THÈSE

présentée à

l'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE I

pour obtenir

LE GRADE DE DOCTEUR DE 3ème CYCLE

par

Alain BERLINET

ESPACES AUTOREPRODUISANTS ET MESURE EMPIRIQUE
MÉTHODES SPLINES EN ESTIMATION FONCTIONNELLE

Membres du jury : MM. Pierre POUZET, Président
Denis BOSQ, Rapporteur
Melle Simone MARQUET
M. Pierre JACOB
Examinateurs

Soutenue le 25 avril 1980

Nous remercions très vivement Monsieur le Professeur Pierre POUZET dont nous avons suivi les enseignements de 3ème cycle avec profit et qui nous fait aujourd'hui l'honneur de présider le jury de cette thèse.

Monsieur le Professeur Denis BOSQ a dirigé notre recherche avec compétence et efficacité, nous prodiguant encouragements et conseils judicieux. Qu'il veuille bien trouver ici l'expression de notre profonde reconnaissance.

Mademoiselle le Professeur Simone MARQUET a bien voulu faire partie de la commission d'examen. Nous lui adressons nos vifs remerciements.

Nous remercions vivement Monsieur le Professeur Pierre JACOB qui a accepté de juger notre travail.

Que soient remerciés tous les membres du Séminaire de Statistique de l'U.E.R. de Mathématiques Pures et Appliquées de Lille I qui nous ont aidé par leurs remarques, et plus particulièrement Jacques BLEUEZ avec qui nous avons eu d'intéressantes conversations au sujet de la troisième partie.

Nous remercions enfin Madame Raymonde BÉRAT, Madame Arlette LENGAIGNE et toutes les personnes qui ont participé à la réalisation matérielle de cette thèse pour le soin, la rapidité et la gentillesse dont elles ont fait preuve.

*

* *

S O M M A I R E .

Les premiers noyaux reproduisants ont été étudiés par S. Zaremba dans ses recherches sur le problème de Dirichlet (1907) puis par S. Bergman en théorie des fonctions analytiques (1929 ; voir la bibliographie de [3]). En 1943 N. Aronszahn présentait la théorie des noyaux reproduisants dans toute sa généralité ([2]) et l'approfondissait en 1950 dans [3]. Le formalisme allait encore être étendu en 1964 par L. Schwartz ([45]). Cette théorie a été largement utilisée dans de nombreux domaines et en particulier en probabilité et statistique ([41],[35],[16],[34],[33],[12],[15],[24], etc...). D'autre part l'étude des variables aléatoires à valeurs dans un espace de Hilbert a fait l'objet de nombreux articles et ouvrages sans hypothèse particulière sur l'espace. Or il est fréquent que les espaces utilisés soient à noyau reproduisant. Nous avons donc tenté de voir ce qu'apportait cette particularité pour l'étude des variables aléatoires.

Le chapitre I est consacré aux généralités sur les variables aléatoires à valeurs dans un autoreproduisant : mesurabilité, intégrabilité, convergence faible des lois, opérateur de covariance. Un espace autoreproduisant étant déterminé (à un isomorphisme de Hilbert près) par la donnée de son noyau, la plupart des propriétés données sont très dépendantes du noyau. Nous donnons en particulier des conditions nécessaires et suffisantes de mesurabilité et d'intégrabilité au sens de Pettis.

Dans le chapitre II nous montrons comment on peut considérer la mesure empirique comme une variable aléatoire à valeurs dans un espace autoreproduisant et déduire des propriétés du noyau un théorème central limite pour cette mesure. Dans un cas particulier la mesure empirique est approximée grâce aux résultats obtenus dans [29]. Les hypothèses faites sur l'espace permettent de le plonger dans un espace L^2. Nous étudions les propriétés qui en résultent et nous caractérisons l'image de son noyau de Schwartz.

Nous pensons que ces deux premiers chapitres sont susceptibles de développements ultérieurs, en particulier en théorie des mesures aléatoires (en utilisant les méthodes de [24]) et en statistique fonctionnelle (en liaison avec le dernier chapitre).

Le chapitre III est consacré aux méthodes splines en estimation fonctionnelle (remarquons que ces méthodes mettent souvent en évidence des variables du type de celles étudiées dans les deux premiers chapitres). L'un de nos soucis majeurs, lors de la rédaction de cette thèse a été d'éviter au lecteur des reports constants à la bibliographie ; nous avons donc fait dans chacun des chapitres certains rappels afin de rendre l'ensemble facilement accessible. En particulier dans celui-ci, donnons-nous quelques résultats classiques et leurs démonstrations afin de permettre au lecteur peu familiarisé avec les splines de suivre aisément notre exposé. En effet, après la publication de [4], qui contient une partie des résultats de ce chapitre, de nombreuses questions nous ont été posées au sujet des estimateurs splines.

Ces estimateurs s'adaptent facilement aux différents problèmes d'estimation fonctionnelle (fonction de répartition, densité, régression,...). Après des considérations d'ordre général, nous particularisons notre étude à la fonction de répartition et à la densité. Le lecteur intéressé par le cas de la régression pourra consulter [39]). Nous définissons les noyaux splines et nous montrons comment les méthodes splines s'intègrent à la théorie générale développée dans [15]. Les majorations obtenues pour les noyaux permettent d'obtenir des conditions suffisantes de convergence ponctuelle en moyenne quadratique et presque complète, et de convergence uniforme presque complète dans le cas de subdivisions quelconques. Dans le cas de subdivisions uniformes des conditions nécessaires et suffisantes de convergence simple ou uniforme suivant différents modes stochastiques sont donnés, ainsi qu'un encadrement de la vitesse de convergence.

Nous terminons par quelques commentaires sur l'emploi des méthodes splines en statistique.

CHAPITRE I

GENERALITES.

I - PRELIMINAIRES - NOTATIONS.-

Nous commençons par rappeler quelques résultats de la théorie des noyaux reproduisants introduite dans toute sa généralité par N. Aronszajn en 1943 ([2]) et qui a été l'objet de nombreux travaux ([3], [45], [35], [20],... etc).

E est un ensemble non vide et H un espace de Hilbert séparable de fonctions réelles définies sur E dont le produit scalaire est noté $(.,.)$, la norme $||.||$ et la tribu borélienne $\mathcal{B}_H$.

On suppose que H admet comme noyau reproduisant K. K est une fonction $E \times E \to \mathbb{R}$ telle que

* $\forall t \in E \quad K(.,t) \in H$

* $\forall f \in H \quad \forall t \in E \quad (f, K(.,t)) = f(t)$ (propriété de reproduction)

où la notation $K(.,t)$ désigne la fonction
$$\begin{aligned} E &\longrightarrow \mathbb{R} \\ s &\longmapsto K(s,t) . \end{aligned}$$

Théorème 1.- Un espace de Hilbert H de fonctions réelles (ou complexes) sur E est à noyau reproduisant si et seulement si toutes les fonctionnelles d'évaluation e_t : $H \longrightarrow \mathbb{R}$ (ou $\mathbb{C}$) $t \in E$
$$f \longmapsto f(t)$$
sont continues.

La propriété de reproduction montre que le représentant dans H de l'évaluation e_t est la fonction $K(.,t)$.

Théorème 2.- Une fonction $K : E \times E \longrightarrow \mathbb{R}$ est le noyau reproduisant d'un espace de Hilbert de fonctions réelles sur E si et seulement si elle possède les propriétés suivantes :

* symétrique : $\forall (t,s) \in E^2 \quad K(t,s) = K(s,t)$

* semi-définie positive :

$\forall n \in \mathbb{N}^* \quad \forall (a_1,\ldots,a_n) \in \mathbb{R}^n \quad \forall (t_1,\ldots,t_n) \in E^n$

$$\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_i\, a_j\, K(t_i,t_j) \geq 0 \, .$$

Les démonstrations détaillées figurent dans [20] et mettent en évidence les points suivants :

a) le sous-espace préhilbertien H_o de H engendré par les fonctions $(K(.,t))_{t \in E}$ est dense dans H (autrement dit la famille $(K(.,t))_{t \in E}$ est totale dans H).

b) H est l'ensemble des applications de E dans $\mathbb{R}$ qui sont limites pour la topologie de la convergence simple sur E de suites d'éléments de H_o qui sont de Cauchy.

$\mathcal{T}$ désignera une tribu sur E, $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une application : $\Omega \longrightarrow H$

$$\omega \longmapsto X(\omega,.)$$

Enfin, pour $g \in H$ on notera $(.,g)$ la forme linéaire continue représentée dans H par g.

II - MESURABILITE DES FONCTIONS A VALEURS DANS H.-

1°) *Lemme*. Il existe un sous-ensemble dénombrable D_o de H_o qui est dense dans H.

Démonstration : Notons $(x_p)_{p\in\mathbb{N}}$ une suite dense dans H.

Soit $n \in \mathbb{N}^*$. Comme $\bar{H}_o = H$, pour tout $p \in \mathbb{N}$ il existe $y_p^n \in H_o$ tel que $||y_p^n - x_p|| < \frac{1}{n}$ alors $D_o = \bigcup_{n\in\mathbb{N}^*} \{y_p^n \mid p \in \mathbb{N}\}$ convient :

* $D_o \subset H_o$

* Soit $y \in H$ et $\varepsilon > 0$ il existe $p \in \mathbb{N}$ tel que $||y-x_p|| < \frac{\varepsilon}{2}$

Soit $n > \frac{2}{\varepsilon}$ $\qquad ||x_p - y_p^n|| < \frac{1}{n} < \frac{\varepsilon}{2}$

et $\qquad ||y - y_p^n|| < \varepsilon$.

2°) *Tribu borélienne de* H.

Théorème 1.- La tribu borélienne d'un Hilbert séparable à noyau reproduisant est engendrée par les évaluations.

Démonstration : Soit $\mathcal{E}$ la tribu engendrée par $(e_t)_{t\in E}$.

Les évaluations sont continues sur H donc $\mathcal{E} \subset \mathcal{B}_H$.

D'autre part si $f \in H$ $\quad ||f|| = \sup\limits_{||g||\leq 1} |(f,g)| = \sup\limits_{\substack{||g||\leq 1 \\ g\in D_o}} |(f,g)|$.

Pour $g \in D_o$ la fonction $(.,g)$ est $\mathcal{E} - \mathcal{B}_{\mathbb{R}}$ mesurable puisque g est combinaison linéaire d'évaluations.

Soit $r \in \mathbb{R}^+$ et $f_o \in H$

$$\{f \mid ||f-f_o|| \leq r\} = \bigcap_{\substack{||g||\leq 1 \\ g\in D_o}} \{f \mid |(f-f_o, g)| \leq r\}$$

Le membre de droite est une intersection dénombrable d'éléments de $\mathcal{E}$, c'est donc un élément de $\mathcal{E}$.

Par suite toute boule fermée est dans $\mathcal{E}$.

H étant séparable, tout ouvert est réunion dénombrable de boules fermées et donc $\mathcal{B}_H \subset \mathcal{E}$.

Il est facile de voir que la démonstration précédente est en fait valable pour toute famille totale :

<u>*Théorème*</u> *2.-* Soit $(g_i)_{i\in I}$ une famille totale dans un Hilbert séparable H. La tribu borélienne de H est engendrée par les applications $((.,g_i))_{i\in I}$.

3°) <u>*Condition de Mesurabilité.*</u>

Du théorème 1 on déduit que $\mathcal{B}_H$ est la trace sur H de la tribu produit sur $\mathbb{R}^E$ ([32]) et les corollaires suivants :

<u>*Corollaire*</u> *1.-* $X : (\Omega,A) \longrightarrow (H,\mathcal{B}_H)$ est mesurable si et
$\omega \longmapsto X(\omega,.)$

seulement si pour tout $t \in E$ l'application $X(.,t)$ est une variable aléatoire réelle (v.a.r.).

<u>*Corollaire*</u> *2.-* La donnée d'une fonction aléatoire à valeurs dans $(H,\mathcal{B}_H)$ est équivalente à la donnée d'un processus $(X_t)_{t\in E}$ dont les trajectoires appartiennent à H et qui est défini sur le même espace (Ω,A).

4°) *Lien entre processus gaussien et mesure gaussienne.*

Dans [40] Rajput et Cambanis ont montré l'équivalence entre processus gaussien et mesure gaussienne dans différents espaces fonctionnels (espaces de fonctions continues, absolument continues et espace L_2). Leurs résultats s'étendent aux espaces à noyau reproduisant ([40] , Remarque 1).

Définitions :

* Un processus $(X_t)_{t\in T}$ est dit gaussien si et seulement si toute combinaison linéaire (finie) des $(X_t)_{t\in T}$ est une v.a.r. gaussienne.

* Une mesure de probabilité μ sur $(H,\mathcal{B}_H)$ est dite gaussienne si et seulement si pour tout $g \in H$ la fonction $(.,g)$ est une v.a.r. gaussienne sur $(H,\mathcal{B}_H,\mu)$.

Théorème 3.-

a) Si $(X_t)_{t\in E}$ est un processus gaussien défini sur $(\Omega,\mathcal{A},P)$ et dont les trajectoires appartiennent à H alors la variable aléatoire (v.a.)

$$X : (\Omega,\mathcal{A},P) \longrightarrow (H,\mathcal{B}_H) \quad \text{induit sur } (H,\mathcal{B}_H) \text{ une mesure } PX^{-1} \text{ qui est}$$
$$\omega \longmapsto X.(\omega)$$

gaussienne.

b) Pour toute mesure gaussienne μ sur $(H,\mathcal{B}_H)$ il existe un processus gaussien $(X_t)_{t\in E}$ à trajectoires dans H tel que $PX^{-1} = \mu$.

Démonstration :

a) Il nous faut montrer que la v.a.r.

$$(.,g) : (H,\mathcal{B}_H,\ PX^{-1}) \longrightarrow (\mathbb{R},\mathcal{B}_{\mathbb{R}}) \quad \text{est gaussienne, pour tout } g \in H.$$
$$f \longmapsto (f,g)$$

Soit $B \in \mathcal{B}_{\mathbb{R}}$ $PX^{-1}((.,g) \in B) = P((X,g) \in B)$.

g est limite dans H d'une suite de fonctions

$$g_n = \sum_{i=0}^{j_n} a_i^n K(.,t_i^n) \quad , \quad n \in \mathbb{N}$$

Donc $(X,g) = \lim_{n\to+\infty} \sum_{i=0}^{j_n} a_i^n X_{t_i^n}$ $(X_t)_{t\in E}$ étant gaussien (X,g) est limite partout d'une suite de v.a.r. gaussiennes, (X,g) est donc gaussienne et par conséquent $(.,g)$ l'est aussi.

b) Pour $t \in E$ et $f \in H$ posons

$$X_t(f) = f(t)$$

Soit $k \in \mathbb{N}^*$ $t_1,\dots,t_k$ dans E

$a_1,\dots,a_k$ des réels

$\sum_{i=1}^{k} a_i X_{t_i}$, $(H,\mathcal{B}_H,\mu) \longrightarrow (\mathbb{R},\mathcal{B}_{\mathbb{R}})$ est la forme linéaire continue sur H représentée par $\sum_{i=1}^{k} a_i K(.,t_i)$ donc c'est une v.a.r. gaussienne.

$(X_t)_{t\in E}$ est un processus gaussien à trajectoires dans H tel que $X_\bullet(f) = f$. Si X est la fonction aléatoire qui lui est associée (corollaire 2) on a

$$\forall B \in \mathcal{B}_H \qquad \mu X^{-1}(B) = \mu\{g \mid X_\bullet(g) \in B\} = \mu(B)$$

et $$\mu X^{-1} = \mu .$$

III - <u>*CONVERGENCE FAIBLE SUR*</u> H.-

1°) <u>Définitions</u> ([13], [8], [37]).

On note $Pr(H)$ l'ensemble des probabilités sur $(H,\mathcal{B}_H)$.

a) Soit $\mathbb{P} \in \mathcal{P}r(H)$ et $(\mathbb{P}_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{P}r(H)$.

On dira que $\mathbb{P}_n$ converge faiblement vers $\mathbb{P}$ $(\mathbb{P}_n \underset{n\to+\infty}{\Longrightarrow} \mathbb{P})$ si et seulement si $\int f \, d\mathbb{P}_n \longrightarrow \int f \, d\mathbb{P}$ pour toute fonction numérique f continue bornée sur H.

b) Pour $(t_1,\ldots,t_k) \in E^k$ on note $\pi_{t_1,\ldots,t_k}$ l'application $H \longrightarrow \mathbb{R}^k$

$$f \longmapsto (f(t_1), \ldots, f(t_k))$$

On dit qu'une partie de H est un ensemble de dimension finie (ou cylindre) s'il est de la forme $\pi^{-1}_{t_1,\ldots,t_k}(B)$ $k \in \mathbb{N}^*$ B borélien de $\mathbb{R}^k$.

2°) *Proposition 1.-*

Si deux éléments de $\mathcal{P}r(H)$ coïncident sur la classe $\mathcal{F}$ des ensembles de dimension finie, ils sont égaux (autrement dit $\mathcal{F}$ est une classe déterminante [8]).

Démonstration : Il suffit de montrer que

a) $\mathcal{F}$ est une algèbre de Boole,

et b) $\mathcal{F}$ engendre $\mathcal{B}_H$ (théorème de prolongement).

a) résulte immédiatement de la définition de $\mathcal{F}$.

b) les évaluations étant continues sur H, les applications $\pi_{t_1,\ldots,t_k}$ le sont également et $\mathcal{F} \subset \mathcal{B}_H$ donc $\sigma(\mathcal{F}) \subset \mathcal{B}_H$.

On a vu (II. 2°) Théorème 1) que toute boule fermée est une intersection dénombrable d'ensembles de la forme $\{f \mid |(f-f_o,g)| \leq r\}$ où $g \in D_o$. Un tel ensemble est de dimension finie : g s'écrit sous la forme $\sum_{i=1}^{k} a_i K(.,t_i)$ $a_i \in \mathbb{R}$ $t_i \in E$ pour $1 \leq i \leq k$.

$$\{f \mid |(f-f_o,g)| \leq r\} = \{f \mid |\sum_{i=1}^{k} a_i(f-f_o, K(.,t_i))| \leq r\}$$

$$= \Pi^{-1}_{t_1,\ldots,t_k}(B)$$

où B est le borélien de $\mathbb{R}^k$, image réciproque par l'application mesurable : $\mathbb{R}^k \longrightarrow \mathbb{R}$ de l'intervalle $[0,r]$.

$$(\alpha_1,\ldots,\alpha_k) \longmapsto |\sum_{i=1}^{k} a_i(\alpha_i - f_o(t_i))|$$

Les boules fermées sont donc dans $\sigma(\mathcal{F})$ et $\mathcal{B}_H \subset \sigma(\mathcal{F})$.

3°) *Remarque* :

En général la classe $\mathcal{F}$ ne détermine pas la convergence (au sens de [8]) : on peut avoir $\lim\limits_{n\to+\infty} \mathbb{P}_n(A) = \mathbb{P}(A)$ pour tout $A \in \mathcal{F}$ sans pour autant que la suite $(\mathbb{P}_n)_{n\in\mathbb{N}}$ de $\mathcal{P}r(H)$ tende vers la probabilité $\mathbb{P}$ au sens de la convergence faible.

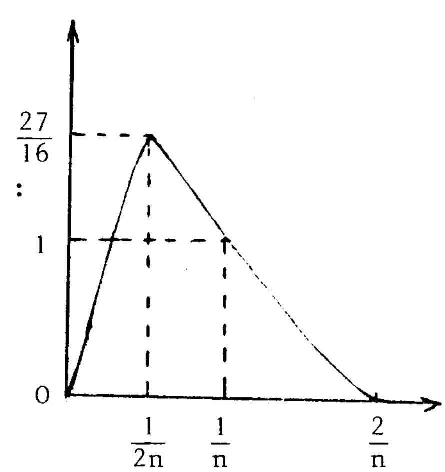
Exemple :

$H = H^2([0,1]) = \{f \mid f$ et f' absolument continues sur $[0,1]$ et $f'' \in L^2([0,1])\}$

$$(f,g) = f(0)\, g(0) + f'(0)\, g'(0) + \int_0^1 f''(x)\, g''(x)\, dx$$

$$K(s,t) = 1 + st + \left(\frac{ts^2}{2} - \frac{s^3}{6}\right) \quad \text{si} \quad s < t$$

$$= 1 + st + \left(\frac{st^2}{2} - \frac{t^3}{6}\right) \quad \text{sinon .}$$



Pour $n \in \mathbb{N}^*$ soit $f_n \in H$ ainsi définie :

sur $[0, \frac{2}{n}]$ $\quad f_n(x) = -\,n^4\, x(x - \frac{2}{n})^3$

sur $[\frac{2}{n}, 1]$ $\quad f_n$ s'annule .

f_n vérifie $f_n(0) = f_n(\frac{2}{n}) = f'_n(\frac{2}{n}) = f''_n(\frac{2}{n}) = 0$ et $f_n(\frac{1}{n}) = 1$.

K étant un noyau borné, la convergence dans H entraîne la convergence uniforme sur $[0,1]$. Cela résulte de l'inégalité de Schwarz : $\forall g \in H \ \forall t \in [0,1] \quad |g(t)| \leq ||g|| \sup_{t \in [0,1]} (K(t,t))^{1/2}$.

Comme $f_n(\frac{1}{n}) = 1$ la suite $(f_n)_{n \in \mathbb{N}^*}$ ne converge pas dans H vers 0 donc $\delta_{f_n} \underset{n\to+\infty}{\not\Longrightarrow} \delta_o$ et pourtant

$$\delta_{f_n}(\pi^{-1}_{t_1,\ldots,t_k}(B)) = \delta_o(\pi^{-1}_{t_1,\ldots,t_k}(B))$$

dès que $\frac{2}{n} \leq \inf\limits_{\substack{1\leq i\leq k \\ t_i>0}} t_i$ si l'un des t_i est non nul ; pour tout n sinon.

4°) *Critère de convergence faible.*

Proposition 2.- Une condition nécessaire et suffisante pour que $\mathbb{P}_n \underset{n\to+\infty}{\Longrightarrow} \mathbb{P}$ est que

a) $(\mathbb{P}_n)_{n\in\mathbb{N}}$ soit équitendue,

et b) $\forall k \in \mathbb{N}^* \ \forall (t_1,\ldots,t_k) \in E^k$

$$\mathbb{P}_n \, \pi^{-1}_{t_1,\ldots,t_k} \underset{n\to+\infty}{\Longrightarrow} \mathbb{P} \, \pi^{-1}_{t_1,\ldots,t_k} .$$

Ce résultat bien connu dans le cas de $C([0,1])$ ([8]) se démontre aisément de la même manière pour H en utilisant les résultats précédents.

a) est équivalente à $(\mathbb{P}_n)_{n\in\mathbb{N}}$ relativement compacte (théorème de Prokhorov). La proposition 2 résulte donc de la continuité des applications $\pi_{t_1,\ldots,t_k}$, du théorème de transfert et de la proposition 1.

