N° d'ordre : 430

THÈSES

présentées à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le grade de

DOCTEUR ES SCIENCES MATHÉMATIQUES

par

Christian GUILBART

Première Thèse

ÉTUDE DES PRODUITS SCALAIRES SUR L'ESPACE DES MESURES. ESTIMATION PAR PROJECTIONS. TESTS A NOYAUX.

Deuxième Thèse

NOYAUX D'UN ESPACE DE BANACH

Thèses soutenues le 11 octobre 1978 devant la Commission d'Examen :

MM. M. PARREAU,	Président,
D. BOSQ,	Directeur de Recherche, Rapporteur,
J. BRETAGNOLLE,	Rapporteurs,
J. GEFFROY	
P. POUZET,	Examinateur.
F. MIGNOT	Invité

UNIVERSITE DES SCIENCES
ET TECHNIQUES DE LILLE

15 Septembre 1978

DOYENS HONORAIRES De l'Ancienne Faculté des Sciences

MM. R.DEFRETIN, H.LEFEBVRE, M.PARREAU.

PROFESSEURS HONORAIRES des Anciennes Facultés de Droit et Sciences Economiques, des Sciences et des Lettres

MM. ARNOULT, Mme BEAUJEU, BONTE, BROCHARD, CHAPPELON, CHAUDRON, CORDONNIER, CORSIN, DECUYPER, DEHEUVELS, DEHORS, DION, FAUVEL, FLEURY, P.GERMAIN, GLACET, GONTIER, HEIM DE BALSAC, HOCQUETTE, KAMPE DE FERIET, KOUGANOFF, LAMOTTE, LASSERRE, LELONG, Mme LELONG, LHOMME, LIEBAERT, MARTINOT-LAGARDE, MAZET, MICHEL, PEREZ, ROIG, ROSEAU, ROUELLE, SAVARO, SCHILTZ, WATERLOT, WIEMAN, ZAMANSKI.

ANCIENS PRESIDENTS DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

MM. R.DEFRETIN, M.PARREAU, J.LOMBARD.

PRESIDENT DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

M. M.MIGEON.

PROFESSEURS TITULAIRES

M. BACCHUS Pierre	Astronomie
M. BEAUFILS Jean-Pierre	Chimie Physique
M. BILLARD Jean	Physique du Solide
M. BIAYS Pierre	Géographie
M. BONNOT Ernest	Biologie Végétale
M. BOUGHON Pierre	Algèbre
M. BOURIQUET Robert	Biologie Végétale
M. CELET Paul	Géologie Générale
M. COEURE Gérard	Analyse
M. CONSTANT Eugène	Electronique
M. CORDONNIER Vincent	Informatique
M. DEBOURSE Jean-Pierre	Gestion des Entreprises
M. DELATTRE Charles	Géologie Générale
M. DELHAYE Michel	Chimie Physique
M. DERCOURT Jean	Géologie Générale
M. DURCHON Maurice	Biologie Expérimentale
M. FAURE Robert	Mécanique
M. FOURET René	Physique du Solide
M. GABILLARD Robert	Electronique
M. GRANELLE Jean-Jacques	Sciences Economiques
M. GRUSON Laurent	Algèbre
M. GUILLAUME Jean	Microbiologie
M. HECTOR Joseph	Géométrie
M. HEUBEL Joseph	Chimie Minérale
M. LABLACHE-COMBIER Alain	Chimie Organique

.../...

M. LACOSTE Louis	Biologie Végétale
M. LANSRAUX Guy	Physique Atomique et Moléculaire
M. LAVEINE Jean-Pierre	Paléontologie
M. LEBRUN André	Electronique
M. LEHMANN Daniel	Géométrie
Mme LENOBLE Jacqueline	Physique Atomique et Moléculaire
M. LHOMME Jean	Chimie
M. LINDER Robert	Biologie et Physiologie Végétales
M. LOMBARD Jacques	Sociologie
M. LOUCHEUX Claude	Chimie Physique
M. LUCQUIN Michel	Chimie Physique
M. MAILLET Pierre	Sciences Economiques
M. MONTREUIL Jean	Biochimie
M. PARREAU Michel	Analyse
M. PAQUET Jacques	Géologie Générale
M. PROUVOST Jean	Minéralogie
M. SALMER Georges	Electronique
Mme SCHWARTZ Marie-Hélène	Géométrie
M. SEGUIER Guy	Electrotechnique
M. STANKIEWICZ François	Sciences Economiques
M. TILLIEU Jacques	Physique Théorique
M. TRIDOT Gabriel	Chimie Appliquée
M. VIDAL Pierre	Automatique
M. VIVIER Emile	Biologie Cellulaire
M. WERTHEIMER Raymond	Physique Atomique et Moléculaire
M. ZEYTOUNIAN Radyadour	Mécanique

PROFESSEURS SANS CHAIRE

M. BELLET Jean	Physique Atomique et Moléculaire
M. BKOUCHE Rudolphe	Algèbre
M. BODARD Marcel	Biologie Végétale
M. BOILLY Bénoni	Biologie Animale
M. CAPURON Alfred	Biologie Animale
M. CARREZ Christian	Informatique
M. CORTOIS Jean	Physique Nucléaire et Corpusculaire
Mme DACHARRY Monique	Géographie
M. DEVRAINNE Pierre	Chimie Minérale
M. GOSSELIN Gabriel	Sociologie
M. GOUDMAND Pierre	Chimie Physique
M. GUILBAULT Pierre	Physiologie Animale
M. HERMAN Maurice	Physique Spatiale
M. JOURNEL Gérard	Physique Atomique et Moléculaire
Mme LEHMANN Josiane	Analyse
M. LENTACKER Firmin	Géographie
M. LOUAGE Francis	Electronique
M. MAIZIERES Christian	Automatique
Mle MARQUET Simone	Probabilités
M. MESSELYN Jean	Physique Atomique et Moléculaire
M. MIGEON Michel	Chimie Physique
M. MONTEL Marc	Physique du Solide
M. RACZY Ladislas	Electronique
M. ROUSSEAU Jean-Paul	Physiologie Animale
M. SLIWA Henri	Chimie Organique
M. WATERLOT Michel	Géologie Générale

.../...

MAITRES DE CONFERENCES (et Chargés d'Enseignement)

M.	AL FAKIR Sabah	Algèbre
M.	ANTOINE Philippe	Analyse
M.	BART André	Biologie Animale
Mme	BATTIAU Yvonne	Géographie
M.	BEGUIN Paul	Mécanique
M.	BOBE Bernard	Sciences Economiques
M.	BONNELLE Jean-Pierre	Chimie
M.	BOSCQ Denis	Probabilités
M.	BREZINSKI Claude	Analyse Numérique
M.	BRUYELLE Pierre	Géographie
M.	CHAMLEY Hervé	Géotechnique
M.	COQUERY Jean-Marie	Psychophysiologie
M.	COURBIS Bernard	Sciences Economiques
M.	COUTURIER Daniel	Chimie Organique
M.	DEBRABANT Pierre	Géologie Appliquée
M.	DEGAUQUE Pierre	Electronique
M.	DELORME Pierre	Physiologie Animale
M.	DE PARIS Jean-Claude	Mathématiques
M.	DHAINAUT André	Biologie Animale
M.	DOUKHAN Jean-Claude	Physique du Solide
M.	DUBOIS Henri	Physique
M.	DUBRULLE Alain	Physique
M.	DUEE Gérard	Géologie
M.	DYMENT Arthur	Mécanique
M.	ESCAIG Bertrand	Physique du Solide
M.	FLAMME Jean-Marie	Technologie de Construction
M.	FONTAINE Hubert	Physique
M.	GAMBLIN André	Géographie
M.	GOBLOT Rémi	Algèbre
M.	GREVET Patrick	Sciences Economiques
M.	JACOB Gérard	Informatique
M.	KREMBEL Jean	Biochimie
M.	LAURENT François	Automatique
Mle	LEGRAND Denise	Algèbre
Mle	LEGRAND Solange	Algèbre
M.	LEVASSEUR Michel	Sciences Economiques
M.	LHENAFF René	Géographie
M.	LOCQUENEUX Robert	Physique Théorique
M.	LOSFELD Joseph	Informatique
M.	MACKE Bruno	Physique
M.	MIGNOT Fulbert	Analyse Numérique
M.	N'GUYEN VAN CHI Régine	Géographie
M.	PARSY Fernand	Mécanique
Mle	PAUPARDIN Colette	Biologie Physiologie Végétales
M.	PERROT Pierre	Chimie Appliquée
M.	PERTUZON Emile	Physiologie Animale
M.	PONSOLLE Louis	Chimie Physique
M.	POVY Lucien	Automatique
M.	RICHARD Alain	Biologie
M.	RIETSCH François	Chimie
M.	ROGALSKI Marc	Analyse
M.	ROY Jean-Claude	Psychophysiologie
M.	SALAMA Pierre	Sciences Economiques
Mme	SCHWARZBACH Yvette	Mathématiques
M.	SIMON Michel	Sociologie
M.	SOMME Jean	Géographie

.../...

Mle SPIK Geneviève	Biochimie
M. STERBOUL François	Informatique
M. TAILLIEZ Roger	Biologie
M. TOULOTTE Jean-Marc	Automatique
M. TREANTON Jean-René	Sociologie
M. VANDORPE Bernard	Chimie Minérale
M. WALLART Francis	Chimie
Mme ZINN-JUSTIN Nicole	Algèbre

à mes parents,

à Annie,

Nous remercions vivement Monsieur le Professeur Michel PARREAU pour l'honneur qu'il nous a fait de présider le jury de cette thèse.

Que Monsieur le Professeur Denis BOSQ, qui a dirigé notre recherche, trouve ici l'expression de notre profonde gratitude pour les encouragements, et les conseils éclairés qu'il n'a jamais cessé de nous prodiguer.

Nous ne remercierons jamais assez Monsieur le Professeur Jean BRETAGNOLLE dont la disponibilité et la rigueur nous ont conduit à une rédaction claire et concise ; et à qui, en particulier, nous devons la généralisation et l'élégante démonstration du "Théorème de Séparabilité (3.A.I.) et la proposition (2.E.I)".

Nous tenons à exprimer notre respectueuse reconnaissance à Monsieur le Professeur Jean GEFFROY pour l'intérêt constant qu'il a porté à notre travail et dont les encouragements nous ont été un stimulant très précieux.

Nous adressons nos vifs remerciements à Monsieur le Professeur Pierre POUZET qui a bien voulu nous proposer un sujet de seconde thèse et faire partie de la Commission d'Examen.

Nous remercions beaucoup Monsieur le Professeur Fulbert MIGNOT qui nous a assisté avec efficacité lors de la rédaction de la seconde thèse.

Nous pensons avec reconnaissance aux collègues probabilistes et statisticiens du Séminaire de l'U.E.R. de Mathématiques Pures et Appliquées de LILLE I, et plus particulièrement à Claude LANGRAND, qui nous ont toujours accordé aide et sympathie.

Nous remercions enfin Madame BERAT, Madame LENGAIGNE et toutes les personnes qui ont participé à la réalisation matérielle de cet ouvrage, pour la compétence, le soin et la célérité dont elles ont fait preuve.

PLAN

INTRODUCTION

PREMIERE PARTIE

ETUDE DES PRODUITS SCALAIRES SUR L'ESPACE DES MESURES

DEUXIEME PARTIE

ESTIMATIONS PAR PROJECTIONS

APPENDICE

TESTS A NOYAUX

PREMIÈRE PARTIE

ÉTUDE DES PRODUITS SCALAIRES SUR L'ESPACE DES MESURES

ETUDE DES PRODUITS SCALAIRES SUR L'ESPACE DES MESURES.
ESTIMATION PAR PROJECTIONS. TESTS A NOYAUX.

Après avoir remarqué que dans de nombreuses méthodes d'estimation on utilise simultanément le théorème de Glivenko-Cantelli et des projections sur des sous-espaces de dimension finie (projections plus ou moins apparentes, s'effectuant sur l'espace des paramètres), nous avons pensé à effectuer d'abord la projection sur l'espace des mesures (muni d'un produit scalaire dont la métrique trace sur l'espace des probabilités vérifie un "théorème de Glivenko-Cantelli") et à lui associer ensuite son image dans l'espace des paramètres (méthode qui donnera souvent des estimateurs sans biais et de bonnes "vitesses" de convergence). Ce qui nous a conduit à la rédaction des deux premières parties auxquelles nous avons ajouté quelques applications à la théorie des tests (qui sont aussi basées sur le "théorème de Glivenko-Cantelli" associé à un (pseudo)-produit scalaire sur l'espace des mesures).

Dans la première partie, à l'aide de la théorie des noyaux reproduisants (Introduite par N. Aronsjazn dans [1] en 1943), nous étudions les (pseudo)-produits scalaires sur l'espace des mesures bornées à signes. Nous montrons (théorème de plongement (2.C.I)) que, muni de certains produits scalaires, l'espace des mesures peut être considéré (avec toutes les commodités que cela comporte) comme un sous-espace d'un espace de Hilbert à noyau reproduisant. Nous identifions (théorème de réintégration (4.C.I)) les produits scalaires, sur l'espace des mesures, qui se réintègrent (i.e. : $\forall \mu , \nu \in M_\Omega , \int <\delta_\cdot,\delta_\cdot> d\mu \otimes \nu = <\mu,\nu>$) et à l'aide de l'exemple (5.C.I) nous montrons qu'il existe des produits scalaires ne se réintégrant pas. Nous déterminons (théorème de caractérisation (6.C.I)) la forme des produits scalaires où l'espace des mesures de Dirac est total dans l'espace des mesures. Nous caractérisons (théorème de faible comparaison (1.D.I)) les produits scalaires dont la topologie trace sur l'espace des mesures de probabilité n'est autre (resp : est moins fine) que la topologie faible. Nous montrons ensuite (théorème d'existence (4.D.I)) que lorsque l'espace de base est métrique séparable, il existe des noyaux reproduisants dont le produit scalaire associé sur l'espace des mesures a pour topologie trace sur les mesures positives, la topologie faible. Le théorème de construction (1.E.I) et le lemme des produits (3.E.I) nous permettent d'exhiber des exemples de noyaux reproduisants (sur R et R^n) d'expression très simple, dont la métrique trace sur les probabilités est équivalente (et uniformément équivalente lorsque Ω est borné dans R ou R^n) à la métrique de Prokorov. Puis nous établissons un "théorème de Glivenko-Cantelli" pour toute une classe de produits scalaires étudiés auparavant ; théorème sur lequel sont basées toutes les applications. Nous montrons enfin que si K_1 est un noyau reproduisant "proche" de K_2, l'estimateur construit à partir d'une projection orthogonale relative au produit scalaire $\int K_1 \, d\mu \otimes \nu$ est "proche" de celui construit à partir d'une projection orthogonale relative au produit scalaire $\int K_2 \, d\mu \otimes \nu$; ce qui d'une certaine façon rend les constructions (de la seconde partie) intrinsèques.

Dans la seconde partie, nous montrons que l'existence d'un estimateur (d'un certain type contenant en particulier les estimateurs usuels) convergent uniformément en probabilité est conditionnée par l'existence d'une métrique dont la topologie faible associée (sur l'espace des mesures) rend le paramètre continu. Nous exhibons une condition nécessaire identique, quant à l'existence d'un estimateur sans biais ; ce qui nous permet en particulier d'affirmer qu'en général la densité n'est estimable sans biais que si et seulement si l'espace vectoriel engendré par $\mathcal{P}_1$ (espace des mesures envisagées) est de dimension finie.

Nous établissons ensuite une condition nécessaire et suffisante pour qu'un paramètre soit estimable sans biais ; condition que nous appliquons en particulier à l'estimation de la loi.

Nous montrons (théorème du sans biais de la loi (2.D.II)) que la loi empirique est un estimateur sans biais de la loi pour le type de produits scalaires que nous avons utilisé le plus souvent. A l'aide des projections orthogonales nous construisons des estimateurs de la loi plus "rapides" que la loi empirique.

Nous montrons (théorème de l'estimation linéaire en dimension finie (1.E.II)) que si le paramètre est linéaire, l'espace paramétré normé (par une norme quelconque) et que si l'on se restreint à l'enveloppe linéaire d'un nombre fini de mesures, on peut construire (à l'aide des projections orthogonales relatives à un certain produit scalaire sur l'espace des mesures) un estimateur sans biais convergent p.s. avec une "vitesse en $\frac{1}{N}$" et en probabilité avec une vitesse exponentielle ; nous prouvons ensuite que ce théorème est optimal.

Lorsque l'on se restreint à l'enveloppe linéaire d'une infinité dénombrable de mesures, le théorème de l'estimation linéaire en dimension dénombrable (2.E.II) nous montre que si le paramètre est linéaire et l'espace paramétré est normé, on peut construire (à l'aide des projections orthogonales) un estimateur séquentiellement sans biais et convergent p.s.

Enfin à l'aide des projections orthogonales, nous établissons de façon simple des résultats en partie connus sur des paramètres limites.

Dans l'appendice, à l'aide des noyaux reproduisants et du "théorème de Glivenko-Cantelli", nous construisons des tests de puissance correcte, de région critique relativement simple et s'appliquant à des situations relativement générales.

Plus précisément, lorsque la distance (relative au pseudo-produit scalaire associé à un noyau reproduisant K) de $\mathcal{P}_o$ à $\mathcal{P}_1$ est strictement positive, nous exhibons un test de puissance satisfaisante de $\mathcal{P}_o$ contre $\mathcal{P}_1$. Puis lorsque $\mathcal{P}_1$ est disjoint de l'adhérence (relative au pseudo-produit scalaire associé à un noyau reproduisant K) de $\mathcal{P}_o$, nous construisons une suite convergente de tests de $\mathcal{P}_o$ contre $\mathcal{P}_1$; nous montrons en particulier, lorsque $\mathcal{P}_o = \{\mu\}$ et $\mathcal{P}_1 = \{\mu\}^C$, que cette suite est formée de tests (de régions critiques élémentaires) dont on peut très facilement maximiser le niveau et minimiser la puissance en un point quelconque.

Enfin, nous proposons des tests s'appliquant à des alternatives multiples.

(A - I)- ETUDE HILBERTIENNE

Dans la première partie de cette thèse, nous allons utiliser la théorie des noyaux reproduisants ; théorie qui a été introduite en 1943 par N. Aronsjazn dans [1], a été complétéedans [2] et a largement utiliséedepuis dans [44] , [47] , [24] , [13] , [31]... (dans [53], l'auteur a approndi les résultats de [1] et [2]).

Soit Ω un espace abstrait, d'éléments x, y, z,... . Une fonction réelle K sur $\Omega \times \Omega$ est dite un noyau reproduisant si $K(x,y) = K(y,x)$ (symétrie) et si $\forall x_i$, $a_i \in R$, $1 \leqslant i \leqslant J$: $\sum\limits_{1 \leqslant i,j \leqslant J} a_i a_j K(x_i, x_j) \geqslant 0$ (semi-définie positive).

Un espace d'applications réelles sur Ω d'éléments f, g,... est dit de Hilbert à noyau reproduisant K, s'il contient toutes les $f_y = K(.,y)$, s'il est de Hilbert pour un produit scalaire (f,g) et si $(f,K(.,y)) = f(y)$ (K est reproduisant).

Théorème (1.A.I.).- Si H est K reproduisant ; K est symétrique semi-défini positif, et H_o, espace engendré par les combinaisons linéaires finies de $K(.,y)$ est dense dans H. Réciproquement, à tout noyau reproduisant K on peut faire correspondre un Hilbert K reproduisant, fermeture pour $(K(.,y),K(.,z)) = K(y,z)$ de H_o (nous considérerons que cet espace K reproduisant est unique et le noterons H_K, la non unicité venant uniquement du choix arbitraire d'un représentant dans les différentes classes d'équivalence). Enfin, si K est borné, la K topologie est plus fine que la norme de la convergence uniforme.

Preuve : Si H est K reproduisant, K est unique, symétrique et semi-défini positif. Si $(K(.,y),f) = 0$ pour tout y de Ω, on a $f = 0$;

$(K(.,y), y \in \Omega)$ est donc total dans H ; en d'autres termes, H_o est dense dans H.

Réciproque : Si K est un noyau reproduisant, H_o est un préhilbertien ayant la propriété de reproduction... (voir [1], page 143-146).

Comme $|f(y)| \leq |(f,K(.,y))| \leq ||f||_K . K(y,y)^{1/2}$ (inégalité de Schwartz) une suite convergente dans H_K converge simplement, uniformément si K est borné. ■

Proposition (2.A.1).- H, Hilbert de fonctions réelles, est reproduisant si et seulement si $\forall\, y \in \Omega$, $f \in H$; $f \to f(y)$ est continue de H dans $\mathbb{R}$.

(poser $K(.,y) = g_y$ définie par $f(y) = (f,g_y)$, $\forall\, f \in H$).

Théorème de séparabilité (3.A.1).- H, Hilbert K reproduisant borné, est séparable si et seulement si il existe une suite de fonctions (f_i) telle que $K = \sum_i f_i \otimes f_i$ (série convergent simplement sur $\Omega \times \Omega$) ; alors les f_i sont bornées et la convergence est uniforme en x pour tout y fixé (dite par la suite convergence simple uniforme).

Preuve : Dans le cas H séparable, on prend (f_i) base orthonormale de H ; alors $(f_i,K(.,y)) = f_i(y)$, de sorte que $K(.,y) = \sum_{i=1}^{\infty} f_i(y) f_i(.)$, série convergente dans H et donc simplement uniformément puisque K est borné. Les f_i sont alors bornées par $||K||_\infty$.

Réciproque : Soit ϕ_i la forme linéaire sur H_o définie par prolongement linéaire de $\phi_i(K(.,y)) = f_i(y)$. Majorons sa norme :

$$[\phi_i(\sum_{j=1}^{p} a_j\, K(.,y_j))]^2 = [\sum_{j=1}^{p} a_j f_i(y_j)]^2 \leq \sum_{j,j',t} a_j a_{j'} f_t(y_j) f_t(y_j)$$

(somme finie de séries convergentes, on peut donc intervertir les sommations)

d'où $[\phi_i(\sum_{j=1}^{p} a_j\, K(.,y_j))]^2 \leq (\sum_{j=1}^{p} a_j\, K(.,y_j), \sum_{j=1}^{p} a_j\, K(.,y_j))_{H_K}$.

ϕ_i étant de norme moindre que 1, elle est représentable par un élément de H ; la formule reproduisante dit alors que c'est f_i. De même on trouve que $\sum_t f_t(x) f_t(.)$ a, comme opérateur sur H_o une norme moindre que $K(x,x)^{1/2}$, et la série du théorème est convergente dans H. Les f_i engendrent donc les $K(.,y)$, et de ce fait sont totales dans H. ■

(B - I) - MESURABLITES

On suppose K borné et strictement positif sur la diagonale, dans toute la suite de ce travail.

H_K étant fermeture, pour une norme plus fine que l'uniforme, des $K(.,y)$ on peut énoncer :

si Ω a une structure mesurable et si $\forall y \in \Omega$, $K(.,y)$ est mesurable, H_K est composé de <u>mesurables</u>.

Si Ω est topologique, si chaque point admet une base dénombrable de voisinages et si pour tout y de Ω, $K(.,y)$ est continu, H_K est composé de continues. Même chose pour la continuité uniforme.

Si K est <u>continu</u> et Ω <u>séparable</u>, H_K est séparable ((y_n) étant dense, les $K(.,y_n)$ sont K-denses dans les $K(.,y)$ qui sont totales dans H_K) ; le développement du théorème de séparabilité est alors composé de continues (resp. uniformément continues). La réciproque de cette dernière affirmation est vraie dans le cas séparable (convergence simple uniforme).

Enfin dans le cas abstrait, si H_K est séparable et si $(K(.,y), y \in \Omega)$ sépare les points, on peut munir Ω d'une topologie métrisable séparable rendant K uniformément continu (prendre la structure uniforme rendant uniformément continues les f_i).

<u>*Lemme*</u> *(1.B.1)*.- Soit $\mathcal{B}$ une tribu de type dénombrable (i.e : engendrable par un nombre dénombrable d'événements) séparant les points de Ω

(i.e : $\forall$ x, y $\in \Omega$, $x \neq y$, $\exists A \in \mathcal{B} : x \in A$, $y \in A^c$). Si H_K est séparable et si $\forall y \in \Omega$, $K(.,y)$ est mesurable, il existe une métrique séparable d sur Ω ayant $\mathcal{B}$ pour tribu borélienne et rendant K $d \otimes d$-continu.

Preuve : H_K étant séparable et K borné, on a $K = \sum_{i=1}^{\infty} f_i \otimes f_i$ (théorème de séparabilité) ; les f_i appartenant à H_K sont mesurables et bornées. $(A_i, i \in \mathbb{N}^*)$ engendrant la tribu $\mathcal{B}$, on pose $A = (f_i^{-1}(]p,q[), (p,q) \in Q^2, p < q, i \in N^*) \cup (A_i, i \in N^*)$; on note A_1 l'algèbre engendrée par A et $\mathcal{T}$ la topologie engendrée par A_1. La tribu borélienne de $\mathcal{T}$ est alors $\mathcal{B}$ et $\mathcal{T}$ rend les f_i continues. Comme $\mathcal{B}$ sépare les points, la famille (A_i) sépare les points et $\mathcal{T}$ est séparée ; A_1 étant dénombrable, $\mathcal{T}$ est séparable ; x étant un point de Ω et O un ouvert de $\mathcal{T}$ contenant x, on a $O = \bigcup_{i \in I} B_i$, avec $B_i \in A_1$; il existe donc $i_o \in I$ tel que $x \in B_{i_o}$, qui est ouvert et fermé dans $\mathcal{T}$; autrement dit $(\Omega, \mathcal{T})$ est "régulier". Le théorème de Uryson (prop. (16) page 125 de [40]) nous permet d'affirmer que $\mathcal{T}$ est métrisable séparable. La convergence de la série étant simple uniforme et les f_i $\mathcal{T}$-continues, $K(.,y)$ est, pour tout y de Ω, $\mathcal{T}$-continu ; K étant borné, la prop. (5.7) page 140 de [1] nous permet d'écrire que K est $\mathcal{T} \otimes \mathcal{T}$-continu. ■

Remarque : $\mathcal{B}$ est la tribu borélienne associée à une métrique séparable, ssi $\mathcal{B}$ vérifie les conditions du lemme ci-dessus.

