



HAL
open science

Indicateurs de dépendance entre deux variables aléatoires fournis par le développement en série de la densité de probabilité

René Blacher

► **To cite this version:**

René Blacher. Indicateurs de dépendance entre deux variables aléatoires fournis par le développement en série de la densité de probabilité. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1983. Français. NNT: . tel-00306714

HAL Id: tel-00306714

<https://theses.hal.science/tel-00306714>

Submitted on 28 Jul 2008

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

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

THESE

présentée à

l'Université Scientifique et Médicale de Grenoble

pour obtenir le grade de
DOCTEUR DE 3ème CYCLE
«mathématiques appliquées»

par

BLACHER René

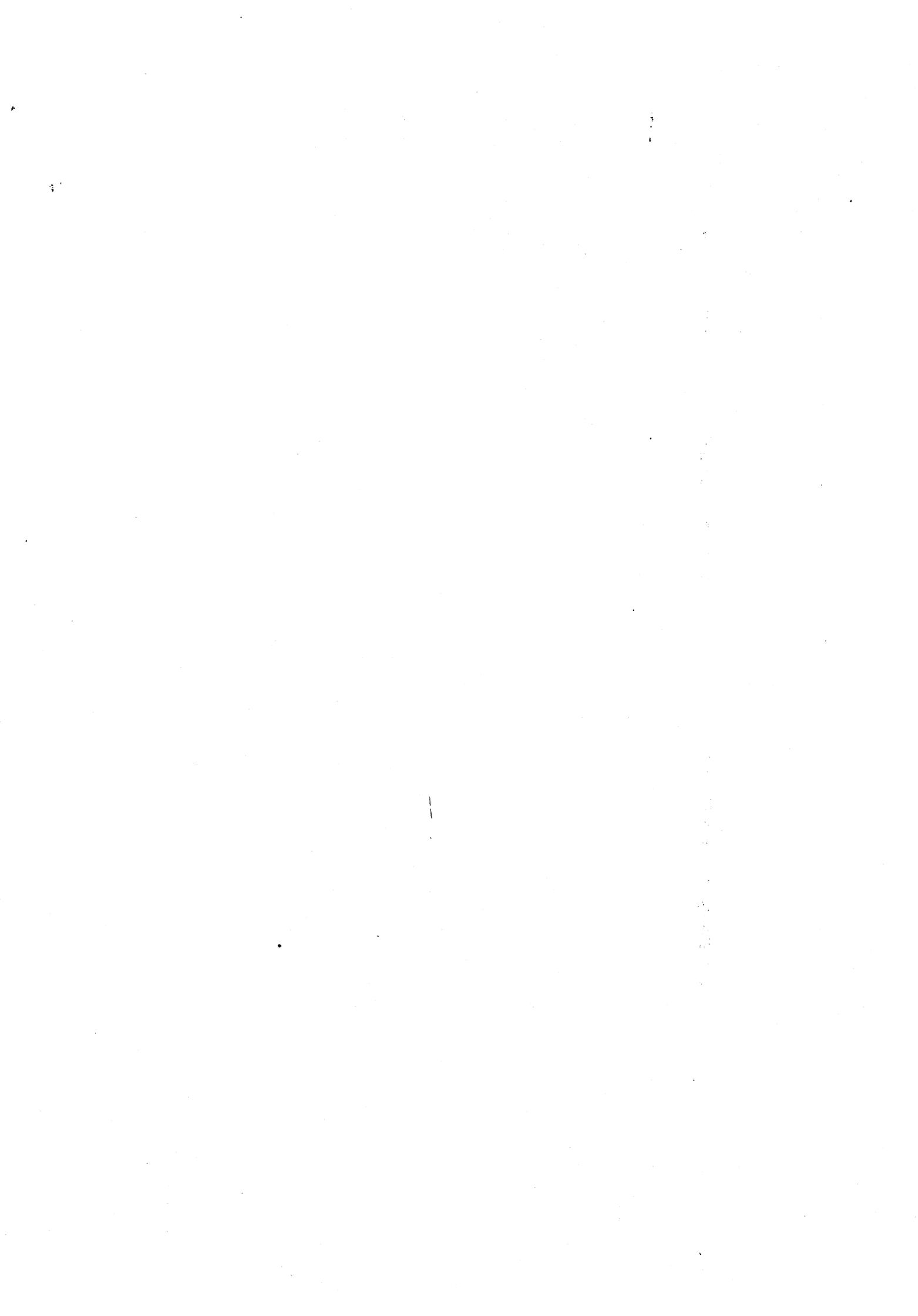


**INDICATEURS DE DEPENDANCE ENTRE DEUX
VARIABLES ALEATOIRES FOURNIS PAR LE
DEVELOPPEMENT EN SERIE DE LA DENSITE
DE PROBABILITE.**



Thèse soutenue le 30 mars 1983 devant la commission d'examen.

J.R. BARRA **Président**
B. VAN CUTSEM **Examinateurs**
A. LE BRETON



UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

année scolaire 1980-1981

Président de l'Université : M. J.J. PAYAN

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

PROFESSEURS DE 1ère CLASSE

Mlle	AGNIUS DELORD Claudine	Biophysique
	ALARY Josette	Chimie analytique
MM.	AMBLARD Pierre	Clinique dermatologie
	AMBROISE THOMAS Pierre	Parasitologie
	ARNAUD Paul	Chimie
	ARVIEU Robert	Physique nucléaire
	AUBERT Guy	Physique
	AYANT Yves	Physique approfondie
Mme	BARBIER Marie-Jeanne	Electrochimie
MM.	BARBIER Jean-Claude	Physique expérimentale
	BARBIER Reynold	Géologie
	BARJON Robert	Physique nucléaire
	BARNOUD Fernand	Biosynthèse de la cellulose
	BARRA Jean-René	Statistiques
	BARRIE Joseph	Clinique chirurgicale A
	BEAUDOING André	Clinique pédiatrie et puériculture