Remarque : Il serait intéressant, pour obtenir un critère plus pratique, de caractériser de façon simple les familles de mesures de probabilité équitendues sur H, comme on peut le faire dans le cas des fonctions continues grâce au théorème d'Ascoli.

IV - INTEGRATION DES VARIABLES ALEATOIRES A VALEURS DANS H.-

1°) Notations - Définitions [13], [38], [26]).

X est une v.a. : $(\Omega,\mathcal{A},P) \longrightarrow (H,\mathcal{B}_H)$.

Soit Z une fonction réelle sur Ω et $A \in \mathcal{A}$ tel que $1_A Z$ soit P-intégrable. On note $E_A(Z) = \int 1_A Z \, dP$ et $E(Z) = E_\Omega(Z)$ (si Z est intégrable).

a) *intégrale faible.*

Soit $A \in \mathcal{A}$. X est faiblement intégrable sur A si et seulement si

* $\forall f \in H$ (X,f) est intégrable sur A.

* il existe $x_A \in H$ tel que

$$\forall f \in H \quad E_A((X,f)) = (x_A,f) .$$

x_A est l'intégrale faible de X sur A. On note $x_A = \int_A X dP$

X est Pettis-intégrable si X est faiblement intégrable sur tout élément de $\mathcal{A}$.

b) *intégrale forte.*

Soit $A \in \mathcal{A}$. X est fortement intégrable sur A si et seulement si $||X||$ est intégrable sur A.

X est Bochner-intégrable si X est fortement intégrable sur Ω.

2°) *Conséquences.*

a) Si X est Bochner intégrable, X est fortement intégrable sur tout élément de $\mathcal{A}$.

b) Si X est fortement intégrable sur A il existe $\bar{x}_A \in H$ tel que $\forall f \in H \quad E_A((X,f)) = (\bar{x}_A, f)$.

En effet : $|(X,f)| \leq ||X|| \; ||f||$.

On notera

$$\bar{x}_A = \bar{\int_A} X dP .$$

en particulier X est alors faiblement intégrable sur A et les intégrales coïncident.

c) Comme dans le cas général, l'intégrabilité au sens de Bochner d'une variable aléatoire à valeurs dans H entraîne son intégrabilité au sens de Pettis (les deux intégrales étant égales) sans que la réciproque soit vraie.

Exemple :

$(\Omega,\mathcal{A},P) = (\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*), P) \quad \forall n \in \mathbb{N}^* \qquad P(n) = \dfrac{1}{2^n}$

$H = \ell^2(\mathbb{R})$ si $x \in H \quad x = (x_n)_{n\in\mathbb{N}^*}$

$y \in H \quad y = (y_n)_{n\in\mathbb{N}^*}$

$$(x,y) = \sum_{n=1}^{\infty} x_n \, y_n$$

On notera $x_n = x(n)$.

La famille des évaluations $(e_i)_{i\in\mathbb{N}^*}$ définies par $e_i(n) = \delta_{in}$ est une base hilbertienne de $\ell^2(\mathbb{R})$.

$$\forall x \in \ell^2(\mathbb{R}) \qquad (x,e_i) = x(i)$$

$\ell_2(\mathbb{R})$ admet donc un noyau reproduisant K donné par

$$K(m,n) = \sum_{i=1}^{\infty} e_i(m)\, e_i(n) = \delta_{mn}$$

Soit $X : (\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*), P) \longrightarrow (\ell^2(\mathbb{R})\ ,\ \mathcal{B}_{\ell^2(\mathbb{R})})$

$$i \longmapsto X(i) = \frac{e_i\ s(i)}{P(i)}$$

où s est l'élément de $\ell^2(\mathbb{R})$ tel que

$$s(i) = \frac{1}{i} \quad \forall i \in \mathbb{N}^*$$

* X est mesurable

* $||X(i)|| = \dfrac{1}{iP(i)} = \dfrac{2^i}{i}$; $\displaystyle\sum_{i=1}^{\infty} ||X(i)||\ P(i) = \sum_{i=1}^{+\infty} \frac{1}{i} = +\infty$.

Donc X n'est pas Bochner - intégrable .

* Soit $h \in \ell^2(\mathbb{R})$

$$(h,X(i)) = (h, \frac{e_i}{iP(i)}) = \frac{h(i)}{iP(i)} .$$

Soit $A \subset \mathbb{N}^*$

$$\int_A (h,X(i))\ dP(i) = \sum_{i\in A} \frac{h(i)}{i} = (h,\ s_A)$$

où s_A est la projection orthogonale de s sur le sous-espace de $\ell^2(\mathbb{R})$ engendré par $(e_i)_{i\in A}$.

$$s_A = \sum_{i\in A} (s,e_i)\ e_i = \sum_{i\in A} \frac{e_i}{i}$$

Donc X est Pettis - intégrable.

3°) *Intégrabilité au sens de Bochner.*

Il est évidemment difficile de trouver un critère plus simple que celui donné en IV 1°) b. La proposition suivante, que l'on peut établir en toute généralité dans le cas séparable, et qui n'a rien à voir avec la propriété de reproduction, peut néanmoins, avec ce qui suit donner une idée comparative des différentes notions d'intégrale.

Proposition 1.- Les conditions suivantes sont équivalentes

a) X est Bochner - intégrable.

b) Il existe une v.a.r. X_o intégrable telle que $\forall g \in H_o$ la v.a.r. (X,g) est majorée en valeur absolue par $X_o\ ||g||$.

Démonstration :

(a) ==> (b) $\forall g \in H_o \quad |(X,g)| \leq ||X||\ ||g||$

(b) ==> (a) X est mesurable car les $(X_t)_{t\in E}$ le sont .

$$||X|| = \sup_{\substack{g\in D_o \\ ||g||\leq 1}} |(X,g)| \qquad \text{et } \forall g \in D_o \text{ telle que } ||g|| \leq 1$$

on a
$$|(X,g)| \leq X_o$$

donc $||X||$ est intégrable ([19] tome II p. 123).

4°) *Intégrabilité des* $(X_t)_{t\in E}$.

Si $X : (\Omega,A,P) \longrightarrow (H,\mathcal{B}_H)$ est intégrable (faiblement ou fortement), la v.a.r. X_t est également intégrable pour tout t

et la fonction $\begin{array}{l} E \longrightarrow \mathbb{R} \\ t \longmapsto E(X_t) \end{array}$ est l'intégrale de X sur Ω, en particulier c'est un élément de H.

Réciproquement, si pour tout $t \in E$, X_t est intégrable, X (qui est donc mesurable) n'est pas nécessairement faiblement intégrable : comme le prouve l'exemple suivant il est possible que la fonction $\begin{array}{l} t \longmapsto E(X_t) \\ E \longrightarrow \mathbb{R} \end{array}$ ne soit pas un élément de H.

Exemple : On reprend les notations de l'exemple du IV 2°) c

$$X : (\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*), P) \longrightarrow (\ell^2(\mathbb{R}), \mathcal{B}_{\ell^2(\mathbb{R})})$$

$$i \longmapsto X(i) = \sum_{j=1}^{i} 2^{(i+j)/2} e_j$$

$\forall n \in \mathbb{N}^*$ $P(n) = \dfrac{1}{2^n}$.

$$\text{Soit } j \in \mathbb{N}^* \quad E((X,e_j)) = \sum_{i=1}^{+\infty} \frac{(X(i), e_j)}{2^i} = \sum_{i \geq j} \frac{2^{(i+j)/2}}{2^i}$$

$$E((X,e_j)) = 2 + \sqrt{2} .$$

La suite constante $(E((X,e_j)))_{j \in \mathbb{N}^*}$ n'est évidemment pas élément de $\ell^2(\mathbb{R})$ et X n'est pas faiblement intégrable sur $\mathbb{N}^*$.

Proposition 2.- Pour tout $A \in \mathcal{A}$ les conditions suivantes sont équivalentes :

a) l'application $\mathcal{E}_A : \begin{array}{l} t \longmapsto E_A X_t \\ E \longrightarrow \mathbb{R} \end{array}$ est un élément de H

b) l'application $\varphi_A : \begin{array}{l} f \longmapsto E_A(X,f) \\ H_o \longrightarrow \mathbb{R} \end{array}$ est une forme linéaire continue sur H_o.

Si elles sont satisfaites $\mathcal{E}_A$ est le représentant dans H de φ_A.

Démonstration :

a) $\Longrightarrow$ b) Soit $f \in H_o$ $\quad f = \sum_{i=1}^{n} a_i \, K(.,t_i)$

$$(X,f) = \sum_{i=1}^{n} a_i \, X_{t_i} \qquad E_A(X,f) = \sum_{i=1}^{n} a_i \, E_A \, X_{t_i} = (\mathcal{E}_A,f)$$

b) $\Longrightarrow$ a) Il résulte du théorème de Hahn-Banach que φ_A se prolonge en une forme linéaire continue Ψ_A sur H de même norme que φ_A et il existe $f_A \in H$ tel que

$$\forall f \in H \qquad \Psi_A(f) = (f_A,f)$$

on a alors

$$f_A(t) = (f_A,K(.,t)) = \Psi_A(K(.,t)) = \varphi_A(K(.,t)) = E_A \, X_t$$

et $\qquad\qquad \mathcal{E}_A = f_A \quad .$

Proposition 3.- Soit $A \in \mathcal{A}$ tel que pour tout t $E_A \, X_t$ existe.

Soit $\quad I_A : \begin{array}{l} H_o \longrightarrow \mathbb{R}^+ \\ f \longmapsto \int_A |(X,f)| \, dP \end{array} \quad .$

Les 3 conditions suivantes sont équivalentes :

a) I_A est continue en 0,

b) I_A est continue.

c) I_A est lipschitzienne.

Démonstration : I_A est bien définie : si $f = \sum_{i=1}^{n} a_i K(.,t_i)$ $(X,f) = \sum_{i=1}^{n} a_i X_{t_i}$ donc (X,f) est intégrable sur A.

On a trivialement c) $\Longrightarrow$ b) $\Longrightarrow$ a) .

I_A étant positivement homogène, sa continuité en 0 entraîne l'existence d'une constante C telle que

$$I_A(f) \leq C \, ||f|| \qquad \forall f \in H_o$$

$\forall f \in H_o \quad \forall g \in H_o$

$$\begin{aligned} |I_A(f) - I_A(g)| &= |E_A(|(X,f)| - |(X,g)|)| \\ &\leq E_A \Big| |(X,f)| - |(X,g)| \Big| \\ &\leq E_A \, |(X,f-g)| \leq C \, ||f-g||. \end{aligned}$$

et donc a) $\Longrightarrow$ c).

5°) *Intégrabilité au sens de Pettis.*

Proposition 1.- Soit $A \in \mathcal{A}$.

Si $E_A : E \longrightarrow \mathbb{R}$ est un élément de H et si l'une des

$$t \longmapsto E_A X_t$$

conditions équivalentes de la proposition précédente est vérifiée X est faiblement intégrable sur A.

Démonstration :

Soit $f \in H$. Il nous faut montrer que (X,f) est intégrable sur A et que $E_A(X,f) = (\mathcal{E}_A, f)$.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de H_o convergeant dans H vers f. $(X,f_n) \longrightarrow (X,f)$ partout.

La suite $((X,f_n))_{n\in\mathbb{N}^*}$ est de Cauchy dans $L_1(A)$:

$$\int_A |(X,f_n) - (X,f_m)| = \mathcal{I}_A(f_n - f_m) \leq C \, ||f_n - f_m||$$

$((X,f_n))_{n\in\mathbb{N}^*}$ converge donc dans $L_1(A)$ vers une v.a. presque sûrement égale à (X,f).

(X,f) est donc intégrable sur A et $E_A(X,f) = \lim E_A(X,f_n)$.

D'après IV 3°) proposition 2 $E_A(X,f_n) = (\mathcal{E}_A, f_n)$

et

$$E_A(X,f) = \lim_n E_A(X,f_n) = \lim_n (\mathcal{E}_A, f_n) = (\mathcal{E}_A, f) \quad .$$

Définition.- On dit qu'une fonction d'ensembles

$$\theta : \begin{array}{l} \mathcal{A} \longrightarrow H \\ A \longmapsto \theta(A) \end{array}$$

est absolument continue sur $(\Omega, \mathcal{A}, P)$ si et seulement si

$$\forall \varepsilon > 0 \ \exists \ \eta_\varepsilon \text{ tel que } (P(A) < \eta_\varepsilon \Longrightarrow ||\theta(A)|| < \varepsilon) \quad .$$

Théorème 1.- Si $X : (\Omega, \mathcal{A}, P) \longrightarrow (H, \mathcal{B}_H)$ est Pettis - intégrable (a fortiori si X est Bochner - intégrable) la fonction d'ensembles

$$\begin{aligned} \mathcal{A} &\longrightarrow H \\ A &\longmapsto x_A = \int_A X dP \end{aligned}$$ est absolument continue sur $(\Omega,\mathcal{A},P)$.

Démonstration :

voir [33] p. 281-284

ou [22] p. 75-78 .

Théorème 2.- (Condition nécessaire et suffisante de Pettis - intégrabilité.

$$X : (\Omega,\mathcal{A},P) \longrightarrow (H,\mathcal{B}_H) .$$

X est Pettis - intégrable si et seulement si les deux conditions suivantes sont vérifiées.

a) $\forall A \in \mathcal{A}$ l'application $\begin{aligned} \mathcal{E}_A : t &\longmapsto E_A X_t \\ E &\longrightarrow \mathbb{R} \end{aligned}$ est un élément de H.

b) La fonction d'ensembles $\begin{aligned} \Omega &\longrightarrow H \\ A &\longmapsto \mathcal{E}_A \end{aligned}$ est absolument continue sur $(\Omega,\mathcal{A},P)$.

Démonstration :

* *si* X *est Pettis - intégrable.*

$$\forall A \in \mathcal{A} \quad \forall t \in E \quad \mathcal{E}_A(t) = \int_A (X,K(.,t))\, dP = x_A(t)$$

donc a) est vérifiée. b) l'est aussi en vertu du théorème 1 ci-dessus.

* *Réciproque.* X est mesurable car les $(X_t)_{t\in E}$ le sont.

Nous allons montrer que pour tout $A \in \mathcal{A}$ l'application I_A est continue en 0. L'intégrabilité de X au sens de Pettis résultera alors de IV 5°) Proposition 1.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de H_o convergeant fortement vers 0.

$(X,f_n) \xrightarrow[n\to+\infty]{} 0$ partout, donc aussi en probabilité.

Soit $\varepsilon > 0$. Il existe $\eta_1(\varepsilon)$ tel que

$\forall B \in \mathcal{A} \quad (P(B) < \eta_1(\varepsilon) \Longrightarrow ||E_B|| < \varepsilon^{1/2})$.

Il existe $N(\varepsilon)$ tel que

$$\bullet \quad n \geq N(\varepsilon) \Longrightarrow \begin{cases} P(|(X,f_n)| > \frac{\varepsilon}{2}) < \eta_1(\varepsilon) \\ \\ \text{et} \quad ||f_n|| < \dfrac{\varepsilon^{1/2}}{4} \end{cases}$$

$$\forall n \in \mathbb{N} \quad I_A(f_n) = \int_A |(X,f_n)| \, dP \leq \int_{A \,\cap\, \{|(X,f_n)|>\frac{\varepsilon}{2}\}} |(X,f_n)| \, dP + \frac{\varepsilon}{2}$$

Soit $\quad A_n^+(\varepsilon) = A \cap \{|(X,f_n)| > \frac{\varepsilon}{2}\} \cap \{(X,f_n) \geq 0\}$

et $\quad A_n^-(\varepsilon) = A \cap \{|(X,f_n)| > \frac{\varepsilon}{2}\} \cap \{(X,f_n) < 0\}$

$$\forall n \in \mathbb{N} \qquad I_A(f_n) \leq \int_{A_n^+(\varepsilon)} (X,f_n) \, dP - \int_{A_n^-(\varepsilon)} (X,f_n) \, dP + \frac{\varepsilon}{2}$$

$$\forall n \in \mathbb{N} \qquad I_A(f_n) \leq ||E_{A_n^+(\varepsilon)}|| \; ||f_n|| + ||E_{A_n^-(\varepsilon)}|| \; ||f_n|| + \frac{\varepsilon}{2}$$

Pour $n \geq N(\varepsilon)$ $\quad ||E_{A_n^+(\varepsilon)}|| < \varepsilon^{1/2}$

$$||E_{A_n^-(\varepsilon)}|| < \varepsilon^{1/2}$$

$$||f_n|| < \frac{\varepsilon^{1/2}}{4}$$

et $\qquad I_A(f_n) \leq \varepsilon$

Remarque au sujet de la condition a).

On trouvera dans [20] des conditions nécessaires et suffisantes pour qu'une fonction réelle définie sur E soit un élément de H.

V - OPERATEUR DE COVARIANCE.-

1°) *Opérateurs de* H.

Proposition : A tout opérateur (endomorphisme continu) L de H est associée une application $\Lambda : E \times E \longrightarrow \mathbb{R}$ telle que :

$$\forall x \in E \qquad \Lambda(.,x) \in H \qquad \text{et}$$

$$\forall f \in H \qquad (Lf)(x) = (f, \Lambda(.,x)) \ .$$

Λ est appelé le noyau de L.

Démonstration : L^* désigne l'adjoint de L ([46]).

On pose pour tout $x \in E$ $\Lambda(.,x) = L^*(K(.,x))$

on a alors, de par les propriétés de l'adjoint :

$$\forall f \in H \quad \forall x \in E \quad (f,\Lambda(.,x)) = (f, L^*(K.,x))) = (Lf,K(.,x)) = (Lf)(x)$$

2°) *Noyau d'un opérateur de covariance.*

Définition : Soit $X : (\Omega,A,P) \longrightarrow (H,\mathcal{B}_H)$ une v.a. du second ordre (i.e. $||X|| \in L^2(P)$).

L'opérateur de covariance de X est l'unique opérateur C_X défini par

$$\forall f \in H \quad \forall g \in H \qquad (C_X f,g) = E((X,f)\ (X,g)) \ .$$

C_X est un opérateur compact symétrique et positif.

Proposition : L'opérateur de covariance de X a pour noyau N_X la fonction : $E \times E \longrightarrow \mathbb{R}$ (fonction des "moments"

$$(t,s) \longmapsto E(X_t \; X_s)$$

ou fonction de covariance dans le cas centré du processus $(X_t)_{t \in E}$ associé).

Démonstration : Soit $(t,s) \in E^2$

$$N_X(t,s) = C_X(K(.,s))\;(t) = (C_X K(.,s),\; K(.,t))$$

$$N_X(t,s) = E((X,\; K(.,s))\;(X, K(.,t))) = E(X_s \; X_t) \; .$$

On notera encore que pour tout $t \in E$ la fonction

$E(X_t \; X_.) : E \longrightarrow \mathbb{R}$ est un élément de H : c'est l'image

$$s \longmapsto E(X_t \; X_s)$$

par C_X de $K(.,t)$ et l'on a

$$\forall f \in H \quad \forall t \in E \quad C_X f(t) = (f,\; E(X_t X_.)) \; .$$

Remarque : Les noyaux des opérateurs de covariance ont bien sûr la même propriété d'additivité que les opérateurs dans le cas de variables du second ordre, centrées et indépendantes.

CHAPITRE II

MESURE EMPIRIQUE.

I - LA MESURE EMPIRIQUE COMME V.A. A VALEURS DANS H.-

1°) *Introduction.*

Soit $Y : (\Omega, \mathcal{A}, P) \longrightarrow (E, \mathcal{T})$ une variable aléatoire (v.a.) de loi μ sur E.

Pour tout $t \in E$ la mesure de Dirac δ_t définit une forme linéaire continue sur H : $f \longmapsto \int f \, d\delta_t = f(t)$ qui est l'évaluation au point t : e_t représentée dans H par la fonction $K(.,t)$.

La variable $K(.,Y)$ est ainsi le représentant dans H de la mesure de Dirac aléatoire δ_Y.

Supposons donné un échantillon $(Y_1,\ldots,Y_n)$ de taille n de $Y(Y_1,\ldots,Y_n$ sont des variables aléatoires $(\Omega,\mathcal{A},P) \longrightarrow (E,\mathcal{T})$ indépendantes et de loi μ sur E). La mesure empirique associée à cet échantillon $\frac{1}{n}\sum_{k=1}^{n} \delta_{Y_k}$ admet pour représentant dans H $\frac{1}{n}\sum_{k=1}^{n} K(.,Y_k)$.

Du corollaire 1 (chapitre I § II 3°) il résulte que l'application $(\Omega,\mathcal{A},P) \longrightarrow (H,\mathcal{B}_H)$ est mesurable pour

$$\omega \longmapsto \frac{1}{n}\sum_{k=1}^{n} K(.,Y_k(\omega))$$

tout $n \in \mathbb{N}^*$ si et seulement si pour tout $t \in E$ $K(t,Y)$ est une v.a.r.

Lorsque cette condition est remplie, on peut donc considérer la mesure empirique comme une v.a. à valeurs dans H. C'est le cas en particulier si l'application

$$F_K : (E,\mathcal{T}) \longrightarrow (H,\mathcal{B}_H) \quad \text{est mesurable.}$$
$$t \longmapsto K(.,t) .$$

2°) *Mesurabilité du noyau reproduisant.*

Proposition.- Les trois conditions suivantes sont équivalentes :

a) K est $(\mathcal{T}^2 - \mathcal{B}_{\mathbb{R}})$ mesurable.

b) F_K est mesurable.

c) $\forall t \in E$ $K(.,t)$ est mesurable.

Si l'une d'elles est vérifiée, H est constituée de fonctions mesurables.

Démonstration :

a) $\Longrightarrow$ b) . Il suffit de montrer que $\forall s \in E$ $t \longmapsto (K(.,t),K(.,s))$ est mesurable, ce qui résulte du fait que cette fonction est la section de $K(s,t)$ en s.

b) $\Longrightarrow$ c) . immédiat .

c) $\Longrightarrow$ a) . La famille $(K(.,t))_{t\in E}$ étant totale dans H et constituée de fonctions mesurables, toute fonction de H est mesurable puisque la convergence dans H entraîne la convergence partout dans E.

Soit $(e_i)_{i\in\mathbb{N}}$ une base orthonormale de H.

La propriété de reproduction entraîne que $K(.,t) = \sum_{i\in\mathbb{N}} e_i(t)\, e_i$ (convergence dans H) et donc que :

$\forall (s,t) \in E^2$ $K(s,t) = \sum_{i\in\mathbb{N}} e_i(s)\, e_i(t)$ (convergence simple dans E^2).