(C - I) - MESURES

Désormais, Ω est muni d'une structure mesurable. On appelle M_Ω l'espace des mesures signées bornées, M_Ω^+ les mesures positives bornées, P_Ω les probabilités, M_o les combinaisons linéaires finies de mesures de Dirac δ_x, $P_o = P_\Omega \cap M_o$.

Un (pseudo)-produit scalaire sur M_Ω est noté $<.,.>$ et l'on pose $<\delta x, \delta y> = K(x,y)$.

Lemme (1.C.I).- Si K est mesurable borné, l'application linéaire de $(\mathcal{M}_o,<.,.>)$ sur $(H_o,(.,.)_{H_K})$ définie par $\phi(\delta_y) = K(.,y)$ est une isométrie entre ces deux espaces.

Théorème de plongement (2.C.I).- On peut alors prolonger ϕ en une application linéaire de $\mathcal{M}_\Omega$ dans H_K par $\phi(\mu) = \int K(.,y)d\mu(y)$; on note alors $(\mu,\nu)_K$ le pseudo-produit scalaire défini sur $\mathcal{M}_\Omega$ par $(\mu,\nu)_K = (\phi(\mu),\phi(\nu))_{H_K}$.

Preuve :

Pour μ fixée, la forme linéaire ϕ_μ définie sur H_o par prolongement de $\phi_\mu(K(.,x)) = \int K(y,x)d\mu(y)$ est continue (en effet, K étant borné, il existe une constante $b > 0$ telle que $\forall f \in H_K$, $||f||_u \leq b||f||_{H_K}$; on a alors $|\phi_\mu(\sum_{i=1}^n a_i\ K(.,x_i))| \leq ||\sum_{i=1}^n a_i K(.,x_i)||_U . ||\mu||_{var} \leq$ $b||\sum_{i=1}^n a_i\ K(.,x_i)||_{H_K} . ||\mu||_{var})$; elle se prolonge donc en une forme linéaire continu sur H_K tout entier ; il existe donc un élément h de H_K tel que $\forall x \in \Omega$, $\int K(x,y)d\mu(y) = (h,K(x,.))_{H_K} = h(x)$; $\int K(.,y)d\mu(y)$ appartient donc à H_K. ■

Remarques :

- Pour ce (pseudo)-produit scalaire, $\mathcal{M}_o$ est dense dans $\mathcal{M}_\Omega$.

- On a $(\mu,\nu)_K = \int K\ d\mu \otimes \nu$ (en effet, pour μ fixée dans $\mathcal{M}_\Omega$, on a : $|\int K\ d\mu \otimes \nu| \leq ||\phi(\nu)||_U \times ||\mu||_{var} \leq b||\phi(\nu)||_{H_K} . ||\mu||_{var}$; d'où $\nu \longrightarrow \int K\ d\mu \otimes \nu$ et $\nu \to (\nu,\mu)_K$ sont deux formes linéaires continues sur $(\mathcal{M}_\Omega,(.,.)_K)$ qui coincident sur $\mathcal{M}_o$; elles sont donc égales).

- $(.,.)_K$ sépare les mesures de Dirac dès que $K(.,x) = K(.,z) \Rightarrow x = z$; il sépare les mesures dès que $K = \sum_{i=1}^\infty c_i\ 1_{A_i} \otimes 1_{A_i}$, où (c_i) est une série convergente de termes strictement positifs, (A_i) engendre la tribu de Ω (de même en remplaçant les A_i par des f_i continues bien choisies dans le cas topologique).

- Si $(.,.)_K$ est non dégénéré, ϕ est une isométrie et $(M_\Omega, (.,.)_K)$ peut être considéré comme un sous-espace de H_K.

Lemme (3.C.I).- Soient $(\Omega,\mathcal{B})$ un espace mesurable, K un noyau reproduisant $\mathcal{B} \otimes \mathcal{B}$ mesurable borné et $T_1 : (\Omega,\mathcal{B}) \to (M_\Omega,(.,.)_K)$. Alors :

$$\omega \to \delta_\omega$$

① T_1 est bornée et faiblement mesurable.

② Si $(.,.)_K$ est séparable, T_1 est mesurable.

③ Si $\mathcal{B}$ est la tribu borélienne d'une topologie (où tout point admet une base dénombrable de voisinages) rendant K continu, T_1 est continu.

④ Si de plus cette topologie est d-métrisable et K $d \otimes d$-uniformément continu, T_1 est uniformément continu.

Preuve :

① Soit e une forme linéaire continu sur $(M_\Omega,(.,.)_K)$; il existe alors une suite (μ_n) de M_o telle que $\forall\, \omega \in \Omega$, $(\delta_\omega,\mu_n) \to e(\delta_\omega)$. Comme $(\delta_\omega,\mu_n)_K = \int K(\omega,x)d\mu_n(x) = \phi(\mu_n)(\omega)$, e o δ est mesurable.

② et ③ découlent du fait que :

$$T_1^{-1}(B(\mu,a)) = \{\omega \in \Omega : K(\omega,\omega) - 2\phi(\mu)(\omega) < a^2 - \int K\, d\mu \otimes \mu\}$$

④ $$||\delta\omega_1 - \delta\omega_2||_K^2 = K(\omega_1,\omega_1) + K(\omega_2,\omega_2) - 2K(\omega_1,\omega_2).$$

Théorème de réintégration (4.C.I).- Soient $(\Omega,\mathcal{B})$ un espace mesurable et $<.,.>$ un produit scalaire sur M_Ω, où $<\delta x,\delta y>$ est $\mathcal{B} \otimes \mathcal{B}$-mesurable et borné. Pour que $<.,.>$ se réintégre (i.e : $\forall \mu, \nu \in M_\Omega, <\mu,\nu> = \int <\delta x,\delta y> d\mu\otimes\nu(x,y)$) il faut et il suffit que $T_1 : \omega \to \delta\omega$ soit un estimateur sans biais (*) d'ordre (1) de la loi.

Preuve : *Condition nécessaire* : Si l'on pose $K(x,y) = <\delta x,\delta y>$ on a alors $<.,.> = (.,.)_K$; comme $<\mu,\delta y> = \int_\Omega <\delta x,\delta y> d\mu(x)$, pour tout ν

(*) Voir définition (1.B.II) de la 2ème partie.

de M_o on a $<\mu,\nu> = \int_\Omega <\delta x,\nu>d\mu(x)$; M_o étant dense dans $(M_o,(.,.)_K)$ (donc dans $(M_\Omega,<.,.>)$), pour toute forme linéaire $<.,.>$-continue, e, il existe une suite (ν_n) de M_o telle que $\forall\ \nu \in M_\Omega$, $<\nu,\nu_n> \to e(\nu)$; on a alors (pour μ fixée) $\int_\Omega <\delta_\omega,\nu_n>d\mu(\omega) \to \int_\Omega e(\delta_\omega)d\mu(\omega)$ (théorème de convergence dominée) et $\int_\Omega <\delta_\omega,\nu_n>d\mu(\omega) = <\mu,\nu_n>$; d'où $e(\mu) = \int_\Omega e \circ \delta_\omega\ d\mu(\omega)$.

Condition suffisante :

T_1 étant un estimateur sans biais de la loi, le théorème de Fubini et la définition de l'intégrale de Pettis nous permettent alors d'écrire :

$$\int_{\Omega^2} <\delta x,\delta y>d\mu \otimes \nu(x,y) = \int_\Omega (\int_\Omega <\delta x,\delta y>d\mu(x))d\nu(y)$$

$$= < \int_\Omega \delta x d\mu(x),\int_\Omega \delta y d\nu(y)> = <\mu,\nu> \blacksquare$$

Remarque : Si K est un noyau reproduisant $\mathcal{B}\otimes\mathcal{B}$-mesurable, borné, non dégénéré (i.e : $\int K\,d\mu \otimes \mu = 0 \Longrightarrow \mu = 0$), et si l'on munit M_Ω de $(.,.)_K$, $T_1 : \omega \to \delta_\omega$ est un estimateur sans biais d'ordre 1 de la loi.

Remarque : Tout produit scalaire sur $M(\mathcal{B})$ ne se réintègre pas ; nous allons donner un exemple de produit scalaire séparable sur $M_{[0,1]}$ ne se réintégrant pas, et où $<\delta x,\delta y>$ est continu et borné.

Exemple (5.C.I).- Si l'on note M_D (resp. M_d) l'ensemble des mesures diffuses (resp. discrètes) sur $[0,1]$, on a $M_{[0,1]} = M_D \oplus M_d$ (somme directe) [en effet, toute mesure sur $[0,1]$ se décompose d'une façon unique en la somme d'une mesure diffuse et d'une mesure discrète].

D'après le théorème d'existence (4.D.I) (que nous démontrerons par la suite) on peut exhiber un noyau reproduisant K borné, continu et non dégénéré (i.e : $\int K d\mu \otimes \mu = 0 \Longrightarrow \mu = 0$) sur $[0,1]$.

Si l'on note h la projection parallèle à M_D et si l'on pose $\langle\mu,\nu\rangle = (\mu,\nu)_K + (h(\mu),h(\nu))_K$, $\langle .,. \rangle$ est non dégénéré et l'on a :

(1) $\langle\delta x,\delta y\rangle = 2(\delta x,\delta y)_K = 2K(x,y)$ est continu et borné.

(2) $\langle .,. \rangle$ est séparable [en effet, si l'on note $\langle\mu,\nu\rangle_o = (h(\mu),h(\nu))_K$ M_o est dense dans $(M_\Omega,\langle .,. \rangle_o)$; $(M_o,\langle .,. \rangle_o)$ étant séparable, $(M_\Omega,\langle .,. \rangle_o)$ l'est ; et comme $(M_\Omega,(.,.)_K)$ est séparable, $(M_\Omega,\langle .,. \rangle)$ l'est].

(3) $\langle .,. \rangle$ ne se réintègre pas [en effet, si μ_o est non nulle et diffuse on a : $\langle\mu_o,\mu_o\rangle = (\mu_o,\mu_o)_K \neq 0$ et

$$\int \langle\delta x,\delta y\rangle d\mu_o \otimes \mu_o = 2(\mu_o,\mu_o)_K].$$

(4) M_o n'est pas dense dans $(M_\Omega,\langle .,. \rangle)$. [En effet, si μ_o est non nulle et diffuse et s'il existe une suite (ν_n) de M_o convergent vers μ_o pour $\langle .,. \rangle$, on a ν_n tend vers μ_o pour $(.,.)_K$ qui est moins fin que $\langle .,. \rangle$. Ce qui nous conduit à :

$$. \ 2(\nu_n,\nu_n)_K = \langle\nu_n,\nu_n\rangle \longrightarrow \langle\mu_o,\mu_o\rangle = (\mu_o,\mu_o)_K$$

$$. \ (\nu_n,\nu_n)_K \longrightarrow (\mu_o,\mu_o)_K \neq 0$$

qui est impossible].

(5) $T_1 : \omega \to \delta_\omega$ n'est pas un estimateur sans biais de la loi [on peut conclure directement à l'aide du théorème de réintégration. Mais on peut aussi le constater directement. Supposons qu'il soit sans biais ; on a alors pour μ_o diffuse non nulle : $\langle\delta x,\int\delta_\omega d\mu_o(\omega)\rangle = (\delta x,\mu_o)_K$, et en utilisant la définition de l'intégrale de Pettis $\langle\delta x,\int\delta_\omega d\mu_o(\omega)\rangle = \int\langle\delta x,\delta_\omega\rangle d\mu_o(\omega) = 2\int(\delta x,\delta_\omega)_K d\mu_o(x) =$ $= 2(\delta x,\mu_o)_K$ qui est contradictoire, dès que $(\delta x,\mu_o)_K \neq 0$.

Il n'existe donc pas d'estimateurs sans biais de la loi pour ce produit scalaire.

Le théorème suivant caractérise les produits scalaires sur M_Ω, où M_o est dense dans M_Ω. (Ce théorème ne s'applique donc pas à l'exemple (5.C.I)).

Théorème de caractérisation (6.C.1).- $(\Omega,\mathcal{B})$ étant un espace mesurable ; si $<.,.>$ est un pseudo-produit scalaire sur M_Ω, borné sur P_o, tel que M_o est dense dans M_Ω et $K(x,y) = <\delta x,\delta y>$ est borné ; il existe une application linéaire h unique de $L_K = \{\phi(\mu) = \int K(.,x)d\mu(x), \mu \in M_\Omega\}$ dans H_K telle que $<\mu,\nu> = (h(\phi(\mu)),h(\phi(\nu)))_{H_K}$.

Preuve : μ étant un élément de M_Ω, il existe une suite (ν_n) de M_o convergent vers μ dans $(M_\Omega,<.,.>)$; et comme $<\nu_n-\nu_p,\nu_n-\nu_p> = (\phi(\nu_n) - \phi(\nu_p),\phi(\nu_n) - \phi(\nu_p))_{H_K}$, la suite $(\phi(\nu_n))$ est une suite de Cauchy de H_K ; il existe donc un élément f unique de H_K tel que $(\phi(\nu_n) \xrightarrow{H_K} f)$; si l'on pose $h(\phi(\mu)) = f$, h est linéaire de L_K dans H_K, et comme $\forall\, \nu \in M_\Omega$, $||h(\phi(\nu))||_{H_K} = <\nu,\nu>^{1/2}$, on a :

$$<\mu,\nu> = \frac{1}{4}\,(<\mu+\nu,\mu+\nu> - <\mu-\nu,\mu-\nu>) = (h(\phi(\mu)),h\phi(\nu))_{H_K} \blacksquare$$

(D - I)- RELATIONS ENTRE PRODUITS SCALAIRES ET TOPOLOGIE FAIBLE

(Ω,d) étant un espace métrique, si l'on munit l'espace, C_b, des fonctions réelles continues bornées de la norme de la convergence uniforme ; C_b est un espace de Banach et M_Ω est un sous-espace du dual topologique, C_b^*, de C_b.

Définition.- On appelle topologie faible (ou encore étroite) de M_Ω (que l'on note F_d) la trace (sur M_Ω) de la topologie affaiblie $\sigma(C_b^*,C_b)$.

Notations : nous écrirons indifféremment $(\mu_n \Longrightarrow \mu)$ ou $(\mu_n \xrightarrow{F_d} \mu)$ pour dire que μ_n tend faiblement vers μ.

Remarque : on a $(\mu_n \xrightarrow{F_d} \mu) \iff (\forall f \in C_b, \int f\, d\mu_n \longrightarrow \int f\, d\mu)$.

Remarque : la trace sur M_Ω de la topologie forte de C_b^* n'est autre que la norme en variation totale.

En effet $\forall \mu \in M_\Omega$, on a

$$||\mu||_{C_b^*} = \sup_{f\in C_b}\left(\left|\int \frac{f}{||f||_u} d\mu\right|\right) \leq \int 1\, d\mu^+ + \mu^- = ||\mu||_{var} .$$

Maintenant comme il existe (*) A tel que $||\mu||_{var} = \int 1_A - 1_{A^C}\, d\mu$ et comme μ est régulière (Ω est métrique) on a : $||\mu||_{var} \leq ||\mu||_{C_b^*}$.

Nous nous placerons désormais dans le cas où (Ω,d) est séparable.

Théorème de faible comparaison (1.D.1).- Soient (Ω,d) un espace métrique séparable et $<.,.>$ un pseudo-produit scalaire sur M_Ω. Si $<\delta x,\delta y>$ est bornée, on peut énoncer :

(1) pour que $<.,.>$ soit moins fin que la topologie faible sur M_Ω^+ il faut et il suffit que $<.,.>$ se réintègre et que $<\delta x,\delta y>$ soit d⊗d-continu.

(2) Si $<.,.>$ vérifie (1), il existe une suite (f_i) uniformément bornée de C_b, telle que $<\delta x,\delta y> = \sum_{i=1}^{\infty} f_i(x)f_i(y)$ (la convergence étant simple uniforme).

(3) Si de plus $<.,.>$ est non dégénéré, pour tout compact A de Ω (f_i/A) est totale dans $U_b(A)$ et $<.,.>$ a pour trace sur M_A^+ la topologie faible. ($U_b(A)$ étant l'ensemble des fonctions réelles, bornées et uniformément continues sur A).

Preuve :

1) Condition nécessaire.

$<.,.>$ étant moins fin sur M_Ω^+ que F_d, on a $(x_n \xrightarrow[n\to\infty]{} x,\ y_n \xrightarrow[n\to\infty]{} y) \Longrightarrow (<\delta x_n,\delta y_n> \to <\delta x,\delta y>)$, et $<\delta x,\delta y>$ est continu. μ étant un élément de P_Ω, il existe une suite (μ_n) incluse dans P_o telle que $(\mu_n \xRightarrow[n\to\infty]{} \mu)$ (théorème 6.3, p. 44 de [49]). Si l'on pose $K(x,y) = <\delta x,\delta y>$ on a alors :

(*) voir [46], page 99.

- $\langle \mu_n, \mu_n \rangle \to \langle \mu, \mu \rangle$
- $\mu_n \otimes \mu_n \xRightarrow[n\to\infty]{} \mu \otimes \mu$ (lemme 1.1. p. 57 de [49]).
- $\int K \, d\mu_n \otimes \mu_n \longrightarrow \int K \, d\mu \otimes \mu$ (K étant continu et borné)
- $\langle \mu_n, \mu_n \rangle = \int K \, d\mu_n \otimes \mu_n$.

D'où $\langle \mu, \mu \rangle = \int K \, d\mu \otimes \mu$, et $\langle .,. \rangle$ se réintègre puisque deux pseudo-produits scalaires ayant même norme sur P_Ω sont égaux.

Condition suffisante : Si l'on pose $K(x,y) = \langle \delta x, \delta y \rangle$ on a $\langle \mu, \nu \rangle = \int K \, d\mu \otimes \nu$ avec K continu et borné. D'où si $(\mu_n \xRightarrow[n\to\infty]{} \mu)$ dans P_Ω on a :

$$\int K d(\mu_n - \mu) \otimes (\mu_n - \mu) \to \int K \, d\mu \otimes \mu + \int K \, d\mu \otimes \mu - 2\int K \, d\mu \otimes \mu = 0.$$

Supposons maintenant que $(\mu_n \xRightarrow[n\to\infty]{} \mu)$ dans M_Ω^+ ; si $\mu = 0$ on a $\mu_n(\Omega) \to 0$ et $\int K \, d\mu_n \otimes \mu_n \xrightarrow[n\to\infty]{} 0$; si $\mu \neq 0$, on a $\left(\frac{\mu_n}{\mu_n(\Omega)} \Longrightarrow \frac{\mu}{\mu(\Omega)}\right)$ qui implique $\left(\frac{\mu_n}{\mu_n(\Omega)} \xrightarrow{\langle .,. \rangle} \frac{\mu}{\mu(\Omega)}\right)$ et $(\mu_n \xrightarrow{\langle .,. \rangle} \mu)$.

2) (P_Ω, F_d) étant séparable, $(P_\Omega, (.,.)_K)$ l'est et de ce fait $(M_\Omega, (.,.)_K)$ l'est ; (2) découle alors directement du théorème de séparabilité (2.A.I) (puisque H_K est séparable).

3) En vertu de (2), il existe une suite (f_i) de C_b telle que $K(x,y) = \langle \delta x, \delta y \rangle = \sum_{i=1}^{\infty} f_i(x) f_i(y)$. A étant compact dans Ω, le théorème classique de Dini nous permet d'affirmer que la convergence de cette série est uniforme sur la diagonale de $A \times A$ et l'inégalité de Schwartz nous permet d'écrire que la convergence de la série $\sum_{i=1}^{\infty} f_i \otimes f_i$ est uniforme sur $A \times A$.

Si (f_i/A) n'est pas totale dans $C_b(A)$, on peut exhiber g dans $C_b(A) - \overline{S_p((f_i))}$ ($\overline{S_p((f_i))}$ étant l'adhérence uniforme de $S_p((f_i))$ enveloppe vectorielle de la suite (f_i)) ; il existe alors (corollaire du théorème de Hahn-Banach) une forme linéaire continue, F, nulle sur $\overline{S_p(f_i)}$ et non nulle en g. Le théorème de Riesz nous permet d'affirmer qu'il existe $\mu \in \mathcal{M}_A$ telle que $\forall f \in C_b(A)$, $F(f) = \int_A f(x)d\mu(x)$; alors si pour tout B de $\mathcal{B}(d)$ on pose $\bar{\mu}(B) = \mu(A \cap B)$, $\bar{\mu} \in \mathcal{M}_\Omega$ et l'on a $\bar{\mu} \neq 0$ et $\langle\bar{\mu},\bar{\mu}\rangle = 0$ qui est contraire à $\langle .,. \rangle$ non dégénéré. (f_i/A) étant totale dans $C_b(A)$ la topologie trace de $\langle .,. \rangle$ sur $\mathcal{P}_A$ est la topologie faible.

Supposons maintenant $(\mu_n) \cup \{\mu\} \subset \mathcal{M}_A^+$; (f_i/A) étant totale dans $C_b(A)$, il existe une combinaison linéaire h_o des f_i telle que $||h_o/A - 1_A||_U < \frac{1}{2}$; on a alors $h_o/A \geq \frac{1}{2} A$, $\int h_o/A \, d\mu_n \geq \frac{1}{2} \mu_n(A)$ et $(||\mu_n||_{var})$ est bornée ; on a donc $(\mu_n \xrightarrow{\langle .,. \rangle} \mu) \Rightarrow (\mu_n(A) \to \mu(A))$.

Si $\mu = 0$, on a $(\mu_n \Longrightarrow \mu) \Longleftrightarrow (\mu_n \xrightarrow{\langle .,. \rangle} \mu)$

Si $\mu \neq 0$, on a :

$$(\mu_n \Longrightarrow \mu) \Longleftrightarrow \left(\frac{\mu_n}{\mu_n(A)} \Longrightarrow \frac{\mu}{\mu_n(A)}\right) \Longleftrightarrow \left(\frac{\mu_n}{\mu_n(A)} \xrightarrow{\langle .,. \rangle} \frac{\mu}{\mu(A)}\right) \Longleftrightarrow (\mu_n \xrightarrow{\langle .,. \rangle} \mu) \blacksquare$$

Remarque :

Si l'on se suppose placé dans les conditions de l'assertion (3) avec en plus Ω compact, on peut énoncer :

. $\mathcal{F}_d$ et $\langle .,. \rangle$ ont même topologie trace sur $\mathcal{M}_\Omega^+$

. $\mathcal{F}_d$ est plus fine que $\langle .,. \rangle$ sur $\mathcal{M}_\Omega$ tout entier. En effet, si $\mathcal{F}_n \xrightarrow{\mathcal{F}_d} \mu$, $(\mu_n) \cup \{\mu\}$ est compact pour $\mathcal{F}_d$, donc borné pour $\mathcal{F}_d$ (prop. 3, p. 6, chap. III § 2, n° 3 de [11]), donc borné pour la topologie forte (prop. (1), p. 86, chap. IV, § 3, n° 3 de [11]), et de ce fait borné pour la norme de la variation totale. Ce qui nous permet de conclure puisque

$\sum_i f_i \otimes f_i$ converge uniformément). [Cette remarque est vraie pour Ω non compact].

. Ω étant compact, on a $C_b^* = M_\Omega$ (théorème de Riesz) ; toute boule en variation $||\mu||_{var} \leq p$ est alors compact pour F_d (prop. 1, page 112 chap. IV, § 5, n° 2 de [11]) puisque $|| \ ||_{C_b^*} = || \ ||_{var}$; la topologie de <.,.> étant moins fine que F_d et séparée, lui est donc égale sur cette boule (qui est donc compacte pour <.,.>).

. P_Ω étant fermé dans la boule $||\mu||_{var} \leq 1$ est aussi compact.

. $(M_\Omega^+, <.,.>)$ est complet. En effet, si (μ_n) est de Cauchy dans $(M_\Omega^+, <.,.>)$, $(\mu_n(\Omega))$ est bornée ; (μ_n) est donc incluse dans une boule en variation qui est compact dans $(M_\Omega, <.,.>)$; ce qui nous permet de conclure .

. Par contre dès que M_Ω est de dimension infinie, $(M_\Omega, <.,.>)$ n'est pas complet ; en effet, il l'est pour $|| \ ||_{C_b^*} = || \ ||_{var}$ qui est strictement plus fine (remarque 2, p. 111, chap. IV, § 5, n° 1 de [11]) que F_d, donc que <.,.>.

. La topologie de <.,.> est en général strictement moins fine que F_d sur M_Ω tout entier. [En effet, $\Omega = \{0\} \cup \{\frac{1}{n}, n \in \mathbb{N}^*, n \geq 2\}$ muni de la métrique de la valeur absolue est compact ; la suite $g_1 = 1$, $g_2 = 1_{\{\frac{1}{2}\}}, \dots$ $g_n = 1_{\{\frac{1}{n}\}}, \dots$ est en vertu du théorème de Stone-Weierstrass totale dans C_b ; et si l'on pose $K_1 = \sum_{i=1}^{\infty} \frac{1}{i^2} g_i \otimes g_i$, $K_2 = \sum_{i=1}^{\infty} \frac{1}{i^4} g_i \otimes g_i$, $(.,.)_{K_1}$ et $(.,.)_{K_2}$ vérifient les conditions de l'assertion (3) du théorème de faible comparaison et pourtant $\mu_n = n^2(\delta_{\frac{1}{n}} - \delta_{\frac{1}{n-1}})$ est une suite de Cauchy de $(M_\Omega,(.,.)_{K_2})$ mais n'est pas une suite de Cauchy de $(M_\Omega,(.,.)_{K_1})]$.