	BELORISKY Elie	Physique
	BENZAKEN Claude	Mathématiques appliquées
Mme	BERIEL Hélène	Pharmacodynamie
M.	BERNARD Alain	Mathématiques pures
Mme	BERTRANDIAS Françoise	Mathématiques pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques pures
	BEZES Henri	Clinique chirurgicale & traumatologie
	BILLET Jean	Géographie
	BONNET Jean-Louis	Clinique ophtalmologique
	BONNET EYMARD Joseph	Clinique Hépato-gastro-entérologie
Mme	BONNIER Jane-Marie	Chimie générale
MM.	BOUCHERLE André	Chimie et toxicologie
	BOUCHET Yves	Anatomie
	BOUCHEZ Robert	Physique nucléaire
	BRAVARD Yves	Géographie

.../...

MM. BUTEL Jean	Orthopédie
CABANEL Guy	Clinique rhumatologie et hydrologie
CARLIER Georges	Biologie végétale
CAU Gabriel	Médecine légale et toxicologie
CAUQUIS Georges	Chimie organique
CHARACHON Robert	Clinique O.R.L.
CHATEAU Robert	Clinique neurologique
CHIBON Pierre	Biologie animale
COEUR André	Chimie analytique et bromotologique
COUDERC Pierre	Anatomie pathologique
CRABBE Pierre	C.E.R.M.O.
DAUMAS Max	Géographie
DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELOBEL Claude	M.I.A.G.
DELORMAS Pierre	Pneumo-phtisiologique
DENIS Bernard	Clinique cardiologique
DEPORTES Charles	Chimie minérale
DESRE Pierre	Electrochimie
DODU Jacques	Mécanique appliquée IUT 1
DOLIQUE Jean-Michel	Physique des plasmas
DUCROS Pierre	Cristallographie
FONTAINE Jean-Marc	Mathématiques pures
GAGNAIRE Didier	Chimie physique
GASTINEL Noël	Analyse numérique
GAVEND Jean-Michel	Pharmacologie
GEINDRE Michel	Electro-radiologie
GERBER Robert	Mathématiques pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
JANIN Bernard	Géographie
JEANNIN Charles	Pharmacie galénique
JOLY Jean-René	Mathématiques pures
KAHANE André	Physique
KAHANE Josette	Physique
KLEIN Joseph	Mathématiques pures
KOSZUL Jean-Louis	Mathématiques pures
LACAZE Albert	Hermodynamique
LACHARME Jean	Biologie cellulaire
LAJZEROWICZ Joseph	Physique