Les $(e_i)_{i\in\mathbb{N}}$ étant mesurables K est $(\mathcal{T}^2 - \mathcal{B}_{\mathbb{R}})$ mesurable.

3°) *Noyaux mesurables tels que* F_K *soit du second ordre.*

Dans la suite de ce paragraphe I, nous ferons les hypothèses suivantes sur K, H et μ :

H_1 K est $(\mathcal{T}^2 - \mathcal{B}_{\mathbb{R}})$ mesurable

H_2 la fonction $\begin{array}{l} E \longrightarrow \mathbb{R}^+ \\ x \longmapsto K(x,x) \end{array}$ appartient à $L_1(\mu)$

H_3 la fonction nulle de H est la seule fonction de H qui est nulle μ-presque partout.

Comme $F_K(x) = K(.,x)$

$||F_K(x)||^2 = K(x,x)$ et l'hypothèse H_2 s'écrit encore :

$$\int K(x,x)\ d\mu(x) = \int ||F_K(x)||^2\ d\mu(x) = \int ||K(.,Y)||^2\ dP < +\infty.$$

La norme et le produit scalaire dans un espace H' différent de H seront notés $||\ \ ||_{H'}$ et $(.,.)_{H'}$.

Proposition 1.- H est un sous-espace hilbertien de $L^2(\mu)$ ([45]) ce qui signifie que H est inclus dans $L^2(\mu)$ et que l'injection canonique est continue.

Démonstration : Soit $g \in H$

$$|g(x)|^2 = |(g, K(.,x))|^2 \leq ||g||^2 K(x,x)$$

donc $g \in L^2(\mu)$ et on a

$$||g||_{L^2(\mu)} \leq ||g|| \left(\int K(x,x)\ d\mu(x)\right)^{1/2}.$$

L'hypothèse H_3 permet de mettre en correspondance bijective l'espace H et le sous-espace H' de $L^2(\mu)$ constitué des classes d'équivalence des fonctions de H pour la relation d'égalité μ-presque partout.

Nous noterons encore H ce sous-espace H' muni de la topologie transportée de H sur H' .

Proposition 2.- Pour tout $g \in L^2(\mu)$, g finie partout, l'application $\begin{array}{l} H \longrightarrow \mathbb{R} \\ f \longmapsto \int fg\, d\mu \end{array}$ est une forme linéaire continue sur H. Elle est représentée dans H par

$$I_{g.\mu} = \bar{\int} K(.,x)\ g(x)\ d\mu(x)$$

La notation $g.\mu$ désigne la mesure de densité g par rapport à μ.

Démonstration : Soit $g \in L^2(\mu)$, finie partout.

La fonction $\begin{array}{l} E \longrightarrow H \\ x \longmapsto g(x)\ K(.,x) \end{array}$ est mesurable ([19] p. 218 T 2) et Bochner intégrable d'après H_2.

Par conséquent $\bar{\int} K(.,x)\ g(x)\ d\mu(x)$ existe et est un élément de H. De plus les propriétés de l'intégrale de Bochner entraînent que :

$$\forall f \in H \quad (f, \bar{\int} K(.,x)\ g(x)\ d\mu(x)) = \int (f, K(.,x)\ g(x))\ d\mu(x) = \int fg\ d\mu.$$

Remarque 1.- Si ℓ est une forme linéaire continue sur H, ℓ est représentée dans H par la fonction $t \longmapsto \ell(K(.,t))$.

En effet, si v_ℓ représente ℓ on a

$$\forall t \in E \quad v_\ell(t) = (v_\ell, K(.,t)) = \ell(K(.,t)) \ .$$

Soit $g \in L^2(\mu)$. La continuité de $\begin{array}{c} f \longmapsto \int fg\, d\mu = (f,g)_{L^2(\mu)} \\ H \longrightarrow \mathbb{R} \end{array}$

résultant de la proposition 1, la remarque précédente permet de conclure que cette forme linéaire continue est représentée dans H par $t \longmapsto \int K(t,x)\, g(x)\, d\mu(x)$ et donc d'affirmer (seulement) que $\int K(.,x)\, g(x)\, d\mu(x)$ existe au sens faible.

Proposition 3.- L'application $L^2(\mu) \longrightarrow L^2(\mu)$

$$N : g \longmapsto I_{g.\mu} = \int K(.,x)\, g'(x)\, d\mu(x)$$

où g' est finie et appartient à la classe g est le noyau de Schwartz ([45]) de H en tant que sous-espace hilbertien de $L^2(\mu)$.

Démonstration : Le noyau L de H est l'unique opérateur linéaire continu symétrique positif de $L^2(\mu)$ dans $L^2(\mu)$ caractérisé en tant qu'application L par :

$$\forall g \in L^2(\mu) \quad \forall f \in H \quad (f,Lg) = (f,g)_{L^2(\mu)} \ .$$

Donc $\qquad \forall g \in L^2(\mu) \ Lg = I_{g.\mu}$

Remarque 2.- Si $g \in H$ $I_{g.\mu}$ est l'image de g par $C_{K(.,Y)}$, opérateur de covariance de $K(.,Y)$ considéré comme v.a. à valeurs dans H.

En effet :

$$\forall (f,g) \in H^2 \quad (C_{K(.,Y)}g,f) = E(f(Y)\ g(Y)) = \int fg\ d\mu = (I_{g.\mu},\ f)$$

Autrement dit l'opérateur de covariance de $K(.,Y)$ est la restriction à H du noyau de Schwartz de H.

Remarque 3.- Dans le cas particulier où $(f,g) = (f,g)_{L^2(\mu)}$ $\forall (f,g) \in H^2$, $I_{g.\mu}$ est bien sûr la projection orthogonale de g sur H. (exemple : estimation par la méthode des fonctions orthogonales).

Proposition 4.- ([24] théorème de plongement 2CI).

Si K est borné la relation $\mu \longmapsto I_\mu = I_{1.\mu} = \int K(.,x)\ d\mu(x)$ définit une application linéaire de M_E : espace des mesures signées bornées ([25] , [42]) sur E, dans H.

Démonstration : K étant borné, l'hypothèse H_2 est vérifiée pour tout $\mu \in M_E$. La proposition 2 entraîne que

$\bar{I}_\mu = \int K(.,x)\ d\mu(x)$ est le représentant dans H de la forme linéaire continue $f \longmapsto \int f\ d\mu$.

Proposition 5.- K est $\mu \otimes \mu$ intégrable et

$$\int K(x,y)\ d\mu(x)\ d\mu(y) = ||I_\mu||^2$$

Démonstration :

Si $(x,y) \in E^2$ $|K(x,y)| \leq ||K(.,x)||\ ||K(.,y)||$

D'où l'intégrabilité de K.

$$||I_\mu||^2 = (\overline{\int} K(.,x)\, d\mu(x), \overline{\int} K(.,y)\, d\mu(y)) = \int K\, d\mu \otimes \mu$$

de par les propriétés de l'intégrale et du produit scalaire et le théorème de Fubini.

4°) *Estimation de* $I_\mu = \overline{\int} K(.,x)\, d\mu(x)$.-

On suppose que la loi μ est inconnue et que pour tout $n \in \mathbb{N}^*$ on dispose d'un échantillon $(Y_1,\ldots,Y_n)$ de taille n de Y.

Les variables aléatoires $K(.,Y_k) : (\Omega,A,P) \longrightarrow (H,\mathcal{B}_H)$ sont intégrables, indépendantes et de même loi. D'après la loi des grands nombres, si l'on pose $S_n = \sum_{k=1}^{n} K(.,Y_k)$ on a $\frac{S_n}{n} \xrightarrow[n\to+\infty]{} I_\mu$ presque sûrement puisque

$$I_\mu = \overline{\int} K(.,Y_k)\, dP = E(K(.,Y_k)) .$$

Pour $n \in \mathbb{N}^*$ soit $\Lambda_n : (\Omega,A,P) \longrightarrow (H,\mathcal{B}_H)$

$$\omega \longrightarrow \Lambda_n(\omega) = \sqrt{n}\left(\frac{S_n(\omega)}{n} - I_\mu\right) .$$

Il est possible de montrer, en utilisant le théorème central limite dans un Hilbert, que la suite $(\Lambda_n)_{n\in\mathbb{N}^*}$ converge faiblement en loi vers une gaussienne centrée ([7]).

Ce résultat peut être démontré facilement sans utiliser ce théorème : c'est l'objet de la suite de ce paragraphe qui généralise donc les résultats de [28], obtenus pour des espaces de distributions sur le tore T^2.

Pour $n \in \mathbb{N}^*$ on notera λ_n la loi de Λ_n sur H.

Proposition 1.- Λ_n est du second ordre et

$$E(||\Lambda_n||^2) = \int K(x,x)\, d\mu(x) - \int K(x,y)\, d\mu(x)\, d\mu(y) \quad .$$

Démonstration : $\Lambda_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (K(.,Y_k) - I_\mu)$.

$$||\Lambda_n||^2 = \frac{1}{n} \sum_{k \neq \ell} (K(.,Y_k) - I_\mu \,,\, K(.,Y_\ell) - I_\mu)$$

$$+ \frac{1}{n} \sum_{k=1}^{n} ||K(.,Y_k) - I_\mu||^2$$

$$(K(.,Y_k) - I_\mu \,,\, K(.,Y_\ell) - I_\mu) = K(Y_k,Y_\ell) - I_\mu(Y_k) - I_\mu(Y_\ell) + ||I_\mu||^2$$

Or $$||I_\mu||^2 = \int K(x,y)\, d\mu(x)\, d\mu(y) = E(K(Y_k,Y_\ell))$$

et $$||I_\mu||^2 = (I_\mu,I_\mu) = \int I_\mu\, d\mu = E(I_\mu(Y)) \quad .$$

Donc Λ_n est du second ordre et

$$E(||\Lambda_n||^2) = E(||K(.,Y) - I_\mu||^2) = E(K(Y,Y)) - ||I_\mu||^2 \quad .$$

Proposition 2.- La suite $(\lambda_n)_{n \in \mathbb{N}^*}$ est relativement compacte dans Pr(H) : espace des mesures de probabilité sur H muni de la topologie de la convergence faible.

Démonstration : Dans le cas où H est de dimension finie, la conclusion résulte de la proposition 1 et de l'inégalité de Tchebychev :

$$\forall R > 0 \quad P(||\Lambda_n|| \geq R) \leq \frac{E(||\Lambda_n||^2)}{R^2}$$

Soit $\varepsilon > 0$. Pour R assez grand la boule fermée $B^*(0,R)$ est donc un compact de λ_n-mesure supérieure à $1-\varepsilon$, pour tout $n \in \mathbb{N}^*$.

Dans le cas où H est de dimension infinie, il suffit de montrer (théorème 2.2 p. 154 de [37]) que

$\sup_{n\in\mathbb{N}^*} \int r_N(x)\, d\,\lambda_n(x) \xrightarrow[N\to+\infty]{} 0$ où r_N est défini comme suit : une base orthonormale $(e_i)_{i\in\mathbb{N}}$ étant fixée dans H

$$r_N(x) = \sum_{i=N}^{\infty} (x,e_i)^2 \quad , \quad \forall x \in H .$$

Soit $n \in \mathbb{N}^*$. Λ_n étant du second ordre $(\Lambda_n,e_i)^2$ est intégrable par rapport à P pour tout i et

$$\sum_{i=0}^{+\infty} E(\Lambda_n,e_i)^2 = E(||\Lambda_n||^2) \qquad (1)$$

$$(\Lambda_n,e_i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (e_i(Y_k) - (I_\mu,e_i)) .$$

si $k \neq \ell$ $(e_i(Y_k) - (I_\mu,e_i))\,(e_i(Y_\ell) - (I_\mu,e_i))$ a une espérance nulle donc

$$\int (\Lambda_n,e_i)^2\, dP = \int [e_i(Y) - (I_\mu,\, e_i)]^2\, dP = \mathrm{Var}(e_i(Y))$$

$$\int (\Lambda_n,e_i)^2\, dP = E(e_i(Y))^2 - (I_\mu,e_i)^2$$

quantité qui est indépendante de n.

$E(r_N(\Lambda_n))$ est donc indépendant de n et tend vers 0 avec N puisque c'est le reste de rang N de la série en (1).

Proposition 3.- La suite $(\lambda_n)_{n\in\mathbb{N}^*}$ des lois des variables aléatoires $(\Lambda_n)_{n\in\mathbb{N}^*}$ converge faiblement vers une loi gaussienne centrée sur H, de fonction de covariance donnée par :

$$C(f,g) = \int fg\, d\mu - \int f\, d\mu \times \int g\, d\mu \quad \forall (f,g) \in H^2$$

Démonstration : Pour $n \in \mathbb{N}^*$ la fonction caractéristique de λ_n est notée $\hat{\lambda}_n$ ([23], [37])

$$\forall g \in H \qquad \hat{\lambda}_n(g) = \int e^{i(f,g)}\, d\lambda_n(f)$$

$$\hat{\lambda}_n(g) = E(e^{i(\Lambda_n,g)}) = E\left(e^{\frac{i}{\sqrt{n}}\left[\sum_{k=1}^{n} g(Y_k) - (I_\mu,g)\right]}\right)$$

$$\hat{\lambda}_n(g) = \left[E\left(e^{\frac{i}{\sqrt{n}}\left[g(Y) - (I_\mu,g)\right]}\right)\right]^n$$

Notons φ_{Z_g} la fonction caractéristique de la variable aléatoire réelle $Z_g = g(Y) - (I_\mu,g)$.

$$E(Z_g) = 0 \quad \text{puisque} \quad E(g(Y)) = \int g\, d\mu = (I_\mu,g)$$

$$E(Z_g^2) = \mathrm{Var}(g(Y)) = E(g(Y)^2) - (I_\mu,g)^2 .$$

Un développement limité au point zéro donne donc :

$$\varphi_{Z_g}(t) = 1 - \frac{\mathrm{Var}(g(Y))}{2}\, t^2 + o(t^2)$$

$$\hat{\lambda}_n(g) = \left[\varphi_{Z_g}\left(\frac{1}{\sqrt{n}}\right)\right]^n = \left(1 - \frac{\mathrm{Var}(g(Y))}{2n}\right)^n + o\left(\frac{1}{n}\right)$$

et $$\lim_{n\to+\infty} \hat{\lambda}_n(g) = \exp\left(- \frac{\mathrm{Var}\ g(Y)}{2}\right)$$

Du lemme 2.1. p. 153 de [37] on déduit alors qu'il existe $\lambda \in \mathcal{P}r(H)$ telle que :

$$\lambda_n \xrightarrow[n\to+\infty]{} \lambda \quad \text{faiblement} .$$

$\forall g \in H \quad \hat{\lambda}(g) = \exp(-\frac{1}{2} \operatorname{Var} g(Y))$ ce qui montre que λ est gaussienne centrée sur H.

Soit C sa fonction de covariance et S son opérateur de covariance.

$$\forall f \in H \quad \forall g \in H \qquad C(f,g) = (Sf,g)$$

$$\forall g \in H \qquad (Sg,g) = \operatorname{Var} g(Y) .$$

Soit $(f,g) \in H^2$

$$(Sf,g) = \frac{1}{2}\left[(S(f+g),\ f+g) - (Sf,f) - (Sg,g)\right]$$

$$C(f,g) = (Sf,g) = \frac{1}{2}\left[\operatorname{Var}\ (f+g)(Y) - \operatorname{Var} f(Y) - \operatorname{Var} g(Y)\right]$$

$$C(f,g) = (Sf,g) = \operatorname{Cov}(f(Y),\ g(Y)) = E(fg(Y)) - E(f(Y)) \times E(g(Y))$$

Evidemment les opérateurs S, C_{Λ_n} et $C_{(K(.,Y)-I_\mu)}$ sont égaux. Leur noyau est la fonction

$$E \times E \longrightarrow \mathbb{R}$$

$$(s,t) \longmapsto C(K(.,t),K(.,s)) = E\left[K(t,Y)\ K(s,Y)\right] - I_\mu(t)\ I_\mu(s)$$

$$= \int K(t,x)\ K(s,x)\ d\mu(x) - \int K(t,x)\ d\mu(x) \times \int K(s,x)\ d\mu(x) .$$

5°) *Caractérisation de l'image de N, noyau de Schwartz de* H.

Tout élément de $\bar{H}_o$ s'écrit sous la forme $\sum_{i=1}^{k} a_i K(.,t_i)$ soit encore $\int K(.,t)\, d\mu_o(t)$ où $\mu_o = \sum_{i=1}^{k} a_i \delta_{t_i}$.

Le sous-espace vectoriel de H constitué des élément de la forme $\int K(.,t)\, d\nu(t)$ où ν est une mesure bornée sur T est donc dense dans H. Le but de ce paragraphe est de caractériser une partie de ces éléments qui constituent l'image de $N : L^2(\mu) \longrightarrow L^2(\mu)$

$$g \longmapsto \bar{\int} K(.,t)\, g(t)\, d\mu(t) \quad .$$

(On note par la même lettre g un élément partout fini de $L^2(\mu)$ et sa classe dans $L^2(\mu)$).

Proposition 1 (*L. Schwartz* [45]).-

L'image de N est un sous-espace dense de H.

Un élément f de H appartient à Im N si et seulement si la forme linéaire $\gamma_f : k \longmapsto (k,f)$ est continue sur H pour la topologie induite par celle de $L^2(\mu)$.

Démonstration : Soit h_o un élément de H orthogonal à Im N

$$\forall g \in L^2(\mu) \qquad (h_o, \bar{\int} K(.,t)\, g(t)\, d\mu(t)) = 0$$

$$\forall g \in L^2(\mu) \qquad (h_o, g)_{L^2(\mu)} = 0$$

donc $h_o = 0$ et Im N est dense dans H.

Soit $f \in$ Im N $f = N(g)$ où $g \in L^2(\mu)$.

$\forall k \in H \quad (k,f) = (k, N(g)) = (k,g)_{L^2(\mu)}$ d'où la continuité de γ_f pour la topologie induite.

Réciproquement si γ_f est continue pour cette topologie sur H, on peut la prolonger en une forme linéaire continue sur $L^2(\mu)$, d'après le théorème de Hahn-Banach.

Il existe donc $g \in L^2(\mu)$ telle que :

$$\forall k \in H \quad (k,f) = (k,g)_{L^2(\mu)} = (k, \mathcal{N}(g))$$

et $$\mathcal{N}(g) = f$$

Hypothèses complémentaires.

On suppose que E est un espace topologique compact, que $\mathcal{T}$ est sa tribu borélienne et que K est continu sur E^2.

μ est une mesure positive bornée sur E telle que $\mathrm{Supp}\ \mu = E$.

Les hypothèses H_1, H_2 et H_3 sont alors évidemment vérifiées.

L'opérateur $\mathcal{N}$ est alors de Hilbert-Schmidt puisque $K \in L^2(\mu \otimes \mu)$ donc (propositions 6.18 et 3.11 de [35]) il existe une base orthonormale $(h_n)_{n\in\mathbb{N}}$ de $L^2(\mu)$ formée de fonctions propres de l'opérateur $\mathcal{N}$ telles que les valeurs propres $(\lambda_n)_{n\in\mathbb{N}}$ associées (positives ou nulles) forment une suite de $\ell_1(\mathbb{R})$.

En outre la famille $(\sqrt{\lambda_n}\, h_n)_{n\in N_o}$ où $N_o = \{n \mid \lambda_n \neq 0\}$ est une base orthonormale de H et le noyau K s'écrit
$\forall (s,t) \in E^2 \quad K(s,t) = \sum_{n\in N_o} \lambda_n h_n(s)\, h_n(t)$ la série convergeant dans $H (\forall t \in E \quad K(.,t) = \sum_{n\in N_o} \lambda_n h_n(t)\, h_n)$ mais aussi uniformément sur E^2 (théorème de Mercer).

Proposition 2.- Un élément f de H appartient à $\operatorname{Im} N$ si et seulement si $\sum_{n\in N_o} (f,h_n)^2 < +\infty$.

Démonstration : Soit $f \in \operatorname{Im} N$ $f = N(g)$ $g \in L^2(\mu)$.

$$\forall n \in N_o \quad (f,h_n) = (N(g),\, h_n) = (g,h_n)_{L^2(\mu)}$$

et
$$\sum_{n\in N_o} (f,h_n)^2 \leqslant ||g||^2_{L^2(\mu)} < +\infty.$$

Réciproquement soit $f \in H$ tel que $\sum_{n\in N_o} (f,h_n)^2 < +\infty$.

Soit g l'élément de $L^2(\mu)$ défini par

$$g = \sum_{n\in N_o} (f,h_n)\, h_n \quad \text{(convergence de la série dans } L^2(\mu)).$$

On a alors :

$$\forall n \in N_o \quad (g,h_n)_{L^2(\mu)} = (N(g),\, h_n) = (f,h_n)$$

donc
$$f = N(g) .$$

Remarques :

a) Si $f \in H$ s'écrit $f = \sum_{n\in N_o} \alpha_n \sqrt{\lambda_n}\, h_n$, la convergence a lieu dans H donc aussi dans $L^2(\mu)$

$$\forall n \in N_o \quad (f,h_n)_{L^2(\mu)} = \alpha_n \sqrt{\lambda_n} = \lambda_n (f,h_n) .$$

La condition nécessaire et suffisante de la proposition précédente s'écrit aussi $\displaystyle\sum_{n\in N_o} \frac{(f,h_n)^2_{L^2(\mu)}}{\lambda_n^2} < +\infty.$

b) $g \in \operatorname{Ker} N \Longleftrightarrow \forall n \in N_o \quad (N(g), h_n) = 0$.

$g \in \operatorname{Ker} N \Longleftrightarrow \forall n \in N_o \quad (g, h_n)_{L^2(\mu)} = 0$.

$g \in \operatorname{Ker} N \Longleftrightarrow \forall s \in E \int K(s,t)\, g(t)\, d\mu(t) = 0$.

Le sous-espace de Hilbert de $L^2(\mu)$ engendré par $(K(.,s))_{s\in E}$ est donc égal à celui engendré par $(h_n)_{n\in N_o}$ et égal à $(\operatorname{Ker} N)^{\perp L^2(\mu)}$.

II - APPROXIMATION FORTE DE LA MESURE EMPIRIQUE.-

1°) *Processus empirique fonctionnel multidimensionnel* (*D'après* [29]).

On suppose que $E = [0,1]^2$, que T est sa tribu borélienne et que μ admet une densité f par rapport à la mesure de Lebesgue telle que :

$$\forall x \in [0,1]^2 \qquad 0 < b \leq f(x) \leq B .$$

On note F_M l'ensemble des fonctions continues sur $[0,1]^2$, bornées en valeur absolue par M $(M > 0)$ et admettant des dérivées partielles d'ordre 1 et 2 bornées en valeur absolue par M et on pose $F = \bigcup_{M>0} F_M$.