Remarque : (Ω,d) étant séparable, si K est continu et si $K = \sum_i f_i \otimes f_i$ (la convergence étant uniforme), F_d est plus fine que $(.,.)_K$ sur M_Ω tout entier.

Remarque : Le théorème de faible comparaison ne peut être amélioré ; (Ω,d) étant séparable et $<.,.>$ étant un produit scalaire (non dégénéré) vérifiant l'assertion (1) du théorème de faible comparaison, la topologie trace de $<.,.>$ sur M_Ω^+ peut être strictement moins fine que F_d. Il suffit pour cela de prendre $\Omega = R$, de choisir (pour tout $i \in N^*$) une suite $f_{i,j}$ de fonctions continues, bornées par 1, nulles sur $]-\infty,-(i+1)[\cup]i+1,+\infty[$, telle que la suite des restrictions $(f_{i,j}/[-i,+i])$ soit totale dans $C_b([-i,+i])$ et de poser $K = \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} f_{i,j} \otimes f_{i,j}$; $(f_{i,j},(i,j) \in N^* \times N^*)$ étant totale dans l'ensemble des fonctions continues à support compact, $(.,.)_K$ est non dégénéré et vérifie l'assertion (1) du théorème de faible comparaison ; et pourtant (δ_n) est une suite tendant vers zéro pour $(.,.)_K$ alors qu'elle ne converge pas vers 0 pour la métrique de Prokorov].

Remarque : (Ω,d) étant séparable si $<.,.>$ à pour trace F_d sur M_Ω^+ (resp. P_Ω), $<.,.>$ est non dégénéré dès que Ω s'écrit comme une réunion dénombrable de compact.

Proposition (2.D.I).- (Ω,d) étant séparable ; si K est un noyau reproduisant borné, $B(d)\otimes B(d)$-mesurable et si H_K est séparable, on peut exhiber une métrique, d', ayant $B(d)$ pour tribu borélienne, telle que $F_{d'}$ soit plus fine que $(.,.)_K$ sur M_Ω^+.

Preuve : découle directement du lemme (1.B.I) et du théorème de faible comparaison. ■

Remarque : Si $K = \sum_i f_i \otimes f_i$ la convergence étant uniforme en (x,y), $F_{d'}$ est plus fine que $(.,.)_K$ sur M_Ω tout entier.

Corollaire (3.D.I).- (Ω,d) étant séparable ; si K est un noyau reproduisant borné, $\mathcal{B}(d)\otimes\mathcal{B}(d)$-mesurable et si H_K est séparable, l'ensemble des combinaisons convexes de mesures de Dirac est dense dans $(P_\Omega,(.,.)_K)$.

Preuve : l'ensemble des combinaisons convexes de mesures de Dirac étant dense dans $(P_\Omega,F_{d'})$, la proposition ci-dessus nous permet de conclure. ■

Théorème d'existence (4.D.I).- Si (Ω,d) est séparable, il existe un noyau reproduisant, K, non dégénéré (i.e. : $\int K\, d\mu \otimes \mu = 0 \Rightarrow \mu = 0$) tel que la topologie trace de $(.,.)_K$ sur M_Ω^+ soit la topologie faible.

Preuve : Il existe alors une métrique d' totalement bornée(*) équivalente à d ; l'espace des applications réelles, bornées, d'-uniformément continues muni de la norme de convergence uniforme (relative à d') étant séparable, il existe une suite (f_i) bornée par 1 et totale dans l'ensemble de ces applications ; si l'on pose $K = \sum_{i=1}^{\infty} \frac{1}{2^i} f_i \otimes f_i$, $(.,.)_K$ est moins fin sur M_Ω^+ que la topologie faible (théorème de faible comparaison). Maintenant si $\mu_n \xrightarrow[n\to\infty]{<.,.>} \mu$ dans P_Ω, on a par linéarité et limite uniforme :

$\int f\, d\mu_n \longrightarrow f\, d\mu$ pour toute application f d'- uniformément continue et bornée ; le théorème de Billingsley (**) nous permet alors de conclure pour P_Ω ; on passe alors de P_Ω à M_Ω^+ comme dans le théorème de faible comparaison. Enfin, si $(\mu,\mu)_K = 0$, on a $(\int f_i d\mu = 0, \forall i \in \mathbb{N}^*)$ qui est équivalent à $(\int f\, d\mu = 0, \forall f$ d'-uniformément continue$)$, d'où $\mu = 0$. ■

Remarque :

. F_d est alors plus fine que $(.,.)_K$ sur M_Ω tout entier.

. K étant $d \otimes d$ continu, si (x_n) est dense dans (Ω,d), (δx_n) est

---

(**) Voir [49] par exemple.

(*) d' totalement bornée est identique à (Ω,d') précompact.

totale dans $(M_\Omega,(.,.)_K)$.

Remarque : (Ω,d) étant séparable ; si Ω_1 est un borélien de Ω tel que $\bar{\Omega}_1 = \Omega$ (resp. $\bar{\Omega}_1$ est compact , Ω_1 est convexe et Ω est de Hilbert), et si $<.,.>$ est un produit scalaire sur M_Ω dont la topologie sur M_Ω^+ est F_d ; $<\mu/C_1,\mu/C_1>$ admet F_d pour trace sur $M_{\Omega_1}^+$.

Remarque :

(Ω,d) étant séparable, A et B deux boréliens de Ω, $<.,.>_A$ et $<.,.>_B$ deux produits scalaires sur M_A et M_B ; si pour tout $\mu,\nu \in M_{\Omega_1}$ $(\Omega_1 = A \cup B)$ on pose

$<\mu,\nu> = <\mu/A,\mu/A>_A + <\mu/B,\nu/B>_B$, on peut énoncer :

(1) $<.,.>$ est un produit scalaire sur M_{Ω_1}.

(2) Si $A \cap B = \emptyset$ et si $<.,.>_A$ (resp. $<.,.>_B$) est plus fin que F_d sur M_A^+ (resp. M_B^+), il est de même pour $<.,.>$ sur $M_{\Omega_1}^+$.

(3) Si $d(A,B) > 0$ et si $<.,.>_A$ (resp. $<.,.>_B$) admet F_d pour trace sur M_A^+ (resp. M_B^+) ; $<.,.>$ admet F_d pour trace sur $M_{\Omega_1}^+$.

Remarque : (Ω,d) étant séparable, localement compact et dénombrable à l'infini ; si (f_i) est totale dans l'ensemble des fonctions continues à support compact et si $K = 1 + \sum_{i=1}^{\infty} f_i \otimes f_i$ est borné, $(.,.)_K$ admet F_d pour trace sur $M_{\Omega_1}^+$.

(E - I)- CONSTRUCTION

Nous allons nous limiter à donner des produits scalaires sur M_R, M_{R^n}, M_A (où $A \subset R$ (resp. $A \subset R^n$)) dont la topologie trace sur M_R^+, $M_{R^n}^+$, M_A^+ est égale à (resp. moins fine que) F_d.

Théorème de construction (I.E.I).- ψ étant un homéomorphisme de R sur $\psi(R) \subset]-a_\psi, a_\psi[$ (avec $a_\psi < +\infty$), si $(\alpha_{p(n)})$ est une suite strictement positive et si $K(x,y) = \sum_{n=0}^{\infty} \alpha_{p(n)} (\psi(x))^{p(n)} (\psi(y))^{p(n)}$ (resp : $K(x,y) = \sum_{n=0}^{\infty} \alpha_{n(p)} e^{-p(n)\psi(x)} e^{-p(n)\psi(y)}$) [où $p(n)$ est une suite croissante de N telle que $(x^{p(n)}, n \in N^*)$, (resp : $(e^{-p(n)x}, n \in N^*)$) soit totale dans $U_b([-a\psi, a\psi])$ (*)] est borné, le produit scalaire $(.,.)_K$ a pour topologie trace, sur M_R^+, la topologie faible.

Preuve : Les deux démonstrations étant identiques, nous établirons uniquement celle relative au premier noyau. R étant muni de la métrique $d(x,y) = |\psi(x)-\psi(y)|$ est précompact (puisque ψ est une isométrie de (R,d) sur $(\psi(R), |\ |)$). Si l'on pose $f_{p(n)}(x) = x^{p(n)}$, $(f_{p(n)})$ est totale dans $U_b(\overline{\psi(R)})$, donc dans $U_b(\psi(R))$; g étant d-uniformément continue, $g \circ \psi^{-1}$ est uniformément continue sur $\psi(R)$; il existe donc $\lambda_1, \dots, \lambda_k$, $f_{p(n_1)}, \dots, f_{p(n_k)}$ tels que $\| g \circ \psi^{-1} - \Sigma \lambda_i f_{p(n_i)} \|_u < \varepsilon$ (dans (R,d)) ; $(\psi^{p(n)})$ est donc totale dans $U_b(R,d)$; si $(\mu_n \xrightarrow[n\to\infty]{Fd} \mu)$ dans P_R, on a par linéarité et limite uniforme $(\forall f \in U_b(R,d), \int f d\mu_n \to \int f\, d\mu)$; d'où $(\forall f \in C_b(R), \int f\, d\mu_n \longrightarrow f\, d\mu)$ (théorème de Billingsley (**)) ; le théorème de faible comparaison nous permet alors d'affirmer que Fd et $(.,.)_K$ ont même trace sur P_R, et l'on passe de P_R à M_R^+ comme à la fin du théorème de faible comparaison. $(.,.)_K$ est bien sûr non dégénéré. ■

Remarque : Si la convergence des séries (ci-dessus) est uniforme, la topologie faible est plus fine que $(.,.)_K$ sur M_R tout entier. En faisant varier ψ et (α_n) dans le théorème ci-dessus on obtient les noyaux suivants :

(*) on peut, par exemple, prendre $(p(n) = n)$ ou $(p(0) = 0$ et $p(n) = 2n+1$, $n \in N)$. Maintenant si l'on remplace "$\psi(x)$" par "$a_\psi+\psi(x)$" il suffit (en vertu du théorème de Müntz p. 22, de [45]) de prendre pour $(p(n))$ une suite croissante telle que $p(0) = 0$ et $\sum_{n>0} \frac{1}{p(n)}$ diverge.

(**) voir 49 (par exemple).

$$K_1(x,y) = \sum_{i=0}^{\infty} \frac{1}{2^n} \left(\frac{x}{1+|x|}\right)^n \left(\frac{y}{1+|y|}\right)^n = \frac{1}{1 - \frac{1}{2}\frac{xy}{(1+|x|)(1+|y|)}}$$

$$K_2(x,y) = \sum_{i=0}^{\infty} \frac{1}{n!} \left(\frac{x}{1+|x|}\right)^n \left(\frac{y}{1+|y|}\right)^n = e^{\left(\frac{x}{1+|x|}\right)\left(\frac{y}{1+|y|}\right)}$$

$$K_3(x,y) = \sum_{i=0}^{\infty} \frac{1}{10^n} \exp\left(-n\left(\frac{x}{1+|x|}\right)\right)\exp\left(-n\frac{y}{1+|y|}\right) = \left(1 - \frac{1}{10}\exp\left(-\frac{x}{1+|x|}\right)\exp\left(-\frac{y}{1+|y|}\right)\right)^{-1}$$

$$K_4(x,y) = \sum_{i=0}^{\infty} \frac{1}{n!} \exp\left(-\frac{n\,x}{1+|x|}\right)\exp\left(-\frac{n\,y}{1+|y|}\right) = \exp\left(\exp\left(-\frac{x}{1+|x|}\right)\exp\left(-\frac{y}{1+|y|}\right)\right)$$

$$K_5(x,y) = 1 + \sum_{i=1}^{\infty} \frac{n}{n!} \left(\frac{x}{1+|x|}\right)^n \left(\frac{y}{1+|y|}\right)^n = 1 + \frac{xy}{(1+|x|)(1+|y|)}\, e^{\left(\frac{x}{1+|x|}\right)\left(\frac{y}{1+|y|}\right)}$$

$$K_6(x,y) = \sum_{i=0}^{\infty} \frac{1}{2^n} \left(\frac{x}{|x|}\, e^{-|x|}\right)^n \left(\frac{y}{|y|}\, e^{-|y|}\right)^n = \frac{1}{1 - \frac{1}{2}\frac{xy}{|xy|}\, e^{-|x|\,|y|}}$$

$$K_7(x,y) = \sum_{i=0}^{\infty} \frac{1}{3^n} (\mathrm{Arctg}(x))^n (\mathrm{Arctg}(y))^n = \frac{1}{1 - \frac{1}{3}\mathrm{Arctg}(x)\mathrm{Arctg}(y)}$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

Lorsque Ω est borné dans R, on peut prendre pour noyau $K(x,y) = e^{xy}, \ldots$, (voir [34], page 93).

Proposition (2.E.1).- Si l'on pose $K(x,y) = e^{-\frac{1}{2}(x-y)^2}$, $\int K d\mu \otimes \mu$ est un <u>produit scalaire</u> sur M_R dont la trace sur M_R^+ est moins fine que la topologie faible, et dont la trace sur M_A^+ est, pour tout compact A, la topologie faible.

Preuve : μ étant un élément de M_R, si $\int e^{-1/2(x-y)^2} d\mu \otimes \mu(x,y) = 0$, on a :

$$\int \exp(iu(x-y)) \exp(-\frac{1}{2}u^2)\, du\, d\mu \otimes \mu(x,y) = 0$$

(puisque $e^{-1/2(x-y)^2} = \int \exp(iu(x-y)) \times \exp(-\frac{1}{2}u^2)\, du$ [théorème de réciprocité de Fourier]), d'où

$$\int \exp(-\frac{1}{2}u^2)\, |f_\mu(u)|^2\, du = 0 \quad ,$$

f_μ étant la transformée de Fourier de μ ; f_μ est donc identiquement nulle , et de ce fait μ aussi. ■

Remarque : Ce que l'on vient de faire pour $e^{-1/2(x-y)^2}$ s'étend immédiatement à $e^{-(x-y)}$ etc...

L'intérêt de ce produit scalaire réside dans sa simplicité et dans le fait que toute mesure de Dirac a la même norme.

Lemme des produits (3.E.1).- ψ étant un homéomorphisme de R sur $\psi(R) =]a\psi,b\psi[$ (avec $\max(|a\psi|,|b\psi|) < +\infty$), si (α_n) est une série strictement positive et si $K(x,y) = \sum_{n=0}^{\infty} \alpha_n(\psi(x))^n(\psi(y))^n$ (resp. $K(x,y) = \sum_{n=0}^{\infty} \alpha_n e^{-n\psi(x)}e^{-n\psi(y)}$) est borné, $(.,.)_K \otimes m$ est un produit scalaire sur M_{R^m} dont la topologie trace sur $M^+_{R^m}$ est la topologie faible.

Preuve : ψ se prolonge de façon unique en une bijection $\bar{\psi}$ de $\bar{R}$ sur $[a\psi,b\psi]$, et si l'on munit $\bar{R}$ de la métrique $\bar{d}(x,y) = |\bar{\psi}(x) - \bar{\psi}(y)|$,

$\bar{\varphi}$ est une isométrie de $\bar{R}$ sur $[a\bar{\varphi},b\bar{\varphi}]$ et $(\bar{R},\bar{d})$ est un complété de $(R,d).((\bar{\varphi}^{\otimes m})^n, n\in N)$ est alors (théorème de Stone-Weierstrass) totale dans $U_b(\bar{R}^m,\bar{d}^{\otimes m})$, et de ce fait $((\varphi^{\otimes m})^n, n \in N)$ est totale dans $U_b(R^m,d^{\otimes m})$. On conclut alors comme dans le théorème de construction. ■

D'où

$$K_1^{\otimes m}((x_1,\dots,x_m),(y_1,\dots,y_m)) =$$

$$= \frac{1}{1-\frac{1}{2}\frac{x_1y_1}{(1+|x_1|)(1+|y_1|)}} \times \dots \times \frac{1}{1-\frac{1}{2}\frac{x_my_m}{(1+|x_m|)(1+|y_m|)}},$$

$$K_2^{\otimes m}((x_1,\dots,x_m),(y_1,\dots,y_m)) =$$

$$= e^{\frac{x_1y_1}{(1+|x_1|))(1+|y_1|)}} \times \dots \times e^{\frac{x_my_m}{(1+x_m)(1+y_m)}}, \dots$$

sont des noyaux reproduisant sur R^m dont les produits scalaires associés $(.,.)_{K_i^{\otimes m}}$ ont pour topologie trace sur $M^+_{R^m}$, la topologie faible.

<u>*Remarque*</u> : Lorsque $\Omega = R^+$ (resp. R,R^n) et lorsque l'on munit M_Ω du produit de convolution usuel (i.e : $\mu * \nu = (\mu \otimes \nu)_S$ où $S : (x,y) \to x+y$) ; il existe (voir [31] page 82-86) des noyaux reproduisants K tels que :

. $(M_\Omega, *, || \ ||_K)$ est une algèbre préhilbertienne.

. La trace de $|| \ ||_K$ sur M^+_Ω est la topologie faible.

($b\,K_3$ et $c\,K_4$ vérifient en particulier cette propriété - [b et c étant des constantes bien choisies]).

$*$ se prolonge alors naturellement à H_K et $(H_K, *, || \ ||_{H_K})$ est alors une algèbre hilbertienne, on peut faire la même remarque pour $\Omega = [0,1]$ lorsque l'on munit M_Ω du produit de convolution $\mu \hat{*} \nu = (\mu \otimes \nu)_\pi$, où $\pi(x,y) = x \times y$.

F - I - "THEOREME DE GLIVENKO-CANTELLI

<u>*Définition*</u>.- $(\Omega,\mathcal{B})$ étant un espace mesurable et K un noyau reproduisant défini sur Ω, nous dirons que K est un <u>noyau normal</u> s'il existe une suite (f_i) de v.a.r. bornées par 1 et une série (α_i) convergente de termes strictement positifs, telles que $K(x,y) = \sum_{i=1}^{\infty} \alpha_i f_i(x) f_i(y)$.

K étant un noyau reproduisant normal, il existe une infinité de séries (α_i) et une infinité de suites (f_i) telles que $K = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$; mais pour simplifier la rédaction, nous ne considérerons que la série (α_i) et la suite (f_i) apparaissant dans l'écriture $K = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$.

Sous ces conditions $K = \sum_i \alpha_i f_i \otimes f_i$ étant un noyau normal et $a > 0$ étant donné, nous désignerons par $n_o(a)$ le plus petit entier tel que $\sum_{i=n_o(a)+1} \alpha_i < \frac{a^2}{2}$ si les f_i sont positives, ou tel que $\sum_{i=n_o(a)+1} \alpha_i < \frac{a^2}{8}$ si les f_i ne sont pas toutes positives.

Nous allons maintenant énoncer trois inégalités, dues à Hoeffding, qui nous seront indispensables dans le théorème suivant. Elles se trouvent page 15 et 16 de [39].

Soient $X_1, X_2, \ldots, X_n$, n v.a.r <u>indépendantes</u> définies sur $(\Omega, \mathcal{A}, P)$ ayant des moments d'ordre (1) et (2), $S = X_1 + \ldots + X_n$, $m = E(\frac{S}{n})$.

(I) Si $0 \leq X_i \leq 1$, pour $i = 1,\ldots,n$, on a pour $0 < a < 1-m$

$$\boxed{P\{|\frac{S}{n} - m| \geq a\} \leq 2\{(\frac{m}{m+a})^{m+a}(\frac{1-m}{1-m-a})^{1-m-a}\}^n \leq 2e^{-na^2 g(m)} \leq 2e^{-2na^2}}$$

où $g(m) = \frac{1}{1-2m} \operatorname{Log}(\frac{1-m}{m})$ pour $0 < m < \frac{1}{2}$ et $g(m) = \frac{1}{2m(1-m)}$ pour $\frac{1}{2} \leq m < 1$

(II) Si $a_i \leqslant X_i \leqslant b_i$ $(i = 1,\dots,n)$, pour tout $a > 0$, on a :

$$\boxed{P\{|\frac{S}{n} - m| > a\} \leq 2\exp\left[-\frac{2n^2a^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right]}$$

(III) Si les X_i ont même loi et si $|X_i| \leqslant b$ $(i = 1,\dots,n)$ on a :

$$P\{|\frac{S}{n} - m| \geqslant a\} \leqslant \{(1+\frac{ba}{\sigma^2})^{-(1+\frac{ba}{\sigma^2})\frac{\sigma^2}{b^2+\sigma^2}}(1-\frac{a}{b})^{-(1-\frac{a}{b})\frac{b^2}{b^2+\sigma^2}}\}^n \leqslant$$

$$\leqslant e^{-\frac{na}{b}|(1+\frac{\sigma^2}{ba})\mathrm{Log}(1+\frac{ba}{\sigma^2})-1|}$$

où $\sigma^2 = \dfrac{(\text{var } S)}{n}$.

Pour tout $\omega = (\omega_1,\dots,\omega_n,\dots) \in \Omega^\infty$, on note $\mu_n^\omega = \frac{1}{n}\sum_{i=1}^{n} \delta\omega_i$, la mesure empirique au point w.

"Théorème de Glivenko-Cantelli" (1.F.I).-

Si $K = \sum_{i=1}^{\infty} \alpha_i\, f_i \otimes f_i$ est un noyau normal et si l'on désigne par $\|\ \|_K$ la semi-norme associée au pseudo-produit scalaire $\int K\, d\mu \otimes \nu$, on peut énoncer :

a) $\displaystyle\sup_{\mu\in P_\Omega} \mu^{\otimes\infty}(\bigcup_{n\geq N_1}\{\omega \in \Omega^\infty : \|\mu_n^\omega - \mu\|_K \geqslant a\}) \leqslant \frac{16n_o(a)\sum_{\ell=1}\alpha_\ell}{a^2.N}$.

b) $\displaystyle\sup_{\mu\in P_\Omega} \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : \|\mu_m^\omega - \mu\|_K \geqslant a\}) \leqslant 2n_o(a)\exp(-\frac{ma^2}{4\sum_{\ell=1}^{\infty}\alpha_\ell})$.

Remarques :

c) Si de plus K est positif (i.e. $K = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$ avec $f_i \geq 0$) on a :

$$\sup_{\mu \in P_\Omega} \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : \|\mu_m^\omega - \mu\|_K \geq a\}) \leq 2n_o(a)\exp(-\frac{ma^2}{\sum_{\ell=1}^{\infty}\alpha_\ell}) .$$

d) Si de plus K est positif et $\alpha_i = 0$ pour $i > p$, on a :

$$\sup_{\mu \in P_\Omega} \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : \|\mu^\omega - \mu\|_K \geq a\}) \leq 2p\exp(-\frac{2ma^2}{\sum_{\ell=1}^{\infty}\alpha_\ell}) .$$

e) Si $K = f \otimes f$ avec $0 \leq f \leq 1$, pour $a < 1$, on peut écrire :

$$\mu^{\otimes\infty}(\{\omega \in \Omega^\infty : \|\mu_m^\omega - \mu\| \geq a\}) \leq 2\left[(1+\frac{a}{\sigma^2})^{-(1+\frac{a}{\sigma^2})\frac{\sigma^2}{1+\sigma^2}}(1-a)^{-\frac{(1-a)}{1+\sigma^2}}\right]^m$$

où $\sigma^2 = \int f^2\,d\mu - (\int f\,d\mu)^2$.

f) Si $K = f \otimes f$ et si $0 < a < 1 - \int f\,d\mu$, on a :

$$\mu^{\otimes\infty}(\{\omega \in \Omega^\infty : \|\mu_m^\omega - \mu\|_K \geq a\}) \leq 2\left[(\frac{b}{b+a})^{b+a}(\frac{1-b}{1-b-a})^{1-b-a}\right]^m$$

où $b = \int f\,d\mu$.

Preuve :

a) pour tout $\mu \in P_\Omega$, on a :

$$\bigcap_{i=1}^{n_o(a)} \{\omega \in \Omega^\infty : |\int f_i\,d\mu_n^\omega - \int f_i\,d\mu| < \frac{a}{(2\sum_{\ell=1}^{\infty}\alpha_\ell)^{1/2}}\}$$ est égal à

$$\{\omega \in \Omega^\infty : \sup_{i\in\{1,\dots,n_o(a)\}} \left|\int f_i d\mu_n^\omega - \int f_i d\mu\right|^2 < \frac{a^2}{2\sum_{\ell=1}^\infty \alpha_\ell}\}$$ qui est inclus dans

$$\{\omega \in \Omega^\infty : \sum_{i=1}^{n_o(\varepsilon)} \alpha_i \left|\int f_i d\mu_n - \int f_i d\mu\right| < \frac{a^2}{2}\}.$$

Et comme :

. $||\mu_n^\omega - \mu||_K^2 = \sum_{i=1}^\infty \alpha_i \left|\int f_i d\mu_n^\omega - \int f_i d\mu\right|^2$

. $||f_i||_u \leqslant 1$

. $\sum_{\ell=n_o(a)+1}^\infty \alpha_\ell < \frac{a^2}{8}$ (ou $\sum_{\ell=n_o(a)+1}^\infty \alpha_\ell < \frac{a^2}{2}$ si les f_i sont positives)

on a :

$$\{\omega \in \Omega^\infty : \sum_{i=1}^{n_o(\varepsilon)} \alpha_i \left|\int f_i d\mu_n - \int f_i d\mu\right|^2 < \frac{\varepsilon^2}{2}\} \subset \{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K < \varepsilon\}.$$

D'où $\forall\mu \in P_\Omega$, $\bigcap_{i=1}^{n_o(a)} \{\omega \in \Omega^\infty : \left|\int f_i d\mu_n^\omega - f_i d\mu\right| < \frac{a}{(2\sum\alpha_\ell)^{1/2}}\} \subset \{\omega \in \Omega^\infty : ||\mu^\omega - \mu||_K < a\}$ (α)

μ étant un élément de P_Ω ; il découle alors de (α) que :

$$\mu^{\otimes\infty}\Big(\bigcup_{n\geqslant N_1} [\omega\in\Omega^\infty : ||\mu_n^\omega - \mu||_K \geqslant a]\Big) \leqslant \sum_{i=1}^{n_o(a)} \mu^{\otimes\infty}\Big(\bigcup_{n\geqslant N_1} \{\omega^\infty : \left|\int f_i d\mu_n^\omega - \int f_i d\mu\right| \geqslant \frac{a}{(2\sum_{\ell=1}^\infty \alpha_\ell)^{1/2}}\}\Big)$$

Si pour $i \in \{1,\dots,n_o(a)\}$ on pose $Y_{ij} = f_i(X_j) - \int f_i d\mu$ (X_j étant la j-ème projection de Ω^∞ sur Ω) ; les Y_{ij} sont - pour i fixé - indépendantes, centrées, de même loi, de $\mu^{\otimes\infty}$-variance inférieure à 1 et $\int f_i d\mu_n^\omega - \int f_i d\mu = \frac{1}{n}\sum_{j=1}^n Y_{ij}$.

