) \geqslant A_3 r^\alpha$

V. $(E, \mathcal{B})$ est un espace métrique muni de sa tribu borélienne et $K_r(x, .)$ vérifie la condition de Lipchitz suivante :

$$\sup_{x \in E} | K_r(x, t') - K_r(x, t) | \leqslant Cr^m | t' - t |^\gamma ;$$

$(t, t') \in E^2$, $c > 0$, $m > 0$, $\gamma > 0$.

VI. $(E, \mathcal{B})$ est un espace métrique muni de sa tribu borélienne, $F \in \mathcal{B}$ est totalement borné et tel que $0 < \mu(F) < +\infty$. De plus, il existe une constante positive h telle que, pour ε positif assez petit, on puisse recouvrir F par $[\varepsilon^{-h}]$ boules de rayon $\leqslant \varepsilon$.

Remarque. — IV $\Rightarrow$ III et IV' $\Rightarrow$ III'

2. Conditions nécessaires et suffisantes de convergence simple

PROPOSITION 4. — *Sous les hypothèses II et IV les trois propriétés suivantes sont équivalentes :*

a) $\dfrac{r(n)^\alpha}{n} \to 0$

b) $E[D_n(t)] \to 0$; $t \in E$, $f \in \mathcal{D}_1$

c) $E[D_n(t)^2] \to 0$; $t \in E$, $f \in \mathcal{D}_1$

Si, de plus, K_r *est positif ces propriétés sont équivalentes à*

d) $D_n(t) \xrightarrow[p]{} 0$

Démonstration. — Sous l'hypothèse II on a évidemment (*a*) $\Rightarrow$ (*c*) $\Rightarrow$ (*b*). Montrons que, sous l'hypothèse IV, on a (*b*) $\Rightarrow$ (*a*) : comme dans [*10*] la démonstration se fait par l'absurde.

On commence par remarquer que, s'il existe $\beta > 0$ et N_0 partie infinie de $\mathbb{N}^*$ tels que $\dfrac{r(n)^\alpha}{n} \geqslant \beta$ pour $n \in N_0$, on a aussi pour n assez grand dans N_0

$$\Pr(E_n) \geqslant \exp[-2M_0 l/\beta] \tag{1}$$

où f_0 est la densité des X_i, $M_0 = \sup_{t \in E} f_0(t)$ et $E_n = \bigcap_{i=1}^{n} \{ X_i \notin B_{r(n)} \}$.

Maintenant, de

$$E_{f_0}[\hat{f}_n^+(t_0)] - E_{f_0}[\hat{f}_n^-(t_0)] \to \bar{f}_0(t_0) > 0$$

et

$$E_{f_0}[D_n(t_0)] \to 0$$

on tire aisément que

$$E_{f_0}[\hat{f}_n^-(t_0)] \to 0 \tag{2}$$

Mais

$$E_{f_0}[\hat{f}_n^-(t_0) \mid E_n] = \frac{1}{\Pr(E_n)} \int_{E_n} \hat{f}_n^-(t_0) d\Pr$$

donc (1) et (2) entraînent

$$\lim_{\substack{n\to\infty \\ n\in N_0}} E_{f_0}[\hat{f}_n^-(t_0) \mid E_n] = 0 \tag{3}$$

Choisissons alors $\delta < \bar{f}_0(t_0)$ et donnons nous ξ et η positifs et tels que $\eta < \delta$ et $\xi < (\delta - \eta)(\bar{f}_0(t_0) + \eta - \delta)^{-1}$.