On pose $\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{Y_k}$ et pour toute fonction g de F

$$\alpha_n(g) = \sqrt{n}\, \left(\int g\, d\mu_n - \int g\, d\mu\right) .$$

(On définit ainsi un processus empirique ([21])).

On suppose enfin que l'espace probabilisé (Ω, A, P) est suffisamment riche pour que l'on puisse y définir les suites de variables et de processus utilisées dans [29]. Cette dernière supposition ne restreint pas la généralité car on peut toujours "agrandir" l'espace $(\Omega,\mathcal{A},P)$.

On a alors le théorème suivant (qui se généralise en dimension $d \geq 1$).

Théorème.- Il existe sur $(\Omega,\mathcal{A},P)$ une suite $(G_n)_{n\in\mathbb{N}^*}$ de processus gaussiens centrés ayant $\mathcal{F}$ pour ensemble des paramètres, de fonction de covariance

$$E(G_n(g)\ G_n(h)) = E(\alpha_n(g)\ \alpha_n(h)) = C(g,h) = \int gh\ d\mu - \int g\ d\mu \times \int h\ d\mu$$

tels que pour tout $\varepsilon > 0$, il existe des constantes c_1 et c_2 ne dépendant que de ε pour lesquelles on ait

$$\forall n \in \mathbb{N}^* \quad P\{\sup_{g\in\mathcal{F}_M} |\alpha_n(g) - G_n(g)| \geq c_1\ M\ B\ n^{-1/12}\ (\mathrm{Log}\ n)^{3/4}\}$$
$$\leq c_2\left[b^{(3\varepsilon+1)/2}\ n^{\varepsilon}\right]^{-1}$$

et le corollaire qui s'en déduit grâce au lemme de Borel-Cantelli :

Corollaire.- On a presque sûrement, quand n tend vers $+\infty$:

$$\sup_{g\in\mathcal{F}_M} |\alpha_n(g) - G_n(g)| = O(n^{-1/12}\ (\mathrm{Log}\ n)^{3/4}) .$$

De plus la suite $(G_n)_{n\in\mathbb{N}^*}$ est telle que les fonctions

$\mathcal{F} \longrightarrow \mathbb{R}$ sont presque sûrement linéaires.

$g \longmapsto G_n(g)$

2°) *Application à l'approximation des variables* $(\Lambda_n)_{n\in\mathbb{N}^*}$.

On suppose qu'il existe $M_o \in \mathbb{R}^+$ tel que la boule unité B_o de H_o soit incluse dans $\mathcal{F}_{M_o}$ (on a alors $H_o \subset \mathcal{F}$).

Proposition.- Il existe sur (Ω,A,P) une suite de variables aléatoires $(\Gamma_n)_{n\in\mathbb{N}^*}$ à valeurs dans H, gaussiennes centrées de fonction de covariance $C(f,g)$ telles que pour tout $\varepsilon > 0$ il existe des constantes c_1 et c_2 ne dépendant que de ε pour lesquelles on ait, pour n assez grand,

$$P\{||\Lambda_n - \Gamma_n|| \geq c_1 M_o B\, n^{-1/12} (\text{Log } n)^{3/4}\} \leq c_2 \left[b^{(3\varepsilon+1)/2} n^{\varepsilon}\right]^{-1}$$

Démonstration : Du corollaire précédent il résulte qu'il existe $n_o \in \mathbb{N}^*$ et $(a_n(M))_{n\geq n_o}$ tels que l'on ait presque sûrement

$$\forall n \geq n_o \quad \sup_{g\in F_M} |\alpha_n(g) - G_n(g)| \leq a_n(M) .$$

Pour $n \geq n_o$ on a alors presque sûrement

$$\sup_{g\in B_o} |G_n(g)| \leq \sup_{g\in F_{M_o}} |G_n(g)| \leq \sup_{g\in F_{M_o}} |\alpha_n(g)| + a_n(M_o)$$

donc $\sup_{g\in B_o} |G_n(g)| \leq 2\, M_o \sqrt{n} + a_n(M_o)$ p.s.

Pour ω appartenant à un ensemble Ω_o de probabilité 1 et pour $n \geq n_o$ l'application $\begin{array}{l} H_o \longrightarrow \mathbb{R} \\ g \longrightarrow G_n(g,\omega) \end{array}$ est donc une forme linéaire continue sur H_o, que l'on peut prolonger sur H grâce au théorème de Hahn-Banach et qui est représentée dans H par un élément $g_n(\omega)$ de norme majorée par $[2\, M_o \sqrt{n} + a_n(M_o)]$.

On pose alors $\Gamma_n(\omega) = g_n(\omega)$ si $\omega \in \Omega_o$

$\Gamma_n(\omega) = 0$ sinon .

Γ_n est bien une v.a. à valeurs dans H puisque

$$\forall t \in [0,1]^2 \quad (\Gamma_n, K(.,t)) = 1_{\Omega_o} G_n(K(.,t))$$

Γ_n est Bochner intégrable, puisque bornée en norme. $(\Gamma_n)_t$ et $G_n(K(.,t))$ sont presque sûrement égales donc Γ_n est gaussienne centrée.

Comme $(\Gamma_n, g) = G_n(g)$ p.s. la fonction de covariance de Γ_n est $C(f,g)$.

L'inégalité de la proposition résulte du théorème précédent compte tenu de :

$$||\Lambda_n - \Gamma_n|| = \sup_{g \in B_o} |(\Lambda_n - \Gamma_n, g)| \leq \sup_{g \in \mathcal{F}_{M_o}} |\alpha_n(g) - G_n(g)| \ .$$

L'application du lemme de Borel-Cantelli entraîne :

Corollaire.- On a presque sûrement pour n tendant vers $+\infty$

$$||\Lambda_n - \Gamma_n|| = O(n^{-1/12} (\mathrm{Log}\ n)^{3/4}) \ .$$

CHAPITRE III

METHODES SPLINES EN ESTIMATION FONCTIONNELLE.

I - INTRODUCTION.-

Soit X une v.a.r. de loi inconnue et $(X_1,\dots,X_n)$ un échantillon de taille n de X.

Le théorème de Glivenko-Cantelli fournit un estimateur presque sûrement uniformément convergent de la fonction de répartition F de la loi de X : c'est la fonction de répartition empirique $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{]-\infty,t[}(X_i)$, $t \in \mathbb{R}$.

Dans le cas où la loi de X admet une densité f à support inclus dans $[0,1]$, nous nous proposons de "lisser" l'estimateur F_n et d'en déduire un estimateur de f.

Ce lissage sera effectué au moyen de splines cubiques.

II - FONCTIONS SPLINES CUBIQUES SUR $[0,1]$ ([1]).-

1°) Définition - Notations. Soit Δ une subdivision de $[0,1]$. $0 = x_0 < x_1 \dots\dots < x_N = 1$ $(N > 1)$.

Une fonction spline cubique (relativement à Δ) est une fonction S de classe C^2 sur $[0,1]$ dont la restriction à $[x_i, x_{i+1}]$ $(0 \leq i \leq N-1)$ est un polynôme de degré 3 au plus.

On pose $\quad h_j = x_j - x_{j-1} \quad , \quad 1 \leqslant j \leqslant N$

et $\quad M_j = S''(x_j) \quad , \quad 0 \leqslant j \leqslant N \quad .$

On va voir que si l'on impose à S des conditions d'interpolation : $S(x_j) = y_j \quad 0 \leqslant j \leqslant N$, la fonction spline S est entièrement déterminée par la donnée de ses "moments" $(M_j)_{0 \leqslant j \leqslant N}$.

2°) Expression de la spline en fonction des (M_j).

Si $\quad t \in [x_{j-1}, x_j] \quad S''(t) = M_{j-1} \dfrac{x_j - t}{h_j} + M_j \dfrac{t - x_{j-1}}{h_j}$

en intégrant il vient :

$$S'(t) = - M_{j-1} \frac{(x_j - t)^2}{2\, h_j} + M_j \frac{(t - x_{j-1})^2}{2\, h_j} + A$$

et

$$S(t) = M_{j-1} \frac{(x_j - t)^3}{6\, h_j} + M_j \frac{(t - x_{j-1})^3}{6\, h_j} + At + B \; .$$

on trouve la valeur des constantes A et B en écrivant que

$$S(x_{j-1}) = y_{j-1}$$

et

$$S(x_j) = y_j \quad .$$

D'où, pour $t \in [x_{j-1}, x_j]$

$$S(t) = M_{j-1} \frac{(x_j - t)^3}{6\, h} + M_j \frac{(t - x_{j-1})^3}{6\, h_j} + M_{j-1} \frac{h_j (t - x_j)}{6}$$

$$- M_j \frac{h_j (t - x_{j-1})}{6} + \frac{y_j (t - x_{j-1}) - y_{j-1} (t - x_j)}{h_j}$$

$$S'(t) = - M_{j-1} \frac{(x_j - t)^2}{2\, h_j} + M_j \frac{(t - x_{j-1})^2}{2\, h_j} + \frac{y_j - y_{j-1}}{h_j} - \frac{M_j - M_{j-1}}{6} h_j$$

La fonction S vérifie les conditions imposées si et seulement si $S'(x_j-) = S'(x_j+) \quad 1 \leq j \leq N-1$,

ce qui équivaut à un système de $(N-1)$ équations linéaires indépendantes en les $(M_j)_{0 \leq j \leq N}$.

L'existence de la spline d'interpolation est donc assurée. Son unicité résultera de deux conditions supplémentaires convenablement choisies.

3°) *Calcul des* (M_j).

Les conditions complémentaires seront de deux types : (conditions aux limites).

$$CL_1 \quad \begin{cases} S'(0) = a \\ S'(1) = b \end{cases} \qquad (a,b) \in \mathbb{R}^2$$

$$CL_2 \quad M_o = M_N = 0 \; .$$

a) *1^er^ cas* CL_1.

Les (M_j) vérifient un système qui est obtenu en remplaçant S' par son expression ci-dessus dans les égalités

$$\begin{cases} S'(x_j-) = S'(x_j+) & 1 \leq j \leq N-1 \\ S'(0) = a \\ S'(1) = b \end{cases}$$

c'est le système suivant :

$$\left\{\begin{array}{l} \dfrac{h_j}{6} M_{j-1} + \dfrac{h_j + h_{j+1}}{3} M_j + \dfrac{h_{j+1}}{6} M_{j+1} = \dfrac{y_{j+1} - y_j}{h_{j+1}} - \dfrac{y_j - y_{j-1}}{h_j} \\[2ex] 1 \leqslant j \leqslant N-1 \\[4ex] 2\, M_o + M_1 = \dfrac{6}{h_1}\left(\dfrac{y_1 - y_o}{h_1} - a\right) = d_o \\[2ex] M_{N-1} + 2\, M_N = \dfrac{6}{h_n}\left(b - \dfrac{y_N - y_{N-1}}{h_N}\right) = d_N \end{array}\right.$$

En posant pour $1 \leqslant j \leqslant N-1$ $\quad \lambda_j = \dfrac{h_{j+1}}{h_j + h_{j+1}} \qquad \mu_j = 1 - \lambda_j = \dfrac{h_j}{h_j + h_{j+1}}$

et $\qquad d_j = \dfrac{6}{h_j + h_{j+1}} \left(\dfrac{y_{j+1} - y_j}{h_{j+1}} - \dfrac{y_j - y_{j-1}}{h_j}\right)$

on obtient pour $1 \leqslant j \leqslant N-1$

$$\mu_j\, M_{j-1} + 2\, M_j + \lambda_j\, M_{j+1} = d_j$$

et le système s'écrit

$$(\Sigma_1) \qquad \underbrace{\begin{bmatrix} 2 & 1 & 0 & & & & & 0 \\ \mu_1 & 2 & \lambda_1 & 0 & & & & 0 \\ 0 & \mu_2 & 2 & \lambda_2 & & & 0 & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & 0 & & & & & \vdots \\ \vdots & & & & \mu_{N-2} & 2 & \lambda_{N-2} & 0 \\ \vdots & & & & & \mu_{N-1} & 2 & \lambda_{N-1} \\ 0 & & & & & & 1 & 2 \end{bmatrix}}_{\substack{M_{CL_1} \\ \text{(d'ordre } N+1)}} \begin{bmatrix} M_o \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ M_N \end{bmatrix} = \begin{bmatrix} d_o \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ d_N \end{bmatrix}$$

La matrice M_{CL_1} du système est bien sûr toujours inversible (elle est à diagonale dominante).

b) 2ème cas CL_2.

$$\begin{cases} M_o = M_N = 0 \\ S'(x_j-) = S'(x_j+) \quad 1 \leqslant j \leqslant N-1 \end{cases}.$$

Le système obtenu est alors

$M_o = M_N = 0$

$$(\Sigma_2) \quad \underbrace{\begin{bmatrix} 2 & \lambda_1 & 0 & & & & 0 \\ \mu_2 & 2 & \lambda_2 & 0 & & & 0 \\ 0 & \mu_3 & 2 & \lambda_3 & & \bigcirc & \\ \vdots & & & \ddots & & & \\ \vdots & & \bigcirc & & \ddots & & \\ \vdots & & & & \mu_{N-2} & 2 & \lambda_{N-2} \\ 0 & & & & & \mu_{N-1} & 2 \end{bmatrix}}_{M_{CL_2}} \begin{bmatrix} M_1 \\ \vdots \\ \vdots \\ M_{N-1} \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ \vdots \\ d_{N-1} \end{bmatrix}$$

(d'ordre N-1 ; que l'on obtient comme sous-matrice de M_{CL_1} en supprimant les lignes et colonnes bordantes).

M_{CL_2} est également à diagonale dominante.

4°) *Lemme II.4.*- Si l'on munit $\mathbb{R}^{N+1}$ et $\mathbb{R}^{N-1}$ de la norme du sup, les normes correspondantes de $M_{CL_1}^{-1}$ et $M_{CL_2}^{-1}$ sont majorées par 1.

On aura donc dans chaque cas

$$\sup_{0\leq j\leq N} |M_j| \leq \sup_{0\leq j\leq N} |d_j| \quad .$$

Démonstration : Elle est faite dans le cas général d'une matrice à diagonale dominante ([1]) :

Soit $M = (m_{ij})_{\substack{1\leq i\leq N \\ 1\leq j\leq N}}$ telle que $\forall i \quad |m_{ii}| > \sum_{j\neq i} |m_{ij}|$

Soit $\alpha = (\alpha_i)_{1\leq i\leq N}$ tel que $||\alpha|| = |\alpha_k| \quad (\neq 0)$ $1\leq k\leq N$

et $\beta = M\alpha$

$$||\beta|| = ||M\alpha|| = \max_i \left|\sum_{j=1}^{N} m_{ij}\, \alpha_j\right|$$

$$||M\alpha|| \geq \left|\sum_{j=1}^{N} m_{kj}\, \alpha_j\right|$$

$$||M\alpha|| \geq |m_{kk}\, \alpha_k| - \left|\sum_{\substack{j=1 \\ j\neq k}}^{N} m_{kj}\, \alpha_j\right|$$

$$||M\alpha|| \geq |m_{kk}\, \alpha_k| - \left(\sum_{\substack{j=1 \\ j\neq k}}^{N} |m_{kj}|\; ||\alpha||\right)$$

$$||M\alpha|| \geq \left(|m_{kk}| - \sum_{\substack{j=1 \\ j\neq k}}^{N} |m_{kj}|\right) ||\alpha||$$

$$||M^{-1}|| = \sup_{\beta\neq 0} \frac{||M^{-1}\beta||}{||\beta||} \leq \frac{1}{\min_i \left(|m_{ii}| - \sum_{j\neq i} |m_{ij}|\right)}$$

et donc $\|M_{CL_1}^{-1}\| \leqslant 1$ $\quad\quad$ $\|M_{CL_2}^{-1}\| \leqslant 1$

puisque pour $1 \leqslant j \leqslant N-1$ $\begin{cases} \lambda_j + \mu_j = 1 \\ \lambda_j > 0 \quad \mu_j > 0 \end{cases}$.

III - ESTIMATION PAR DES FONCTIONS SPLINES.-

X v.a.r. dont la densité f est à support inclus dans $[0,1]$.

$(X_1,\ldots,X_n)$ échantillon de taille n de X.

Soit $\Delta_n : 0 = x_o < x_1 ,\ldots, < x_N = 1$ une subdivision de $[0,1]$.

1°) Cas CL_1.

Soit $\hat{F}_n$ la spline cubique interpolant la fonction de répartition empirique F_n sur les noeuds $(x_i)_{0\leqslant i\leqslant N}$ avec les conditions

$$\hat{F}'_n(0) = \hat{a}$$

$$\hat{F}'_n(1) = \hat{b}$$

où $\hat{a}$ et $\hat{b}$ sont également estimés à partir de F_n : étant donné $m \in \mathbb{N}^*$ (fixé une fois pour toutes) , $m \leqslant N$.

$\hat{a}$ est la dérivée en 0 du polynôme de degré m interpolant F_n en $x_o, x_1 ,\ldots, x_m$.

$\hat{b}$ est la dérivée en 1 du polynôme de degré m interpolant F_n en $x_{N-m} ,\ldots, x_N$.

On pose alors $\hat{f}_n = \hat{F}'_n$, dérivée de $\hat{F}_n$ (méthode introduite par G. Wahba [47] dans le cas particulier d'une subdivision uniforme et pour lequel l'auteur ne donne qu'une condition suffisante de convergence ponctuelle en moyenne quadratique).

2°) <u>*Cas*</u> CL_2.

On n'a pas à estimer de paramètres complémentaires : Soit encore $\hat{F}_n$ la spline cubique interpolant F_n sur les noeuds $(x_i)_{0\leq i\leq N}$ avec $M_o = M_N = 0$.

On pose de la même façon $\hat{f}_n = \hat{F}'_n$.

3°) <u>*Définition des noyaux splines.*</u>

Pour $x \in [0,1]$ soit S^1_x la spline cubique interpolant $\varphi_x : t \longmapsto 1_{[0,t]}(x)$ sur les noeuds $(x_i)_{0\leq i\leq N}$ avec les conditions CL_1 : $(S^1_x)'(0) = a(x)$

$$(S^1_x)'(1) = b(x)$$

$a(x)$: dérivée en 0 du polynôme de degré m interpolant φ_x en $x_o , x_1 ,\dots, x_m$.

$b(x)$: dérivée en 1 du polynôme de degré m interpolant φ_x en $x_{N-m} ,\dots, x_N$.

Comme $F_n(t) = \frac{1}{n} \sum_{k=1}^{n} 1_{[0,t]}(X_k)$, $t \in [0,1]$. On a, par linéarité de l'application qui, à une fonction associe la spline qui l'interpole,

$$\forall t \in [0,1] \quad \hat{F}_n(t) = \frac{1}{n} \sum_{k=1}^{n} S^1_{X_k}(t)$$

et

$$\hat{f}_n(t) = \frac{1}{n} \sum_{k=1}^{n} (S^1_{X_k})'(t)$$

en posant $$\frac{d}{dt} S_x^1(t) = K_{\Delta_n}^1(x,t)$$

on a alors $$\hat{f}_n(t) = \frac{1}{n} \sum_{k=1}^{n} K_{\Delta_n}^1(X_k,t)$$

et $\hat{f}_n$ appartient à une classe connue d'estimateurs de la densité.

Pour $x \in [0,1]$ soit S_x^2 la spline cubique interpolant ψ_x sur les noeuds $(x_i)_{0 \leq i \leq N}$ avec les conditions CL_2 $(S_x^2)''(0) = (S_x^2)''(1) = 0$.

La même remarque montre que, dans ce cas en posant $\frac{d}{dt} S_x^2(t) = K_{\Delta_n}^2(x,t)$

$$\hat{f}_n(t) = \frac{1}{n} \sum_{k=1}^{n} K_{\Delta_n}^2(X_k,t) \ .$$

Plus généralement ces remarques seront valables pour le lissage d'estimateurs qui sont déjà de la forme

$$\frac{1}{n} \sum_{k=1}^{n} K_{r_n}(X_k,t) \ .$$

Avant de donner des théorèmes de convergence pour les estimateurs splines, voyons comment ces estimateurs s'intégrent dans la théorie hilbertienne des splines ([31]).

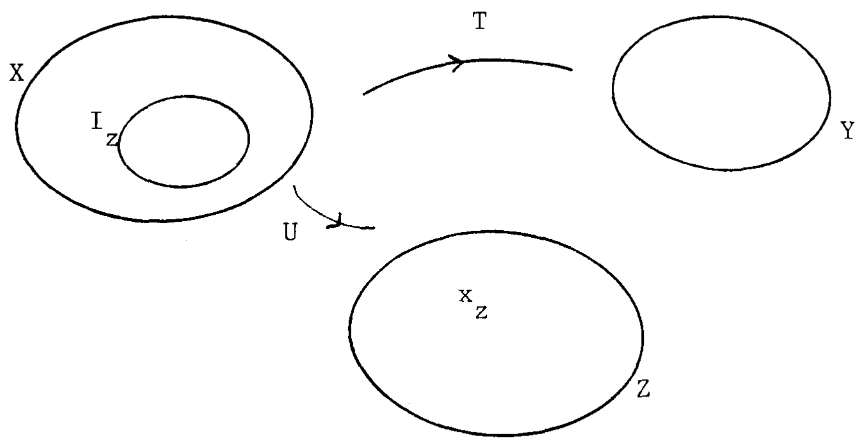
IV - THEORIE HILBERTIENNE DES SPLINES D'INTERPOLATION.-

1°) Position du problème.

Soit X, Y, Z 3 Hilbert réels.

$U : X \longrightarrow Z$ et $T : X \longrightarrow Y$ deux applications linéaires continues surjectives.

Soit $z \in Z$ et $I_z = U^{-1}(\{z\})$



Problème (P) : On cherche $\sigma \in I_z$ tel que $||T(\sigma)||_Y = \min\limits_{x \in I_z} ||T(x)||_Y$.

Un tel σ s'il existe est appelé fonction spline d'interpolation relativement à T, U et z.

2°) *Existence - Unicité.*

Théorème 1.- Il existe au moins une fonction spline pour tout $z \in Z$ si et seulement si $\operatorname{Ker} T + \operatorname{Ker} U$ est fermé.

Théorème 2.- Si $\operatorname{Ker} T + \operatorname{Ker} U$ est fermé.

On a unicité si $\operatorname{Ker} T \cap \operatorname{Ker} U = \{0\}$.

3°) *Propriétés de minimisation.*

Théorème 3.- Si $\operatorname{Ker} T + \operatorname{Ker} U$ fermé et $\operatorname{Ker} T \cap \operatorname{Ker} U = \{0\}$.

Si $S = \{s \in X \mid \langle T(s), T(x)\rangle_Y = 0 \ \forall x \in \operatorname{Ker} U\}$ (espace des fonctions splines).