Si $N_1 = 2^h$, on a :

$$\mu^{\otimes\infty}\left[\bigcup_{n\geqslant 2^h} \{|\frac{1}{n}\sum_{i=1}^n Y_{ij}| \geqslant \frac{a}{(2\sum_{\ell=1}^\infty \alpha_\ell)^{1/2}}\}\right] = \mu^{\otimes\infty}\left[\bigcup_{k=0} \bigcup_{2^{h+k}\leqslant n\leqslant 2^{h+k+1}} \{|\frac{1}{n}\sum_{j=1}^n Y_{ij}| \geqslant \frac{a}{(2\sum_{\ell=1}^\infty \alpha_\ell)^{1/2}}\}\right]$$

$$\leqslant \sum_{k=0}^{\infty} \mu^{\otimes\infty}\left[\bigcup_{2^{h+k}\leqslant n\leqslant 2^{h+k+1}} \{|\sum_{j=1}^{n} Y_{ij}| \geqslant \frac{a.2^{h+k}}{(2\sum_{\ell=1}^{\infty}\alpha_\ell)^{1/2}}\}\right]$$

$$\leqslant \sum_{k=0}^{\infty} \mu^{\otimes\infty}\left[\bigcup_{n\leqslant 2^{h+k+1}} \{|\sum_{j=1}^{n} Y_{ij}| \geqslant \frac{a.2^{h+k}}{(2\sum_{\ell=1}^{\infty}\alpha_\ell)^{1/2}}\}\right]$$

$$\leqslant \sum_{k=0}^{\infty} \frac{(2\sum_{\ell=1}^{\infty}\alpha_\ell).2^{h+k+1}}{a^2(2^{h+k})^2} \quad \text{(d'après l'inégalité de Kolmogorov (*)).}$$

On obtient donc :

$$\mu^{\otimes\infty}\left[\bigcup_{n\geqslant 2^h}\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K \geqslant a\}\right] \leqslant \frac{8n_o(a) \times \sum_{\ell=1}^{\infty}\alpha_\ell}{a^2.2^h} .$$

Maintenant si N_1 est quelconque, il existe h tel que $2^h \leqslant N_1 \leqslant 2^{h+1}$ et l'on a alors :

$$\mu^{\otimes\infty}\left[\bigcup_{n\geqslant N_1}\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K \geqslant a\}\right] \leqslant \frac{2\times 8n_o(a)\sum_{\ell=1}^{\infty}\alpha_\ell}{2\times a^2\times 2^h} \leqslant \frac{16n_o(a)\sum_{\ell=1}^{\infty}\alpha_\ell}{a^2.N_1} .$$

b) En vertu de (α), pour tout μ de P_Ω, on peut écrire :

$$\mu^{\otimes\infty}[\{\omega \in \Omega^\infty : ||\mu_m^\omega - \mu||_K \geqslant a\}] \leqslant \sum_{i=1}^{n_o(a)} \mu^{\otimes\infty}\left[\{|\frac{1}{m}\sum_{j=1}^{m} Y_{ij}| \geqslant \frac{a}{(2\sum_{\ell=1}^{\infty}\alpha_\ell)^{1/2}}\}\right]$$

(où $Y_{ij} = f_i(X_j) - \int f_i(X_j)d\mu^{\otimes\infty}$ et X_j est la j-ème projection). Ce qui nous permet de conclure en vertu du fait que $-1 \leqslant f_i(X_j) \leqslant 1$ et de l'inégalité (II) de Hoeffding précitée).

(*) Inégalité de Kolmogorov usuelle.

c) Comme $0 \leq f_i(X_j) \leq 1$, on conclut de façon analogue à (b).

d) Il découle de l'inclusion (α) que :

$$\forall \mu \in P_\Omega, \quad \bigcap_{i=1}^{p} \{\omega \in \Omega^\infty : \left|\int f_i d\mu_n^\omega - \int f_i \, d\mu\right| < \frac{a}{\left(\sum_{\ell=1}^{p} \alpha_\ell\right)^{1/2}}\} \subset \{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K < a\} (\beta)$$

ce qui implique :

$$\forall \mu \in P_\Omega, \quad \mu^{\otimes\infty}\left[\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K \geq a\}\right] \leq \sum_{i=1}^{p} \mu^{\otimes\infty}\left[\{|\frac{1}{m} \sum_{j=1}^{m} Y_{ij}| \leq \frac{a}{\left(\sum_{\ell=1}^{p} \alpha_\ell\right)^{1/2}}\}\right]$$

et comme $0 \leq f_i(X_j) \leq 1$, l'inégalité (II) de Hoeffding (précitée) nous permet de conclure.

e) Pour tout μ de P_Ω l'inclusion (β) nous permet d'écrire :

$$\mu^{\otimes\infty}\left[\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K \geq a\}\right] \leq \mu^{\otimes\infty}\left[\{|\frac{1}{m} \sum_{j=1}^{m} Y_{1j}| \geq a\}\right]$$

(où $Y_{1j} = f(X_j) - \int f \, d\mu$) et comme $|f(X_j) - \int f \, d\mu| \leq 1$, l'inégalité (III) de Hoeffding (précitée) nous permet de conclure.

f) On conclut comme en (e) en remplaçant l'inégalité (III) par l'inégalité (I). ■

<u>*Remarque*</u> : **Dans la démonstration de l'assertion (b), si l'on remplace l'inégalité (II) de Hoeffding par l'inégalité de "Cramer-Chernov", on obtient une inégalité équivalente au voisinage de 0.**

Par contre ceci n'est plus vrai dans l'assertion (c) ; le résultat obtenu avec l'inégalité de Hoeffding étant nettement plus fin qu'avec Cramer-Chernov.

Corollaire (2.F.I).- (Ω,d) étant séparable ; si H_K est séparable et si K est $\mathcal{B}(d) \otimes \mathcal{B}(d)$ mesurable et borné, pour tout $a > 0$ et tout $\mu \in \mathcal{P}_\Omega$ il existe une constante $C_1(a,\mu)$ et une constante $C_2(a,\mu)$ telles que :

$$. \ \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K < a, \ \forall n \geq N_1\}) \geq 1 - \frac{C_1(a,\mu)}{N_1} .$$

$$. \ \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K < a\}) \geq 1 - e^{-n\,C_2(a,\mu)} .$$

Preuve : Il existe alors prop. (2.D.I) une métrique d' telle que $\mathcal{F}_{d'}$ soit plus fine que $(.,.)_K$ sur $\mathcal{P}_\Omega$; de même il existe (théorème d'existence) un noyau reproduisant normal K' dont le produit scalaire $(.,.)_{K'}$ associé admet $\mathcal{F}_{d'}$ pour trace sur $\mathcal{P}_\Omega$. μ et a étant donnés, il existe alors $b > 0$ tel que :

$$\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_{K'} < b\} \subset \{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_K < a\} \quad ;$$

le "théorème de Glivenko-Cantelli" nous permet alors de conclure.

Corollaire (3.F.I).- (Ω,d) étant compact, si $<.,.>$ est un produit scalaire sur $\mathcal{M}_\Omega$ dont la topologie trace sur $\mathcal{P}_\Omega$ est moins fine que la topologie faible, pour tout $a > 0$, il existe une constante $C_1(a)$ et une constante $C_2(a)$ telles que :

$$. \ \sup_{\mu\in\mathcal{P}_\Omega} \mu^{\otimes\infty}(\bigcup_{n\in N_1} \{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_{<.,.>} \geq a\}) \leq \frac{C_1(a)}{N_1} .$$

$$. \ \sup_{\mu\in\mathcal{P}_\Omega} \mu^{\otimes\infty}(\{\omega \in \Omega^\infty : ||\mu_n^\omega - \mu||_{<.,.>} \geq a\}) \leq 2e^{-n\,C_2(a)}$$

Preuve : D'après le théorème d'existence, il existe un noyau reproduisant normal, K, dont la topologie trace (de $(.,.)_K$) sur $\mathcal{P}_\Omega$ est $\mathcal{F}_d$;

la métrique associée à $<.,.>$ est alors uniformément moins fine sur P_Ω que celle associée à $(.,.)_K$ (puisque (P_Ω, F_d) est compact). Le "théorème de Glivenko-Cantelli" nous permet alors de conclure.

Remarque : Le corollaire précédent s'étend immédiatement (en passant au complété) au cas où (Ω,d) est totalement borné (i.e. précompact) et borélien de son complété.

(G - I) - CONTINUITE DES PROJECTIONS, SUR UN SOUS-ESPACE HILBERTIEN DE MESURES, EN FONCTION DES NOYAUX REPRODUISANTS.

Dans les paragraphes (D.II), (E.II), (F.II), lors de la construction de certains estimateurs, nous serons souvent amenés à projeter des mesures de probabilité sur des sous-espaces vectoriels de dimension finie de M_Ω, muni d'un produit scalaire de la forme $\int K d\mu \otimes \nu$; le paragraphe (G - I) rend d'une certaine façon nos prochains résultats intrinsèques. En effet, si K_1 est un noyau reproduisant "proche" de K_2, l'estimateur construit à partir d'une projection orthogonale relative au produit scalaire $\int K_1 d\mu \otimes \nu$ sera proche de celui construit à partir d'une projection orthogonale relative au produit scalaire $\int K_2 d\mu \otimes \nu$.

$(\Omega,\mathcal{B})$ étant un espace mesurable, on note $K_o(\Omega^2,R^+)$ l'ensemble des noyaux reproduisants K bornés, mesurables et non dégénérés sur Ω (i.e : $(\mu,\nu)_K = \int K d\mu \otimes \nu$ est non dégénéré sur M_Ω).

$\mu_1,\dots,\mu_k$ étant des éléments linéairement indépendants de M_Ω, pour tout $\nu \in M_\Omega$ et tout $K \in K_o(\Omega^2,R^+)$, on note $pr^K_{E(\mu_1,\dots,\mu_k)}(\nu)$ la projection de ν sur l'espace vectoriel $E(\mu_1,\dots,\mu_k)$ engendrée par $\mu_1,\dots,\mu_k$ (la projection étant prise suivant le produit scalaire $\int K d\mu \otimes \nu$).

d_u désignant la métrique de la convergence uniforme sur $K_o(\Omega^2,R^+)$; sous ces hypothèses, on peut énoncer :

Proposition (1.G.1).- Si $\|\ \|$ est une norme sur l'espace vectoriel $E(\mu_1,\dots,\mu_k)$, on a :

1) l'application

$$P^1_\nu : (K_o(\Omega^2,R^+),d_u) \to (E(\mu_1,\dots,\mu_k),\|\ \|)$$
$$K \to pr^K_{E(\mu_1,\dots,\mu_k)}(\nu)$$

est continue pour tout ν de M_Ω.

2) si H est une partie de $K_o(\Omega^2,R^+)$ telle que pour tout K appartenant à H la matrice $M_K = (\int K d\mu_i \otimes \mu_j)$ vérifie $\det(M_K) \geq a > 0$; l'application

$$P^2_\nu : (H,d_u) \longrightarrow (E(\mu_1,\dots,\mu_k),\|\ \|) \text{ est}$$
$$K \longrightarrow pr^K_{E(\mu_1,\dots,\mu_k)}(\nu)$$

lipschitzienne pour tout ν appartenant à M_Ω.

3) si de plus M_1 est une partie de M_Ω bornée en variation totale, il existe une constante A de R^+ telle que :

$$\sup_{\nu\in M_1} \|pr^K_{E(\mu_1,\dots,\mu_k)}(\nu) - pr^{K'}_{E(\mu_1,\dots,\mu_k)}(\nu)\| \leq A\ d_u(K,K'),\ \forall (K,K') \in H\times H.$$

4) si K_n est une suite de H telle que $K_n(x,y) \xrightarrow[n\to\infty]{} K(x,y)$ pour tout couple (x,y) de $\Omega\times\Omega$, on a :

$$\forall\, \nu \in M(B),\ \|pr^{K_n}_{E(\mu_1,\dots,\mu_k)}(\nu) - pr^K_{E(\mu_1,\dots,\mu_k)}(\nu)\| \xrightarrow[n\to\infty]{} 0.$$

Preuve.- $E(\mu_1,\dots,\mu_k)$ étant de dimension finie, il est complet ; la projection orthogonale $pr^K_{E(\mu_1,\dots,\mu_k)}$ est donc définie. Toutes les normes sur $E(\mu_1,\dots,\mu_k)$ étant équivalentes, nous prendrons pour norme $\|\ \|$, la norme $\|\sum_{i=1}^{k} a_i\mu_i\| = \sup(|a_i|, i \in \{1,\dots,k\})$.

En vertu de la prop. 3, page 346 de [2] nous avons :

$$pr^K_{E(\mu_1,\dots,\mu_k)}(\nu) = \sum_{i=1}^{k} \left(\sum_{j=1}^{n} \left[B_{ij}\int K d\nu\otimes\mu_j\right]\mu_i \right) \quad ;$$

(B_{ij}) étant la matrice inverse de la matrice $M = \left(\int K d\mu_i\otimes\mu_j\right)$ (plus explicitement $B_{ij} = \dfrac{(-1)^{i+j}\det(M_{ji})}{\det M}$, M_{ji} étant la matrice déduite de M en supprimant la j-ème ligne et la i-ème colonnne).

La démonstration des assertions 1, 2, 3 et 4 est alors triviale.

. Nous nous servirons souvent de l'expression de la projection d'une mesure sur un espace vectoriel de dimension finie dont l'expression apparaît dans la démonstration ci-dessus.

Remarque : Lorsque l'espace vectoriel E sur lequel on est amené à projeter n'est pas de dimension finie, on ne peut énoncer une proposition du type ci-dessus. En effet, l'espace E peut être complet pour le produit scalaire $\int K d\mu\otimes\nu$ et ne pas l'être pour le produit scalaire $\int K' d\mu\otimes\nu$, K' étant pourtant aussi proche que l'on veut de K.

DEUXIÈME PARTIE

ESTIMATIONS PAR PROJECTIONS

(A - II) - CONDITION (TOPOLOGIQUE) NECESSAIRE POUR QU'UN PARAMETRE SOIT UNIFORMEMENT ESTIMABLE EN PROBABILITE.

De nombreuses conditions nécessaires (resp. nécessaires et suffisantes) d'existence d'estimateurs convergents ont déjà été établies (voir [4], [13], [28], [44],[41 bis],[40 bis]) ; mais généralement ces conditions sont pratiquement invérifiables puisque portant sur des topologies ou structures uniformes définies sur des espaces produits infinis.

L'intérêt de ce paragraphe vient de ce que la condition nécessaire exhibée s'avère élémentaire (puisqu'elle porte sur la continuité du paramètre relativement à la topologie faible associée à une métrique séparable), et de la simplicité de la démonstration.

Définition (1.A.II.).- Soit ϕ une application de $\mathcal{P}_1$ dans un espace métrique (Θ, d_Θ). On dit que ϕ est estimable uniformément en probabilité s'il existe un estimateur (T_n) tel que :

$$\forall \varepsilon > 0 \ , \ \sup_{P \in \mathcal{P}_1} P^{\otimes\infty}(d_\Theta(T_n, \phi(P)) \geq \varepsilon) \xrightarrow[n\to\infty]{} 0 \ .$$

On dit alors que (T_n) converge uniformément en probabilité vers ϕ.

Proposition (1.A.II.).- Soit (Ω,d) un espace métrique séparable et soit ϕ une application de $\mathcal{P}_1$ dans un espace métrique séparable (Θ, d_Θ). Alors si (T_n) est un estimateur convergent uniformément en probabilité vers ϕ et si (T_n) est régulier (i.e. s'il existe une métrique d' sur Ω telle que :

. $B(d') = B(d)$,

. $\forall n \in \mathbb{N}^*$, T_n est $d'^{\otimes n}$-continue).

L'application ϕ est continue de $(\mathcal{P}_1, \mathcal{F}_{d'})$ dans (Θ , d_Θ).

<u>*Preuve*</u> : (Θ, d_Θ) étant séparable, il existe une métrique d'_Θ topologiquement équivalente à d_Θ telle que $d'_\Theta \leq d_\Theta$ et (Θ, d'_Θ) est précompact. Comme (T_n) converge uniformément en probabilité vers ϕ dans (Θ, d_Θ), il converge aussi uniformément en probabilité vers ϕ dans (Θ, d'_Θ). (Θ, d'_Θ) étant précompact, il existe une suite $(f_i, i \in N^*)$ de fonctions réelles d'-uniformément continues et bornées par 1, telle que la topologie engendrée par la suite $(f_i, i \in N^*)$ soit celle de d'_Θ, donc aussi celle de d_Θ.

Il nous reste à montrer que si $(\nu_\ell \xrightarrow{F_{d'}} \nu)$ on a $\phi(\nu_\ell) \xrightarrow[\ell\to\infty]{} \phi(\nu)$. Or $(\nu_\ell \xrightarrow[\ell\to\infty]{F_{d'}} \nu)$ entraîne $(\nu_\ell^{\otimes n} \xrightarrow[\ell\to\infty]{F_{d'^{\otimes n}}} \nu^{\otimes n})$; ce qui implique :

$$(\forall i \in N^*, \int f_i \circ T_n \, d\nu_\ell^{\otimes n} \longrightarrow \int f_i \circ T_n \, d\nu^{\otimes n}) \qquad \text{(I)} \quad .$$

Maintenant pour tout $n \in N^*$, on peut écrire :

$$|f_i \circ \phi(\nu_\ell) - f_i \circ \phi(\nu)| \leq \left|\int f_i \circ T_n \, d\nu_\ell^{\otimes n} - f_i \circ \phi(\nu_\ell)\right| + \left|\int f_i \circ T_n \, d\nu_\ell^{\otimes n} - \int f_i \circ T_n \, d\nu^{\otimes n}\right| + \left|\int f_i \circ T_n \, d\nu^{\otimes n} - f_i \circ \phi(\nu)\right| \quad \text{(II)}$$

f_i étant d'_Θ-uniformément continue et bornée par 1, on a :

. $\forall (n,p) \in N^* \times P_1$, $|f_i \circ T_n - f_i \circ \phi(P)| \leq 2$

. $\exists N_i : n \geq N_i \Longrightarrow \sup\limits_{P \in P_0} P^{\otimes\infty}\{|f_i \circ T_n - f_i \circ \phi(P)| \geq \frac{\varepsilon}{2}\} < \frac{\varepsilon}{4}$.

Pour $n \geq N_i$, on a alors :

$$\sup_{P\in P_0} \{\left|\int f_i \circ T_n \, dP^{\otimes\infty} - f_i \circ \phi(P)\right|\} \leq \sup_{P\in P_0} \{\left|\int_{\{|f_i \circ T_n - f_i \circ\phi(P)| \geq \frac{\varepsilon}{2}\}} f_i \circ T_n - f_i \circ \phi(P)\right| dP^{\otimes\infty}$$

$$+ \int_{\{|f_i \circ T_n - f_i \circ \phi(P)| < \frac{\varepsilon}{2}\}} |f_i \circ T_n - f_i \circ \phi(P)| \, dP^{\otimes\infty}\} < \frac{2\varepsilon}{4} + \frac{\varepsilon}{2} = \varepsilon .$$

Ce qui en vertu de (I) et (II) implique ($\forall\, i \in \mathbb{N}^*$, $f_i \circ \phi(\nu_\ell) \xrightarrow[\ell\to\infty]{} f_i \circ \phi(\nu)$) et nous permet de conclure. ■

Lemme (2.A.II.).- Soient (Ω, d) et (Θ, d_Θ) deux espaces métriques séparables.

Si pour tout $n \in \mathbb{N}^*$, T_n est un estimateur d'ordre n de la forme :

$T_n(w_1,\ldots,w_n) = S_{p(n)}(h_{(1,n)}(w_{i_1}),\ldots,h_{(p(n),n)}(w_{i_{p(n)}}))$, où $p(n)$ est un entier, $S_{p(n)}$ une application continue de $(\Theta^{p(n)}, d^{\otimes p(n)})$ dans (Θ, d), $h_{(1,n)},\ldots,h_{(p(n),n)}$ $p(n)$ applications mesurables de $(\Omega, \mathcal{B}(d))$ dans $(\Theta, \mathcal{B}(d_\Theta))$, $i_1,\ldots,i_{p(n)}$ sont fixés et $\{w_{i_1},\ldots,w_{i_{p(n)}}\} \subset \{w_1,\ldots,w_n\}$; (T_n) est régulier.

Preuve :

(Θ, d_Θ) étant séparable, il existe d'après la démonstration précédente une suite $(f_i, i \in \mathbb{N}^*)$ de fonctions réelles d_Θ-continues et bornées par 1 qui engendre la topologie de (Θ, d_Θ). $\bigcup_{n\in\mathbb{N}^*}\{h_{(1,n)},\ldots,h_{(p(n),n)}\}$ étant dénombrable, on peut l'écrire sous la forme d'une suite $(g_p, p \in \mathbb{N}^*)$; Ω étant séparable et $(f_i \circ g_p, (i,p) \in \mathbb{N}^* \times \mathbb{N}^*)$ étant dénombrable, il existe une métrique d' (voir démonstration du lemme (1.B.I.)) sur Ω plus fine que d telle que :

. $\mathcal{B}(d') = \mathcal{B}(d)$

. $\forall (i,p) \in \mathbb{N}^* \times \mathbb{N}^*$, $f_i \circ g_p$ est d'-continue.

Les fonctions $h_{(m,n)}$, $m \in \{1,\ldots,p(n)\}$, $n \in \mathbb{N}^*$ sont donc continues de (Ω, d) dans (Θ, d_Θ) et de ce fait T_n est continu de $(\Omega^n, d'^{\otimes n})$ dans (Θ, d_Θ). ■

Remarque :

. D'après le lemme ci-dessus, la plupart des estimateurs usuels sont réguliers, d'où la généralité de la proposition (1.A.II).

. Si maintenant Θ est un e.v.t.l.c.s. b-séparable, on peut énoncer une proposition du même genre que la prop. (1.A.II), en ajoutant (T_n) borné et en remplaçant "si (T_n) est un estimateur convergent uniformément en probabilité" par "si (T_n) est un estimateur convergent faiblement uniformément en probabilité" (i.e. : "si $(e \circ T_n)$ est pour tout $e \in \Theta^*$ un estimateur convergent uniformément en probabilité") et en remplaçant "ϕ continue" par "ϕ faiblement continue". Il suffit, en effet, de remplacer dans la démonstration la suite $(f_i, i \in N^*)$ par une suite $(e_i, i \in N^*)$, de Θ^*, qui engendre la topologie affaiblie sur les bornés de Θ.

Applications (3.A.II.).-

① *à l'estimation de l'espérance et de la variance.*

La prop. (1.A.II) nous permet de retrouver très facilement la non existence d'estimateur régulier convergent uniformément en probabilité vers l'espérance ou la variance, et qu'en particulier (lemme (1.A.II)) les estimateurs usuels de la variance et de l'espérance ne convergent pas uniformément en probabilité.

En effet, si d est une métrique sur R dont la tribu borélienne associée est la tribu usuelle, on a :

$$(1 - \frac{1}{n})\delta_o + \frac{1}{n}\delta_n \xrightarrow{\mathcal{F}_d} \delta_o, \text{ alors que :}$$

$$\int x \, d\left[(1 - \frac{1}{n})\delta_o + \frac{1}{n}\delta_n\right](x) \not\longrightarrow \int x \, d\delta_o(x)$$

$$\int (x-1)^2 d(1 - \frac{1}{n})\delta_o(x) + \frac{1}{n}\delta_n(x) \not\longrightarrow \int x^2 \, d\delta_o(x).$$

② *à l'estimation de la densité.*

Remarquons tout d'abord que l'on a :

. $(x_n \xrightarrow[n\to\infty]{} x,\ x_n \neq x) \Longrightarrow (\delta x_n \underset{n\to\infty}{\Longrightarrow} \delta x)$ alors que

$||\delta x_n - \delta x||_{var} \underset{n\to\infty}{\not\longrightarrow} 0.$

. $(P_n \simeq N(0, \frac{1}{n}),\ P = \delta_o) \Longrightarrow (P_n \Longrightarrow P)$ alors que $||P_n - P||_{var} \underset{n\to\infty}{\not\longrightarrow} 0.$

D'où en général si la dimension de l'enveloppe vectoriel, $E(\mathcal{P}_1)$, de $\mathcal{P}_1$ est infinie la topologie trace de la norme en variation sur $\mathcal{P}_1$ est strictement plus fine que la topologie faible.