Pour n assez grand on aura les inégalités

$$[\Pr(X_i \in B^c_{r(n)})]^{-1} < 1 + \xi$$

$$\int_{B_{r(n)}} K_{r(n)}(x, t_0) f_0(x) d\mu(x) \geqslant \delta$$

et

$$| f_n(t_0) - \bar{f}_0(t_0) | < \eta$$

Mais comme

$$E_{f_0}[\hat{f}_n(t_0) \mid E_n] = [\Pr(X_1 \in B^c_{r(n)})]^{-1} \int_{B^c_{r(n)}} K_{r(n)}(x, t_0) f_0(x) d\mu(x)$$

on en déduit alors que

$$E_{f_0}[\hat{f}_n(t_0) \mid E_n] < (1 + \xi)(\bar{f}_0(t_0) + \eta - \delta) < \bar{f}_0(t_0) \tag{4}$$

Donnons nous maintenant $\gamma \in](1 + \xi)(\bar{f}_0(t_0) + \eta - \delta), \bar{f}_0(t_0)[$ et considérons l'inégalité

$$\Pr[\hat{f}_n^+(t_0) > \gamma \mid E_n] \leqslant \frac{1}{\gamma}[E_{f_0}(\hat{f}_n(t_0) \mid E_n) + E_{f_0}(\hat{f}_n^-(t_0) \mid E_n)]$$

Compte tenu de (3), (4) et du choix de γ on en déduit pour n assez grand dans N_0 :

$$\Pr[\hat{f}_n^+(t_0) > \gamma \mid E_n] \leqslant \varepsilon < 1$$

où ε ne dépend que de $\bar{f}_0(t_0)$ et de γ.

Maintenant, comme $\gamma < \bar{f}_0(t_0)$, on a l'implication

$$\hat{f}_n^+(t_0) \leqslant \gamma \Rightarrow | \hat{f}_n^+(t_0) - \bar{f}_0(t_0) | \geqslant \bar{f}_0(t_0) - \gamma$$

d'où

$$\Pr\left[\,|\hat{f}_n^+(t_0) - \bar{f}_0(t_0)| \geqslant \bar{f}_0(t_0) - \gamma\right] \geqslant \Pr\,[\hat{f}_n(t_0) \leqslant \gamma]$$
$$\geqslant \Pr\,(E_n)\Pr\left[\hat{f}_n^+(t_0) \leqslant \gamma \mid E_n\right]$$

donc, pour n assez grand dans N_0

$$\Pr\left[\,|\hat{f}_n^+(t_0) - \bar{f}_0(t_0)| \geqslant \bar{f}_0(t_0) - \gamma\right] \geqslant \exp\,[-\,2M_0 l/\beta](1-\varepsilon) > 0$$

ce qui contredit $\hat{f}_n^+(t_0) \underset{p}{\rightarrow} f_0(t_0)$ donc aussi (*b*).

Il reste à montrer que, si K_r est positif, (*d*) entraîne (*a*) ; pour cela il suffit de reprendre la démonstration précédente.

En effet, l'hypothèse (*b*) n'est utilisée que pour montrer (3) et cette propriété est triviale quand K_r est positif. ■

Le lemme suivant donne une majoration de la vitesse de convergence de $\hat{f}_n(t)$ vers $\bar{f}(t)$:

LEMME 1. — *Sous les hypothèses I et II, on a :*

$$\Pr\,[D_n(t) \geqslant \varepsilon] \leqslant 2\exp\left[-\,\rho(t)\,\varepsilon\,\frac{n}{r(n)^\alpha}\right] \tag{5}$$

$t \in E$, $\varepsilon > 0$, $f \in \mathscr{D}_1$, *n assez grand, où*

$$\rho(t) = \frac{1}{4A_1(t)}\left[\left(1 + \frac{MB(t)}{A_1(t)}\right).\ \mathrm{Log}\left(1 + \frac{A_1(t)}{MB(t)}\right) - 1\right]$$

et $M = \sup_{t\in E} f(t)$.

Démonstration. — Elle consiste à appliquer à $\hat{S}_n(t)$ une inégalité de type exponentiel due à Hœffding (cf. [7], p. 58) en utilisant le fait que la fonction $(1+u)\,\mathrm{Log}\,(1+u^{-1}) - 1$ est positive et décroissante sur $]0, +\infty[$. ■

Remarque. — Pour que l'inégalité (5) soit vérifiée il suffit que $(K_{r(n)})$ vérifie I et II et que $|f_n(t) - \bar{f}(t)| < \dfrac{\varepsilon}{2}$.