Si z est fixé dans Z . Alors

* il existe un unique σ dans S tel que $U(\sigma) = z$,

* $\forall x \in I_z \quad ||T(\sigma-x)||_Y = \min\limits_{s\in S} ||T(s-x)||_Y$

et tout autre $\tilde{\sigma}$ ayant cette propriété diffère de σ par un élément de Ker T.

* $\forall s \in S \quad ||T(\sigma-s)||_Y = \min\limits_{x\in I_z} ||T(x-s)||_Y$

et σ est l'unique élément de I_z ayant cette propriété.

4°) *Cas où* $Z = \mathbb{R}^n$. Alors Ker T + Ker U est fermé.

Si on a $\text{Ker } T \cap \text{Ker } U = \{0\}$ (donc unicité) on a $\dim \text{Ker } T = q \leq n$.

$U(x) = (\langle k_1,x\rangle_X , \ldots\ldots, \langle k_n,x\rangle_X)$ où $k_1 , \ldots, k_n$ sont linéairement indépendants.

Théorème de caractérisation.-

Soit $z \in Z = \mathbb{R}^n$ et $\sigma \in I_z$

(σ est fonction spline) $\Longleftrightarrow$ ($\exists (\lambda_i)_{1\leq i\leq n}$ tels que $T^*T(\sigma) = \sum\limits_{i=1}^{n} \lambda_i \, k_i$

où les λ_i satisfont $\sum\limits_{i=1}^{n} \lambda_i \langle k_i,p\rangle_X = 0$

$\forall p \in \text{Ker } T$) .

Tous ces résultats sont démontrés dans [31].

Les résultats obtenus jusqu'à présent en estimation au sujet des fonctions splines l'ont été par des méthodes très "calculatoires" et il semble difficile (comme le dit G. Wahba dans [47]) de les étendre à des splines de degré supérieur à 3 à cause de la complexité des expressions locales des splines.

On peut pourtant penser que pour des fonctions à estimer "très régulières", des splines de degré supérieur donnent de meilleurs résultats. Le cadre hilbertien devrait permettre de donner des résultats plus généraux. D'autre part l'opérateur T qui est égal à D^2 (D : opérateur de dérivation) dans le cas des splines cubiques pourra être choisi suivant le problème à traiter et les informations que l'on a sur la fonction à estimer (fonction de répartition, densité, régression, etc ...).

Le lecteur intéressé pourra consulter [30] et [44] qui contient une importante bibliographie.

Il est à noter également que pour certains problèmes les splines "d'ajustement" ou les "splines dans un convexe" [31] peuvent être mieux adaptées.

Nous espérons revenir prochainement sur ces questions.

V - SPLINES ALEATOIRES.-

1°) *Cas général.*

Soit (Ω,A,P) un espace probabilisé et z une variable aléatoire sur (Ω,A) à valeurs dans (Z,B_Z).

On suppose que $\operatorname{Ker} T + \operatorname{Ker} U$ fermé et $\operatorname{Ker} T \cap \operatorname{Ker} U = \{0\}$.

Pour tout $\omega \in \Omega$ il existe une unique spline $S(\omega)$ telle que $U(S(\omega)) = z(\omega)$.

Théorème.- L'application $S : (\Omega, A) \longrightarrow (X, B_X)$ est

$$\omega \longmapsto S(\omega)$$

mesurable (spline aléatoire).

Démonstration : S est un Hilbert car il est fermé dans X.

∗ Soit $u = U_{|S}$

u est linéaire, continue, bijective. Il en est donc de même de u^{-1} (théorème de l'inverse continu).

∗ $S = u^{-1} \circ z$ est donc mesurable.

Dans les applications l'espace X sera souvent à noyau reproduisant. La spline aléatoire est alors une variable aléatoire du type étudié dans la première partie.

2°) *Espérance de la spline aléatoire.*

Il résulte des propriétés de l'intégrale au sens de Pettis ou de Bochner (chapitre I) la :

Proposition.- Si z est intégrable S l'est aussi et

$$E(S) = u^{-1}(E_z) \ .$$

Ce qui montre que l'on pourra construire des estimateurs splines asymptotiquement sans biais d'un élément de X dès que l'on aura des théorèmes de convergence pour les splines et que l'on disposera d'observations convenables dans $\mathbb{R}^n$.

3°) *Cas des splines cubiques avec les conditions* CL_1.

$X = H^2([0,1])$ $\quad Y = L^2([0,1])$ $\quad Z = \mathbb{R}^{N+3}$.

$T = D^2 \qquad <x,y>_X = \int_0^1 xy + \int_0^1 x'y' + \int_0^1 x''y''$

$$<x,y>_Y = \int_0^1 xy \ .$$

$U(x) = (x'(0),\ x(0),\ x(x_1)\ ,\ldots\ldots,\ x(1),\ x'(1))$.

Si $z = (a,\ y_o,\ y_1\ ,\ldots\ldots,\ y_N\ ,\ b)$ la solution du problème (P) est la spline cubique obtenue en résolvant $(\sum_1)$ et donc (propriétés de minimisation IV 3°) c) cette spline minimise $\int_0^1 (f'')^2$ parmi toutes les fonctions de H^2 qui interpolent z.

Propriété de l'estimateur $\hat{f}_n$ *correspondant*.

$\hat{f}_n$ minimise $\int_0^1 g'^2$ parmi toutes les fonctions g de $H^1([0,1])$ dont une primitive interpole F_n en $x_o, x_1, \ldots, x_N$ et qui vérifient

$$g(0) = \hat{a}$$
$$g(1) = \hat{b} \ .$$

4°) *Cas des splines cubiques avec les conditions* CL_2.

$\hat{f}_n$ peut être obtenu directement comme solution du problème (P) si $H = H^1([0,1])$ $\quad Y = L^2([0,1])$ $\quad Z = \mathbb{R}^N$

$T = D \qquad <x,y>_X = \int_0^1 xy + \int_0^1 x'y'$

$$<x,y>_Y = \int_0^1 xy$$

$$U(x) = \left(\int_0^{x_1} x \;,\; \int_{x_1}^{x_2} x \;, \ldots\ldots, \int_{x_{N-1}}^{1} x \right) \quad .$$

On estime $\displaystyle\int_{x_{i-1}}^{x_i} f(t)\,dt$ par la fréquence des observations tombant dans $[x_{i-1}, x_i]$. C'est un cas particulier des splines par moyenne locale ([31]).

Propriété de l'estimateur $\hat{f}_n$ *correspondant.*

$\hat{f}_n$ minimise $\displaystyle\int_0^1 g'^2$ parmi toutes les fonctions g de $H^1([0,1])$ dont une primitive interpole F_n en $x_o, \ldots, x_N$.

Les conditions CL_2 conduisent donc à un estimateur meilleur du point de vue "lissage" et bien plus simple à calculer pratiquement.

VI - THEOREMES DE CONVERGENCE POUR LES SPLINES CUBIQUES.-

Les démonstrations classiques ([1]) s'appliquent (cas CL_2) ou s'adaptent facilement (cas CL_1).

1°) Conditions CL_1.

F fonction continue sur $[0,1]$. S spline cubique d'interpolation de F avec CL_1 $\begin{cases} S'(0) = a \\ S'(1) = b \end{cases}$.

Notations : On posera dans toute la suite $H = \max\limits_{1 \leq j \leq N} h_j$

$$h = \min_{1 \leq j \leq N} h_j$$

si $\delta \in \mathbb{R}^{+*}$ $\mu(F,\delta) = \sup_{|t-t'|\leq\delta} |F(t) - F(t')|$.

Théorème VI.1.-

$\forall t \in [0,1]$

$$|F(t) - S(t)| \leq \frac{2}{3\sqrt{3}} \frac{H^2}{h^2} (\mu(F,H) + h \max(|a|, |b|)) + 2\mu(F,H) .$$

Démonstration : si $t \in [x_{j-1}, x_j]$

$$|S(t)-F(t)| = \left| M_{j-1} \frac{(x_j-t)^3 + h_j^2(t-x_j)}{6\, h_j} + M_j \frac{(t-x_{j-1})^3 - h_j^2(t-x_{j-1})}{6\, h_j} \right.$$

$$\left. + \frac{y_j(t-x_{j-1}) - y_{j-1}(t-x_j)}{h_j} - F(t) \right|$$

La valeur absolue des coefficients de M_j et M_{j-1} est majorée par $\frac{h_j^2}{3}$. Une étude plus précise montrerait qu'elle est majorée par $\frac{h_j^2}{9\sqrt{3}}$.

Majoration de $\max_{0\leq j\leq N} |M_j|$.

Pour $1 \leq j \leq N-1$ $d_j = \frac{6}{h_j + h_{j-1}} \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j} \right)$

$$|d_j| \leq \frac{6}{h^2} \mu(F,H)$$

$$d_o = \frac{6}{h_1}\left(\frac{y_1 - y_o}{h_1} - a\right) \qquad |d_o| \leqslant \frac{6}{h^2}\,\mu(F,H) + \frac{6\,|a|}{h}$$

$$d_N = \frac{6}{h_N}\left(b - \frac{y_N - y_{N-1}}{h_N}\right) \qquad |d_N| \leqslant \frac{6}{h^2}\,\mu(F,H) + \frac{6\,|b|}{h} \quad .$$

Du lemme II.4 il résulte que $\sup\limits_{0\leqslant j\leqslant N} |M_j| \leqslant \frac{6}{h^2}\,(\mu(F,H) + h\max(|a|, |b|)$.

D'autre part

$$\left|\frac{y_j(t-x_{j-1}) - y_{j-1}(t-x_j)}{h_j} - F(t)\right| = \left|\frac{t-x_{j-1}}{h_j}\,(y_j - F(t)) - \frac{t-x_j}{h_j}\,(y_{j-1} - F(t))\right|$$

$$\leqslant 2\,\mu(F,H) \ .$$

D'où le théorème.

On supposera qu'il existe une constante C_o telle que pour toutes les subdivisions considérées on ait $1 \leqslant \frac{H}{h} \leqslant C_o$.

Corollaire VI.1.- Si $|a|$ et $|b|$ sont majorés par s_o on a $\forall t \in [0,1]\ |F(t) - S(t)| \leqslant \frac{2}{3\sqrt{3}}\,C_o^2(\mu(F,H) + h\,s_o) + 2\mu(F,H)$

et $S \longrightarrow F$ uniformément si $H \longrightarrow 0$.

Théorème VI.2.- Si F est de classe C^{m+1} , $m \in \mathbb{N}^*$ (on note $f = F'$).

Si a est la dérivée en 0 du polynôme P_m de degré m interpolant F en $x_o, \ldots, x_m$.

Si b est la dérivée en 1 du polynôme Q_m de degré m interpolant F en $x_{N-m}, \ldots, x_N$.

Alors 1 - $S \longrightarrow F$ uniformément si $H \longrightarrow 0$.

2 - $\sup\limits_{t\in[0,1]} |S(t) - F(t)| = O(H^2)$.

3 - $S' \longrightarrow F'$ uniformément si $H \longrightarrow 0$

et $\sup\limits_{t\in[0,1]} |S'(t) - F'(t)| = O(H)$.

On a posé $M'' = \sup\limits_{t\in[0,1]} |F''(t)|$ $M^{(m+1)} = \sup\limits_{t\in[0,1]} |F^{(m+1)}(t)|$

$M' = \sup\limits_{t\in[0,1]} |f(t)|$

Démonstration :

a) *expression de a et b et démonstration de 1* -

$a = P'_m(0)$.

Pour $1 \leqslant i \leqslant m$ on a $(F-P_m)(x_{i-1}) = (F-P_m)(x_i) = 0$

donc il existe $\theta_i \in]x_{i-1}, x_i[$ tel que $(f - P'_m)(\theta_i) = 0$.

P'_m est donc le polynôme d'interpolation de f

en $\theta_1, \theta_2, \ldots, \theta_m$.

f étant de classe C^m, il existe $\theta \in [0, \theta_m]$ tel que

$$f(0) = P'_m(0) = f(0) - a = \frac{V_{m-1}(0)}{m!} f^{(m)}(\theta)$$

(expression de l'erreur d'interpolation)

où $V_{m-1}(x) = \prod\limits_{i=1}^{m} (x-\theta_i)$.

on a : $\quad |V_{m-1}(0)| = \prod_{i=1}^{m} \theta_i \leqslant \prod_{i=1}^{m} i\,H \quad$ car $\quad \theta_i \in]x_{i-1}, x_i[$

$$\begin{cases} |V_{m-1}(0)| \leqslant m!\ H^m \\ a = f(0) - \dfrac{V_{m-1}(0)}{m!} f^{(m)}(\theta) \end{cases}$$

$$b = Q'_m(1) \ .$$

Pour $\quad N-m+1 \leqslant i \leqslant N \quad$ on a $\quad (F-Q_m)(x_{i-1}) = (F-Q_m)(x_i) = 0$

donc il existe $\quad \tau_i \in]x_{i-1} \ , \ x_i[\quad$ tel que $\quad (f-Q'_m)(\tau_i) = 0$

Q'_m est donc le polynôme d'interpolation de f en $\tau_{N-m+1}, \ldots, \tau_N$.

f étant de classe C^m il existe $\tau \in [\tau_{N-m+1}, 1]$ tel que

$$f(1) - Q'_m(1) = f(1) - b = \frac{W_{m-1}(1)}{m!} f^{(m)}(\tau)$$

où
$$W_{m-1}(x) = \prod_{i=N-m+1}^{N} (x-\tau_i)$$

$$\begin{cases} b = f(1) - \dfrac{W_{m-1}(1)}{m!} f^{(m)}(\tau) \\ \text{et on a} \\ \left|W_{m-1}(1)\right| = \displaystyle\prod_{i=N-m+1}^{N} (1-\tau_i) \leqslant \prod_{i=N-m+1}^{N} (N-i+1)H = H^m\ m! \end{cases}$$

Donc
$$\begin{cases} |a| \leqslant |f(0)| + H^m\ M^{(m+1)} \\ |b| \leqslant |f(1)| + H^m\ M^{(m+1)} \end{cases}$$

et le corollaire VI.1 s'applique .

On a donc 1- et la majoration :

$\forall t \in [0,1]$

$$|F(t)-S(t)| \leq \frac{2}{3\sqrt{3}} C_o^2 (HM' + h(M' + H^m M^{(m+1)}) + 2 HM' \quad .$$

b) *majoration de* d_o , d_N *et* $(d_j)_{1 \leq j \leq N-1}$.

$$d_o = \frac{6}{h_1}\left(\frac{y_1 - y_o}{h_1} - a\right)$$

il existe $\alpha_1 \in]x_o, x_1[$ tel que $\dfrac{F(x_1) - F(x_o)}{h_1} = f(\alpha_1)$

donc $$d_o = \frac{6}{h_1} (f(\alpha_1) - f(0) + \frac{V_{m-1}(0)}{m!} f^{(m)}(\theta))$$

Or $$|f(\alpha_1) - f(0)| \leq \alpha_1 M''$$

donc $$|d_o| \leq 6(M'' + C_o H^{m-1} M^{(m+1)})$$

$$d_N = \frac{6}{h_N}\left(b - \frac{y_N - y_{N-1}}{h_N}\right)$$

il existe $\alpha_N \in]x_{N-1}, x_N[$ tel que $\dfrac{F(x_N) - F(x_{N-1})}{h_N} = f(\alpha_N)$

donc $$d_N = \frac{6}{h_N}\left(f(1) - f(\alpha_N) - \frac{W_{m-1}(1)}{m!} f^{(m)}(\tau)\right) .$$

Or $$|f(1) - f(\alpha_N)| \leq (1 - \alpha_N) M''$$

donc $$|d_N| \leq 6 (M'' + C_o H^{m-1} M^{(m+1)}) .$$

Pour $1 \leqslant j \leqslant N-1$

$$d_j = \frac{6}{h_j + h_{j+1}} \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j} \right)$$

$\frac{d_j}{6}$ est la différence divisée d'ordre 2 de F en

x_{j-1} , x_j , x_{j+1} donc il existe $\zeta_{j+1} \in [x_{j-1}, x_{j+1}]$ tel que

$$d_j = 6 \, \frac{F''(\zeta_{j+1})}{2}$$

d'où
$$|d_j| \leqslant 3 \, M''$$

D'après le lemme II.4 $\sup\limits_{0 \leqslant j \leqslant N} |M_j| \leqslant 6(M'' + C_o \, H^{m-1} \, M^{(m+1)})$

c) *majoration de* $|S(t) - F(t)|$ (2-).

si $t \in [x_{j-1}, x_j]$

$$|S(t) - F(t)| \leqslant \frac{4}{3\sqrt{3}} H^2 (M'' + C_o \, H^{m-1} \, M^{(m+1)})$$

$$+ \left| \frac{t-x_{j-1}}{h_j} (y_j - F(t)) - \frac{(t-x_j)}{h_j} (y_{j-1} - F(t)) \right|$$

Or $\quad y_j - F(t) = (x_j - t) \; f(\beta_j^1) \qquad \beta_j^1 \in [x_{j-1}, x_j]$

$$y_{j-1} - F(t) = (x_{j-1} - t) \; f(\beta_j^2) \qquad \beta_j^2 \in [x_{j-1}, x_j]$$

Le deuxième terme ci-dessus est donc

$$\frac{1}{h_j} \left| (t-x_{j-1}) \; (x_j - t) \; f(\beta_j^1) - (x_j - t) \; (t - x_{j-1}) \; f(\beta_j^2) \right|$$

et il est majoré par $H^2 \, M''$ puisque $|f(\beta_j^1) - f(\beta_j^2)| \leqslant h_j \, M''$.

Donc

$$\sup_{t\in[0,1]} |S(t) - F(t)| \leqslant \left[\left(\frac{4}{3\sqrt{3}} + 1\right) M'' + \frac{4}{3\sqrt{3}} C_o H^{m-1} M^{(m+1)}\right] H^2$$

D'où 2- .

d) majoration de $|S'(t) - f(t)|$ (3-).

Si $t \in [x_{j-1}, x_j]$

$$|S'(t) - f(t)| = \left| - M_{j-1} \frac{(x_j-t)^2}{2 h_j} + M_j \frac{(t-x_{j-1})^2}{2 h_j} - \frac{M_j - M_{j-1}}{6} h_j + \frac{y_j - y_{j-1}}{h_j} - f(t)\right|$$

$$|S'(t) - f(t)| \leqslant \max_{0\leqslant j\leqslant N} |M_j| \times H(C_o + \frac{1}{3}) + \left|\frac{y_j - y_{j-1}}{h_j} - f(t)\right| .$$

Il existe $\gamma_j \in |x_{j-1}, x_j|$ tel que $y_j - y_{j-1} = h_j\, f(\gamma_j)$

$$\left|\frac{y_j - y_{j-1}}{h_j} - f(t)\right| = |f(\gamma_j) - f(t)| \leqslant H M''$$

$$\sup_{t\in[0,1]} |S'(t) - f(t)| \leqslant H\left[6(C_o + \frac{1}{3}) (M'' + C_o H^{m-1} M^{(m+1)}) + M''\right]$$

d'où 3 -

2°) Conditions CL_2.

Théorème VI.3.- Soit $F \in H^2([0,1])$ et S la spline d'interpolation de F avec les conditions CL_2 alors :

$$||F-S||_\infty \leq \frac{1}{2}\, ||F''||_{L^2} \times H^{3/2}$$

et $$||F'-S'||_\infty \leq ||F''||_{L^2} \times H^{1/2} \quad .$$

Démonstration : Pour $1 \leq i \leq N$ $(F-S)\,(x_{i-1}) = (F-S)\,(x_i) = 0$ donc il existe $\xi_i \in \,]x_{i-1}, x_i[$ tel que $(F'-S')\,(\xi_i) = 0$.

Soit $t \in \,]x_{i-1}, x_i[$

$$(F'-S')\,(t) = \int_{\xi_i}^{t} (F''-S'')\,(u)\,du$$

$$|(F'-S')(t)| = \left(\int_{\xi_i}^{t} [(F''-S'')(u)]^2\,du\right)^{1/2} \times |t-\xi_i|^{1/2}$$

D'après les propriétés d'optimalité de S (IV 3°) b).

$$\int_0^1 (F''-S'')^2 \leq \int_0^1 F''^2$$

donc $$||F'-S'||_\infty \leq ||F''||_{L^2} \times H^{1/2} \quad .$$

D'autre part $$(F-S)\,(t) = \int_{\bar{x}}^{t} (F'-S')\,(u)\,du$$

où $\bar{x} = x_{i-1}$ si $t \leq \dfrac{x_{i-1}+x_i}{2}$ et x_i sinon.

D'où $$||F-S||_\infty \leq ||F''||_{L^2} \times H^{1/2} \times \frac{H}{2} \quad .$$

VII - ESPERANCE DES ESTIMATEURS SPLINES.-

Si $t \in [x_{i-1}, x_i]$

$$ES(t) = EM_{j-1}\frac{(x_j-t)^3}{6h_j} + EM_j\frac{(t-x_{j-1})^3}{6h_j} + EM_{j-1}\frac{h_j(t-x_j)}{6}$$

$$- EM_j h_j \frac{(t-x_{j-1})}{6} + \frac{Ey_j(t-x_{j-1}) - Ey_{j-1}(t-x_j)}{h_j}$$

$$ES'(t) = - EM_{j-1}\frac{(x_j-t)^2}{2h_j} + EM_j\frac{(t-x_{j-1})^2}{2h_j} + \frac{Ey_j - Ey_{j-1}}{h_j} - \frac{EM_j - EM_{j-1}}{6}h_j$$

où les $(EM_j)_{0\leq j\leq N}$ vérifient dans chaque cas les systèmes obtenus à partir de $(\sum_1)$ et $(\sum_2)$ en remplaçant d_j par Ed_j $0 \leq j \leq N$.

Calcul de Ed_j $0 \leq j \leq N$.

si $1 \leq j \leq N-1$ $\quad d_j = \frac{6}{h_j + h_{j+1}}\left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j}\right)$

$$Ed_j = \frac{6}{h_j + h_{j+1}}\left(\frac{Ey_{j+1} - Ey_j}{h_{j+1}} - \frac{Ey_j - Ey_{j-1}}{h_j}\right)$$

Pour $0 \leq j \leq N$ $\quad y_j = \frac{1}{n}\sum_{i=1}^{n} 1_{[0,x_j]}(X_i) = F_n(x_j)$

$$Ey_j = F(x_j) \ .$$

Pour le cas CL_1

$$Ed_o = \frac{6}{h_1}\left(\frac{Ey_1 - Ey_o}{h_1} - E\hat{a}\right)$$

$$Ed_N = \frac{6}{h_N}\left(E\hat{b} - \frac{Ey_N - Ey_{N-1}}{h_N}\right)$$

Calcul de $E\hat{a}$ *et* $E\hat{b}$.