D'où si $\mathcal{P}_1$ est un ensemble de mesures de probabilité absolument continues par rapport à une mesure σ-finie (resp. finie) μ, si la dimension de $E(\mu_1)$ est infinie et si l'on munit l'ensemble des densités (par rapport à μ) appartenant à L^1_μ ou L^2_μ muni de la norme de L^1_μ ou de L^2_μ, il n'existe pas d'estimateur régulier de la densité convergent uniformément en probabilité. En particulier, les estimateurs usuels de la densité ne converge pas uniformément en probabilité.

En effet, si $f_n \xrightarrow{L^1_\mu} f$, on a $\sup\limits_{A\in\mathcal{B}(d)} \left|\int_A f_n\, d\mu - \int_A f\, d\mu\right| \underset{n\to\infty}{\longrightarrow} 0.$

Par contre dès que la dimension de l'enveloppe vectoriel, $E(\mathcal{P}_1)$, de $\mathcal{P}_1$ est finie, nous montrons dans le théorème (1.E.II) qu'il en existe de très simples et qui sont de plus sans biais.

(B - II) - CONDITION (TOPOLOGIQUE) NECESSAIRE POUR QU'UN PARAMETRE SOIT ESTIMABLE SANS BIAIS.

Le célèbre "Théorème de Bickel-Lehmann" donne une condition nécessaire quant à l'existence d'un estimateur sans biais ; mais bien souvent cette condition nécessaire est vérifiée alors qu'il n'existe pas d'estimateur sans biais.

La condition topologique nécessaire que nous exhibons est un complément (indépendant) du "théorème de Bickel-Lehmann" elle se révèle d'expression très simple et se montre souvent plus sélective que ce théorème.

Définition (1.B.II.).- Lorsque Θ est un e.v.t.l.c.s. muni de sa tribu borélienne, on dit que T_n est un estimateur faiblement sans biais d'ordre n de ϕ, si pour tout $\mu \in P_1$, on a :

1) T_n est $\mu^{\otimes n}$-intégrable Pettis (*)

2) $\int T_n \, d\mu^{\otimes n} = \phi(\mu)$.

Lorsqu'il existe un tel estimateur T_n, on dit que ϕ est faiblement estimable sans biais d'ordre n ; si, de plus, $n = 1$, on dit que ϕ est de degré (1) et il est alors estimable sans biais de tout ordre.

Définition (2.B.II.).- Lorsque de plus Θ est un espace de Banach séparable, on dit que T_n est un estimateur fortement sans biais d'ordre n de ϕ, si pour tout $\mu \in P_1$, on a :

1) T_n est $\mu^{\otimes n}$-intégrable Bochner (**)

2) $\int T_n \, d\mu^{\otimes n} = \phi(\mu)$.

(*) voir [12] page 21 (par exemple), ou encore [37] page 77

(**) voir 12 page 22 [37] pages 78-80 ou [25] ou encore [9].

Proposition (1.B.II.).- Soient (Ω,d) un espace métrique séparable $\mathcal{P}_1$ un sous-espace de $\mathcal{P}_\Omega$ et ϕ une application de $\mathcal{P}_1$ dans un e.v.t.l.c.s. Θ

Si T_n est un estimateur faiblement borné, faiblement sans biais d'ordre n de ϕ et si T_n est faiblement régulier (i.e : s'il existe une métrique d' sur Ω telle que $\mathcal{B}(d') = \mathcal{B}(d)$ et telle que T_n soit $d'^{\otimes n}$-faiblement continue).

ϕ est faiblement continue de $(\mathcal{P}_1, \mathcal{F}_{d'})$ dans Θ .

Preuve :

Si $(\mu_\ell \xrightarrow{\mathcal{F}_{d'}} \mu)$, on a $(\mu_\ell^{\otimes n} \xrightarrow{\mathcal{F}_{d'}^{\otimes n}} \mu^{\otimes n})$; $e \circ T_n$

étant $d'^{\otimes n}$-continue et bornée, on a :

$(\int e \circ T_n \, d\mu_\ell^{\otimes n} \xrightarrow[\ell\to\infty]{} \int e \circ T_n \, d\mu^{\otimes n})$ qui est équivalent à

$(e \circ \phi(\mu_\ell) \xrightarrow[\ell\to\infty]{} e \circ \phi(\mu))$. ■

Lemme (2.B.II).- Soient (Ω,d) un espace métrique séparable et Θ un e.v.t.l.c.s. b-séparable. Si T_n est un estimateur (d'ordre n) borné et de la forme :

$T_n(w_1,\ldots,w_n) = S_{p(n)}(h_1(w_{i_1}),\ldots,h_{p(n)}(w_{i_{p(n)}}))$, où $p(n)$ est un entier, $S_{p(n)}$ une application continue de $\Theta^{p(n)}$ dans Θ , $h_1,\ldots,h_{p(n)}$ $p(n)$ applications mesurables de $(\Omega,\mathcal{B}(d))$ dans Θ , $i_1,\ldots,i_{p(n)}$ sont fixés et $\{w_{i_1},\ldots,w_{i_{p(n)}}\} \subset \{w_1,\ldots,w_n\}$; T_n est faiblement régulier.

Preuve :

Soient $(e_i, i \in N^*)$ une base de la topologie affaiblie sur les

---

(*) Θ est un e.v.t.l.c.s. b-séparable s'il existe une suite $(e_i, i \in N^*)$ de Θ^* qui engendre la topologie affaiblie sur les bornés de Θ. $(e_i, i \in N^*)$ est alors appelé base de la topologie affaiblie sur les bornés de Θ .

bornés de Θ . L'ensemble $\bigcup_{j=1}^{p(n)} \{e_i \circ h_j^n, i \in N^*\}$ étant dénombrable, il existe une métrique d' sur Ω plus fine que d telle que :

. $\mathcal{B}(d') = \mathcal{B}(d)$

. $\forall (i,j) \in N^* \times \{1,\dots,p(n)\}$ $e_i \circ h_j^n$ est d'-continue

T_n est alors continu de $(\Omega^n, d'^{\otimes n})$ dans Θ muni de la topologie associée à la suite $(e_i, i \in N^*)$. T_n étant borné il est alors faiblement continu de $(\Omega^n, d'^{\otimes n})$ dans Θ. ■

Proposition (3.B.II.).- Soient (Ω, d) un espace métrique séparable et Θ un espace de Hilbert séparable. Si T_n est un estimateur faiblement sans biais d'ordre n de ϕ vérifiant les conditions du lemme (2.B.II) ci-dessus, il existe une métrique d'' telle que :

. $\mathcal{B}(d) = \mathcal{B}(d'')$

. ϕ est continue lorsque l'on munit $\mathcal{P}_1$ de la topologie $\mathcal{F}_{d''}$.

Preuve :

D'après le lemme (2.B.II), T_n est alors régulier ; ce qui en vertu de la prop. (1.B.II) nous permet d'affirmer que ϕ est faiblement continue de $(\mathcal{P}_1, \mathcal{F}_d)$ dans Θ . Θ étant alors un espace métrique séparable il existe une suite (f_i) de fonctions réelles, continues, bornées par 1, telle que la topologie engendrée par la suite (f_i) soit celle de Θ muni de son produit scalaire.

$\bigcup_{j=1}^{p(n)} (f_i \circ h_j, i \in N^*)$ étant dénombrable, il existe une métrique d'' plus fine que d' telle que :

. $\mathcal{B}(d'') = \mathcal{B}(d')$

. $\forall (i,j) \in N^* \times \{1,\dots,p(n)\}$, $f_i \circ h_j$ est d''-continue.

Alors $\forall i \in \mathbb{N}^*$, T_n est $d''^{\otimes n}$-continue ; autrement dit, T_n est $d''^{\otimes n}$-continu et borné, et ϕ est faiblement continue de $(\mathcal{P}_1, \mathcal{F}_{d''})$ dans Θ. Il nous reste à montrer que $(\nu_\ell \xrightarrow{\mathcal{F}_{d''}} \nu)$ dans $(\mathcal{P}_1, \mathcal{F}_{d''})$, entraîne $||\phi(\nu_\ell)|| \xrightarrow[\ell\to\infty]{} ||\phi(\nu)||_\Theta$ or $(\nu_\ell \xrightarrow{\mathcal{F}_{d''}} \nu)$ entraîne

$(\nu_\ell^{\otimes n} \otimes \nu_\ell^{\otimes n} \xrightarrow{\mathcal{F}_{d''^{\otimes n} \otimes d''^{\otimes n}}} \nu^{\otimes n} \otimes \nu^{\otimes n})$, d'où :

$$\int <T_n, T_n> \quad d\nu_\ell^{\otimes n} \otimes \nu_\ell^{\otimes n} \xrightarrow[\ell\to\infty]{} \int <T_n, T_n> \quad d\nu^{\otimes n} \otimes \nu^{\otimes n} \quad ;$$

et comme pour tout μ de $\mathcal{P}_1$, $\int <T_n, T_n> \quad d\mu^{\otimes n} \otimes \mu^{\otimes n} = <\int T_n d\mu^{\otimes n}, \int T_n \, d\mu^{\otimes n}>_\Theta$,

on peut écrire $||\int T_n d\nu_\ell^{\otimes n}||_\Theta^2 \xrightarrow[\ell\to\infty]{} ||\int T_n d\nu^{\otimes n}||_\Theta^2$, qui est équivalent à $||\phi(\nu_\ell)||_\Theta \longrightarrow ||\phi(\nu)||_\Theta$ ∎

Application à l'estimation sans biais de la densité (4.B.II).-

M. Rosenblatt en 1956 dans [51] a été un des premiers à étudier l'estimation de la densité, Bickel-Lehmann lui ont succédé en 1969 dans [5], enfin D. Bosq en 1977 dans [15] (p. 15) a établi un résultat beaucoup plus précis ; en travaillant dans L^2_μ avec des densités bornées et en supposant $\mathcal{P}_1$ convexe, il a montré qu'une condition nécessaire et suffisante pour que la densité soit estimable sans biais, est que l'ensemble des densités associées à $\mathcal{P}_1$ soit inclus dans un espace vectoriel de dimension finie.

Avec des hypothèses différentes, nous retrouvons très simplement ce résultat à partir des prop. (3.B.II) et (1.B.II).

a) On considère un espace métrique séparable (Ω, d), une mesure μ sur $(\Omega, \mathcal{B}(d))$ et $\mathcal{P}_1$ un ensemble de mesures de probabilité absolument continue par rapport à μ, dont les densités (par rapport à μ) appartiennent à $L^1_\mu \cap L^2_\mu$

On munit $L^1_\mu \cap L^2_\mu$ de la norme de $L^2\mu$.

Alors si la dimension de l'enveloppe linéaire, $E(\mathcal{P}_1)$, de $\mathcal{P}_1$ est infinie, il n'existe en général pas d'estimateur borné sans biais d'ordre n de la densité.

Preuve :

Si $\mathcal{P}_1$ est convexe et s'il existe un estimateur borné sans biais d'ordre n de la densité ; la densité étant linéaire sur $\mathcal{P}_1$, il existe d'après la démonstration du théorème de Bickel-Lehmann un estimateur T_1 borné sans biais d'ordre 1 de la densité. Comme T_1 est (Lemme (2.B.II)) faiblement régulier, il existe (prop. (3.B.II)) une métrique d' telle que :

. $\mathcal{B}(d') = \mathcal{B}(d)$

. l'application $(\mathcal{P}_1, \mathcal{F}_{d'}) \to (L^2, \| \ \|_2)$ est faiblement continue.

$$\nu \to \frac{d\nu}{d\mu} .$$

Ce qui est impossible puisque $(f_n \xrightarrow[n\to\infty]{L^2_\mu} f) \Longrightarrow$ $(\sup_{A\in\mathcal{B}(d)} \left|\int_A f_n d\mu - \int_A f\, d\mu\right| \xrightarrow[n\to\infty]{} 0)$ et que la dimension de $E(\mathcal{P}_1)$ étant infinie la norme en variation totale est en général sur $\mathcal{P}_1$ strictement plus fine que la topologie faible. ■

Par contre lorsque la dimension de $E(\mathcal{P}_1)$ est finie, pour tout $n \in \mathbb{N}^*$ nous construisons dans le théorème (1.E.II) d'une façon très simple un estimateur d'ordre n de la densité ; estimateur qui de plus converge uniformément presque surement avec une "vitesse en $\frac{1}{N}$ ".

Remarque : Si $(x_n \xrightarrow[n\to\infty]{d} x)$ avec $x_n \neq x$, on a $(\delta x_n \xrightarrow{\mathcal{F}_d} \delta x)$, alors que $\delta x_n(\{x\}) \not\longrightarrow \delta x(\{x\})$; de même si $(\mu_n \simeq \mathcal{N}(0,\frac{1}{n}),\ \mu = \delta_0)$ on a $(\mu_n \xRightarrow[n\to\infty]{} \mu)$ alors que $\mu_n(]0,+\infty[) \not\longrightarrow \delta_0(]0,+\infty[)$; autrement dit, bien souvent lorsque l'espace vectoriel engendré par $\mathcal{P}_1$ est de dimension infinie, la topologie affaiblie de la norme en variation totale est strictement plus fine que la topologie faible, sur $\mathcal{P}_1$.

b) On considère un espace métrique séparable (Ω,d), une mesure μ α-finie sur $(\Omega,\mathcal{B}(d))$ et $\mathcal{P}_1$ un ensemble de mesures de probabilité absolument continues par rapport à μ. Alors si l'on munit L^1_μ de sa norme usuelle et si la dimension de l'enveloppe linéaire, $E(\mathcal{P}_1))$, de $\mathcal{P}_1$ est infinie, il n'existe en général pas d'estimateur borné sans biais de la densité.

Preuve :

Identique à la précédente en utilisant la remarque ci-dessus et en remplaçant la prop. (3.B.II) par la prop. (1.B.II).

c) En utilisant le lemme de Scheffé, on peut énoncer une assertion analogue à b) en remplaçant la norme de L^1_μ par la topologie de la convergence simple.

(C - II) - CONDITIONS NECESSAIRES ET SUFFISANTES POUR QU'UN PARAMETRE SOIT ESTIMABLE SANS BIAIS.

Proposition (1.C.II.).- Soient (Ω,d) un espace métrique séparable, P_1 un sous-espace de mesures de probabilité contenant l'ensemble des combinaisons convexes de mesures de Dirac (i.e : P_1 de type F), Θ un e.v.t.l.c.s. b-séparable et ϕ une application bornée de P_1 dans Θ .

① Pour qu'il existe un estimateur T_n sans biais d'ordre n de ϕ de la forme $T_n(w_1,\dots,w_n) = \frac{1}{n}\sum_{i=1}^{n} h_i(w_i)$ (où h_i est pour tout i une v.a. bornée) il faut et il suffit que ϕ transforme linéairement les combinaisons convexes de P_1 et qu'il existe une métrique d' telle que :

. $B(d') = B(d)$

. ϕ est faiblement continue lorsque l'on munit P_1 de $F_{d'}$.

Maintenant si ϕ transforme linéairement les combinaisons convexes de P_1, on peut énoncer :

② Pour que ϕ soit estimable sans biais il faut et il suffit qu'il existe une métrique d' telle que :

. $B(d') = B(d)$

. ϕ est faiblement continue lorsque l'on munit P_1 de $F_{d'}$.

1) *Condition nécessaire* :

La linéarité de ϕ sur P_1 découle de la linéarité de l'espérance et du théorème de Fubini. L'existence de la métrique d' est une conséquence de la proposition (1.B.II) et du lemme (2.B.II).

Condition suffisante : Si l'on pose $T_1(\omega) = \phi(\delta\omega)$, T_1 est faiblement continu et borné. De plus, pour toute combinaison convexe $\nu_o = \sum_{i=1}^{p} \alpha_j \, \delta x_j$ de mesures de Dirac, on a : $\int T_1(\omega) d\nu_o(\omega) = \sum_{i=1}^{p} \alpha_j \phi(\delta x_j) = \phi(\sum_{i=1}^{p} \alpha_j \delta x_j) = \phi(\nu_o)$.

Si l'on note $\mathcal{F}_{d'}$ la topologie faible associée à d', pour tout $\nu \in \mathcal{P}_1$, il existe une suite (ν_ℓ) de combinaisons convexes de mesures de Dirac telle que $(\nu_\ell \xrightarrow{\mathcal{F}_{d'}} \nu)$. Pour tout $e \in \Theta^*$, on a alors :

$$\begin{array}{ccc} \int e \circ T_1 \, d\nu_\ell & \longrightarrow & \int e \circ T_1 \, d\nu \\ \| & & \\ e \circ \phi(\nu_\ell) & \longrightarrow & e \circ \phi(\nu). \end{array}$$

D'où $\int T_1 d\nu = \phi(\nu)$. Ce qui en vertu de la linéarité de l'espérance et du théorème de Fubini entraîne $(\forall \nu \in \mathcal{P}_1, \int_{\Omega^n} \frac{1}{n} \sum_{i=1}^{n} T_1(\omega_i) d\nu^{\otimes n}(\omega_1,\ldots,\omega_n) = \phi(\nu)$.

2) *Condition nécessaire* :

Si T_n est un estimateur sans biais d'ordre n de ϕ, le symétrisé T_n^* de T_n est encore un estimateur sans biais d'ordre n de ϕ sur $\mathcal{P}_1$, donc sur l'enveloppe convexe $<\mathcal{D}>$ des mesures de Dirac. $<\mathcal{D}>$ étant convexe et ϕ linéaire sur $<\mathcal{D}>$ le théorème de Bickel-Lehmann (appliqué à $e \circ \phi$ et $e \circ T_n$, $\forall e \in \Theta^*$) nous permet d'affirmer que $T_1(\omega) = \phi(\delta_\omega)$ est le seul estimateur sans biais d'ordre 1 de ϕ sur $<\mathcal{D}>$. $\frac{1}{n} \sum_{i=1}^{n} T_1(\omega_i)$ est alors un

un estimateur symétrique sans biais d'ordre n de ϕ sur $<\mathcal{D}>$. $<\mathcal{D}>$ étant de type F, le théorème d'Halmos (appliqué à $e \circ \phi$ et $e \circ T_n^*$ pour tout $e \in \Theta^*$) nous permet d'écrire que pour tout $(\omega_1,\ldots,\omega_n) \in \Omega^n$,
$T_n^*(\omega_1,\ldots,\omega_n) = \frac{1}{n} \sum_{i=1}^{n} T_1(\omega_i)$.

La condition nécessaire découle alors de celle de 1

Condition suffisante : La démonstration est identique à celle de la condition suffisante de ① ■.

Remarque :

- $\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i)$ est le seul estimateur symétrique, sans biais d'ordre n de ϕ (P_1 étant de type F, il découle du théorème de Halmos que pour tout $e \in \Theta^*$, $e(\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i)$ est le seul estimateur symétrique sans biais d'ordre n de $e \circ \phi$. Ce qui nous permet de conclure en vertu de la définition de l'intégrale de Pettis et du fait que Θ^* sépare les points de Θ).

- P_1 étant de type F, $\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i)$ est faiblement optimal parmi les estimateurs sans biais d'ordre n de ϕ (i.e : $\forall e \in \Theta^*$, $e(\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i))$ est optimal - au sens usuel - dans la classe des estimateurs sans biais d'ordre n de $e \circ \phi$).

- Si de plus Θ est un espace de Banach séparable, $\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i)$ est un estimateur fortement sans biais d'ordre n de ϕ, de plus, il est optimal dans la classe des estimateurs fortement sans biais d'ordre n de ϕ.

- $(\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i))$ est en vertu du théorème de "Glivenko-Cantelli" un estimateur convergent faiblement p.s. vers ϕ avec une "vitesse" en $\frac{1}{N}$ (i.e. : $\forall e \in \Theta^*$, $e(\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i))$ est un estimateur convergent p.s. (avec une vitesse en $\frac{1}{N}$) vers $e \circ \phi$).

- Si, de plus Θ est un espace de Banach séparable, il découle de la loi forte des grands nombres dans un espace de Banach que $\frac{1}{n}\sum_{i=1}^{n}(\delta\omega_i)$ est un estimateur convergent p.s. vers ϕ.

• Si, de plus Θ est un espace de Hilbert séparable, on peut, en vertu de la prop. (3.B.II), remplacer dans la proposition ci-dessus "ϕ est faiblement continue" par "ϕ est continue".

• L'optimalité des estimateurs ci-dessus est encore vrai pour une relation d'ordre plus fine que l'usuelle (voir [34]) pages 123-130).

Remarque : Il suffit de regarder la démonstration précédente pour pouvoir énoncer la proposition suivante :

Proposition (2.C.II.).- Pour que l'estimateur borné T_n soit sans biais d'ordre n pour ϕ, sur $\mathcal{P}_1$, il suffit qu'il existe une métrique d' sur Ω telle que :

- $\mathcal{B}(d') = \mathcal{B}(d)$
- ϕ est faiblement continue lorsque l'on munit $\mathcal{P}_1$ de $\mathcal{F}_{d'}$
- T_n est faiblement $d'^{\otimes n}$-continue
- T_n est sans biais d'ordre n pour ϕ sur un sous-ensemble $\mathcal{P}_2$ dense dans $\mathcal{P}_1$ muni de $\mathcal{F}_{d'}$.

De plus, la condition ci-dessus est en vertu de la prop. (1.B.II) nécessaire dès que T_n est "régulier".

En particulier, lorsque $\Omega = R$ est muni de sa tribu borélienne et $\mathcal{P}_1$ est de type C (i.e : contient l'ensemble des mesures de probabilité qui ont une densité uniforme par morceaux sur une réunion finie d'intervalles disjoints deux à deux) la proposition ci-dessus nous permet d'énoncer :

• Une condition nécessaire et suffisante pour que l'estimateur T_n continu et borné soit sans biais d'ordre n pour ϕ sur $\mathcal{P}_1$ est que :

① ϕ soit continue lorsque l'on munit $\mathcal{P}_1$ de la topologie faible.

② Pour tout μ admettant une densité uniforme par morceaux on ait : $\int T_n \, d\mu^{\otimes n} = \phi(\mu)$.

(D - II) - ESTIMATION DE LA LOI ET APPLICATIONS

Proposition (1.D.II.).- Soient (Ω,d) un espace métrique séparable, $\mathcal{P}_1$ un sous-espace de type F et $\mathcal{T}$ une topologie sur $\mathcal{M}_\Omega$ faisant de $\mathcal{M}_\Omega$ un e.v.t.l.c.s. où $\mathcal{P}_1$ est borné.

Alors pour que la loi soit estimable sans biais, il suffit qu'il existe une métrique d' sur Ω telle que :

. $\mathcal{B}(d') = \mathcal{B}(d)$

. $\mathcal{F}_{d'}$ est plus fine sur $\mathcal{P}_1$ que la topologie affaiblie de $(\mathcal{M}_\Omega,\mathcal{T})$.

Si de plus $\mathcal{M}_\Omega$ est b-séparable cette condition est nécessaire et suffisante.

Preuve : L'application $\phi : \begin{array}{ccc}\mathcal{P}_1 & \longrightarrow & \mathcal{M}_\Omega \\ \mu & \longrightarrow & \mu\end{array}$ étant linéaire sur $\mathcal{P}_1$, nous sommes dans les conditions de l'assertion (2) de la prop. (1.C.II).

Remarque : Comme à la suite de la prop. (1.C.II), on peut énoncer : pour tout $n \in \mathbb{N}^*$, $T_n(\omega_1,\ldots,\omega_n) = \frac{1}{n}\sum_{i=1}^{n} \delta\omega_i$ est le seul estimateur symétrique sans biais d'ordre n de la loi, il est faiblement optimal, (T_n) converge faiblement p.s. (avec une vitesse en $\frac{1}{N}$) vers la loi...

Remarque : Il suffit de regarder l'exemple (5.C.I) pour s'apercevoir qu'il existe des topologies (associées à des produits scalaires) d'aspect sympathique pour lesquelles il n'existe pas d'estimateur sans biais de la loi.

Le théorème de réintégration (4.C.I) nous permet d'énoncer le théorème suivant :

Théorème du sans biais de la loi (2.D.II.).-

Soit $(\Omega,\mathcal{B})$ un espace mesurable. Si K est un noyau reproduisant borné (sur Ω), $\mathcal{B}\otimes\mathcal{B}$-mesurable, non dégénéré (i.e : $\int K\, d\mu\otimes\nu$ est non dégénéré), et si l'on munit M_Ω de $(.,.)_K$, on peut énoncer :

1) $\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i$ est un estimateur sans biais d'ordre n de la loi.

2) Si ϕ est une application linéaire continue de $(M_\Omega, \|\ \|_K)$ dans un e.v.t.l.c.s. Θ, $\frac{1}{n}\sum_{i=1}^{n}\phi(\delta\omega_i)$ est un estimateur sans biais de ϕ.

Preuve : découle directement du théorème de réintégration (4.C.I.)

Ces estimateurs sont bien sûr faiblement optimaux, convergents p.s. et...

Dans toute la suite de ce paragraphe et dans le paragraphe suivant (sauf mention contraire), nous nous supposerons placés sous les conditions suivantes :

. (Ω,d) est un espace métrique séparable, P_1 un sous-espace de P_Ω, K un noyau reproduisant normal non dégénéré (i.e : $K = \sum_{i=1}^{\infty}\alpha_i f_i\otimes f_i$, où $\alpha_i \in \ell_1^+$, f_i est mesurable et bornée par 1 et $(\mu,\nu)_K = \int K d\mu\otimes\nu$ est non dégénéré sur M_Ω).

. pr_A^K désigne la projection orthogonale - associée au produit scalaire $(.,.)_K$ - de M_Ω sur le sous-espace convexe complet A de M_Ω.

. (μ_n) est une suite dense (ou dont l'ensemble des combinaisons convexes est dense) dans $(P_1,(.,.)_K)$.

. $<\mu_1,\dots,\mu_n>$ et $E(\mu_1,\dots,\mu_n)$ désignent respectivement l'enveloppe convexe et vectoriel de $\{\mu_1,\dots,\mu_n\}$.