Nous allons maintenant donner une condition nécessaire et suffisante de convergence simple presque complète et en moyenne d'ordre 1 ou 2.

Rappelons tout d'abord que l'on dit que (Y_n) suite de v. a. r. converge presque complètement (p. co.) vers Y si

$$\forall\varepsilon > 0, \qquad \sum_{\mathbb{N}} P_r(\,|Y_n - Y| \geqslant \varepsilon) < +\infty\,.$$

La convergence presque complète entraîne la convergence presque sûre.

PROPOSITION 5. — *Sous les hypothèses I, II, IV les trois propriétés suivantes sont équivalentes :*

a) $$\forall \beta > 0, \quad \sum_{\mathbb{N}^*} \exp\left[-\beta \frac{n}{r(n)^\alpha}\right] < +\infty .$$

b) $\mathrm{D}_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$ *et* $\mathrm{E}[\mathrm{D}_n(t)] \to 0$; $\quad t \in \mathrm{E},\ f \in \mathscr{D}_1$

c) $\mathrm{D}_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$ *et* $\mathrm{E}[\mathrm{D}_n(t)^2] \to 0$; $\quad t \in \mathrm{E},\ f \in \mathscr{D}_1$.

Si, de plus, K_r *est positif, a) et d) sont équivalentes, avec*

d) $\mathrm{D}_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$; $\quad t \in \mathrm{E},\ f \in \mathscr{D}_1$.

Démonstration. — *a)* entraîne *c)* d'après le lemme 1 et la proposition 4. D'autre part, il est clair que *c)* implique *b)*. Enfin, si K_r est positif, *b)* et *d)* sont équivalents d'après la proposition 4.

Il suffit donc de montrer que *b)* implique *a)*. Pour cela, on suppose que la densité des X_i est f_0 et on remarque que *b)* implique $\dfrac{r(n)^\alpha}{n} \to 0$.

On peut alors appliquer à $\hat{\mathrm{S}}_n(t)$ une minoration exponentielle due à Kolmogorov (cf. [*8*], p. 378-384); il vient pour n assez grand et a positif bien choisi

$$\Pr[\hat{\mathrm{S}}_n(t) \geqslant \varepsilon] > \exp\left[-\mathrm{R}\varepsilon^2 \frac{n}{r(n)^\alpha} \ ;\right] \quad 0 < \varepsilon < a \tag{6}$$

où R est une constante positive.

L'inégalité (6) permet de conclure. ■

Remarque. — (6) permet d'obtenir une minoration de la vitesse de convergence de $\hat{f}_n(t)$ vers $\bar{f}(t)$ pour toute densité de $\mathscr{D}_1$ vérifiant III.

On a, en effet, pour n assez grand

$$\Pr[\mathrm{D}_n(t) \geqslant \varepsilon] > \exp\left[-4\mathrm{R}\varepsilon^2 \frac{n}{r(n)^\alpha}\right]; \quad 0 < \varepsilon < a \tag{7}$$

3. Conditions nécessaires et suffisantes de convergence uniforme

Donnons d'abord une condition suffisante de convergence uniforme :

LEMME 2. — *Sous les hypothèses I', II', V, VI on a, si* $f \in \mathscr{D}_\mathrm{F}$,

$$\frac{r(n)^\alpha}{n} \operatorname{Log} n \to 0 \Rightarrow \Delta_n \underset{\text{p.co.}}{\longrightarrow} 0 \quad \textit{et} \quad \mathrm{E}\Delta_n \to 0 .$$

Démonstration. — Montrons d'abord que $\Delta_n \underset{\text{p.co.}}{\longrightarrow} 0$; comme $f \in \mathscr{D}_\mathrm{F}$ il suffit de montrer que $\sup_{t \in \mathrm{F}} |\hat{\mathrm{S}}_n(t)| \underset{\text{p.co.}}{\longrightarrow} 0$.