Soit $\ell_{o,\nu}$ le polynôme de degré m interpolant 1 en
$0\leq\nu\leq m$

x_ν et 0 en x_j $\begin{array}{l} 0 \leq j \leq m \\ j \neq \nu \end{array}$

et $\ell_{1,\nu}$ le polynôme de degré m interpolant 1 en
$N-m\leq\nu\leq N$

x_ν et 0 en x_j ; $\begin{array}{l} N-m \leq j \leq N \\ j \neq \nu \end{array}$

$$\hat{a} = \sum_{\nu=0}^{m} \ell'_{0,\nu}(0) \quad F_n(x_\nu) \qquad\qquad \hat{b} = \sum_{\nu=N-m}^{N} \ell'_{1,\nu}(1) \quad F_n(x_\nu)$$

Donc $$E\hat{a} = \sum_{\nu=0}^{m} \ell'_{0,\nu}(0) \quad F(x_\nu) \qquad\qquad E\hat{b} = \sum_{\nu=N-m}^{N} \ell'_{1,\nu}(1) \; F(x_\nu) \; .$$

D'où le

Théorème VII.1.- Dans chacun des cas CL_1 et CL_2 l'espérance de la spline est la spline de même type interpolant F sur les mêmes noeuds.

l'espérance de la dérivée de la spline est la dérivée de la spline de même type interpolant F sur les mêmes noeuds.

On retrouve ainsi les résultats de V 2°).

Du théorème VI.3 on tire alors le :

Théorème VII.2.- Si $F \in H^2([0,1])$, si S est l'estimateur du type CL_2 de F, S est un estimateur uniformément asymptotiquement sans biais (E.U.A.S.B.) de F et S' est un E.U.A.S.B. de f.

Du théorème VI.2 on tire alors le

Théorème VII.3.- Si $F \in C^{m+1}([0,1])$, si S est du type CL_1 .

S est un E.U.A.S.B. de F.

S' est un E.U.A.S.B. de f.

Dans le cas où F est seulement de classe C^1 sur $[0,1]$, pour déduire du corollaire V.1 que S est un E.U.A.S.B. de F il suffit de majorer $|E\hat{a}|$ et $|E\hat{b}|$.

Majoration de $|\ell'_{0,\nu}(0)|$ $\quad 0 \leq \nu \leq m$

et $|\ell'_{1,\nu}(1)|$ $\quad N-m \leq \nu \leq N$.

$$\ell_{0,\nu}(t) = \prod_{\substack{0\leq j\leq m \\ j\neq\nu}} (t-x_j) \times \Big[\prod_{\substack{0\leq j\leq m \\ j\neq\nu}} (x_\nu - x_j)\Big]^{-1}$$

si $m \geq 2$
$$\ell'_{0,\nu}(t) = \Big[\prod_{\substack{0\leq j\leq m \\ j\neq\nu}} (x_\nu-x_j)\Big]^{-1} \sum_{\substack{0\leq k\leq m \\ k\neq\nu}} \Big(\prod_{\substack{0\leq j\leq m \\ j\neq k \\ j\neq\nu}} (t-x_j)\Big)$$

$$|\ell'_{0,\nu}(0)| \leq \prod_{\substack{0\leq j\leq m \\ j\neq\nu}} |x_\nu-x_j|^{-1} \sum_{\substack{0\leq k\leq m \\ k\neq\nu}} \underbrace{\Big(\prod_{\substack{0\leq j\leq m \\ j\neq k \\ j\neq\nu}} x_j\Big)}_{\text{produit nul sauf si } \nu = 0 \text{ ou } k = 0}$$

$$|x_\nu - x_j| \geq |\nu-j| \; h \geq h \quad \text{si } \nu \neq j$$

$$|x_j| \leq j\,H .$$

α) si $\nu \neq 0$.

$$|\ell'_{0,\nu}(0)| \leq \frac{1}{h^m} H^{m-1} \; m!$$

β) <u>si $\nu = 0$</u> .

$$|\ell'_{0,0}(0)| \leqslant \left[h^m \prod_{0<j<m} j\right]^{-1} \sum_{0<k\leqslant m} \left(\prod_{\substack{0<j\leqslant m \\ j\neq k}} x_j \right)$$

$$|\ell'_{0,0}(0)| \leqslant m \frac{H^{m-1}}{h^m} .$$

<u>si $m = 1$</u> $\quad \ell_{0,0}(x) = \dfrac{x - x_1}{- x_1} \qquad \ell_{0,1}(x) = \dfrac{x}{x_1}$

$$|\ell'_{0,0}(0)| = \frac{1}{x_1} \leqslant \frac{1}{h} \qquad |\ell'_{0,1}(0)| = \frac{1}{x_1} \leqslant \frac{1}{h}$$

et donc $\quad |\ell'_{0,\nu}(0)| \leqslant m! \dfrac{H^{m-1}}{h^m} \leqslant \dfrac{C_o^{m-1} \; m!}{h}$

$$\forall m \geqslant 1 \qquad \forall \nu \qquad 0 \leqslant \nu \leqslant m .$$

Pour $N-m \leqslant \nu \leqslant N$ soit $L_{0,N-\nu}$ défini comme $\ell_{0,N-\nu}$ mais sur les noeuds $0 , 1 - x_{N-1} , 1 - x_{N-2} ,\ldots, 1 - x_{N-m}$ on a alors

$$\ell_{1,\nu}(t) = L_{0,N-\nu}(1-t)$$

$$\ell'_{1,\nu}(t) = - L'_{0,N-\nu}(1-t)$$

et $\quad \ell'_{1,\nu}(1) = - L'_{0,N-\nu}(0)$

On a donc

$$|\ell'_{1,\nu}(1)| \leqslant m! \frac{H^{m-1}}{h^m} \quad \frac{C_o^{m-1} \; m!}{h}$$

$$\forall m \geqslant 1 \qquad \forall \nu \qquad N-m \leqslant \nu \leqslant N .$$

Majoration de $|\hat{a}|$ *et* $|\hat{b}|$:

$$\hat{a} = \sum_{\nu=0}^{m} \ell'_{0,\nu}(0)\ F_n(x_\nu) \qquad \hat{b} = \sum_{\nu=N-m}^{N} \ell'_{1,\nu}(1)\ F_n(x_\nu)$$

comme $0 \leqslant F_n(t) \leqslant 1 \quad \forall t$

$$|\hat{a}| \leqslant (m+1)!\ \frac{H^{m-1}}{h^m} \leqslant \frac{C_o^{m-1}\ (m+1)!}{h}$$

$$|\hat{b}| \leqslant (m+1)!\ \frac{H^{m-1}}{h^m} \leqslant \frac{C_o^{m-1}\ (m+1)!}{h}$$

Si on pose $\qquad C_o^{m-1}\ (m+1)! = C \qquad (C \geqslant 1)$

on a $\qquad |\hat{a}| \leqslant \frac{C}{h}$ et $|\hat{b}| \leqslant \frac{C}{h}$.

Majoration de $|E(\hat{a})|$ *et* $|E(\hat{b})|$:

$|\hat{a}|$ ne peut prendre une valeur non nulle que si au moins l'un des X_i appartient à $]0,x_m]$.

$|\hat{b}|$ ne peut prendre une valeur non nulle que si au moins l'un des X_i appartient à $]x_{N-m}\ ,\ x_N]$.

$$P(\bigcup_{i=1}^{n} (X_i \in]0,\ x_m])) \leqslant \sum_{i=1}^{n} P(X_i \in]0,x_m]) \leqslant n\ (F(x_m) - F(x_o))$$

$$|E(\hat{a})| \leqslant E(|\hat{a}|) \leqslant \frac{C}{h} \times n \times mH\ M'$$

$$|E(\hat{a})| \leqslant n\ C\ C_o\ mM'$$

$$P\left(\bigcup_{i=1}^{n} (X_i \in]x_{N-m}, x_N])\right) \leqslant \sum_{i=1}^{n} P(X_i \in]x_{N-m}, x_N]) \leqslant n(F(x_N) - F(x_{N-m}))$$

$$|E(\hat{b})| \leqslant E(|\hat{b}|) \leqslant \frac{C}{h} \times n \times mHM'$$

$$|E(\hat{b})| \leqslant n\ C\ C_o\ mM'$$

VIII - MAJORATION DES NOYAUX SPLINES ET CONDITIONS SUFFISANTES DE CONVERGENCE DE L'ESTIMATEUR SPLINE (SUBDIVISION QUELCONQUE).-

1°) Majoration du noyau spline $K^1_\Delta(x,t)$ *(défini au III 3°).*

Pour $1 \leqslant j \leqslant N-1$ $\quad d_j = \dfrac{6}{h_j + h_{j+1}} \left(\dfrac{y_{j+1} - y_j}{h_{j+1}} - \dfrac{y_j - y_{j-1}}{h_j} \right)$

les éventualités pour y_{j-1}, y_j et y_{j+1} sont les suivantes :

y_{j-1}	y_j	y_{j+1}	$y_{j+1} - y_j$	$y_j - y_{j-1}$
0	0	0	0	0
0	0	1	1	0
0	1	1	0	1
1	1	1	0	0

et $|d_j| \leqslant \dfrac{3}{h^2}$

$$|d_o| \leqslant \frac{6}{h^2} + \frac{6\ |\hat{a}|}{h} \qquad |d_N| \leqslant \frac{6}{h^2} + \frac{6\ |\hat{b}|}{h}$$

$$|d_o| \leqslant \frac{6(C+1)}{h^2} \qquad |d_N| \leqslant \frac{6(C+1)}{h^2}$$

Donc $\sup\limits_{0\leq j\leq N} |M_j| \leqslant \dfrac{6(C+1)}{h^2}$ (lemme II.4).

et $\forall x \in [0,1]$ $\forall t \in [0,1]$ $\quad |K^1_\Delta(x,t)| \leqslant \dfrac{6(C+1)}{h^2}\,\dfrac{H^2}{h} + \dfrac{1}{h} + \dfrac{C+1}{h^2} \times 2H$

En posant $D = 6\ C_o^2(C+1) + 2\ C_o(C+1) + 1$

$$\sup_{\substack{x\in[0,1]\\ t\in[0,1]}} |K^1_\Delta(x,t)| \leqslant \frac{D}{h}$$

2°) *Majoration du noyau spline* $K^2_\Delta(x,t)$ (*défini au* III 3°).

De même

$\forall x \in [0,1]$ $\forall t \in [0,1]$ $\quad |K^2_\Delta(x,t)| \leqslant \dfrac{3}{h^2}\,\dfrac{H^2}{h} + \dfrac{1}{h} + \dfrac{1}{h^2}\,H$

$$\sup_{\substack{x\in[0,1]\\ t\in[0,1]}} |K^2_\Delta(x,t)| \leqslant \frac{D}{h}$$

$$D = 3\ C_o^2 + C_o + 1\ .$$

3°) *Théorème VIII.1.-*

Si $i \in \{1,2\}$ $\quad \sup\limits_{\substack{x\in[0,1]\\ t\in[0,1]}} |K^i_\Delta(x,t)| \leqslant \dfrac{D}{h}$

D est une constante ne dépendant que de C_o si $i = 2$

de C_o et m si $i = 1$.

4°) Notations.-

Pour tout $n \in \mathbb{N}^*$ on se donne une subdivision Δ_n de $[0,1]$

$$0 = x_o^n < x_1^n \ldots\ldots < x_{N_n}^n = 1$$

On pose $h_j^n = x_j^n - x_{j-1}^n \quad 1 \leqslant j \leqslant N_n \qquad h_n = \min_{1\leqslant j\leqslant N_n} h_j^n$

$$H_n = \max_{1\leqslant j\leqslant N_n} h_j^n$$

on fait l'hypothèse suivante $\begin{cases} \forall n \in \mathbb{N}^* \quad \dfrac{H_n}{h_n} \leqslant C_o \\ \\ \lim\limits_{n\to+\infty} h_n = 0 . \end{cases}$

On adoptera les notations de [9], [14] et [15] où des résultats très généraux sont obtenus notamment à partir des méthodes de [11], [6] et [22].

$\mathcal{D}_1$ est l'ensemble des densités f définies sur $[0,1]$, bornées et telles que $\lim\limits_{n\to+\infty} \int_0^1 K_{\Delta_n}(x,t)\, f(x)\, dx$ existe $\forall t \in [0,1]$.

Cette limite est notée $\bar{f}(t)$. $\hat{f}_n(t) = \frac{1}{n} \sum\limits_{i=1}^{n} K_{\Delta_n}(X_i,t)$.

On posera $D_n(t) = |\hat{f}_n(t) - \bar{f}(t)| \qquad t \in [0,1] \qquad n \in \mathbb{N}^*$

$$E(\hat{f}_n(t)) = f_n(t) = \int_0^1 K_{\Delta_n}(x,t)\, f(x)\, dx \qquad t \in [0,1] \quad n \in \mathbb{N}^*$$

$$\hat{S}_n(t) = \hat{f}_n(t) - f_n(t) \qquad (\text{donc } E(\hat{S}_n(t)) = 0)$$

$$\delta_n = \sup_{t\in[0,1]} |\hat{f}_n(t) - f(t)|$$

$$\mathcal{D}_2 = \{g \mid g \in \mathcal{D}_1 \text{ et } \lim_{n\to+\infty} \sup_{t\in[0,1]} \left|\int_0^1 K_{\Delta_n}(x,t)\, f(x)\, dx - f(t)\right| = 0\}$$

Remarques sur les ensembles $\mathcal{D}_1$ et $\mathcal{D}_2$:

$\int_0^1 K_{\Delta_n}(x,t)\, f(x)dx$ est l'espérance de la dérivée de l'estimateur spline de F.

Dans le cas CL_1, d'après les théorèmes V.2 et VI.1., on a, si F est de classe C^{m+1},

$$\sup_{t\in[0,1]} \left|\int_0^1 K^1_{\Delta_n}(x,t)\, f(x)\, dx - f(t)\right| = O(H)$$

et donc $$C^m([0,1]) \subset \mathcal{D}_2 \subset \mathcal{D}_1 \quad .$$

Dans le cas CL_2, d'après les théorèmes V.3 et VI.1 on a, si $F \in H^2([0,1])$

$$\sup_{t\in[0,1]} \left|\int_0^1 K^2_{\Delta_n}(x,t)\, f(x)\, dx - f(t)\right| = O(H^{1/2})$$

et donc $$H^1([0,1]) \subset \mathcal{D}_2 \subsetneq \mathcal{D}_1 \quad .$$

On peut aussi remarquer que, d'après les majorations de $|E(\hat{a})|$ et $|E(\hat{b})|$ et les théorèmes VI.1 et VI.3, si F est seulement de classe C^1 on a $\sup_{t\in[0,1]} \left|\int_0^1 S_{\Delta_n}(x,t)\, f(x)\, dx - F(t)\right| \xrightarrow[n\to+\infty]{} 0$ où $S_{\Delta_n}(x,.)$ est défini comme S^1_x ou S^2_x (§ III.3°).

Rappel des conditions de convergence figurant dans [15] *et utilisées dans la suite* : $\alpha \in \mathbb{R}^{+*}$.

I - (resp. I') Il existe une fonction réelle définie sur $[0,1]$ (resp. définie et bornée sur $[0,1]$) telle que

$$\sup_{x\in[0,1]} |K_{\Delta_n}(x,t)| \leq \frac{A_1(t)}{h_n^\alpha}$$

II - (resp. II') Il existe une fonction positive B sur $[0,1]$ (resp. définie et bornée sur $[0,1]$) telle que

$$\int_0^1 [K_{\Delta_n}(x,t)]^2\, dx \leq \frac{B(t)}{h_n^\alpha}$$

III - Il existe $g_1 \in \mathcal{D}_1$, il existe $t_1 \in [0,1]$, il existe $A_2 > 0$ tel que

$$\int_0^1 [K_{\Delta_n}(x,t_1)]^2 \, g_1(x) \, dx \geq \frac{A_2}{h_n^\alpha}$$

IV - Il existe $f_o \in \mathcal{D}_1$, il existe $t_o \in [0,1]$ et une famille (B_{Δ_n}) d'ensembles λ-mesurables tels que ;

a) $\bar{f}_o(t_o) > 0$

b) $\lambda(B_{\Delta_n}) = \ell \, h_n^\alpha (\ell > 0)$ λ : mesure de Lebesgue sur $[0,1]$

et

$$\int_{B_{\Delta_n}} K_{\Delta_n}(x,t_o) \, f_o(x) \, dx \geq \delta > 0 \quad .$$

V - $$\sup_{x \in [0,1]} |K_{\Delta_n}(x,t') - K_{\Delta_n}(x,t)| \leq \frac{C}{h_n^{m_1}} |t'-t|^\gamma$$

$t \in [0,1]$ $t' \in [0,1]$

$C > 0$ $m_1 > 0$ $\gamma > 0$ (condition de Lipschitz).

VI - Est une condition sur l'espace des observations vérifiée par $[0,1]$.

5°) *Conditions suffisantes de convergence ponctuelle.*

$$\int_0^1 [K_{\Delta_n}(x,t)]^2 \, dx \leq \frac{D^2}{h_n^2}$$ (condition II de [15] vérifiée avec $\alpha = 2$).

Les démonstrations de [9] s'appliquent et on a le théorème suivant :

Théorème 1.- Soit $t \in [0,1]$ $f \in \mathcal{D}_1$.

Si $\lim\limits_{n\to+\infty} n\, h_n^2 = +\infty$ alors

a) $E\left[\left[D_n(t)\right]^2\right] \xrightarrow[n\to+\infty]{} 0$

et donc b) $E(D_n(t)) \xrightarrow[n\to+\infty]{} 0$.

D'autre part $\sup\limits_{x\in[0,1]} |K_{\Delta_n}(x,t)| \leqslant \frac{D}{h_n} \leqslant \frac{D}{h_n^2}$

(condition I de (1) vérifiée pour $\alpha = 2$) et donc :

Théorème 2.- Si pour tout $\beta > 0$

$$\sum_{n\in\mathbb{N}^*} \exp(-\beta\, n\, h_n^2) < +\infty$$

alors $D_n(t) \xrightarrow[n\to+\infty]{} 0$ presque complètement

$$E((D_n(t))^2) \xrightarrow[n\to+\infty]{} 0$$

et $E(D_n(t)) \xrightarrow[n\to+\infty]{} 0$.

la démonstration de ce théorème est basée sur une inégalité de Hoeffding [9] et met en évidence la majoration :

$$P(\hat{S}_n(t) \geqslant \varepsilon) \leqslant \exp(-\beta_1\, \varepsilon\, n\, h_n^2)$$

où β_1 est une constante, et ceci pour ε suffisamment petit.

6°) *Condition suffisante de convergence uniforme.*

Théorème.- Si $f \in \mathcal{D}_2$

et si
$$\frac{\operatorname{Log} n}{n\, h_n^2} \xrightarrow[n\to+\infty]{} 0 \quad .$$

Alors a) $\delta_n \xrightarrow[n\to+\infty]{} 0$ presque complètement.

et b) $E\,\delta_n \xrightarrow[n\to+\infty]{} 0$.

La démonstration de [9] s'adapte immédiatement :

$$\frac{d}{dt} K_{\Delta_n}(x,t) = M_{j-1} \frac{x_j - t}{h_j} + M_j \frac{t - x_{j-1}}{h_j} \quad \text{si} \quad t \in [x_{j-1}, x_j]$$

donc
$$\left|\frac{d}{dt} K_{\Delta_n}(x,t)\right| \leqslant \frac{12(C+1)}{h_n^2} \qquad \text{(page 69)}$$

donc
$$\sup_{x\in[0,1]} |K_{\Delta_n}(x,t') - K_{\Delta_n}(x,t)| \leqslant \frac{12(C+1)}{h_n^2} |t'-t|$$

et les conditions I', II', V, VI de [15] sont vérifiées.

IX - CAS DES SUBDIVISIONS UNIFORMES.-

Dans ce cas les matrices des systèmes $(\sum_1)$ et $(\sum_2)$ sont des cas particuliers de la matrice régulière :

$$M_n(\lambda,\mu) = \begin{bmatrix} 2 & \lambda & & & & & \\ \frac{1}{2} & 2 & \frac{1}{2} & & & \bigcirc & \\ & \frac{1}{2} & 2 & \frac{1}{2} & & & \\ & & \ddots & \ddots & \ddots & & \\ & \bigcirc & & \frac{1}{2} & 2 & \frac{1}{2} & \\ & & & & \frac{1}{2} & 2 & \frac{1}{2} \\ & & & & & \mu & 2 \end{bmatrix} \quad \begin{matrix} (|\lambda| < 2 \\ |\mu| < 2 \,) \end{matrix}$$

(d'ordre n).

dont on se propose de calculer l'inverse ([1]) .

Pour $n \in \mathbb{N}^*$ on pose $D_n = \det M_n(\frac{1}{2}, \frac{1}{2})$

$$D_o = 1$$

si $n \geqslant 2$ $\qquad D_n = 2\, D_{n-1} - \frac{1}{4} D_{n-2}$

équation aux différences finies dont l'équation caractéristique est

$$X^2 - 2X + \frac{1}{4} = 0 \qquad \text{d'où} \qquad X_1 = 1 - \frac{\sqrt{3}}{2}$$

$$X_2 = 1 + \frac{\sqrt{3}}{2}$$

Donc $\qquad D_n = \alpha\, X_1^n + \beta\, X_2^n$.

$$\left.\begin{array}{l} D_o = 1 \\ D_1 = 2 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} \alpha + \beta = 1 \\ \alpha(1 - \frac{\sqrt{3}}{2}) + \beta(1 + \frac{\sqrt{3}}{2}) = 2 \end{array}\right.$$

$$\alpha = -\frac{1}{\sqrt{3}}(1 - \frac{\sqrt{3}}{2})$$

$$\beta = \frac{1}{\sqrt{3}}(1 + \frac{\sqrt{3}}{2})$$

D'où $D_n = -\frac{1}{\sqrt{3}}(1 - \frac{\sqrt{3}}{2})^{n+1} + \frac{1}{\sqrt{3}}(1 + \frac{\sqrt{3}}{2})^{n+1}$

On pose $$r = 1 + \frac{\sqrt{3}}{2}$$

$$s = 1 - \frac{\sqrt{3}}{2}$$

$\forall n \in N$ $$D_n = \frac{r^{n+1} - s^{n+1}}{\sqrt{3}} \qquad (D_n > 0)$$

Pour $\lambda \in \mathbb{R}$, $n \in \mathbb{N}^*$ on pose $Q_n(\lambda) = \det M_n(\lambda, \frac{1}{2})$

(on a aussi $Q_n(\lambda) = \det M_n(\frac{1}{2}, \lambda)$) .

si $n \geqslant 2$ $Q_n(\lambda) = 2\,D_{n-1} - \frac{\lambda}{2}\,D_{n-2}$

$Q_1(\lambda) = 2$ $\quad Q_2(\lambda) = 4 - \frac{\lambda}{2}$ $\quad Q_o(\lambda) = 2\lambda$.