Proposition (3.D.11.).-

(a) $(pr^K_{<\mu_1,\dots,\mu_n>} \frac{1}{n}\sum_{i=1}^{n} \delta\omega_i))$ est un estimateur convergent presque sûrement vers la loi sur P_1 (qui de plus est strict lorsque P_1 est convexe).

(b) $(pr^K_{E(\mu_1,\dots,\mu_n)} (\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i))$ est un estimateur asymptotiquement sans biais

(i.e. : $\forall\, \mu \in P_1$, $\int_{\Omega^n} pr^K_{E(\mu_1,\dots,\mu_n)} (\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i)\ d\mu^{\otimes n}(\omega_1,\dots,\omega_n) \xrightarrow[n\to\infty]{} \mu)$

convergent presque sûrement vers la loi.

Si, de plus, P_1 est inclus dans l'enveloppe convexe de la suite (μ_n), on peut énoncer :

(c) $(pr^K_{<\mu_1,\dots,\mu_n>} (\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i))$, resp. $(pr^K_{E(\mu_1,\dots,\mu_n)} (\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i))$ est un estimateur de la loi séquentiellement plus rapide que la loi empirique

(i.e : $\forall\, \mu \in P_1$,

$$\exists\, N_\mu \ : n \geq N_\mu \implies \|pr^K_{<\mu_1,\dots,\mu_n>}(\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i) - \mu\|_K \leq$$

$$\leq \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K).$$

(d) pour toute application linéaire $\phi : E((\mu_n : n \in N)) \to (\mathcal{G}, \|\ \|_{\mathcal{G}})$, $(\phi \circ pr^K_{E(\mu_1,\dots,\mu_n)} (\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i))$ est un estimateur séquentiellement sans biais de ϕ (i.e : $\forall\, \mu \in P_1$, $\exists\, N_\mu \ : n \geq N_\mu \implies$

$$\int_{\Omega^n} \phi \circ \mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}\left(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\right)d\mu^{\otimes n}(\omega_1,\dots,\omega_n) = \phi(\mu)).$$

Preuve :

a) les projections étant contractantes, pour tout $\mu \in P_1$, on a :

$$||\mathrm{pr}^K_{<\mu_1,\dots,\mu_n>}\left(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\right)-\mu||_K \leq ||\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i-\mu||_K + ||\mathrm{pr}^K_{<\mu_1,\dots,\mu_n>}\left(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\right)-\mu|| \quad (I)$$

ce qui nous permet de conclure en vertu du "théorème de Glivenko-Cantelli".

b) On montre comme en (a) que $\left(\mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}\left(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\right)\right)$ est un estimateur convergent p.s. vers la loi.

Maintenant $\mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}$ étant linéaire et continue, le théorème (2.D.II) nous permet d'affirmer que :

$$\forall \mu \in P_\Omega, \quad \int_{\Omega^n} \mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}\left(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\right)d\mu^{\otimes n}(\omega_1,\dots,\omega_n) = \mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}(\mu)$$

ce qui nous permet de conclure, puisque pour tout μ de P_1, on a, dans ces conditions :

$$\mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}(\mu) \xrightarrow[n\to\infty]{} \mu.$$

c) Comme $P_1 \subset <(\mu_n, n \in N^*)> = \bigcup_{n\in N^*} <\mu_1,\dots,\mu_n>$; pour tout $\mu \in P_1$, on a $\mathrm{pr}^K_{<\mu_1,\dots,\mu_n>}(\mu) = \mu$ (resp. $\mathrm{pr}^K_{E(\mu_1,\dots,\mu_n)}(\mu) = \mu$) à partir d'un certain rang. La démonstration de cette assertion est alors terminée en vertu de (I).

d) ϕ étant linéaire, la restriction de ϕ à $E(\mu_1,\dots,\mu_n)$ est continue ; $\phi \circ pr^K_{E(\mu_1,\dots,\mu_n)}$ est donc une application linéaire continue ; le théorème (2.D.II) nous permet alors d'affirmer que $\forall\, n \in N^*$,

$$\int_{\Omega^n} \phi \circ pr^K_{E(\mu_1,\dots,\mu_n)} \left(\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i\right) d\mu^{\otimes n}(\omega_1,\dots,\omega_n) = \phi \circ pr^K_{E(\mu_1,\dots,\mu_n)}(\mu) \quad ;$$

et comme à partir d'un certain rang (dépendant de μ) on a $pr^K_{E(\mu_1,\dots,\mu_n)}(\mu) = \mu$, la démonstration est terminée.

Remarque : Sous ces conditions, on peut énoncer :

. Si (μ_n) est une suite de mesures absolument continues par rapport à ν, si $P_1 \subset <(\mu_n)>$, si l'on pose $\phi(\nu) = \frac{d\nu}{d\mu}$ et si l'on munit l'espace des densités d'une norme quelconque, $\left(\phi \circ pr^K_{E(\mu_1,\dots,\mu_n)}\left(\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i\right)\right)$ est un estimateur séquentiellement sans biais de la densité.

On suppose maintenant P_1 convexe.

. Si ϕ est continue de $(P_1, \|\ \|_K)$ dans l'espace topologique Θ, $\left(\phi \circ pr^K_{<\mu_1,\dots,\mu_n>}\left(\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i\right)\right)$ est un estimateur convergent p.s. vers ϕ.

. Si ϕ est une application de P_1 dans un espace métrique séparable et s'il existe un estimateur (T_n) régulier convergent uniformément en probabilité vers ϕ, on peut exhiber un noyau reproduisant normal non dégénéré K tel que pour toute suite (μ_n) dense dans $(P_1, \|\ \|_K)$, $\left(\phi \circ pr^K_{<\mu_1,\dots,\mu_n>}\left(\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i\right)\right)$ soit un estimateur convergent p.s. vers ϕ.

Si ϕ est une application de P_1 dans un e.v.t.l.c.s. b-séparable (resp. dans un espace de Hilbert séparable) et s'il existe un estimateur T_n faiblement régulier sans biais d'ordre n de ϕ, on peut exhiber un noyau reproduisant normal non dégénéré K tel que pour toute suite (μ_n) dense dans $(P_1, \| \|_K)$ (ou dont l'enveloppe convexe est dense dans $(P_1, \| \|_K)$),

$(\phi \text{ o } pr^K_{<\mu_1,\ldots,\mu_n>}(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i))$ est un estimateur convergent faiblement p.s. (resp. p.s.) vers ϕ.

(E - II) - ESTIMATION DE PARAMETRES LINEAIRES ET DE PARAMETRES CONVEXES

Remarque :

Le théorème (2.D.II) du sans biais de la loi nous permet d'affirmer que si ϕ est une application linéaire continue d'un sous-espace vectoriel E de M_Ω muni de $(.,.)_K$ (où K est un noyau reproduisant $\mathcal{B}\otimes\mathcal{B}$-mesurable, borné et non dégénéré), dans un e.v.t.l.c.s., Θ, ϕ est estimable sans biais sur $P_\Omega \cap E$. Plus précisément, si l'on note $\bar{E}$ l'adhérence de E dans le complété $(\widehat{M_\Omega}, \widehat{(.,.)}_K)$ de $(M_\Omega,(.,.)_K)$, $\bar{\phi} \circ pr^K_{\bar{E}}(\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i)$ est un estimateur sans biais d'ordre n de ϕ, qui converge faiblement p.s. avec une "vitesse en $\frac{1}{N}$".

La première assertion du théorème suivant est un cas particulier de la remarque ci-dessus ; la seconde découle directement du "théorème de Glivenko-Cantelli". Nous allons montrer que ces deux assertions sont optimales (dans le sens où l'on ne peut énoncer une telle proposition dès que la dimension de l'espace vectoriel engendré par P_1 n'est pas finie.

Théorème de l'estimation linéaire en dimension finie (1.E.II.).

Soient (Ω,d) un espace métrique séparable, $K = \sum \alpha_i f_i \otimes f_i$ un noyau reproduisant normal, non dégénéré, défini sur Ω et $(\Theta, \|\ \|_\Theta)$ un espace vectoriel normé. Si ϕ est une application linéaire de $E(\mu_1,\dots,\mu_k)$ dans Θ on peut énoncer :

(1) $\phi \circ pr^K_{E(\mu_1,\dots,\mu_k)}(\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i)$ est un estimateur fortement sans biais d'ordre n la restriction de ϕ à $P_\Omega \cap E(\mu_1,\dots,\mu_k)$.

$$(2)\quad \inf_{\mu\in P_\Omega \cap E(\mu_1,\dots,\mu_k)} \mu^{\otimes\infty}\Big[\omega\in\Omega^\infty : \Big\|\phi \circ pr^K_{E(\mu_1,\dots,\mu_k)}\Big(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\Big) - \phi(\mu)\Big\|_{\Theta} < a\varepsilon,$$

$$\forall\, n \geq N_1\Big] \geq 1 - 16 n_o(a\varepsilon)\,\frac{\sum_{i=1}^{\infty}\alpha_i}{a^2\varepsilon^2 N_1},$$

où a est la norme de ϕ en tant qu'application linéaire continue de $E(\mu_1,\dots,\mu_k)$ dans Θ et $n_o(a\varepsilon)$ est le plus petit entier tel que $\sum_{i=n_o(a\varepsilon)+1}^{\infty} \alpha_i < \frac{a^2\varepsilon^2}{2}$ si les f_i sont positives, ou tel que $\sum_{i=n_o(a\varepsilon)+1}^{\infty} \alpha_i < \frac{a^2\varepsilon^2}{8}$ si les f_i ne sont pas toutes positives.

Preuve : La restriction de ϕ à $E(\mu_1,\dots,\mu_k)$ étant continue, l'assertion (1) découle directement du théorème (2.D.II) du sans biais de la loi. La projection sur $E(\mu_1,\dots,\mu_k)$ étant contractante, (2) découle directement de l'assertion (a) du "théorème de Glivenko-Cantelli".

Remarque : $P_\Omega \cap E(\mu_1,\dots,\mu_k)$ étant relativement petit, l'estimateur exhibé n'est sûrement pas optimal. En utilisant [7], on peut dans ce cas montrer qu'il existe une minoration de (2).

Remarque : Sous ces conditions, on peut remplacer l'assertion (2) par

$$\sup_{\mu\in P_\Omega} \mu^{\otimes\infty}\Big[\omega\in\Omega^\infty : \Big\|\phi \circ pr^K_{E(\mu_1,\dots,\mu_k)}\Big(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i\Big) - \phi(\mu)\Big\|_{\Theta} \geq a\varepsilon\Big]$$

$$\leq 2n_o(a\varepsilon)\exp\Big(-\frac{n^2 a^2}{\sum_{i=1}^{\infty}\alpha_i}\Big).$$

Remarque : En vertu de l'application (4.B.II) à l'estimation sans biais de la densité, la première assertion est optimale.

Maintenant l'application (3.A.II.2) à l'estimation convergente uniforme en probabilité de la densité nous permet d'affirmer que la seconde assertion de la prop. (1.E.II) est aussi optimale.

Plus précisément, si la dimension de l'espace vectoriel engendré par $\mathcal{P}_1$ n'est pas finie, on ne peut en toute généralité trouver d'estimateurs vérifiant les assertions (1) et (2) de la prop. (1.E.II).

Remarque : Si, dans le théorème (1.E.II), on remplace l'espace normé Θ par un espace vectoriel top.l.c.s., (1) est encore vrai dès que l'on remplace "fortement sans biais" par faiblement sans biais" et (2) est vrai dès que l'on remplace "ϕ" par "e o ϕ" où $e \in \Theta^*$.

. Si dans le théorème (1.E.II) $\Theta = R$ et ϕ est convexe sur $E(\mu_1,\dots,\mu_k)$; $<\mu_1,\dots,\mu_n>$ étant compact dans $(M_\Omega, \| \ \|_K)$, le corollaire de la prop. (21) du chap. (II), §(2), n° 10 de [10] nous permet d'affirmer que $\phi \circ pr^K_{<\mu_1,\dots,\mu_k>}$ est uniformément continue ; ce qui en vertu du "théorème de Glivenko-Cantelli" implique :

$$\inf_{\mu \in \mathcal{P}_1} \mu^{\otimes\infty}\left[\omega^\infty : \left|\phi \circ pr^K_{<\mu_1,\dots,\mu_k>}\left(\frac{1}{n}\sum_{i=1}^n \delta\omega_i\right) - \phi(\mu)\right| < \varepsilon, \forall n \geq N_1\right] \geq \frac{C(\varepsilon)}{N_1}$$

où $C(\varepsilon)$ est une constante dépendant de ε.

. Si $\mu_1,\dots,\mu_k$ sont linéairement indépendantes, on a (prop. (1.G.I))

$$\phi \circ pr^K_{E(\mu_1,\dots,\mu_k)}\left(\frac{1}{n}\sum_{i=1}^n \delta\omega_i\right) = \frac{1}{n}\,\phi\left(\sum_{i=1}^k\left[\sum_{j=1}^k B_{ij}\left(\sum_{\ell=1}^n\int_\Omega K(x,\omega_\ell)d\mu_j(x)\right]\mu_i\right)\right.$$

où $B_{ij} = (-1)^{i+j}\dfrac{\det M_{ij}}{\det M}$, $M = \alpha_{ij}$, $\alpha_{ij} = \int K\, d\mu_i \otimes \mu_j$ et M_{ij} est la matrice déduite de M en supprimant la j-ème ligne et la i-ème colonne.

De plus, la norme de a de l'application linéaire continue $\phi \circ pr^K_{E(\mu_1,\ldots,\mu_k)}$ – qui est aussi celle de $\phi/_{E(\mu_1,\ldots,\mu_k)}$ – est majorée par : (voir remarque 2, page 190 de [34])

$$(\sum_i^k \|\phi(\mu_i)\|_\theta) \times (k!) \times |\mu_1|(\Omega) \times \ldots \times |\mu_k|(\Omega) \times \frac{1}{\det(M)} .$$

Théorème de l'estimation linéaire en dimension dénombrable (2.E.II.).-

Soient (Ω,d) un espace métrique séparable, K un noyau reproduisant normal non dégénéré défini sur Ω, $(\theta,\|\ \|_\theta)$ un espace vectoriel normé et (μ_n) une suite de M_Ω.

Si ϕ est une application linéaire définie sur l'espace vectoriel $E((\mu_n))$ engendré par la suite (μ_n), on peut exhiber (à l'aide des projections orthogonales sur des sous-espaces vectoriels de dimension finie) un estimateur (S_n) séquentiellement fortement sans biais et convergent presque sûrement de la restriction de ϕ à $P_\Omega \cap E((\mu_n))$.

Preuve : Afin de simplifier la rédaction, nous poserons :

$$T_\ell(\omega^\infty) = \frac{1}{n}\sum_{i=1}^{\ell} \delta\omega_i = T_\ell(\omega_1,\ldots,\omega_n).$$

Pour tout $n \in N^*$, la restriction de ϕ à l'espace vectoriel $E(\mu_1,\ldots,\mu_n)$ est continue et de norme a_n ; il existe alors (assertion (a) du théorème de Glivenko-Cantelli) une suite (N_n), que l'on choisit strictement croissante, telle que :

$$\inf_{\mu\in P_\Omega} (\mu^{\otimes\infty}[\ \|T_\ell - \mu\|_K < \frac{1}{n\ a_n}\ ,\ \forall \ell \geq N_n) \geq 1 - \frac{1}{2^n . n} .$$

Posons $\phi_n = \phi \circ pr^K_{E(\mu_1,\ldots,\mu_n)}$ et construisons l'estimateur (S_n) de la façon suivante :

$$\ldots, S_{N_n} = \phi_n \circ T_{N_n},\ S_{N_n+1} = \phi \circ T_{N_n+1}, \ldots, S_{N_{n+1}-1} = \phi_n \circ T_{N_{n+1}-1},$$

$$S_{N_{n+1}} = \phi_{n+1} \circ T_{N_{n+1}}, \ldots$$

$\varepsilon > 0$ étant donné, il existe $N(\varepsilon)$ tel que $n \geq N(\varepsilon)$ entraîne $\frac{1}{n} < \varepsilon$, et pour tout μ de $P_\Omega \cap E(\mu_n))$, il existe $N(\mu)$ tel que $n \geq N(\mu)$ entraîne μ appartient à $E(\mu_1,\ldots,\mu_n)$; pour $m \geq \max(N(\varepsilon),N(\mu))$, on a alors :

$$\mu^{\otimes\infty}\left[\,||S_{N_n} - \phi(\mu)||_\Theta < \varepsilon, \ldots, ||S_{N_{n+1}-1} - \phi(\mu)||_\Theta < \varepsilon, \quad \forall\, n \geq m\right]$$

$$\geq \mu^{\otimes\infty}\left[a_n||T_{N_n} - \mu||_K < \frac{1}{n}, \ldots, a_n||T_{N_{n+1}-1} - \mu||_K < \frac{1}{n}, \quad \forall\, n \geq m\right]$$

$$\geq \mu^{\otimes\infty}\left[\,||T_\ell - \mu||_K < \frac{1}{n\, a_n}, \quad \forall \ell \geq N_n,\ \forall\, n \geq m\right]$$

$$\geq 1 - \sum_{n=m}^{\infty} \frac{1}{2^n \times n} \geq 1 - \frac{1}{m} \geq 1 - \varepsilon.$$

D'où $\forall\, \mu \in P_\Omega \cap E((\mu_n))$, $\mu^{\otimes\infty}\left[\,||S_\ell - \phi(\mu)||_\Theta < \varepsilon, \forall\, \ell \geq N\right] \xrightarrow[N\to\infty]{} 1.$

Maintenant pour ℓ assez grand il existe $(n,p) \in \mathbb{N} \times \mathbb{N}$ tel que que $S_\ell = \phi_n(T_{N_n+p})$; ϕ_n étant une application continue de $(M(\mathcal{B}),||\ \ ||_K)$ dans $(\Theta,||\ \ ||_\Theta)$ le théorème (2.D.II) nous permet d'écrire :

$$\forall\, \mu \in P_\Omega,\quad \int S_\ell d\mu^{\otimes\ell} = \int \phi_n \circ T_{N_n+p} d\mu^{\otimes N_n+p} = \phi_n(\mu) \quad ;$$

ce qui nous permet de conclure puisque pour tout $\mu \in P_\Omega \cap E(\mu_n)$ on a $\phi_n(\mu) = \phi(\mu)$ à partir d'un certain rang.

Remarque :

. Si dans le théorème ci-dessus $\Theta = R$ et ϕ est convexe sur $E((\mu_n))$, on peut construire un estimateur S_ℓ convergent presque sûrement de la restriction de ϕ à $P_o \cap E((\mu_n))$.

En effet $\phi \circ pr^K_{<\mu_1,\ldots,\mu_n>}$ étant uniformément continue, il suffit de remplacer dans la démonstration précédente :

$$\phi_n = \phi \circ pr^K_{E(\mu_1,\ldots,\mu_n)} \quad \text{par} \quad \phi \circ pr^K_{<\mu_1,\ldots,\mu_n>}$$

et $\frac{1}{a_n}$ par α_n où $\alpha_n > 0$ et vérifie :

$$\forall(\mu,\nu) \in <\mu_1,\ldots,{}_n{}^2, \ ||\mu - \nu||_K < \alpha_n \implies |\phi(\mu) - \phi(\nu)| < 1.$$

. Si dans le théorème ci-dessus la suite (μ_n) est formée d'éléments linéairement indépendants et si $K = \sum_{i=1}^{\infty} \frac{1}{2^i} f_i(x) f_i(y)$ (où $\sup_{i\in N^*} ||f_i||_U \leq 1$), un simple calcul montre que la suite :

$$N_n = 4n^2 \times 2^{n+8} \times E(\sup(b_i^2, i \in \{1,\ldots,n\}) + 1)) \quad \text{convient.}$$

E désignant la partie entière et

$b_n = \sum_{i=1}^{n} ||\phi(\mu_i)||_\Theta \times n! \times |\mu_1|(\Omega) \times \ldots \times |\mu_n|(\Omega) \times \frac{1}{\det(M_n)}$ où M_n est la matrice carrée d'ordre n ayant pour coefficients $(\alpha_{ij}) = (\mu_i,\mu_j)_K$.

. La finesse des deux théorèmes précédents résulte du fait que l'on peut choisir $||\ \ ||_\Theta$ aussi fine que l'on veut.

(F - II) - *ESTIMATION DE PARAMETRES LIMITES*

La proposition suivante n'est pas nouvelle ; en effet, le théorème 6 page 31 de [14] et le théorème (1.A.8), page 20, de [44] sont par exemple du même genre.

Dans ces deux théorèmes (Θ étant supposé séparable) les auteurs ont plutôt cherché à montrer l'existence d'un estimateur convergent qu'à en proposer un ; de ce fait l'estimateur exhibé est inextricable.

L'intérêt de notre proposition vient du fait que :

. L'estimateur proposé étant construit à partir des projections orthogonales, il est d'expression très simple.

. Il est d'-continu (resp. d'-uniformément continu) dès que K est $d'\otimes d'$ continu (resp : $d'\otimes d'$-uniformément continu).

. La séparabilité de Θ n'est pas nécessaire.

. La démonstration est très simple.

<u>*Proposition*</u> *(1.F.II.)*.- Soient $(\Omega, \mathcal{B})$ un espace mesurable, $\mathcal{P}_1$ un sous-espace convexe de $\mathcal{P}_\Omega$, K un noyau reproduisant normal non dégénéré, ϕ une application de $\mathcal{P}_1$ dans un espace métrique Θ et (ϕ_k) une suite d'applications uniformément continues de $(\mathcal{P}_1, \| \ \|_K)$ dans Θ.

Si ϕ est limite simple de la suite (ϕ_k), il existe un estimateur à valeurs dans $\bar{\Theta}$ et convergent presque sûrement vers ϕ sur $\mathcal{P}_1$ ($\bar{\Theta}$ désignant le complété de Θ).

Preuve : L'application ϕ_n étant uniformément continue, elle se prolonge en une application uniformément continue $\bar{\phi}_n$ sur le complété $\overline{P_1}$ de P_1 qui est lui aussi convexe.

Si l'on note $pr^K_{P_1}$ la projection - associée au produit scalaire $(.,.)_K$ - de M_Ω sur $\overline{P}_1$ et si l'on pose $\phi^*_n = \bar{\phi}_n \circ pr^K_{\overline{P}_1}$, ϕ^*_n est une application uniformément continue de $(P_\Omega, ||\ ||_K)$ dans $(\Theta, \bar{d}_\Theta)$; de plus, pour tout μ de P_1, on a :

$$\phi^*_n(\mu) = \phi_n(\mu) \xrightarrow[n\to\infty]{} \phi(\mu) \qquad (I).$$

ϕ^*_n étant uniformément continue, il existe α_n - que l'on choisit inférieur à 1 - tel que :

$$\forall \mu_1, \mu_2 \in P_\Omega : ||\mu_1 - \mu_2||_K < \alpha_n \implies \bar{d}_\Theta(\phi^*_n(\mu_1), \phi^*_n(\mu_2)) < \frac{1}{n}.$$

Il existe alors en vertu du "théorème de Glivenko-Cantelli" une suite N_n - que l'on choisit strictement croissante - telle que :

$$\inf_{\mu \in P_\Omega} \mu^{\otimes\infty}\Big[(\omega_1,\ldots,\omega_\ell,\ldots) : ||\frac{1}{\ell}\sum_{i=1}^{\ell} \delta\omega_i - \mu||_K < \alpha_n,\ \forall \ell \geq N_n\Big] \geq 1 - \frac{1}{n\,2^n}.$$

Posons $T_\ell(\omega_1,\ldots,\omega_\ell) = T_\ell(\omega_1,\ldots,\omega_\ell,\ldots) = \frac{1}{\ell}\sum_{i=1}^{\ell} \delta\omega_i$ et

$$\ldots\ S_{N_n} = \phi^*_n \circ T_{N_n},\ S_{N_n+1} = \phi^*_n \circ T_{N_n+1}, \ldots, S_{N_{n+1}-1} = \phi^*_n \circ T_{N_{n+1}-1},$$

$$S_{N_{n+1}} = \phi^*_{n+1} \circ T_{N_{n+1}}, \ldots$$

$\varepsilon > 0$ étant donné, il existe $N(\varepsilon)$ tel que $n \geq N(\varepsilon)$ entraîne $\frac{1}{n} < \frac{\varepsilon}{2}$, et pour toute mesure de probabilité μ de P_1 il existe (d'après (I)) $N(\mu)$ tel que $n \geq N(\mu)$ entraîne $d(\phi_n^*(\mu),\phi(\mu)) < \frac{\varepsilon}{2}$.

Pour $m \geq \max(N(\varepsilon),N(\mu))$, on a :

$$\mu^{\otimes\infty}\left[\bar{d}(S_{N_n},\phi(\mu)) < \varepsilon,\ \bar{d}(S_{N_n+1},\phi(\mu)) < \varepsilon,\dots,\ \bar{d}(S_{N_{n+1}-1},\phi(\mu)) < \varepsilon,\quad \forall\, n \geq m\right]$$

$$\geq \mu^{\otimes\infty}\left[\bar{d}(S_{N_n},\phi_n^*(\mu)) < \frac{\varepsilon}{2},\ \bar{d}(S_{N_n+1},\phi_n^*(\mu)) < \frac{\varepsilon}{2},\dots,\ \bar{d}(S_{N_{n+1}-1},\phi_n^*(\mu)) < \frac{\varepsilon}{2},\quad \forall n \geq m\right]$$

$$\geq \mu^{\otimes\infty}\left[\,||T_{N_n}-\mu||_K < \alpha_n,\ ||T_{N_n+1}-\mu||_K < \alpha_n,\dots,\ T_{N_{n+1}-1}-\mu||_K < \alpha_n,\quad \forall\, n \geq m\right]$$

$$\geq \mu^{\otimes\infty}\left[\,||T_\ell-\mu||_K < \alpha_n,\quad \forall \ell \geq N_n,\quad \forall\, n \geq m\right] \geq 1 - \sum_{n=m}^{\infty} \frac{1}{2^n\times n} \geq 1 - \frac{1}{m}\ .$$

D'où $\forall\, \mu \in P_1$, $\quad \mu^{\otimes\infty}\left[d(S_\ell,\phi(\mu)) < \varepsilon,\ \forall\, \ell \geq N\right] \xrightarrow[N\to\infty]{} 1$;

ce qui termine cette démonstration.