D'après VI, on peut trouver $v > 0$ tel que, pour n assez grand, il existe un recouvrement de F par $[r(n)^v]$ boules de rayon $\leqslant r(n)^{-\frac{m+1}{\gamma}}$ et de centres respectifs $t_j (1 \leqslant j \leqslant [r(n)^v])$.

Il suffit de choisir $v = (m+1)h\gamma^{-1}$.

Alors, si t appartient à la boule de centre t_j, il vient d'après V

$$|\hat{S}_n(t) - \hat{S}_n(t_j)| \leqslant \frac{2C}{r(n)}\,; \qquad j = 1, \ldots, [r(n)^v]$$

on en tire l'inégalité

$$\sup_{t \in F} |\hat{S}_n(t)| \leqslant \sup_{1 \leqslant j \leqslant [r(n)^v]} |\hat{S}_n(t_j)| + \frac{2C}{r(n)}$$

et pour n assez grand

$$\Pr[\sup |\hat{S}_n(t)| \geqslant \eta] \leqslant \Pr\left[\sup_{1 \leqslant j \leqslant [r(n)^v]} |\hat{S}_n(t_j)| \geqslant \frac{\eta}{2}\right]$$

$$\leqslant \sum_{j=1}^{[r(n)^v]} \Pr\left[|\hat{S}_n(t_j)| \geqslant \frac{\eta}{2}\right]$$

Maintenant, d'après le lemme 1 et l'hypothèse I′

$$\Pr[\sup_{t \in F} |\hat{S}_n(t)| \geqslant \eta] \leqslant 2r(n)^v \exp\left[-\beta \frac{n}{r(n)^\alpha}\right] \tag{8}$$

où β est une constante. Mais

$$\frac{r(n)^\alpha}{n} \operatorname{Log} n \to 0 \Rightarrow \sum_{\mathbb{N}^*} r(n)^v \exp\left[-\beta \frac{n}{r(n)^\alpha}\right] < +\infty$$

ce qui donne le résultat.

Maintenant, pour montrer que $E\Delta_n \to 0$, il suffit de vérifier que $E[\sup_{t \in F} |\hat{S}_n(t)|] \to 0$.

Or, d'après une inégalité bien connue (cf. [9], p. 58)

$$E[\sup_{t \in F} |\hat{S}_n(t)|] \leqslant 2 . \sup_{t \in F} A_1(t) . \Pr[\sup_{t \in F} |\hat{S}_n(t)| > \varepsilon] + \varepsilon\,; \qquad \varepsilon > 0\,.$$

La première partie permet alors de conclure. ■

Pour l'étude des conditions nécessaires et suffisantes de convergences uniformes, nous distinguerons trois cas. Donnons d'abord une définition :

Définition 2. — *Nous dirons que $r(.)^\alpha$ est asymptotiquement concave s'il existe une fonction g concave, $a > 0$ et $b > 0$ tels que, pour n assez grand*

$$ar(n)^\alpha \leqslant g(n) \leqslant br(n)\,.$$

L'intérêt de cette notion réside dans le fait que, si $r(.)^\alpha$ est asymptotiquement concave, les conditions « $\frac{r(n)^\alpha}{n} \operatorname{Log} n \to 0$ » et

$$\forall \beta > 0, \qquad \sum_{\mathbb{N}^*} \exp\left[-\beta \frac{n}{r(n)^\alpha}\right] < +\infty \text{ »}$$

sont équivalentes

Remarquons que, dans les cas usuels, le statisticien sera souvent amené à choisir $r(.)^\alpha$ asymptotiquement concave pour éviter des variations brutales de la forme de l'estimateur quand la taille de l'échantillon augmente.

Nous allons maintenant envisager successivement le cas où $r(.)^\alpha$ est asymptotiquement concave, le cas d'un noyau positif, puis le cas général.