La matrice des mineurs de $M_n(\lambda,\mu)$ est la matrice $(C_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$ déterminée par :

$$C_{ij} = (\frac{1}{2})^{j-i} \quad Q_{n-j}(\mu)\ Q_{i-1}(\lambda) \quad \text{pour} \quad 1 < i \leqslant j < n$$

$$\text{ou} \quad (1 < i < n \quad \text{et} \quad j = n)$$

$$C_{ij} = (\frac{1}{2})^{i-j} \quad Q_{j-1}(\lambda)\ Q_{n-i}(\mu) \quad \text{pour} \quad 1 < j \leqslant i < n$$

$$\text{ou} \quad (1 < i < n \quad \text{et} \quad j = 1)$$

$$C_{1j} = (\frac{1}{2})^{j-1} \quad Q_{n-j}(\mu) \qquad \forall j \quad 1 \leqslant j \leqslant n$$

$$C_{nj} = (\frac{1}{2})^{n-j} \quad Q_{j-1}(\lambda) \qquad \forall j \quad 1 \leqslant j \leqslant n\ .$$

Dans le cas CL_1 on a $\lambda = \mu = 1$ et on posera

$$Q_n(1) = E_n \quad \text{pour} \quad n \in \mathbb{N}\ .$$

Dans le cas CL_2 on a $\lambda = \mu = \frac{1}{2}$ et les résultats ci-dessus se condensent de la façon suivante : $(Q_n(\frac{1}{2}) = D_n)$.

$$\text{si} \quad 1 \leqslant i \leqslant j \leqslant n \qquad C_{ij} = (\frac{1}{2})^{j-i} \ D_{n-j}\ D_{i-1}$$

$$\text{si} \quad 1 \leqslant j \leqslant i \leqslant n \qquad C_{ij} = (\frac{1}{2})^{i-j} \ D_{j-1}\ D_{n-i}$$

on a alors $C_{ij} = C_{ji}$ (ce qui provient du fait que $M_n(\frac{1}{2}\ ,\ \frac{1}{2})$ est une matrice symétrique).

X - *CONDITIONS NECESSAIRES ET SUFFISANTES DE CONVERGENCE DANS LE CAS CL2 AVEC SUBDIVISIONS UNIFORMES.-*

si $t \in [x_{j-1}]$, si $x \in [x_{k-1}, x_k]$

$$K_\Delta^2(x,t) = - M_{j-1}^k \frac{(x_j-t)^2}{2h} + M_j^k \frac{(t-x_{j-1})^2}{2h} + \frac{y_j^k - y_{j-1}^k}{h} - \frac{M_j^k - M_{j-1}^k}{6} h$$

si $x = 0$ $\qquad K_\Delta^2(0,t) = 0 \qquad \forall\, t \in [0,1]$.

1°) *Expression des* M_j^k.

On notera $M_{CL_2}^{-1} = (A_{ij})_{\substack{1\leq i\leq N-1\\ 1\leq j\leq N-1}}$

si $2 \leq k \leq N-1$ seuls d_k et d_{k-1} sont non nuls $d_k = \frac{-3}{h^2}$, $d_{k-1} = \frac{3}{h^2}$

si $2 \leq j \leq N-1$ $\qquad M_j^k = (A_{j,k-1} - A_{j,k}) \times \frac{3}{h^2}$;

si $x \in]x_o, x_1]$ seul d_1 est non nul $d_1 = - \frac{3}{h^2}$

$M_j^1 = A_{j,1}\, d_1$;

si $x \in]x_{N-1}, x_N]$ seul d_{N-1} est non nul $d_{N-1} = \frac{3}{h^2}$

$M_j^N = A_{j,N-1}\, d_{N-1}$

si $i \leq j$, on a

$$A_{ij} = \frac{(-1)^{i+j}}{D_{N-1}} c_{ij} = \frac{(-1)^{i+j}}{D_{N-1}} \left(\frac{1}{2}\right)^{j-i} D_{N-1-j}\, D_{i-1}$$

et $A_{ij} = A_{ji} \quad \forall i \quad \forall j$.

Donc $\displaystyle A_{j,k-1} - A_{j,k} = \frac{(-1)^{j+k-1}}{D_{N-1}} (c_{j,k-1} + c_{j,k})$.

$$M_j^1 = \frac{(-1)^{j+1}}{D_{N-1}} \left(\frac{1}{2}\right)^{j-1} D_{N-1-j} \times \left(-\frac{3}{h^2}\right)$$

donc

$$(M_j^1)^2 = \frac{D_{N-1-j}^2}{D_{N-1}^2} \times \left(\frac{1}{4}\right)^{j-1} \times \frac{9}{h^4}$$

$$M_j^N = \frac{(-1)^{N+j-1}}{D_{N-1}} \left(\frac{1}{2}\right)^{N-j-1} D_{j-1} \times \frac{3}{h^2}$$

donc $$(M_j^N)^2 = \frac{D_{j-1}^2}{D_{N-1}^2} \left(\frac{1}{4}\right)^{N-j-1} \times \frac{9}{h^4}$$

Si $2 \leqslant k \leqslant N-1$ et $1 \leqslant j \leqslant N-1$

<u>Si $j \leqslant k-1$</u> $$c_{j,k-1} + c_{j,k} = \left(\frac{1}{2}\right)^{k-j} D_{j-1}(2D_{N-k} + D_{N-k-1})$$

et $$M_j^k = \frac{3}{h^2} = \frac{(-1)^{j+k-1}}{D_{N-1}} \left(\frac{1}{2}\right)^{k-j} D_{j-1}(2D_{N-k} + D_{N-k-1})$$

donc $$(M_j^k)^2 = \frac{9}{h^4} \frac{D_{j-1}^2}{D_{N-1}^2} \left(\frac{1}{4}\right)^{k-j} (2D_{N-k} + D_{N-k-1})^2.$$

<u>Si $j \geqslant k$</u> $$c_{j,k-1} + c_{j,k} = \left(\frac{1}{2}\right)^{j-k+1} D_{N-1-j}(2D_{k-1} + D_{k-2})$$

et $$M_j^k = \frac{3}{h^2} \frac{(-1)^{j+k-1}}{D_{N-1}} \left(\frac{1}{2}\right)^{j-k+1} D_{N-1-j}(2D_{k-1} + D_{k-2})$$

donc $$(M_j^k)^2 = \frac{9}{h^4} \frac{D_{N-1-j}^2}{D_{N-1}^2} \left(\frac{1}{4}\right)^{j-k+1} (2D_{k-1} + D_{k-2})^2$$

2°) *Minoration de* $\int_o^h K_\Delta^2(x,0)\,dx$.

si $x \in]0,h]$ $\quad K_\Delta^2(x,0) = \frac{1}{h} - \frac{M_1^1 h}{6}$

$$A_1^1 = \frac{D_{N-2}}{D_{N-1}} \quad \text{et} \quad M_1^1 = -\frac{D_{N-2}}{D_{N-1}} \times \frac{3}{h^2}$$

donc $\quad K_\Delta^2(x,0) = \frac{1}{h} + \frac{D_{N-2}}{D_{N-1}} \times \frac{1}{2h} > \frac{1}{h}$

d'où $$\int_o^h K_\Delta^2(x,0)\,dx > 1$$

et la condition IV de [15] est vérifiée. La condition III l'est donc aussi puisqu'en appliquant l'inégalité de Cauchy-Schwarz, on obtient :

$$1 < h \times \int_o^1 [K_\Delta^2(x,0)]^2\,dx.$$

Nous allons voir que cette condition III est en fait vérifiée pour tout $t \in [0,1]$, avec $g_1 \equiv 1$ $\quad A_2 = \frac{1}{4}$, ceci pour n assez grand.

3°) *Etude de* $I(t) = h \int_o^1 [K_\Delta^2(x,t)]^2 dx.$

si $t \in [x_{j-1}, x_j]$

$$I(t) = h^2 \sum_{k=1}^{N} \left(M_{j-1}^k \frac{h^2 - 3(x_j - t)^2}{6h} + M_j^k \frac{3(t-x_{j-1})^2 - h^2}{6h} + \frac{y_j^k - y_{j-1}^k}{h} \right)^2$$

a) Calcul de $I(0)$:

si $x \in]x_{k-1}, x_k]$

$$K_\Delta^2(x,0) = \frac{y_1^k - y_o^k}{h} - \frac{M_1^k h}{6}$$

$$\text{si} \quad k = 1 \qquad K_\Delta^2(x,0) = \frac{1}{h} - \frac{M_1^1 h}{6}$$

$$\text{si} \quad k > 1 \qquad K_\Delta^2(x,0) = -\frac{M_1^k h}{6}$$

$$I(0) = h^2\left(\frac{1}{h^2} - \frac{2}{h}\,\frac{M_1^1 h}{6} + \sum_{k=1}^{N} \frac{(M_1^k)^2 h^2}{36}\right)$$

$$I(0) = 1 - \frac{M_1^1 h^2}{3} + \frac{h^4}{36} \sum_{k=1}^{N} (M_1^k)^2$$

$$I(0) = 1 + \frac{D_{N-2}}{D_{N-1}} + \frac{h^4}{36} \sum_{k=1}^{N} (M_1^k)^2$$

$$\text{car} \quad M_1^1 = -\frac{D_{N-2}}{D_{N-1}} \times \frac{3}{h^2}\,.$$

b) <u>Calcul de</u> $I(1)$:

si $x \in \left]x_{k-1}, x_k\right]$

$$K_\Delta^2(x,1) = \frac{y_N^k - y_{N-1}^k}{h} + \frac{M_{N-1}^k\, h}{6}$$

$$\text{si} \quad k = N \qquad K_\Delta^2(x,1) = \frac{1}{h} + \frac{M_{N-1}^N h}{6}$$

$$\text{si} \quad k < N \qquad K_\Delta^2(x,1) = \frac{M_{N-1}^k h}{6}$$

$$I(1) = h^2\left(\frac{1}{h^2} + \frac{2}{h}\,\frac{M_{N-1}^N h}{6} + \sum_{k=1}^{N} \frac{(M_{N-1}^k)^2 h^2}{36}\right)$$

$$I(1) = 1 + \frac{M_{N-1}^N h^2}{3} + \frac{h^4}{36} \sum_{k=1}^{N} (M_{N-1}^k)^2$$

$$I(1) = 1 + \frac{D_{N-2}}{D_{N-1}} + \frac{h^4}{36} \sum_{k=1}^{N} (N_{N-1}^k)^2$$

$$\text{car} \quad M_{N-1}^{N} = \frac{D_{N-2}}{D_{N-1}} \times \frac{3}{h^2}$$

Si $t \in]0,1[$ on aura, pour n assez grand,

$$t \in [x_{j-1}, x_j] \quad \text{avec} \quad 2 \leq j \leq N-2$$

$$\text{et} \quad I(t) = h^2 \sum_{k=1}^{N} \left(M_{j-1}^k \frac{h^2 - 3(x_j - t)^2}{6h} + M_j^k \frac{3(t - x_{j-1})^2 - h^2}{6h} + \frac{y_j^k - y_{j-1}^k}{h} \right)^2$$

$$I(t) = h^2 \left[\sum_{k=1}^{N} (M_{j-1}^k)^2 \left(\frac{h^2 - 3(x_j - t)^2}{6h} \right)^2 + (M_j^k)^2 \left(\frac{3(t - x_{j-1})^2 - h^2}{6h} \right)^2 \right.$$

$$\left. + 2M_{j-1}^k M_j^k \left(\frac{h^2 - 3(x_j - t)^2}{6h} \right) \left(\frac{3(t - x_{j-1})^2 - h^2}{6h} \right) \right]$$

$$+ 1 + 2h \left[M_{j-1}^j \frac{h^2 - 3(x_j - t)^2}{6h} + M_j^j \frac{3(t - x_{j-1})^2 - h^2}{6h} \right]$$

c) <u>Notations</u>.

On pose $u = \dfrac{s}{r} = \dfrac{1 - \frac{\sqrt{3}}{2}}{1 + \frac{\sqrt{3}}{2}}$.

On a alors $\forall n \in \mathbb{N}$ $\quad D_n = \dfrac{r^{n+1} - s^{n+1}}{\sqrt{3}} = \dfrac{r^{n+1}}{\sqrt{3}} (1 - u^{n+1})$

on pose $T(n,p) = \dfrac{1 - u^{n-p+1}}{1 - u^{n+1}}$

$$\frac{D_m}{D_n} = r^{m-n} \frac{1 - u^{m+1}}{1 - u^{n+1}}$$

$$\frac{D_m}{D_n} = \frac{1}{r^{n-m}} T(n, n-m).$$

On a de façon évidente $\lim\limits_{n\to+\infty} u^n = 0$

et $\forall\, p \quad \lim\limits_{n\to+\infty} T(n,p) = 1.$

<u>Calcul de</u> $(2D_n + D_{n-1})^2$.

$$2D_n + D_{n-1} = \frac{1}{\sqrt{3}}\,(r^n(2r+1) - s^n(2s+1)).$$

Comme $rs = \frac{1}{4}$ (on a $r^2 = \frac{7}{4} + \sqrt{3}$ et $s^2 = \frac{7}{4} - \sqrt{3}$),

$$(2D_n + D_{n-1})^2 = \frac{1}{3}\left(r^{2n}(2r+1)^2 + s^{2n}(2s+1)^2 - 2\,\left(\frac{1}{4}\right)^n(2r+1)(2s+1)\right)$$

d) <u>Etude de</u> $\frac{h^4}{9}\sum\limits_{k=1}^{N}(M_1^k)^2$.

$$\frac{h^4}{9}\sum_{k=1}^{N}(M_1^k)^2 = \frac{D_{N-2}^2}{D_{N-1}^2} + \frac{1}{D_{N-1}^2}\left(\frac{1}{4}\right)^{N-2}$$

$$+ \sum_{k=2}^{N-1}\frac{1}{D_{N-1}^2}\left(\frac{1}{4}\right)^{k-1}\frac{1}{3}\left(r^{2(N-k)}(2r+1)^2 + s^{2(N-k)}(2s+1)^2 - 2\left(\frac{1}{4}\right)^{N-k}(2r+1)(2s+1)\right)$$

$$\frac{h^4}{9}\sum_{k=1}^{N}(M_1^k)^2 = \frac{1}{r^2}\,T(N-1,1)^2 + \frac{1}{D_{N-1}^2}\left(\frac{1}{4}\right)^{N-2} + \frac{1}{3D_{N-1}^2}(2r+1)^2\times r^{2(N-1)}\times\frac{1-\left(\frac{1}{4r^2}\right)^{N-2}}{4r^2-1}$$

$$+\frac{1}{3D_{N-1}^2}(2s+1)^2\, s^{2(N-1)}\,\frac{1-\left(\frac{1}{4s^2}\right)^{N-2}}{4s^2-1} - \frac{2}{3D_{N-1}^2}\times\frac{N-2}{4^{N-1}}\times(2r+1)\times(2s+1)$$

tenant compte du fait que $\frac{1}{D_{N-1}^2} = \frac{3}{r^{2N}}\times\frac{1}{(1-u^N)^2}$

on obtient $\lim\limits_{n\to+\infty}\frac{h^4}{9}\sum\limits_{k=1}^{N}(M_1^k)^2 = \frac{4}{r(2r-1)}$.

On en déduit immédiatement

$$\lim_{n\to+\infty} I(0) = \frac{2r+1}{2r-1} = \sqrt{3}.$$

e) Etude de $\frac{h^4}{9}\sum_{k=1}^{N}(M_{N-1}^{k})^2$.

De la même façon, on montre que

$$\lim_{n\to+\infty}\frac{h^4}{9}\sum_{k=1}^{N}(M_{N-1}^{k})^2 = \frac{4}{r(2r-1)}.$$

On en déduit immédiatement

$$\lim_{n\to+\infty} I(1) = \frac{2r+1}{2r-1} = \sqrt{3}.$$

f) Etude de $\frac{h^4}{9}\sum_{k=1}^{N}(M_j^{k})^2 \qquad 2 \leqslant j \leqslant N-2.$

j est tel que $t \in [x_{j-1},x_j]$, il tend donc vers l'infini avec n.

$$\frac{h^4}{9}\sum_{k=2}^{j}(M_j^k)^2 = \sum_{k=2}^{j}\frac{1}{D_{N-1}^2}\left(\frac{1}{4}\right)^{j-k+1}D_{N-1-j}^2 \times\frac{1}{3}\times\Big(r^{2(k-1)}(2r+1)^2+s^{2(k+1)}(2s+1)^2 - 2\left(\frac{1}{4}\right)^{k-1}(2r+1)(2s+1)\Big)$$

$$\frac{h^4}{9}\sum_{k=2}^{j}(M_j^k)^2 = \frac{1}{3}\times\frac{1}{r^{2j}}\,T(N-1,j)^2\left[(2r+1)^2\times\frac{\left(\frac{1}{4r^2}\right)^{j-1}-1}{1-4r^2}\times r^{2j}\right.$$

$$\left. + (2s+1)^2\times\left(\frac{1}{4}\right)^{j+1}\times 4^2\times s^2\times\frac{1-(4s^2)^{j-1}}{1-4s^2} - (2r+1)(2s+1)\times\frac{j-1}{4^j}\right]$$

$$\lim_{n\to+\infty}\frac{h^4}{9}\sum_{k=2}^{j}(M_j^k)^2 = \frac{2r+1}{3(2r-1)} = \frac{\sqrt{3}}{3}$$

Comme $\lim\limits_{n\to+\infty} (M_j^1) \dfrac{h^4}{9} = 0$

et que $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} (M_j^j)^2 = \dfrac{1}{12} \dfrac{(2r+1)^2}{r^2}$

on a $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=1}^{j} (M_j^k)^2 = \dfrac{\sqrt{3}}{3}$

et $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=1}^{j-1} (M_j^k)^2 = \dfrac{\sqrt{3}}{12r^2}$

$$\frac{h^4}{9} \sum_{k=j+1}^{N-1} (M_j^k)^2 = \sum_{k=j+1}^{N-1} \frac{1}{D_{N-1}^2} \left(\frac{1}{4}\right)^{k-j} D_{j-1}^2 \times \frac{1}{3} \times \left(r^{2(N-k)}(2r+1)^2 + s^{2(N-k)}(2s+1)^2 - 2\left(\frac{1}{4}\right)^{N-k}(2r+1)(2s+1)\right)$$

$$\frac{h^4}{9} \sum_{k=j+1}^{N-1} (M_j^k)^2 = \frac{1}{3} \times \frac{1}{r^{2(N-j)}} T(N-1,N-j)^2 \left((2r+1)^2 \times \frac{1}{4} \times r^{2(N-j-1)} \times \frac{1-\left(\frac{1}{4r^2}\right)^{N-j-1}}{1-\frac{1}{4r^2}} + (2s+1)^2 \times \frac{1}{4} \times s^{2(N-j-1)} \frac{1-\left(\frac{1}{4s^2}\right)^{N-j-1}}{1-\frac{1}{4s^2}} - 2(2r+1)(2s+1)\frac{N-j-1}{4^{N-j}}\right)$$

et $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=j+1}^{N-1} (M_j^k)^2 = \dfrac{2r+1}{3(2r-1)} = \dfrac{\sqrt{3}}{3}$.

Comme $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} (M_j^N)^2 = 0$

on a $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=j+1}^{N} (M_j^k)^2 = \dfrac{\sqrt{3}}{3}$.

D'où $\lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=1}^{N} (M_j^k)^2 = \dfrac{2\sqrt{3}}{3}$

et pour $3 \leqslant j \leqslant N-1$ $\quad \lim\limits_{n\to+\infty} \dfrac{h^4}{9} \sum\limits_{k=1}^{N} (M_{j-1}^k)^2 = \dfrac{2\sqrt{3}}{3}$.

g) <u>Calcul de</u> $Q = \lim_{n\to+\infty} \frac{h^4}{9} \sum_{k=1}^{N} M_{j-1}^k M_j^k$

<u>pour $3 \leq j \leq N-2$</u>

De l'expression des (M_j^k), il résulte que

<u>si $k \leq j-1$</u> :

$$M_{j-1}^k = \mp M_j^k \times 2 \times \frac{D_{N-j}}{D_{N-1-j}}$$

<u>si $k = j$</u> :

$$M_j^j = -\frac{3}{h^2} \times \frac{1}{2} \times \frac{D_{N-1-j}}{D_{N-1}} (2D_{j-1} + D_{j-2})$$

$$M_{j-1}^j = \frac{3}{h^2} \times \frac{1}{2} \times \frac{D_{j-2}}{D_{N-1}} (2D_{N-j} + D_{N-j-1}).$$

<u>si $k \geq j+1$</u> :

$$M_{j-1}^k = -M_j^k \times \frac{1}{2} \times \frac{D_{j-2}}{D_{j-1}}$$

et $$\sum_{k=1}^{N} M_{j-1}^k M_j^k = -2 \frac{D_{N-j}}{D_{N-1-j}} \sum_{k=1}^{j-1} (M_j^k)^2 + M_{j-1}^j M_j^j - \frac{1}{2} \frac{D_{j-2}}{D_{j-1}} \sum_{k=j+1}^{N} (M_j^k)^2$$

et $$Q = -\frac{2}{r} \frac{\sqrt{3}}{12r^2} + \lim_{n\to\infty} \frac{h^4}{9} M_{j-1}^j M_j^j - \frac{1}{2r} \times \frac{\sqrt{3}}{3}$$

$$\lim_{n\to+\infty} \frac{h^4}{9} M_{j-1}^j M_j^j = -\frac{1}{12} \frac{(2r+1)^2}{r^2}.$$

Donc $$Q = \frac{2\sqrt{3} - 6}{3}.$$

En posant $$T_1 = \frac{h^4}{9} \sum_{k=1}^{N} (M_{j-1}^k)^2 \left(\frac{h^2 - 3(x_j - t)^2}{6h}\right)^2 \times \frac{9}{h^2}$$

$$+ \frac{h^4}{9} \sum_{k=1}^{N} (M_j^k)^2 \left(\frac{h^2 - 3(x_{j-1} - t)^2}{6h}\right)^2 \times \frac{9}{h^2}$$

$$T_2 = \left(\frac{h^4}{9} \sum_{k=1}^{N} M_{j-1}^k M_j^k\right) \times 2 \times \frac{9}{h^2} \left(\frac{h^2 - 3(x_j - t)^2}{6h}\right) \left(\frac{3(t - x_{j-1})^2 - h^2}{6h}\right)$$

$$T_3 = 2h\left[M_{j-1}^j \frac{h^2 - 3(x_j-t)^2}{6h} + M_j^j \frac{3(t-x_{j-1})^2 - h^2}{6h}\right]$$

On a : $I(t) = T_1 + T_2 + T_3 + 1$.

h) Encadrement de T_3.

$$\lim_{n\to+\infty} h^2 M_{j-1}^j = \frac{\sqrt{3}}{2}\left(2 + \frac{1}{r}\right) = \ell = 3(\sqrt{3} - 1).$$

$$\lim_{n\to+\infty} h^2 M_j^j = -\ell \quad .$$

Soit $\varepsilon > 0$, $\varepsilon < 0{,}01$

pour n assez grand $\ell - \varepsilon < h^2 M_{j-1}^j < \ell + \varepsilon$

$$-\ell - \varepsilon < h^2 M_j^j < -\ell + \varepsilon.$$

$$T_3 = \frac{M_{j-1}^j h^2 - M_j^j h^2}{3} + M_j^j (t-x_{j-1})^2 - M_{j-1}^j (x_j-t)^2$$

$$\frac{2(\ell-\varepsilon)}{3} < \frac{M_{j-1}^j h^2 - M_j^j h^2}{3} < \frac{2(\ell+\varepsilon)}{3}$$

$$\frac{(-\ell-\varepsilon)\left[(t-x_{j-1})^2+(x_j-t)^2\right]}{h^2} < M_j^j(t-x_{j-1})^2 - M_{j-1}^j(x_j-t)^2 < \frac{\left[(x_j-t)^2+(t-x_{j-1})^2\right]}{h^2}(-\ell+\varepsilon)$$

Or $(t-x_{j-1})^2 + (x_j-t)^2 = 2\left(t - \frac{x_j+x_{j-1}}{2}\right)^2 + \frac{1}{2} h^2$

et

$$\frac{2(\ell-\varepsilon)}{3} - (\ell+\varepsilon) < T_3 < \frac{2(\ell+\varepsilon)}{3} + \frac{-\ell+\varepsilon}{2}$$

$$\frac{-\ell-5\varepsilon}{3} < T_3 < \frac{7\ell+\varepsilon}{6}$$

$$\frac{1}{4} < (2 - \sqrt{3}) - \frac{5\varepsilon}{3} < 1 + T_3 < \frac{7\ell+\varepsilon+6}{6} < 4.$$

i) <u>Minoration de</u> $I(t)$.