Remarques :

1) Si $K(x,y) = \sum_{i=1}^{\infty} \frac{1}{2^i} f_i(x)f_i(y)$, où $(||f_i||_U \leq 1,\ \forall\, i \in N^*)$;

un simple calcul que la suite $N_n = 4n \times 2^n \times E(\frac{1}{\alpha_n} + 1)$ convient.

2) Lorsque dans la prop. (1.F.II) ci-dessus, $P_o = \langle(\mu_n, n \in N^*)\rangle$, - i.e. : P_o est l'enveloppe convexe de la suite $(\mu_n, n \in N^*)$ -, on peut construire l'estimateur S_n à partir de projections sur des sous-espaces convexes de dimension finie. (Projections que l'on peut alors facilement déterminer).

En effet, si l'on pose $\widehat{\phi_n} = \phi_n \circ \mathrm{pr}^K_{<\mu_1,\ldots,\mu_n>}$, $\widehat{\phi_n}$ est uniformément continue sur $\mathcal{P}_\Omega$, et l'on a : $\widehat{\phi_n}(\mu) \xrightarrow[n\to\infty]{} \phi(\mu)$, pour tout μ de $\mathcal{P}_o$. Ce qui nous permet de conclure en vertu de la démonstration ci-dessus.

Cette deuxième partie conduit à de nombreuses applications ; notamment à l'estimation convergente et sans biais d'un paramètre ensembliste (en plongeant naturellement l'espace des sous-ensembles convexes compacts muni de la métrique de Hausdorff dans un espace de Banach - plongement qui est dû à C. Debreux [18]), à l'estimation d'un paramètre <u>fonctionnel ensembliste</u> en employant la même méthode, au modèle de contamination de Tuckey et, enfin, à l'estimation de quelques paramètres usuels (densité, $\int_\Omega K(x,y)d\mu(x),\ldots$).

Ces applications sont données dans [34], pages 187-214 ; nous y exhibons des estimateurs sans biais, optimaux et convergents. (En particulier, pour l'estimation sans biais d'un paramètre ensembliste, nous prenons l'intégrale de Bochner au sens usuel, qui est beaucoup plus maniable que l'intégrale introduite dans [16] et [54]).

De nombreuses études sur les tests asymptotiques ont déjà été effectuées ; dans [3], [38], [36], [8] (par exemple) les auteurs exhibent à l'aide du rapport de vraisemblance des tests asymptotiquement optimaux ; tests qui s'appliquent donc surtout aux modèles exponentiels (et ne fonctionnent pas en particulier pour tester $\{\mu\}$ contre son complémentaire dans $\mathcal{P}_{\Omega}$).

Dans [44] l'auteur établit une condition nécessaire et suffisante très fine d'existence de tests convergents ; malheureusement cette condition porte sur la séparation asymptotique uniforme des deux familles de lois, et de ce fait ne permet pas de construire très explicitement les régions critiques.

Les tests que nous exhibons ont l'avantage d'être de région critique très simple, d'avoir de bonnes "vitesses" de convergence et de s'appliquer à des situations relativement générales ; de plus on peut les étendre aux cas d'alternatives multiples.

APPENDICE

TESTS A NOYAUX

(A - III) TESTS A DEUX ALTERNATIVES.

Définition (1.A.III.).-

. Soient $(\Omega,\mathcal{B})$ un espace mesurable, $\mathcal{P}_o$ et $\mathcal{P}_1$ deux sous-espaces disjoints de $\mathcal{P}_\Omega$ et n un entier strictement positif.

. On appelle test d'ordre n de $\mathcal{P}_o$ contre $\mathcal{P}_1$ toute application mesurable ϕ_n de $(\Omega^n, \mathcal{B}^{\otimes n})$ dans $\{0,1\}$ muni de sa tribu discrète.

. $\phi_n^{-1}(\{1\})$ est la région critique de ϕ_n.

. $\alpha_{\phi_n} = \sup_{\mu\in\mathcal{P}_o} (\int \phi_n \, d\mu^{\otimes n})$ est le niveau du test ϕ_n.

. L'application $\mathcal{B}_{\phi_n} : \mathcal{P}_1 \longrightarrow \mathbb{R}$, $\mu \longrightarrow \int \phi_n \, d\mu^{\otimes n}$ est la puissance du test ϕ_n.

Proposition (1.A.III.).- Soient $K = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$ un noyau reproduisant normal sur $(\Omega,\mathcal{B})$, d_K la pseudo-métrique associée à $\| \ \|_K$, $\mathcal{P}_o$ et $\mathcal{P}_1$ deux sous-espaces disjoints de $\mathcal{P}_\Omega$ et b un réel strictement plus grand que 1.

Si $d_K(\mathcal{P}_o,\mathcal{P}_1) \geq a$, le test d'ordre n de région critique :

$W_n = \bigcup_{\mu\in\mathcal{P}_1} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{b}\}$ est de niveau inférieur à $2n_o (\frac{a(b-1)}{b}) \exp(-\frac{na^2(b-1)^2}{4\sum_{i=1}^{\infty}\alpha_i b^2})$ et de puissance supérieure à

$$1 - 2n_o(\frac{a}{b}) \exp(-\frac{na^2}{4\sum_{i=1}^{\infty}\alpha_i b^2}).$$

($n_o(\frac{a}{b})$ étant le plus petit entier tel que $\sum_{i=n_o(\frac{a}{b})+1}^{\infty} \alpha_i < \frac{a^2}{2b^2}$ si les f_i sont positives ou tel que $\sum_{i=n_o(\frac{a}{b})+1}^{\infty} \alpha_i < \frac{a^2}{8b^2}$ sinon).

Preuve : Si (μ_j) est une suite dense de $(P_1, \|\ \|_K)$, on a : $W_n = \bigcup_{j\in N^*} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_j\|_K < \frac{a}{b}\}$; W_n est donc $\mathcal{B}$-mesurable.

Pour tout $\mu \in P_1$, on a :

$$\mu^{\otimes n}(W_n) \geq \mu(\{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{b}\}) \quad ;$$

ce qui en vertu de l'assertion (b) du "théorème de Glivenko-Cantelli" entraîne $\mu^{\otimes n}(W_n) \geq 1 - 2\, n_o(\frac{a}{b}) \exp(-\frac{n\, a^2}{4\sum_{i=1}^{\infty} \alpha_i\, b^2})$.

Maintenant si $\mu \in P_o$, on a :

$$\mu^{\otimes n}(W_n^C) \geq \mu^{\otimes n}(\{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a(b-1)}{b}\}) \quad ;$$

En effet si $\|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a(b-1)}{b}$ et si $(\omega_1,\ldots,\omega_n) \in W_n$, il existe $\nu \in P_1$ tel que $\|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \nu\|_K < \frac{a}{b}$ et l'on a alors $\|\mu-\nu\|_K < a$ qui est contraire à $d_K(P_o,P_1) \geq a$.

D'où $$\sup_{\mu\in P_o} \mu^{\otimes n}(W_n) \leq 2\, n_o(\frac{a(b-1)}{b}) \exp(-\frac{n\, a^2(b-1)^2}{4\sum_{i=1}^{\infty} \alpha_i\, b^2})$$

(assertion (b) du "théorème de Glivenko-Cantelli").

Remarque :

. On peut énoncer une proposition analogue à la précédente en remplaçant W_n par $W_n' = \bigcap_{\nu\in P_o} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{j=1}^{n} \delta\omega_i - \nu\|_K \geq \frac{a(b-1)}{b}\}$.

D'où suivant la forme de P_o et de P_1 on choisira le test d'ordre n de région critique W_n ou W'_n de façon à avoir une région critique d'expression la plus simple possible.

. Si de plus (dans la proposition ci-dessus) K est non dégénéré P_1 est convexe et $\overline{P}_1$ est complet, on peut remplacer la région critique W_n par $\{(\omega_1,\dots,\omega_n) : ||\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i - pr^K_{\overline{P}_1}(\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i)||_K < \frac{a}{b}\}$. Si les f_i sont ≥ 0, on peut remplacer dans les expressions ci-dessus "$4\sum_{i=1}^{\infty}\alpha_i$" par "$\sum_{i=1}^{\infty}\alpha_i$".

. Si $K = \sum_{i=1}^{p}\alpha_i f_i \otimes f_i$ on peut remplacer

$$2\, n_o(\frac{a(b-1)}{b}) \exp(-\frac{n\, a^2(b-1)^2}{4\sum\alpha_i\, b^2}) \text{ (resp : } 1-2n_o(\frac{a}{b}) \exp(-\frac{n\, a^2}{4\sum\alpha_i\, b^2}) \text{) par}$$

$$2p \exp(-\frac{2n\, a^2(b-1)^2}{\sum_{i=1}^{p}\alpha_i\, b^2}) \text{ (resp : } 1 - 2p \exp(-\frac{2\, n^2}{\sum_{i=1}^{p}\alpha_i\, b^2}) \text{) (assertion (d)}$$

du "théorème de Glivenko-Cantelli").

. Si de plus $K = 1_A \otimes 1_A$, on peut remplacer

$$2\, n_o(\frac{a(b-1)}{b}) \exp(-\frac{n\, a^2(b-1)^2}{4\sum\alpha_i\, b^2}) \text{ (resp : } 1-2n_o(\frac{a}{b}) \exp(-\frac{n\, a^2}{4\sum\alpha_i\, b^2}))$$

par $2 \exp(-\frac{2n\, a^2(b-1)^n}{b^2})$ (resp : $1 - 2 \exp(-\frac{2n\, a^2}{b^2})$) .

. Le théorème de séparation asymptotique uniforme est un corollaire de la proposition précédente.

. L'intérêt de ces tests est qu'on ne perd pas d'information et qu'en général la région critique est d'expression très simple.

. Si P_o et P_1 sont simplement disjoints le test d'ordre n de région critique

$$W_n = \bigcap_{\mu \in P_o} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K \geq \frac{a(b-1)}{b}\}$$

ou de région d'acceptation.

$$W_n = \bigcup_{\mu \in P_o} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a(b-1)}{b}\}$$

est de niveau inférieur à $2n_o \left(\frac{a(b-1)}{b}\right) \exp\left(- \frac{n\, a^2(b-1)^2}{4 \sum \alpha_i\, b^2}\right)$.

Dans le cas de test d'hypothèses simples le test de Neymann-Pearson est parfois (lorsque les densités sont très compliquées) inextricable et l'on ne peut donner une minoration de la puissance. Dans ce cas le test donné par la prop. (1.A.III.) est d'expression beaucoup plus simple et l'on a immédiatement une minoration de la puissance, comme le montre l'exemple suivant.

Exemple : On désigne par μ_o la loi normale centrée réduite, δ_o la mesure de Dirac en zéro, U_o la loi uniforme sur $\left[-\frac{1}{2}, \frac{1}{2}\right]$, μ_1 la loi normale réduite d'espérance 2, U_1 la loi uniforme sur $[-2, 1] \cup [1, 2]$, ν_1 la loi de Poisson d'espérance 3, Q_1 la loi de Laplace, h_1 la loi Γ_1 ; et l'on se propose de tester $P_o = \{\frac{1}{3}\mu_o + \frac{1}{5}\delta_o + \frac{7}{15}U_o\}$ contre $P_1 = \{\frac{2}{5}\mu_1 + \frac{1}{5}U_1 + \frac{1}{5}\nu_1 + \frac{1}{10}Q_1 + \frac{1}{10}h_1\}$.

Si l'on pose $A =]-2, -1[\cup]\frac{1}{2}, +\infty[$ et $K = 1_A \otimes 1_A$, on a :

$$d_K(P_o,P_1) = |\frac{2}{5}\mu_1(A) + \frac{1}{5}U_1(A) + \frac{1}{5}\nu_1(A) + \frac{1}{10}Q_1(A) + \frac{1}{5}h_1(A) - \frac{1}{3}\mu_o(A) + \frac{1}{5}S_o(A) + \frac{7}{15}U(A)| \geq 0{,}72 .$$

D'où si dans la prop. (1.A.III.) on prend $b = 2$, le test d'ordre n de région critique :

$$W_n = \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - (\frac{2}{5}\mu_1 + \frac{1}{5}U_1 + \frac{1}{5}\nu_1 + \frac{1}{10}Q_1 + \frac{1}{10}h_1)\|_K < 0{,}36\}$$

$$= \{(\omega_1,\dots,\omega_n) : -\frac{n}{2} < \sum_{i=1}^{n} \delta\omega_i(A) < n \times (1{,}22)\}$$

est pour :

$n = 10$, de niveau inférieur à $2e^{-2,6}$

et de puissance supérieure à $1 - 2e^{-2,6}$

$n = 20$, de niveau inférieur à $2e^{-5,18}$

et de puissance supérieure à $1 - 2e^{-5,18}$

$n = 30$, de niveau inférieur à $2e^{-7,7}$

et de puissance supérieure à $1 - 2e^{-7,7}$

$n = 50$, de niveau inférieur à $2e^{-13}$

et de puissance supérieure à $1 - 2e^{-13}$

$n = 100$, de niveau inférieur à $2e^{-26}$

et de puissance supérieure à $1 - 2e^{-26}$

Proposition (2.A.III.).- Soient $K = \sum \alpha_i f_i \otimes f_i$ un noyau reproduisant normal sur $(\Omega,\mathcal{B})$, $\mathcal{P}_o$ et $\mathcal{P}_1$ deux sous-espaces de $\mathcal{P}_\Omega$.

1) Si b,c,d sont trois réels positifs tels que $\frac{1}{b} + \frac{1}{c} + \frac{1}{d} \leq 1$ si $d_K(\mathcal{P}_o,\mathcal{P}_1) \geq a > 0$, si $\bigcup_{i=1}^{\infty} \{\mu_i\} \subset \mathcal{P}_1$ et si $\mathcal{P}_1 \subset \bigcup_{j=1}^{\infty} B_{d_K}(\mu_j, \frac{a}{b})$, le test d'ordre n de région critique :

$$W_n = \bigcup_{j=1}^{\infty} (\{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_j\|_K < \frac{a}{c} + \frac{a}{b}\},$$

est de puissance supérieure à $1 - 2n_o(\frac{a}{c}) \exp(-\frac{n\,a^2}{4\sum \alpha_i\, c^2})$ et de niveau inférieur à $2n_o(\frac{a}{d}) \exp(-\frac{n\,a^2}{4\sum \alpha_i\, d^2})$.

2) Si b, c, d sont trois réels tels que $\frac{2}{b} + \frac{1}{c} + \frac{1}{d} \leq 1$ et si dans (1) on remplace "$\bigcup_{j=1}^{\infty} \{\mu_j\} \subset P_1$" par $\forall j \in N^* \ldots$ $B_{d_K}(\mu_j, \frac{a}{b}) \cap P_1 \neq \phi$, le test d'ordre n de région critique W_n est de puissance supérieure à $1 - 2\, n_o(\frac{a}{c}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, c^2})$ et de niveau inférieur à $2\, n_o\, (\frac{a}{d}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, d^2})$.

Preuve :

1) Si $\mu \in P_1$ on a :

$$\mu^{\otimes n}(W_n) \geq \mu(\{(\omega_1,\ldots,\omega_n) : ||\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i - \mu||_K < \frac{a}{c})$$

En effet si $|| \frac{1}{n} \sum_{i=1}^{n} \delta\omega_i - \mu||_K < \frac{a}{c}$, comme il existe $j \in N^*$ tel que $||\mu_j - \mu||_K < \frac{a}{b}$, on a $(\omega_1,\ldots,\omega_n) \in W_n$.

L'assertion (b) du théorème de Glivenko-Cantelli nous permet alors d'affirmer que $\mu^{\otimes n}(W_n) \geq 1 - 2n_o(\frac{a}{c}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, c^2})$.

Maintenant si $\mu \in P_o$, on a :

$$\mu^{\otimes n}(W_n^C) \geq \mu^{\otimes n}(\{(\omega_1,\ldots,\omega_n) : ||\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i - \mu||_K < \frac{a}{d}\}) .$$

En effet si $||\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i - \mu||_K < \frac{a}{d}$ et si $(\omega_1,\ldots,\omega_n) \in W_n$; il existe $j \in N^*$ tel que $||\frac{1}{n} \sum_{i=1}^{n} \delta\omega_i - \mu_j||_K < \frac{a}{c} + \frac{a}{b}$ et l'on a $||\mu - \mu_j||_K < \frac{a}{c} + \frac{a}{b} + \frac{a}{d}$ qui est contraire à $d_K(P_o,P_1) \geq a$.

D'où $\sup_{\mu\in P_o} \mu^{\otimes n}(W_n) \leq 2n_o(\frac{a}{d}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, d^2})$ (assertion (b) du "théorème de Glivenko-Cantelli").

2) La première partie découle de (1).

Maintenant si $\mu \in P_o$, on a :

$$\mu^{\otimes n}(W_n^C) \geq \mu^{\otimes n}(\{(\omega_1,\ldots,\omega_n) : ||\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i - \mu||_K < \frac{a}{d}\}) \quad .$$

En effet si $||\frac{1}{n}\sum_{i=1}^{n}\delta\omega_i - \mu||_K < \frac{a}{d}$ et si $(\omega_1,\ldots,\omega_n) \in W_n$, il existe $j \in N^*$ tel que $||\mu-\mu_j||_K < \frac{a}{c} + \frac{a}{b} + \frac{a}{d}$ et il existe $\nu \in P_1$ tel que $||\mu_j-\nu||_K < \frac{a}{b}$; on a alors $||\mu-\nu||_K < \frac{2a}{b} + \frac{a}{c} + \frac{a}{d} \leq a$ qui est contraire à $d_K(P_o,P_1) \geq a$.

D'où
$$\sup_{\mu\in P_o} \mu^{\otimes n}(W_n) \leq 2n_o(\frac{a}{d}) \exp(-\frac{n\,a^2}{4\sum \alpha_i\, d^2}) \quad .$$

Remarque : On peut bien sur remplacer $j \in N^*$ par $j \in \{1,\ldots,k\}$.

Remarque : Comme à la suite de la prop. (1.A.III.), suivant la forme de K, on peut donner une meilleure majoration (resp : minoration) du niveau (resp : de la puissance) du test d'ordre n de région critique W_n. En particulier si $K = 1_A \otimes 1_A$, on peut en (1) et (2) remplacer $2n_o(\frac{a}{c}) \exp(-\frac{n\,a^2}{4\sum \alpha_i\, c_i^2})$ par $2\exp(-\frac{2n\,a^2}{c^2})$ et $2n_o(\frac{a}{c}) \exp(-\frac{n\,a^2}{4\sum \alpha_i\, d^2})$ par $2\exp(-\frac{2n\,a^2}{d^2})$.

Exemple : On considère μ_o la loi normale réduite d'espérance 2, U_o la loi uniforme sur $[-2,-1{,}8] \cup [1{,}2{,}8]$, ν_o la loi de Poisson d'espérance 2, μ_1 la loi normale centrée réduite, $H_1 = \{\delta_{-i},\ i \in N\}$, $F_1 = \{\delta_i,\ i \geq 5\}$, et l'on se propose de tester P_o enveloppe convexe de $\{\mu_o\} \cup \{U_o\} \cup \{\nu_o\}$ contre P_1 enveloppe convexe de $H_1 \cup F_1 \cup \{\mu_1\}$.

Si l'on pose $A =]1, +\infty[- \{i \in N^*,\ i \geq 5\}$ et $K = 1_A \otimes 1_A$, on a $d_K(P_o,P_1) \geq 0{,}64$ et $P_1 \subset B_{d_K}(\mu_1, 0{,}16)$.

D'où si (dans la prop. (2.A.III.)) l'on prend $a = 0{,}64$ $b = 4$ (puisque $\frac{a}{b} = 0{,}16$) , $c = 3$ et $d = 2{,}5$, le test d'ordre n de région critique :

$$W_n = \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_1\|_K < \frac{0,64}{3} + \frac{0,64}{4}\}$$

$$= \{(\omega_1,\dots,\omega_n) : -0,21 \times n \leqslant \sum_{i=1}^{n} \delta\omega_i(A) \leqslant 0,53\ n\}$$

est de niveau inférieur à $2\exp(-\frac{2n\ (0,41)}{6,3})$

et de puissance supérieure à $1 - 2\exp(-\frac{2n\ (0,41)}{9})$

pour :

$n = 20$ il est donc de niveau inférieur à $\frac{14}{100}$

et de puissance supérieure à $1 - \frac{32}{100}$

$n = 50$ il est de niveau inférieur à $\frac{45}{10000}$

et de puissance supérieure à $1 - \frac{2}{100}$

$n = 100$ il est de niveau inférieur à 12×10^{-6}

et de puissance supérieure à $1 - 24 \times 10^{-5}$

(B - III) TESTS CONVERGENTS.

Définition (1.B.III.).- Soient $(\Omega,\mathcal{B})$ un espace mesurable, $\mathcal{P}_o$ et $\mathcal{P}_1$ deux sous-espaces disjoints de $\mathcal{P}_\Omega$. Une suite de test (ϕ_n) (où ϕ_n est pour tout n d'ordre n) de $\mathcal{P}_o$ contre $\mathcal{P}_1$ est dite convergente si $\alpha_{\phi_n} \xrightarrow[n\to\infty]{} 0$ et si pour tout $\mu \in \mathcal{P}_1$, $\int \phi_n \, d\mu^{\otimes n} \xrightarrow[n\to\infty]{} 1$.

Proposition (1.B.III.).- Soient $(\Omega,\mathcal{B})$ un espace mesurable et $K = \sum \alpha_i f_i \otimes f_i$ un noyau reproduisant normal sur Ω. Si $\mathcal{P}_1$ est disjoint de l'adhérence de $\mathcal{P}_o$ dans $(\mathcal{P}_\Omega, \|\ \|_K)$, il existe une suite convergente de tests de $\mathcal{P}_o$ contre $\mathcal{P}_1$.

Preuve : $(\mathcal{P}_o, \|\ \|_K)$ étant séparable, il existe une suite (μ_j) dense dans $\mathcal{P}_o$. (a_k) et (b_k) étant deux suites strictement positives convergent vers 0, si l'on pose :

$$W_{p(k)} = \bigcap_{j\in\mathbb{N}^*} \{(\omega_1,\dots,\omega_{p(k)}) : \Big\|\frac{1}{p(k)} \sum_{i=1}^{p(k)} \delta\omega_i - \mu_j\Big\|_K \geq a_k\} .$$

(où $p(k)$ est une suite strictement croissante telle que $2n_o(a_k)\exp(-\frac{p(k)\ a_k}{4\sum \alpha_i}) \leq b_k$). On a :

$$\inf_{\mu\in\mathcal{P}_o} \mu^{\otimes p(k)}(W^C_{p(k)}) = \inf_{\mu\in\mathcal{P}_o} \mu^{\otimes p(k)}\Big(\bigcup_{\nu\in\mathcal{P}_o} \{(\omega_1,\dots,\omega_{p(k)}) : \Big\|\frac{1}{p(k)} \sum_{i=1}^{p(k)} \delta\omega_i - \nu\Big\|_K < a_k\}\Big)$$

$$\geq \inf_{\mu\in\mathcal{P}_o} \mu^{\otimes p(k)}\Big((\omega_1,\dots,\omega_{p(k)}) : \Big\|\frac{1}{p(k)} \sum_{i=1}^{p(k)} \delta\omega_i - \mu\Big\|_K < a_k\Big) \geq 1-b_k$$

d'après l'assertion (b) du théorème de "Glivenko-Cantelli".

D'où si pour $n = p(k)$ on désigne par ϕ_n le test d'ordre n de $\mathcal{P}_o$ contre $\mathcal{P}_1$ ayant pour région critique $W_{p(k)}$ et si pour $p(k) < n < p(k+1)$ on désigne par ϕ_n le test d'ordre n de $\mathcal{P}_o$ contre $\mathcal{P}_1$ ayant $W_{p(k)} \times \Omega^{n-p(k)}$ pour région critique , on a $\alpha_{\phi_n} \xrightarrow[n\to\infty]{} 0$.

Maintenant si $\nu \in \mathcal{P}_1$, comme $d_K(\mathcal{P}_o,\nu) > 0$, il existe k_o tel que $k \geq k_o$ entraîne $d_K(\mathcal{P}_o,\nu) > 2a_k$. Pour $k \geq k_o$ on a alors

$$\left\|\frac{1}{p(k)}\sum_{i=1}^{p(k)} \delta\omega_i - \nu\right\|_K < a_k \quad \text{implique} \quad \forall\mu \in \mathcal{P}_o \;,\; \left\|\mu - \frac{1}{p(k)}\sum_{i=1}^{p(k)} \delta\omega_i\right\|_K \geq a_k$$

et $(\omega_1,\ldots,\omega_{p(k)}) \in W_{p(k)}$. Pour $k \geq k_o$ on a donc :

$$\nu^{\otimes p(k)}(W_{p(k)}) \geq \nu^{\otimes p(k)}(\{(\omega_1,\ldots,\omega_n) : \left\|\frac{1}{p(k)}\sum_{i=1}^{p(k)} \delta\omega_i - \nu\right\|_K < a_k\}) \geq 1-b_k \;.$$

D'où $$\int \phi_n \, d\nu^{\otimes n} \xrightarrow[n\to\infty]{} 1 \;. \blacksquare$$

Remarque : Nous avons utilisé dans cette démonstration l'assertion (b) du "théorème de Glivenko-Cantelli" qui est vrai pour tous les noyaux ; mais si K est un noyau particulier par exemple de la forme $K = 1_A \otimes 1_A$, on choisit (en vertu de l'assertion (d) du "théorème de Glivenko-Cantelli") la suite $p(k)$ telle que $2e^{-2p(k)a_k^2} \leq b_k$.