1) *$r(.)^\alpha$ est asymptotiquement concave.*

PROPOSITION 6. — *Sous les hypothèses I', II', IV', V, VI et si $r(.)^\alpha$ est asymptotiquement concave, les propriétés suivantes sont équivalentes*

a) $\frac{r(n)^\alpha}{n} \operatorname{Log} n \to 0$

b) $D_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$, $ED_n(t) \to 0$; $\quad t \in F$, $f \in \mathscr{D}_F$

c) $\Delta_n \underset{\text{p.co.}}{\longrightarrow} 0$, $E\Delta_n \to 0$; $\quad f \in \mathscr{D}_F$

Si, de plus, K_r est positif elles sont encore équivalentes à

d) $D_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$; $\quad t \in F$, $f \in \mathscr{D}_F$

e) $\Delta_n \underset{\text{p.co.}}{\longrightarrow} 0$; $\quad f \in \mathscr{D}_F$

Démonstration. — Elle découle immédiatement de la proposition 5 et du lemme 2. ■

2) *Le noyau K_r est positif.*

On suppose ici que $E = \mathbb{R}^s$, que μ est la mesure de Lebesgue sur $\mathbb{R}^s$ et que F est une partie bornée de E telle que $\mathring{F} \neq \emptyset$.

PROPOSITION 7. — *Si $K_r(x, t) \geqslant 0$, $(x, t) \in E^2$, si les hypothèses I', II', IV'', V sont vérifiées et si de plus*

$$m \leqslant \alpha(1 + \gamma s^{-1})$$

alors les conditions suivantes sont équivalentes

a) $\frac{r(n)^\alpha}{n} \operatorname{Log} n \to 0$

b) $\Delta_n \xrightarrow[p]{} 0$; $f \in \mathscr{D}_F$
c) $\Delta_n \xrightarrow[\text{p.s.}]{} 0$; $f \in \mathscr{D}_F$
d) $\Delta_n \xrightarrow[\text{p.co.}]{} 0$; $f \in \mathscr{D}_F$
e) $E\Delta_n \to 0$; $f \in \mathscr{D}_F$.

Démonstration. — D'après ce qui précède, il suffit de montrer que *b*) $\Rightarrow$ *a*). La démonstration de cette implication est tout à fait analogue à celle que J. Geffroy et M. Bertrand-Retali ont donné dans [*10*] pour un noyau de convolution. Pour les détails nous renvoyons à [*11*].

3) *Le cas général.*

Pour le cas général, nous aurons besoin d'une hypothèse supplémentaire :

VII. $\forall \eta > 0$, $\exists E_1 \in \mathscr{B} \cap F$, $\exists m \in]0, \eta[\mu(E_1)]^{-1}[$, $\exists l(E_1) > 0$ tels que

a) $\exists g_0 \in \mathscr{D}_F$ tel que $g_0(x) = m$ pour $x \in E_1$

b) Pour r assez grand, $\sup_{t \in F_1} \left| \int_{E_1} K_r(x, t) d\mu(x) \right| \leqslant l(E_1)$

c) Il existe une suite $\{ B_{n,t_l} ; 1 \leqslant l \leqslant [r(n)^\alpha] \}$, $n \in \mathbb{N}^*$ de partitions de E_1 en ensembles de même mesure et un nombre $\delta_1 > 0$ tel que, pour $\varepsilon \in]0, 1]$, n assez grand et j vérifiant $1 \leqslant j \leqslant \varepsilon[r(n)^\alpha]$, les inégalités

$$\int_{\bigcup_{l=1}^{j} B_{n,t_l}} K_{r(n)}(x, \xi) d\mu(x) \geqslant \delta_1 \ ; \qquad 1 \leqslant t_1 < \ldots < t_j \leqslant [r(n)^\alpha]$$

pour ξ convenablement choisi dans E_1, soient vérifiées pour une proportion de j-uples $(t_1, \ldots, t_j)$, $1 \leqslant t_1 < \ldots < t_j \leqslant [r(n)^\alpha]$, plus grande qu'un certain nombre $p > 0$ indépendant de n et de j.