Si $T_2 \geq 0$, on a alors $I(t) > \frac{1}{4}$.

Soit $\varepsilon > 0$, $\varepsilon < 0{,}01$. Pour n assez grand

$$Q - \varepsilon < \frac{h^4}{9} \sum_{k=1}^{N} M_{j-1}^k \, M_j^k < Q + \varepsilon < 0.$$

$$Q + 2 - \varepsilon < \frac{h^4}{9} \sum_{k=1}^{N} (M_j^k)^2 < Q + 2 + \varepsilon$$

$$Q + 2 - \varepsilon < \frac{h^4}{9} \sum_{k=1}^{N} (M_{j-1}^k)^2 < Q + 2 + \varepsilon.$$

Si $T_2 < 0$, on a alors :

$$T_1 + T_2 > (Q+2-\varepsilon) \times \frac{9}{h^2} \left[\left(\frac{h^2 - 3(x_j - t)^2}{6h} \right)^2 + \left(\frac{3(t - x_{j-1})^2 - h^2}{6h} \right)^2 \right]$$

$$- (-Q+\varepsilon) \times \frac{9}{h^2} \times 2 \left(\frac{h^2 - 3(x_j - t)^2}{6h} \right) \left(\frac{3(t - x_{j-1})^2 - h^2}{6h} \right).$$

Or $Q + 2 - \varepsilon > - Q + \varepsilon > 0$.

Donc

$$T_1 + T_2 > (-Q+\varepsilon) \frac{9}{h^2} \left[\frac{h^2 - 3(x_j - t)^2}{6h} - \frac{3(t - x_{j-1})^2 - h^2}{6h} \right]^2 \geq 0.$$

Donc $I(t) > \frac{1}{4}$ pour n assez grand.

j) <u>Majoration de</u> $I(t)$.

Soit $\varepsilon > 0$.

Pour n assez grand

$I(t) \leq 2(Q+2+\varepsilon) + 2(|Q|+\varepsilon) + 4 = 8 + 4\varepsilon$.

k) Choix du pas pour une étude ponctuelle de $I(t)$.

Si on choisit h de la forme $\frac{t}{j-1}$ $(\forall n \quad x_{j-1} = t)$.

On a convergence de $I(t)$:

$$I(t) = \frac{h^4}{9} \sum_{k=1}^{N} (M_{j-1}^k)^2 + \frac{h^4}{9} \sum_{k=1}^{N} (M_j^k)^2 \times \frac{1}{4}$$

$$+ \frac{h^4}{9} \sum_{k=1}^{N} M_{j-1}^k M_j^k + 1 + 2h\left[-\frac{M_{j-1}^j h}{3} - \frac{M_j^j h}{6}\right]$$

et $$\lim_{n\to+\infty} I(t) = \frac{2\sqrt{3}}{3} + \frac{2\sqrt{3}}{12} + \frac{2\sqrt{3}}{3} - 2 + 1 - 2\left(\frac{\sqrt{3}}{6}\left(2 + \frac{1}{r}\right) - \frac{\sqrt{3}}{12}\left(2 + \frac{1}{r}\right)\right)$$

$$\lim_{n\to+\infty} I(t) = \frac{\sqrt{3}}{2},$$

Si on choisit h de la forme $\frac{t}{j - \frac{1}{2}}$ $\left(\forall n \quad t = \frac{x_j + x_{j-1}}{2}\right)$

$$\lim_{n\to+\infty} I(t) = \frac{13\sqrt{3}}{24} + \frac{7}{16}.$$

Si h est quelconque, $I(t)$ n'a, en général, pas de limite.

Les conditions de [15] étant vérifiées, on a les théorèmes suivants :

4°) *Théorèmes de convergence simple.*

a) *théorème 1.- convergence en moyenne d'ordre 1 et 2.*

Les 3 propositions suivantes sont équivalentes :

α) $n\, h_n \xrightarrow[n\to+\infty]{} +\infty$

β) $\forall \; f \in \mathcal{D}_1 \quad \forall \; t \in [0,1] \quad E[D_n(t)] \xrightarrow[n\to+\infty]{} 0$

γ) $\forall \; f \in \mathcal{D}_1 \quad \forall \; t \in [0,1] \quad E[D_n(t))^2] \xrightarrow[n\to+\infty]{} 0.$

b) *théorème 2.- convergence presque complète et en moyenne d'ordre 1 et 2.*

Les 3 propositions suivantes sont équivalentes :

α) $\forall \varepsilon > 0 \quad \sum_{n\in\mathbb{N}^*} \exp[-\varepsilon n h_n] < +\infty$

β) $\forall t \in [0,1] \quad \forall f \in \mathcal{D}_1 \quad D_n(t) \xrightarrow[n\to+\infty]{} 0$ presque complètement et $E(D_n(t)) \xrightarrow[n\to+\infty]{} 0$.

γ) $\forall t \in [0,1] \quad \forall f \in \mathcal{D}_1 \quad D_n(t) \xrightarrow[n\to+\infty]{} 0$ presque complètement et $E[(D_n(t))^2] \xrightarrow[n\to+\infty]{} 0$.

5°) *Cas où* $(\frac{1}{h_n})$ *est asymptotiquement concave.*

Définition.- On dira que la suite $(\frac{1}{h_n})_{n\in\mathbb{N}^*}$ est asymptotiquement concave si

(C) $\left\{\begin{array}{l}\text{il existe } a_1 > 0,\ b_1 > 0 \text{ et une fonction concave } g_1 \text{ telle} \\ \text{que } \dfrac{a_1}{h_n} \leq g_1(n) \leq \dfrac{b_1}{h_n} \text{ pour } n \text{ assez grand.}\end{array}\right.$

La condition (C) est en particulier vérifiée

pour $h_n = \dfrac{1}{n^{\alpha_1}} \quad \alpha_1 \in]0,1[$

pour $h_n = \dfrac{1}{(\text{Log } n)^{\beta_1}} \quad \beta_1 > 0$

etc...

Théorème (équivalence entre convergence simple et convergence uniforme).-

Si $(h_n)_{n\in\mathbb{N}^*}$ vérifie (C) les propositions suivantes sont équivalentes.

α) $\dfrac{\text{Log } n}{n\, h_n} \xrightarrow[n\to+\infty]{} 0.$

β) $\forall\, t \in [0,1]\quad \forall\, f \in \mathcal{D}_2 \quad D_n(t) \xrightarrow[n\to+\infty]{} 0$ presque complètement

et $E[D_n(t)] \xrightarrow[n\to+\infty]{} 0.$

γ) $\forall\, f \in \mathcal{D}_2 \qquad \delta_n \xrightarrow[n\to+\infty]{} 0$ presque complètement

et $E(\delta_n) \xrightarrow[n\to+\infty]{} 0.$

6°) *Vitesse de convergence.*

Théorème.- Si $f \in \mathcal{D}_2$, s'il existe $t_1 \in [0,1]$ et $E > 0$ tels que $\int_0^1 [K_{\Delta_n}^{\;2}(x,t_1)]^2 f(x)dx \geqslant \dfrac{E}{h_n}$ pour $n \in \mathbb{N}^*$, alors il existe $b_2 > 0$ tel que pour n assez grand

$$\exp[-\rho''\eta^2\, n\, h_n] \leqslant P(\delta_n \geqslant \eta) \leqslant 2\exp[-\rho'\eta\, n\, h_n]$$

pour $0 < \eta < b_2$ où ρ' et ρ'' sont des constantes positives ne dépendant que de la suite $(h_n)_{n\in\mathbb{N}^*}$ et de f.

De la minoration de $I(t)$, il résulte que cet encadrement est valable pour toute densité de $\mathcal{D}_2$ qui est minorée sur $[0,1]$ par un réel strictement positif.

7°) *Remarques.*

a) L'encadrement obtenu pour $I(t)$, qui sera utilisé pour une étude ultérieure de la variance, permet d'étendre les résultats du 5°) et du 6°) aux éléments de

$$\mathcal{D}_F = \{f \mid f \in \mathcal{D}_1 \text{ et } \lim_{n\to+\infty} \sup_{t\in F} \left| \int_0^1 K_{\Delta_n}^2(x,t)f(x)dx - f(t)\right| = 0\}$$

où F est un sous-ensemble relativement compact de $[0,1]$ tel que

* $\lambda(F) > 0$
* la condition IV est vérifiée pour $t_o \in F$ et $f_o \in \mathcal{D}_F$.

On remplace alors δ_n par $\delta_n^F = \sup_{t \in F} |\hat{f}_n(t) - f(t)|$ et on ne considère la convergence de $D_n(t)$ que pour $t \in F$.

b) Il semble que les conditions suffisantes données dans VIII (5°/ et 6°/) ne soient pas nécessaires et que les conditions nécessaires et suffisantes données dans la partie X pour le cas (CL_2) soient valables dans les cas (CL_1) et (CL_2) pour des subdivisions quelconques.

c) On trouvera dans [17] un théorème de convergence uniforme presque complète avec subdivision uniforme en dimension 2, obtenu à partir des résultats de [5] sur l'histogramme. Pour l'estimation en dimension 2, on pourra consulter [43] et son intéressante bibliographie

d) Les noyaux splines définis en III.3°) sont susceptibles de convenir à d'autres problèmes d'estimation : par exemple l'estimation des densités d'un processus stationnaire à temps continu ([18]).

XI - CONCLUSION.-

Nous avons vu que les méthodes splines s'intégraient dans un schéma général d'estimation étudié dans [9] et englobant les méthodes du noyau et des fonctions orthogonales. Nous ne reviendrons pas sur ces deux dernières méthodes qui ont été largement utilisées et ont fait l'objet de nombreuses études.

Il nous paraît important d'insister sur le fait que les méthodes splines utilisées jusqu'à présent ne sont pas à proprement parler des méthodes statistiques : elles consistent à interpoler la fonction de répartition empirique et à dériver, ce qui revient à interpoler l'histogramme par une courbe "lisse" , réaction habituelle du praticien qui se trouve en présence d'une courbe discontinue (remarque analogue dans le cas du regressogramme).

Du point de vue statistique, le fait de substituer à la fonction de répartition empirique la spline qui l'interpole se traduit par une perte d'information, sauf si les noeuds de la spline sont les valeurs des observations (pour chaque échantillon la subdivision est obtenue à partir de l'image de la statistique d'ordre). Ce dernier cas pour lequel les calculs théoriques sont compliqués par le caractère aléatoire des noeuds, peut conduire à des résultats pratiques très décevants, à cause de la "proximité" ou de "l'éloignement" relatifs de certains noeuds.

L'estimation par les splines n'est évidemment pas stricte (l'estimateur obtenu n'est pas en général une densité : on n'a pas nécessairement $\hat{f}_n \geq 0$ et $\int_0^1 \hat{f}_n(t)\, dt = 1$). Il est même très fréquent que l'estimateur prenne des valeurs négatives sur des intervalles relativement grands : le lecteur pourra s'en rendre compte en consultant [10] où sont représentés nombre d'"histosplines" et où l'on trouvera une discussion sur les programmes employés. (On pourra comparer avec les courbes de [36]). Ce défaut n'est cependant pas majeur, il est partagé par beaucoup d'estimateurs dont le noyau n'est pas positif et on peut y remédier aisément. Il est surtout constaté dans le cas où des sous-intervalles de $[0,1]$ contiennent très peu d'observations, par exemple des intervalles du type $[0,\alpha]$, $[\beta,1]$ pour une densité nulle en 0 à croissance "lente" sur $[0,\alpha]$ et à décroissance "lente" ves zéro sur $[\beta,1]$. On a alors un mauvais comportement de la spline aux bornes de l'intervalle.

Dans le cas inverse : croissance "rapide" sur $[0,\alpha]$ et décroissance "rapide" sur $[\beta,1]$ le comportement de la spline est en général bon. Dans le cas CL_1 l'estimation de $f(0)$ et de $f(1)$ avant lissage permet (théoriquement) une meilleure connaissance de la densité au voisinage des bornes, mais cette contrainte nuit quelque peu au lissage. L'étude de l'estimateur aux bornes mériterait d'ailleurs une étude plus poussée (voir [27] pour le cas de la "fenêtre mobile"), en particulier dans le cas où des observations tombent en dehors de $[0,1]$. Le problème de l'estimation sur la totalité de la droite réelle se pose également.

Pour terminer nous évoquerons le problème le plus important qui se pose à l'utilisateur : la détermination du pas de la subdivision. Les théorèmes de convergence nous renseignent sur le comportement asymptotique et la vitesse de convergence des estimateurs mais ne donnent pas d'indication précise sur le problème pratique suivant : étant donné un échantillon de taille n, quel pas choisir ? A défaut de connaître le pas idéal correspondant à un échantillon observé nous pensons que plusieurs essais seront nécessaires pour obtenir une bonne idée de la densité à estimer en comparant les courbes obtenues, compte tenu des phénomènes classiques du lissage : si le pas est trop grand, la courbe est très lisse mais en général très "lointaine" de la courbe idéale (sur-lissage) ; si le pas est trop petit, la courbe contient de nombreuses oscillations (sous-lissage). (Pour ces problèmes pratiques voir également [10]).

Malgré leurs qualités essentielles (simplicité de mise en oeuvre et obtention de courbes très régulières) les méthodes splines actuellement utilisées en statistique ne semblent pas devoir supplanter les autres méthodes qui ont fait leurs preuves. Il faut noter cependant que l'on est loin d'avoir exploité dans ce domaine l'ensemble des possibilités offertes par la théorie des splines.

BIBLIOGRAPHIE

[1] J.H. AHLBERG, E.N. NILSON et J.L. WALSH -
The Theory of splines and their Applications,
Academic Press, 1967.

[2] N. ARONSZAJN -
Théorie des noyaux reproduisants et appl.
dans Proceedings of the Cambridge Philosophical Society, 39, 1943,
p. 133-153.

[3] N. ARONSZAJN -
The Theory of Reproducing kernels,
Transactions of the American Mathematical Society, 68, 1950, p. 307-404.

[4] A. BERLINET -
Sur les méthodes splines en estimation de la densité,
Comptes Rendus, 288, série A, 1979, p. 847-850.

[5] M. BERTRAND-RETALI -
Convergence uniforme stochastique d'un estimateur d'une densité de
probabilité dans R^s,
Comptes rendus, 278, série A, 1974, p. 451.

[6] M. BERTRAND-RETALI et J. GEFFROY -
Sur la convergence uniforme des estimateurs d'une densité de probabilité,
Publications internes de l'Université de Paris VI.

[7] A.T. BHARUCHA-REID -
Random Integral Equations,
Academic Press, 1972.

[8] P. BILLINGSLEY -
Convergence of Probability Measures,
John Wiley and Sons, 1968.

[9] J. BLEUEZ -
Conditions nécessaires et suffisantes de convergence pour une classe d'estimateurs de la densité,
Thèse de 3ème cycle, 1976, Université de Lille I.

[10] L. BONEVA, D. KENDALL et I. STEFANOV -
Spline transformations. Three new diagnostic aids for the Statistical Data-analyst,
Journal of the Royal Statistical Society, vol. 33, n° 1, 1971.

[11] D. BOSQ -
Contribution à la théorie de l'estimation fonctionnelle,
Thèse - Publications de l'ISUP, octobre 1970.

[12] D. BOSQ -
Sur l'estimation sans biais d'un paramètre fonctionnel,
Publication interne de l'U.E.R. de Math. Pures et Appliquées de Lille I, n° 104, 1977.

[13] D. BOSQ -
Cours de DEA rédigé par L. NOBECOURT et A. WIBAUX, 1974,
Université de Lille I.

[14] D. BOSQ et J. BLEUEZ -
Conditions nécessaires et suffisantes de convergence pour une classe d'estimateurs de la densité,
Comptes Rendus, 282, série A, 1976, p. 63-66.

[15] D. BOSQ et J. BLEUEZ -
Etude d'une classe d'estimateurs non-paramétriques de la densité,
Ann. Inst. Henri Poincaré XIV, n° 4, 1978, p. 479-498.

[16] P. CHEMARIN -
Quelques applications statistiques des espaces autoreproduisants,
Thèse de 3ème cycle, 1970, Université de Lyon.

[17] M. DELABROYE -
Convergence uniforme de certains estimateurs de la densité par la méthode spline,
Comptes Rendus, 289, Série A, 1979, p. 825-827.

[18] M. DELECROIX -
Estimation des densités d'un processus stationnaire à temps continu,
Comptes Rendus, 290, Série A, 1980, p. 45-48.

[19] J. DIEUDONNE -
Eléments d'Analyse, Gauthier-Villars, 1974.

[20] M. DUC-JACQUET -
Approximation des fonctionnelles linéaires sur les espaces hilbertiens auto-reproduisants,
Thèse, 1973, Université Scientifique et Médicale de Grenoble.

[21] P. GAENSSLER et W. STUTE -
Empirical processes : A survey of results for independant and identically distributed random variables,
The Annals of Probability, 7, 1979, p. 193-243.

[22] J. GEFFROY -
Sur l'estimation d'une densité dans un espace métrique,
Comptes Rendus, 278, série A, 1974, p. 1449.

[23] I.I. GIHMAN et A.V. SKOROHOD -
The Theory of Stochastic Processes, Springer-Verlag, 1974.

[24] C. GUILBART -
Etude des produits scalaires sur l'espace des mesures.
Estimation par projections. Tests à noyaux.
Thèse d'Etat, 1978, Université de Lille I.

[25] P.L. HENNEQUIN et A. TORTRAT
Théorie des Probabilités et quelques applications,
MASSON, 1965.

[26] E. HILLE et R.S. PHILLIPS -
Functional Analysis and semi-groups,
American Mathematical Soc. Colloquium, publications, 1957.

[27] P. HOMINAL -
Estimation non paramétrique de la densité compte tenu d'information sur le support,
Thèse de 3ème cycle, Université de Paris VI (1979).

[28] M. IBERO -
Déviation de la loi empirique d'une variable bidimensionnelle,
Comptes Rendus, 281, série A, p. 1059-1062.

[29] M. IBERO -
Approximation forte du processus empirique fonctionnel multidimensionnel,
Bull. Sc. Math., 2ème série, 103, 1979.

[30] S. KIMELDORF et G. WAHBA -
A correspondence between bayesian estimation on stochastic processes and smoothing by splines,
Annals of Math. Statistics, 41, 1970, p. 495-502.

[31] P.J. LAURENT -
Approximation et Optimisation, Hermann, 1972.

[32] METIVIER -
Notions fondamentales de la théorie des probabilités,
DUNOD, 1972.

[33] R. MOCHÉ -
Décantation et séparation asymptotiques uniformes...
Thèse d'Etat, 1977, Université de Lille I.

[34] G.F. de MONTRICHER, R.A. TAPIA et J.R. THOMPSON -
Non parametric maximum likelihood estimation of probability densities by penalty function methods,
The Annals of Statistics, 1975, Vol. 3, p. 1329-1348.

[35] J. NEVEU -
Processus aléatoires gaussiens,
Séminaire de Math. Sup., Les Presses de l'Université de Montréal, 1968.

[36] M. OULEDCHEIKH -
Construction d'estimateurs de la densité par la méthode des fonctions orthogonales,
Thèse de 3ème cycle, 1979, Université de Lille I.

[37] K.R. PARTHASARATHY -
Probability measures on metric spaces,
Academic Press, 1967.

[38] B.J. PETTIS -
On integration in vector spaces.
Transvections of the American Mathematical Society,44, 1938.

[39] N. RAHMANIA -
Estimation non paramétrique de la régression par les fonctions splines,
Thèse de 3ème cycle, à paraître.

[40] B.S. RAJPUT et S. CAMBANIS -
Gaussian processes and gaussian measures,
Ann. of Math. Stat., 1972, 43, p. 1944-52.

[41] J. RICATTE -
Applications de la théorie des n.r. à des problèmes de détection et d'estimation de signaux,
Thèse de 3ème cycle, 1966, Paris.

[42] W. RUDIN -
Analyse réelle et complexe,
Masson, 1975.

[43] P. SABLONNIERE -
Interpolation d'Hermite par des surfaces de classe C^1 quadratiques par morceaux,
Publication ANO 16 (novembre 1979)
UER IEEA Université de Lille I.

[44] M.H. SCHULTZ et R.S. VARGA -
L - splines,
Numer. Math. 10, 1967, p. 345-69.

[45] L. SCHWARTZ -
Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants),
J. Anal. Math. of Jerusalem XIII, 1964, p. 115-256.

[46] L. SCHWARTZ -
Analyse II - Topologie générale et analyse fonctionnelle,
Hermann, 1970.

[47] G. WAHBA -
Interpolating spline methods for density estimation I,
Equi-spaced knots,
Ann. of Stat., 3, 1975, p. 30-48.