On considère $\mathbb{R}$ muni de sa trivu borélienne usuelle $\mathcal{B}_{\mathbb{R}}$, μ une mesure sur $(\mathbb{R},\mathcal{B}_{\mathbb{R}})$ et l'on se propose de tester $\mathcal{P}_o = \{\mu\}$ contre $\mathcal{P}_1 = \{\mu\}^C$.

Si l'on pose $K(x,y) = \exp\left(\frac{x}{1+|x|}\right)\left(\frac{y}{1+|y|}\right) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{x}{1+|x|}\right)^n\left(\frac{y}{1+|y|}\right)^n$,

K est (théorème de construction (I.E.I.)) un noyau reproduisant normal et non dégénéré ; $\mathcal{P}_o$ étant fermé dans $(\mathcal{P}_{\mathbb{R}}, \|\;\|_K)$ on peut appliquer la prop. (I.B.II.) en prenant $a_k = \frac{3,3}{k^{1/4}}$ et $b_k = \frac{1}{k}$; pour $k \geq 10$ on peut choisir $n_o(a_k) \leq k^{1/4}$ et $p(k) = k$; la suite de tests de région critique (d'ordre n)

$$W_n = \{(\omega_1,\ldots,\omega_n) : \left\|\mu - \frac{1}{n}\sum_{i=1}^{n} \delta\omega_i\right\|_K \geq \frac{3,3}{n^{1/4}}\} \;.$$

$$= \{(\omega_1,\ldots,\omega_n) : \left|\int K\, d\mu \otimes \mu + \sum_{i,j=1}^{n} K(\omega_i,\omega_j) - 2\,\frac{1}{n}\sum_{i=1}^{n}\int K(\omega_i,\omega)\, d\mu(\omega)\right| \geq \frac{3,3}{n^{1/4}}\}$$

est alors convergente de $\mathcal{P}_o$ contre $\mathcal{P}_1$; plus précisément pour $n \geq 10$ le niveau du test de région critique W_n est inférieur à $\frac{1}{n}$; maintenant si $\nu \in \mathcal{P}_1$, il existe h_o tel que $d_K(\nu,\mu) \geq \frac{6,6}{h_o^{1/4}}$ et pour $n \geq h_o$.

La puissance au point ν du test d'ordre n de région critique W_n est supérieur à $1 - \frac{1}{n}$.

Remarque : Si l'on s'intéresse à de grands échantillons (de taille supérieure à 150 par exemple), on peut dans l'exemple ci-dessus prendre $b(k) = k^2$.

Le niveau du test est alors (pour $n \geq 150$) inférieur à $\frac{1}{n^2}$ et pour $n \geq h_o$, la puissance au point ν du test d'ordre n de région critique W_n est supérieure à $1 - \frac{1}{n^2}$.

(C - III) TESTS A ALTERNATIVES MULTIPLES.

Définition.-

. Soient $(\Omega,\mathcal{B})$ un espace mesurable $\mathcal{P}_1,\ldots,\mathcal{P}_p$ p sous-espaces disjoints de $\mathcal{P}_\Omega$ et n un entier strictement positif.

On appelle test d'ordre n entre $\mathcal{P}_1,\ldots,\mathcal{P}_p$ toute partition ordonnée $\mathcal{B}^{\otimes n}$-mesurable $A_1,\ldots,A_p$ de Ω^n.

. Si l'on note ψ_n le test d'ordre n entre $\mathcal{P}_1,\ldots,\mathcal{P}_p$ associé à la partition ordonnée $A_1,\ldots,A_p$ et si $0 < k \leqslant p$, le k-uple $\alpha^1_{\psi_n} = \sup\limits_{\mu\in\mathcal{P}_1} \mu^{\otimes n}(A_1^C)$,..., $\alpha^k_{\psi_n} = \sup\limits_{\mu\in\mathcal{P}_k} \mu^{\otimes n}(A_k^C)$ est appelé k-niveau du test ψ_n.

Si $0 < \ell \leqslant p$, la ℓ-puissance du test ψ_n est l'application qui à μ appartenant à $\mathcal{P}_i$, $p - \ell+1 < i \leqslant p$ fait correspondre $\mu^{\otimes n}(A_i)$.

. Nous dirons que le k-niveau de ψ_n est inférieur à $(\mathcal{B}_1,\ldots,\mathcal{B}_k)$ si $\forall i \in \{1,\ldots,k\}$, $\alpha^i_{\psi_n} \leqslant \mathcal{B}_i$; de même nous dirons que la ℓ-puissance de ψ_n est supérieure à $(\mathcal{B}_{p-\ell+1},\ldots,\mathcal{B}_p)$ si $\forall i \in \{p-\ell+1,\ldots,p\}$, $\forall \mu \in \mathcal{P}_i$, on a $\mu^{\otimes n}(A_i) \geqslant \mathcal{B}_i$.

Remarque : A toute partition (à p éléments $\mathcal{B}^{\otimes n}$-mesurable) de Ω^n correspond p! partitions ordonnées donc p! tests d'ordre n entre $\mathcal{P}_1,\ldots,\mathcal{P}_p$.

Proposition (1.C.III.).- Soit K un noyau reproduisant normal, d_K la pseudo-métrique associée à $\|\ \|_K$, $\mathcal{P}_1,\ldots,\mathcal{P}_p$ p sous-espaces disjoints de $\mathcal{P}_\Omega$ et $b_1,\ldots,b_k$ k-réels positifs tels que pour $i \neq j$, $i, j \in \{1,\ldots,k\}$, $\frac{1}{b_i} + \frac{1}{b_j} \leqslant 1$ $(1<k<p)$.

1) Si pour $i \neq j$, $i , j \in \{1,\dots,k\}$, $d_K(P_i,P_j) \geq a$, si l'on pose :

$$A_1 = \bigcup_{\mu\in P_1} \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{b_1}\}$$

$\vdots$

$$A_k = \bigcup_{\mu\in P_k} \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{b_k}\}$$

et si $A_{k+1},\dots,A_p$ est une partition $\mathcal{B}^{\otimes n}$-mesurable de $\Omega^n - \bigcup_{i=1}^{k} A_i$, le test associé d'ordre n entre $P_1,\dots,P_p$ associée à la partition ordonnée $(A_1,\dots,A_p)$ est de k-niveau inférieur à

$$2n_o(\frac{a}{b_1}) \exp(-\frac{n a^2}{4\sum \alpha_i b_1^2}) ,\dots, 2n_o(\frac{a}{b_k}) \exp(-\frac{n a^2}{4\sum \alpha_i b_k^2}) .$$

2) Si $C_1,\dots,C_p$ sont p réels positifs tels que pour $i \neq j$, $\frac{1}{C_i} + \frac{1}{C_j} \leq 1$ et si pour $i \neq j$, $d_K(P_i,P_j) \geq a$, le test d'ordre n entre $P_1,\dots,P_p$ associé à la partition ordonnée :

$$A_1 = \bigcup_{\mu\in P_1} \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{C_1}\}$$

$\vdots$

$$A_{p-1} = \bigcup_{\mu\in P_{p-1}} \{(\omega_1,\dots,\omega_n) : \|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu\|_K < \frac{a}{C_{p-1}}\}$$

$$A_p = \Omega^n - \bigcup_{j=1}^{p-1} A_j$$

est de p-niveau inférieur à

$$2n_o(\frac{a}{C_1}) \exp(-\frac{n a^2}{4\sum \alpha_i C_1^2}) ,\dots, 2n_o(\frac{a}{C_k}) \exp(-\frac{n a^2}{4\sum \alpha_i C_k^2})$$

et de p-puissance supérieure à

$$1 - 2n_o(\frac{a}{C_1}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, c_1^2}) , \ldots , 1-2n_o(\frac{a}{C_k}) \exp(- \frac{n\, a^2}{4 \sum \alpha_i\, c_k^2})$$

Preuve :

1) . Soit $j \in \{1,\ldots,k\}$; $(P_j, \|\ \|_K)$ étant séparable, il existe une suite (μ_m) dense dans P_j et l'on a

$$A_j = \bigcup_{m \in \mathbb{N}^*} \{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n} \sum_{i=1}^n \delta\omega_i - \mu_m\|_K < \frac{a}{b_j}\} .$$

Pour tout $j \in \{1,\ldots,p\}$, A_j est donc $\mathcal{B}^{\otimes n}$-mesurable.

. Soient $p \neq q$, $p, q \in \{1,\ldots,k\}$ et supposons qu'il existe $(\omega_1,\ldots,\omega_n)$ appartenant à $A_p \cap A_q$; on peut alors exhiber μ dans A_p et ν dans A_q tels que : $\|\frac{1}{n} \sum_{i=1}^n \delta\omega_i - \mu\|_K < \frac{a}{b_p}$ et $\|\frac{1}{n} \sum_{i=1}^n \delta\omega_i - \mu\|_K < \frac{a}{b_q}$, qui est impossible puisque $d_K(P_p,P_q) \geq a$ et $\frac{a}{b_q} + \frac{a}{b_p} \leq a$.

Maintenant pour tout $j \in 1,\ldots,k$ et tout $\nu \in P_j$ on a :

$$\nu^{\otimes n}(A_j^C) \leq \nu^{\otimes n}(\{(\omega_1,\ldots,\omega_n) : \|\frac{1}{n} \sum_{i=1}^n \delta\omega_i - \nu\|_K \geq \frac{a}{b_j}\}) ;$$

ce qui nous permet de conclure en vertu de l'assertion (b) du théorème de Glivenko-Cantelli.

2) On montre comme en (1) que $A_1,\ldots,A_p$ est une partition $\mathcal{B}^{\otimes n}$-mesurable et l'on conclut de la même façon.

Comme à la suite des propositions précédentes, suivant la forme du noyau K, on peut affiner les minorations et majorations des k-puissances et des k-niveaux.

En particulier si $K = 1_A \otimes 1_A$, on peut dans la prop. (I.C.III.) remplacer $2n_o(\frac{a}{b_j}) \exp(- \frac{4\, a^2}{4 \sum \alpha_i\, b_j^2})$ (resp : $2n_o(\frac{a}{c_j}) \exp(- \frac{4\, a^2}{4 \sum \alpha_i\, c_j^2})$) par $2\exp(- \frac{2n\, a^2}{b_j^2})$ (resp : $2\exp(- \frac{2n\, a^2}{c_j^2})$).

Exemple :

1) On considère μ_1 la loi normale centrée réduite, μ_2 la loi normale réduite d'espérance 2, μ_3 la loi uniforme sur $\left]-\frac{1}{2}, \frac{1}{2}\right[$, μ_4 la loi de Poisson d'espérance 3, et si l'on se propose d'effectuer un test entre $\{\mu_1\}$, $\{\mu_1\}$, $\{\mu_3\}$, $\{\mu_4\}$.

On pose $A = \left]\frac{1}{2}, +\infty\right[- \{n \in N \; ; \; n \geq 5\}$, $K = 1_A \otimes 1_A$ et $C_1 = C_2 = C_3 = C_4 = 2$.

Comme $d_K(\mu_i,\mu_j) \geq 0{,}30$ la prop. (1.C.III.) nous permet d'affirmer que le test d'ordre n entre $\{\mu_1\},\ldots, \{\mu_4\}$ associé à la partition ordonnée :

$$A_1 = \{(\omega_1,\ldots,\omega_n) \; : \; ||\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_1||_K < 0{,}15\}$$

$$= \{(\omega_1,\ldots,\omega_n) \; : \; n(0{,}94 - 0{,}15) \leq \sum_{i=1}^{n} \delta\omega_i(A) \leq n(0{,}94 + 0{,}15)\}$$

$$A_2 = \{(\omega_1,\ldots,\omega_n) \; : \; ||\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_2||_K < 0{,}15\}$$

$$A_3 = \{(\omega_1,\ldots,\omega_n) \; : \; ||\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_3||_K < 0{,}15\} .$$

$$A_4 = \Omega^n - (A_1 \cup A_2 \cup A_3)$$

est de 4 - niveau inférieur à $(2\exp(-\frac{n \times 45}{1000}), 2\exp(-\frac{n \times 45}{1000}), \ldots, 2\exp(-\frac{n \times 45}{1000})$

et de 4 - puissance supérieure à

$(1 - 2\exp(-\frac{n \times 45}{1000}), \ldots, 1 - 2\exp(-\frac{n \times 45}{1000})$

pour $n = 50$ on a $2\exp(-\frac{n \times 45}{1000}) \leq \frac{21}{100}$,

pour $n = 100$ $2\exp(-\frac{n \times 45}{1000}) \leq \frac{22}{1000}$, pour $n = 150$ $2\exp(-\frac{n \times 45}{1000}) \leq 48 \times 10^{-4}$.

2) On considère μ_1 la loi normale centrée réduite, μ_2 la loi normale réduite d'espérance 2, μ_3 la loi réduite d'espérance - 2 et l'on se propose d'effectuer un test entre $\{\mu_1\}$, $\{\mu_2\}$, $\{\mu_3\}$.

On pose $A =]0, +\infty[$, $K = 1_A \otimes 1_A$ et $C_1 = C_2 = C_3 = 2$.

Comme $d_K(\mu_i,\mu_j) \geq 0{,}5$ la prop. (I.C.III.) nous permet d'affirmer que le test d'ordre n entre $\{\mu_1\}$, $\{\mu_2\}$, $\{\mu_3\}$ associée à la partition ordonnée

$$A_1 = \{(\omega_1,\ldots,\omega_n) : \left\|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_1\right\|_K < 0{,}25\}$$

$$A_2 = \{(\omega_1,\ldots,\omega_n) : \left\|\frac{1}{n}\sum_{i=1}^{n} \delta\omega_i - \mu_2\right\|_K < 0{,}25\}$$

$$A_3 = \Omega^n - (A_1 \cup A_2)$$

est de 3 - niveau inférieur à $\left(2\exp\left(-\frac{n \times 125}{1000}\right), \ldots, 2\exp\left(-\frac{n \times 125}{1000}\right)\right)$

et de 3 - puissance supérieure à $\left(1 - 2\exp\left(-\frac{n \times 125}{1000}\right), \ldots, 2\exp\left(-\frac{n \times 125}{1000}\right)\right)$.

pour $n = 20$, $2\exp\left(-\frac{n \times 125}{1000}\right) \leq 0{,}16$, pour $n = 50$, $2\exp\left(-\frac{n \times 125}{1000}\right) \leq 0{,}005$

pour $n = 100$, $2\exp\left(-\frac{n \quad 125}{1000}\right) \leq 0{,}000012$.

(D - III) TESTS CONVERGENTS A ALTERNATIVES MULTIPLES.

Définition (1.D.III.).- Soient $P_1,\dots,P_k$ k-sous-espaces disjoints de P_Ω. Une suite $(\varphi_n = (A_{1,n},\dots,A_{k,n}))$ de tests entre $P_1,\dots,P_k$ est dite convergente si $\alpha^1_{\varphi_n} \xrightarrow[n\to\infty]{} 0,\dots,\alpha^{k-1}_{\varphi_n} \longrightarrow 0$ et si $\forall\mu \in P_k$, $\mu^{\otimes n}(A_{k,n}) \xrightarrow[n\to\infty]{} 1$.

Proposition (2.C.III.).- Soient $K = \sum \alpha_i\, f_i \otimes f_i$ un noyau reproduisant normal sur Ω et $P_1,\dots,P_k$ k-sous-espaces disjoints de P_Ω. Si pour $i \neq j$, $i, j \in \{1,\dots, k-1\}$ $d_K(P_i,P_j) > 0$ et si l'adhérence de $\bigcup_{i=1}^{k-1} P_i$ est disjointe de l'adhérence de P_k dans $(P_\Omega, \|\ \|_K)$, il existe une suite convergente de tests entre $P_1,\dots,P_k$.

Preuve : Identique à la prop. (1.B.III.).

Remarque : Il suffit de prendre

$$A_{(1,pk)} = \bigcup_{\mu\in P_o} \{(\omega_1,\dots,\omega_{p(k)}) : \|\frac{1}{p(k)} \sum_{i=1}^{p(k)} \delta\omega_i - \mu\|_K < \inf(a_k, \frac{C}{2})\},$$

(où $\forall i, j \in \{1,\dots,k-1\}$, $i \neq j$, $d_K(P_i,P_j) \geq C > 0$),... Cette proposition s'applique en particulier quand on veut effectuer un test entre $\{\mu_1\},\dots, \{\mu_{k-1}\}$ et $P_\Omega - \bigcup_{i=1}^{k-1} \{\mu_i\}$.

CONCLUSION.

Comme nous l'avons vu au cours de cet article, l'introduction des produits scalaires sur l'espace des mesures permet de simplifier notablement les démonstrations et ouvre avec les projections orthogonales et le "théorème de Glivenko-Cantelli" de nouvelles méthodes de construction d'estimateurs et de tests.

De plus, ce travail est susceptible de nombreux développements ultérieurs ; aussi bien dans l'étude de la robustesse qu'en théorie des mesures aléatoires.

BIBLIOGRAPHIE

[1] ARONSZAJN N. - *La théorie des noyaux reproduisants et applications,* dans Proceedings of the Cambridge philosophical Society (vol. 39), 1943 (page 133-153).

[2] ARONSZAJN N. - *Theory of reproducting kernels,* dans transactions of the American mathematical Society, n° 68, 1950, (p. 307-404).

[3] BAHADUR R.R. - *An optimal property of the likelihood ratio. statistic.,* Proc. Fifth Berkeley Symp. Math. Stat. Prob. (1), 13-26 (1966).

[3 bis] BERGER A. - *On uniformly consistent test,* A.M.S. Vol. 22 (1951) p. 289-293.

[4] BERTRAND-RETALI et J. GEFFROY - *Sur la convergence uniforme des estimateurs d'une densité de probabilité,* Publication interne de l'Université de Paris VI.

[5] BICKEL P.J. - LEHMANN E.L. - *Unbiased estimation in convex families,* Ann. Math. Stat. Vol. 40, n° 5, p. 1523-35 (1969).

[6] BLEUEZ J. - *Condition n. et s. de convergence d'une classe d'estimateurs de la densité,* Université des Sciences et Techniques de Lille I, Soutenue le 27 février 1976.

[7] BRETAGNOLLE J. - C. HUBER - *Estimation des densités : risques minimaux,* Publication interne de l'Université de Paris-Nord (Villetaneuse).

[8] BROWN L.D. - *Non-local asymptotic optimality of appropriate likelihood ratio tests,* The annals of Math. Stat. (1971), Vol. 42, n° 4, 1206-1240.

[9] BOCHNER S. - *Intégration Von Funktionen, deven Werte die elemente eines Vektorraümes Sind,* Fond. Math. ; 20 (1933) 262-276.

[10] BOURBAKI N. - *Espace vectoriel topologique,* Chap. II (1ère édition).

[11] BOURBAKI N. - *Espace vectoriel topologique,* Chap. III et IV (1ère édition).

[12] BOSQ D. - Cours de D.E.A. de Probabilités, rédigé par L. NOBECOURT et A. WIBAUX, année 1974-75, UNIVERSITE DE LILLE I.

[13] BOSQ D. - *Sur l'estimation d'un paramètre à valeurs dans un espace de Hilbert,* Publ. int. de l'U.E.R. de Math. Pures et Appliquées de Lille I, n° 70 - janvier 1976.

[14] BOSQ D. - *Contribution à la théorie de l'estimation fonctionnelle,* Thèse - Publications de l'Inst. de Stat. de l'Université de Paris - Octobre 1970.

[15] BOSQ D. - *Sur l'estimation sans biais d'un paramètre fonctionnel,* Publication interne n° 104, de l'Université de Lille I, (1977).

[16] CASTAING - *Sur les multi-applications mesurables,* Thèse, Caen (1967).

[17] CHOQUET G. - Cours d'Analyse (tome II), 2ème édition (1973) - Masson et Cie (Paris).

[18] DEBREUX C. - *Intégration of correspondance,* Fifth Berkeley Symposium.

[19] DELECROIX M. - *Sur l'estimation des densités marginales et de transition d'un processus stationnaire et mélangeant,* Thèse de 3ème cycle (Université des Sciences et Techniques de Lille I - soutenue le 15 janvier 1975).

[20] DEVOLDER J. - *Une inégalité de type exponentiel,* Publication interne de l'U.E.R. de Mathématiques Pures et Appliquées de Lille I, n° 98.

[21] DIEUDONNE J. - Eléments d'Analyse (Tome 1) - 1969 - Gauthier-Villars.

[22] DIEUDONNE J. - Eléments d'Analyse (Tome 2) - 1969 - Gauthier-Villars.

[23] DIXMIER J. - Cours de Math. Premier Cycle (2ème année), Ed. 1968 - Gauthier-Villars.

[24] DUC-JACQUET M. - *Approximation des fonctionnelles linéaires sur les espaces hilbertiens auto-reproduisants,* Thèse, Université Scientifique et Médicale de Grenoble, 1973.

[25] DUNFORD N. - *Integration in general analysis,* Trans. Amer. Math. Soc. 37 (1935) 441-453 ; 38 (1935), 600-601.

[26] DUNFORD N. et SCHWARTZ J. - *Linear operators,* Interscience Pub., New-York (1958).

[27] GEFFROY J. et MOCHÉ R. - *Sur la séparation asymptotique uniforme des produits de loi de probabilité,* C.R. Série A, t. 278, (1974), p. 969-971.

[28] GEFFROY J. et MOCHE R. - *Construction d'estimateurs uniformément convergents à partir d'observations non nécessairement équidistribuées,* C.R. Série A, t. 280, (1975), p. 133-135.

[29] GEFFROY J. - Comptes Rendus, Série A, 1974, p. 1449.

[29 bis] GEFFROY J. - *Inégalités pour le niveau de signification et la puissance de certains tests reposant sur des données quelconques,* C.R.A.S. t. 278 (1976), p. 1299-1301.

[30] GRADSHTEYN I.S. and RYZHIK I.W. - Table on integrals, series and products, Academic Press.

[31] GUILBART C. - *Caractérisation des produits scalaires sur l'espace des mesures bornées à signes dont la topologie trace sur l'espace des mesures de probabilité est la topologie faible et théorème de Glivenko-Cantelli associé,* Publ. Int. de l'U.E.R. de Mathématiques Pures et Appliquées de Lille I, n° 108, mars 1977.

[32] GUILBART C. - *Contribution à l'estimation du support d'une mesure de probabilité,* Thèse de 3ème cycle, Université de Lille I, (1973).

[33] GUILBART C. - *Une topologie sur l'ensemble des mesures de probabilité définies sur une tribu borélienne à base dénombrable et théorème de Glivenko-Cantelli associé,* N.C.R.Acad.Sc. t. 271 (9 juillet 1973), Série A, (129-132).

[34] GUILBART C. - *Etude des produits scalaires sur l'espace des mesures et applications à l'estimation,* Publication interne de l'U.E.R. de Mathématiques Pures et Appliquées de Lille I, n° 139, (1978).

[35] HAJEK et SIDAK - *Theory of Rank tests,* Acad. Press. (1967).

[36] HERR D.G. - *Asymptotically optimal tests for multivariate normal distributions,* Ann. Math. Stat. 38, 1829-1844 (1967).

[37] HILLE et PHILLIPPS - *Fonctional Analysis and semi-groups,* American Math. Society col. pub. 1957.

[38] HOEFFDING W. - *Asymptotically optimal tests for multinomial distributions,* Ann. Math. Stat. 36, pages 369-408.

[39] HOEFFDING W. - *Probability inequalities for sums of bounded random variables,* J.A.S.A. 58 (1963), pages 13-30.

[40] KELLEY J.L. - *General topology,*
The University series in Higher mathematics.

[40 bis] KRAFT C. - *Some conditions for consistancy and uniform consistency of statistical procedures,*
Univ. of California, Pub. in Statistics, Berkeley, vol. 2 (1955), p. 125-142.

[41] KUELBS J. - *Gaussian Measures on Banach space,*
Journal of Fonctional analysis 5, 354-367 (1970).

[41 bis] LECAM L. et SCHWARTZ L. - *A necessary and suffisen condition for the existence of consistent estimates,*
A.M.S. volume 31, 1960, p. 140-150.

[42] LEHMANN E. - *Testing statistical Hypotheses,*
Wiley New York (1954).

[43] MEYNART C. - *Les séries et leur application à la résolution de divers problèmes pratiques d'analyse mathématique,*
(Tome 1) (Eyrolles).

[44] MOCHÉ R. - *Décantation et séparation asymptotiques uniformes...*
Thèse d'Etat, Université des Sciences et Techniques de Lille I, soutenue le 10 janvier 1977.

[45] MÜNTZ - *Ueber den Approximationssatz von Weierstrass,*
Mathematische Abhandlungen,
(Schwarzes Festschrift), Berlin, 1914, p. 302-312.

[46] NEVEU J. - *Bases mathématiques du calcul des probabilités,*
Masson, Paris, 1964.

[47] NEVEU J. - *Processus aléatoires gaussiens,*
Séminaire Math. Sup. (Les Presses de l'Université de Montréal).

[48] PARREAU M. - Compléments d'Analyse (Cours).

[49] PARTHASARATHY K.R.- *Probability measures on metric spaces,*
Academic press (1967).

[50] PROKOROV YU V. - *Theory of probability and its applications,*
(an extension of s.n. Bernstein's inequalities to multi-dimensional distribution, 1968).

[51] ROSENBLATT M. - Ann. Math. Stat., 27, 1956, p. 832-837.

[52] RUDIN R. - *Real and complex analysis* (2ème édition).
Mc Graw-Hill series in Higher Mathematics.

[53] SCHWARTZ L. - *Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants),*
J. Anal. Math. of Jerusalem XIII (1964), p. 115-256.

[54] VALADIER M. - Thèse, Paris, 1970.