Remarque. — L'hypothèse VII est vérifiée dans les cas usuels comme nous le verrons dans le paragraphe suivant.

D'autre part, sous les hypothèses de la proportion 7, VII est une conséquence de IV″ et V.

Dans ces conditions, on obtient la

PROPOSITION 8. — *Sous les hypothèses I′, II′, V, VI, VII, les conditions suivantes sont équivalentes :*

a) $\dfrac{r(n)^\alpha}{n} \operatorname{Log} n \to 0$

b) $E\Delta_n \to 0$; $f \in \mathscr{D}_F$

c) $E\Delta_n \to 0$; $\Delta_n \xrightarrow[\text{p.co.}]{} 0$; $f \in \mathscr{D}_F$

Démonstration. — *a*) $\Rightarrow$ *c*) (lemme 2) ; *c*) $\Rightarrow$ *b*) (évident) ; pour montrer que *b*) $\Rightarrow$ *a*) on reprend encore la méthode de J. Geffroy et M. Bertrand-Retali ([*10*]). Pour une démonstration détaillée nous renvoyons à [*11*] p. 13-30. ■

Le corollaire suivant résume les résultats obtenus sur la vitesse de convergence de Δ_n vers 0.

COROLLAIRE. — *Sous les hypothèses du lemme 2, on a pour tout* $\eta > 0$, *toute densité* $f \in \mathcal{D}_F$ *et n assez grand*

$$\Pr\left[\Delta_n \geqslant \eta\right] \leqslant 2 \exp\left[-\rho\eta\frac{n}{r(n)^\alpha}\right] \tag{9}$$

où

$$\rho = \frac{1}{5A_1}\left[\left(1 + \frac{MB}{A_1\eta}\right)\operatorname{Log}\left(1 + \frac{A_1\eta}{MB}\right) - 1\right]$$

avec $A_1 = \sup_{t\in F} A_1(t)$; $B = \sup_{t\in F} B(t)$; $M = \sup_{t\in E} f(t)$.

En outre, pour toute densité de $\mathcal{D}_F$ *vérifiant III', il existe b tel que, pour n assez grand*

$$\exp\left[-\rho''\eta^2\frac{n}{r(n)}\right] \leqslant \Pr\left[\Delta_n \geqslant \eta\right] \leqslant 2\exp\left[-\rho'\eta^2\frac{n}{r(n)^\alpha}\right], \quad 0 < \eta < b, \tag{10}$$

où ρ' *et* ρ'' *sont des nombres positifs qui ne dépendent que de* (K_r) *et de la densité considérée.*

Démonstration. — (9) se déduit aisément de (8) et de (5). Quant à (10) c'est une conséquence de (7) et de l'implication

$$D_n(t) \geqslant \varepsilon \Rightarrow \Delta_n \geqslant \varepsilon . \quad ■$$

4. Exemples et cas particuliers

Dans ce paragraphe, nous donnons des exemples d'estimateurs qui vérifient les hypothèses précédentes. Nous indiquons également des cas particuliers où l'on peut améliorer certains résultats obtenus précédemment.

a) *La méthode du noyau.*

Nous dirons peu de choses sur cette méthode qui a été abondamment traitée par ailleurs.

Signalons seulement que si $E = \mathbb{R}^s$, si μ est la mesure de Lebesgue et si K est une densité de carré intégrable, la proposition 4 s'applique. Il en est de même pour la proposition 5 si l'on suppose en outre que K est borné. Si, de plus, K est lipchitzienne, la proposition 7 s'applique. Dans tous les cas on a $\alpha = 1$.

Bien entendu, les résultats obtenus pour la convergence uniforme dans [*10*] sont meilleurs que ceux de la proposition 7 puisque la méthode de démonstration utilise directement le fait que K_r est une fonction positive de $x - t$.

b) *La méthode des fonctions orthogonales.*

Les propositions précédentes s'appliquent à de nombreux systèmes orthogonaux avec parfois des améliorations.

Nous allons indiquer rapidement les résultats obtenus dans trois cas particuliers importants :

— les fonctions trigonométriques
— les fonctions d'Hermite
— les bases de Haar (cf. [*12*], p. 51).

A propos de ce dernier exemple, nous signalons une application des propositions précédentes à un problème d'estimation dans lequel E est de dimension infinie.

PROPOSITION 9. — *Soit $\mathcal{D}$ l'ensemble des densités f définies sur $[a, b]$ intervalle compact de $\mathbb{R}$, bornées et possédant une série de Fourier qui converge vers $f(t)$ en tout point t de $[a, b]$. Alors :*

1° Les propriétés suivantes sont équivalentes :

a) $\dfrac{r(n)}{n} \to 0$

b) $D_n(t) \underset{p}{\to} 0\,; \qquad t \in [a, b],\ f \in \mathcal{D}$

c) $ED_n(t) \to 0\,; \qquad t \in [a, b],\ f \in \mathcal{D}$

d) $E[D_n(t)^2] \to 0\,; \qquad t \in [a, b],\ f \in \mathcal{D}$

e) $E \displaystyle\int_b^a D_n(t)^2 dt \to 0\,; \qquad f \in \mathcal{D}$

2° Une condition nécessaire et suffisante pour que $D_n(t) \underset{\text{p.co.}}{\longrightarrow} 0$ $t \in [a, b]$, $f \in \mathcal{D}$ est que l'on ait :

$$\sum_{\mathbb{N}^*} \exp\left[-\beta \frac{n}{r(n)}\right] < +\infty\,, \qquad \beta > 0$$

3° Les propriétés suivantes sont équivalentes :

a) $\frac{r(n)}{n} \operatorname{Log} n \to 0$

b) $\Delta_n \underset{p}{\to} 0\,; \quad f \in \mathscr{D}_{[a,b]}$

c) $\Delta_n \underset{\text{p.s.}}{\to} 0\,; \quad f \in \mathscr{D}_{[a,b]}$

d) $\Delta_n \underset{\text{p.co.}}{\longrightarrow} 0\,; \quad f \in \mathscr{D}_{[a,b]}$

e) $E\Delta_n \to 0\,; \quad f \in \mathscr{D}_{[a,b]}$

PROPOSITION 10. — *Si $\hat{f}_n$ est construit à partir des fonctions d'Hermite on a les résultats suivants :*

1° Les propriétés suivantes sont équivalentes :

a) $\frac{r(n)}{n^2} \to 0$.

b) $D_n(t) \underset{p}{\to} 0\,; \quad t \in \mathbb{R},\ f \in \mathscr{D}_1$

c) $ED_n(t) \to 0\,; \quad t \in \mathbb{R},\ f \in \mathscr{D}_1$

d) $E[D_n(t)^2] \to 0\,; \quad t \in \mathbb{R},\ f \in \mathscr{D}_1$

2° Une condition nécessaire et suffisante pour que l'on ait à la fois $\Delta_n \underset{\text{p.co.}}{\longrightarrow} 0$ et $E\Delta_n \to 0$; $f \in \mathscr{D}_{[a,b]}$ ($[a, b]$ intervalle compact de $\mathbb{R}$) est que

$$\frac{r(n)^{1/2}}{n} \operatorname{Log} n \to 0 .$$

3° Une condition suffisante pour que

$$\sup_{t\in\mathbb{R}} | \hat{f}_n(t) - f(t) | \underset{\text{p.co.}}{\longrightarrow} 0\,; \qquad f \in \mathscr{D}_{\mathbb{R}}$$

est que $\frac{r(n)}{n} \operatorname{Log} n \to 0$.

PROPOSITION 11. — *Soit $\mathscr{D}$ l'ensemble des densités définies et bornées sur $(E, \mathscr{B}, \mu)$ où $\mathscr{B}$ est engendrée par un système de Haar et μ est une probabilité sur $(E, \mathscr{B})$.*

Alors si l'estimateur $\hat{f}_n$ est construit à partir de la base de Haar associée à ce système on a les mêmes résultats que dans la proposition 9 à ceci près que

— dans 1° et 2°, il faut remplacer « $t \in [a, b]$ » par « pour μ presque tous les $t \in E$ » et e) ne figure pas.

— dans 3°, il faut remplacer $\mathscr{D}_{[a,b]}$ par $\mathscr{D}_E$.

Démonstrations. — Pour les démonstrations, nous renvoyons à [*11*] ou [*13*].

Remarque. — Considérons un processus à trajectoires continues dont on observe une suite de « copies » sur des intervalles de longueur T. Si la loi du processus (considéré comme une v. a. à valeurs dans C[0, T]) admet une densité par rapport à une mesure sur C[0, T] la proposition 11 permet de construire un estimateur de cette densité.

BIBLIOGRAPHIE

[*1*] D. BOSQ, Sur l'estimation d'un paramètre à valeurs dans un espace de Hilbert. *Publ. int. univ. Lille I*, n° 70, 1976.
[*2*] D. BOSQ, Sur l'estimation sans biais d'un paramètre fonctionnel. *Publ. int. univ. Lille I*, n° 104, 1977.
[*3*] P. HALMOS, The theory of unbiased estimation. *Ann. Math. Stat.*, t. **17**, 1946, p. 34-43.
[*4*] P. HALMOS, *A Hilbert space Problem Book*. Van Nostrand, 1967.
[*5*] P. J. BICKEL, E. L. LEHMANN, Unbiased estimation in convex families. *A. Math. Stat.*, Vol. **40**, n° 5, 1969, p. 1523-1535.
[*6*] N. ARONSZAJN, Theory of reproducing kernels. *Trans. A. M. S.*, t. **68**, 1950, p. 337-404.
[*7*] PETROV, *Sums of independent random variables*. Springer Verlag, 1975.
[*8*] P. A. P. MORAN, *An introduction to Probability theory*. Oxford, 1968.
[*9*] M. LOEVE, *Probability Theory*. Van Nostrand, 1963.
[*10*] M. BERTRAND-RETALI, Convergence uniforme d'un estimateur de la densité par la méthode du noyau. *Revue Roumaine de Mathématiques Pures et Appliquées*, Tome **XXIII**, n° 3, 1978.
[*11*] J. BLEUEZ, *Conditions nécessaires et suffisantes de convergence pour une classe d'estimateurs de la densité*. Thèse de 3ᵉ cycle (Université de Lille I), février 1976.
[*12*] J. NEVEU, *Martingales à temps discret*. Masson, 1972.
[*13*] J. BLEUEZ, D. BOSQ, *Estimation de la densité par la méthode des fonctions orthogonales* (à paraître).
[*14*] D. BOSQ, Contribution à la théorie de l'estimation fonctionnelle (1ʳᵉ partie). *Publ. Inst. Stat. Univ. Paris*, fasc. 2, 1970.
[*15*] J. BLEUEZ, D. BOSQ, *C. R. Acad. Sci.*, t. **282**, sér. A, 1976, p. 63.
[*16*] J. BLEUEZ, D. BOSQ, *C. R. Acad. Sci.*, t. **282**, sér. A, 1976, p. 1023.
[*17*] A. FOLDES, P. REVESZ, A general method for density estimation. *Studia Scientiarum Mathematicarum Hungarica*, t. **9**, 1974, p. 81-92.
[*18*] J. GEFFROY, Sur l'estimation de la densité dans un espace métrique. *C. R. Acad. Sci.*, sér. A, 1974, p. 1449.
[*19*] M. ROSENBLATT, Remark on some non parametric estimates of a density function. *Ann. Math. Stat.*, t. **27**, 1956, p. 832-837.

(Manuscrit reçu le 29 juin 1978)