Mme	LAJZEROWICZ Jeannine	Physique
MM.	LATREILLE René	Chirurgie thoracique
	LATURAZE Jean	Biochimie pharmaceutiques
	LAURENT Pierre	Mathématiques appliquées
	LE NOC Pierre	Bactériologie virologie
	LLIBOUTRY Louis	Géophysique
	LOISEAUX Jean-Marie	Sciences nucléaires
	LOUP Jean	Géographie
	LUU DUC Cuong	Chimie générale et minérale
	MALINAS Yves	Clinique obstétricale
Mlle	MARIOTTE Anne-Marie	Pharmacognosie
MM.	MAYNARD Roger	Physique du solide
	MAZARE Yves	Clinique médicale A
	MICHEL Robert	Minéralogie et pétrographie
	MIGOUD Max	Clinique maladies infectieuses
	MOURIQUAND Claude	Histologie
	NEGRE Robert	Mécanique IUT 1
	MOZIERES Philippe	Spectrométrie physique
	OMONT Alain	Astrophysique
	OZENDA Paul	Botanique
	PAYAN Jean-Jacques	Mathématiques pures
	PEBAY PEYROULA Jean-Claude	Physique
	PERRET Jean	Sémiologie médicale (neurologie)
	PERRIER Guy	Géophysique
	PIERRARD Jean-Marie	Mécanique
	RACHAIL Michel	Clinique médicale B
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
Mme	RENAUDET Jacqueline	Bactériologie
M.	REVOL Michel	Urologie
Mme	RINAUDO Marguerite	Chimie CERMAV
MM.	DE ROUGEMONT Jacques	Neuro-chirurgie
	SARRAZIN Roger	Clinique chirurgicale B
Mme	SEIGLE MURANDI Françoise	Botanique et crytogamie
MM.	SENGEL Philippe	Biologie animale
	SIBILLE Robert	Construction mécanique IUT 1
	SOUTIF Michel	Physique
	TANCHE Maurice	Physiologie
	VAILLANT François	Zoologie
	VALENTIN Jacques	Physique nucléaire

MM. VAN CUTSEM Bernard
 VAUQUOIS Bernard
 VERAÏN Alice
 VERAÏN André
 VIGNAIS Pierre

Mathématiques appliquées
 Mathématiques appliquées
 Pharmacie galénique
 Biophysique
 Biochimie médicale

PROFESSEURS DE 2ème CLASSE

MM. ARNAUD Yves
 AURIAULT Jean-Louis
 BEGUIN Claude
 BOITET Christian
 BOUTHINON Michel
 BRUGEL Lucien
 BUISSON Roger
 CASTAING Bernard
 CHARDON Michel
 CHEHIKIAN Alain
 COHEN Henri
 COHENADDAD Jean-Pierre
 COLIN DE VERDIERE Yves
 CONTE René
 CYROT Michel
 DEPASSEL Roger
 DOUCE Roland
 DUFRESNOY Alain
 GASPARD François
 GAUTRON René
 GIDON Maurice
 GIGNOUX Claude
 GLENAT René
 GOSSE Jean-Pierre
 GROS Yves
 GUITTON Jacques
 HACQUES Gérard
 HERBIN Jacky
 HICTER Pierre
 IDELMAN Simon
 JOSELEAU Jean-Paul
 JULLIEN Pierre
 KERCKOVE Claude

Chimie IUT 1
 Mécanique IUT 1
 Chimie organique
 Mathématiques appliquées
 E.E.A. IUT 1
 Energétique IUT 1
 Physique IUT 1
 Physique
 Géographie
 E.E.A. IUT 1
 Mathématiques pures
 Physique
 Mathématiques pures
 Physique IUT 1
 Physique du solide
 Mécanique des fluides
 Physiologie végétale
 Mathématiques pures
 Physique
 Chimie
 Géologie
 Sciences nucléaires
 Chimie organique
 E.E.A. IUT 1
 Physique IUT 1
 Chimie
 Mathématiques appliquées
 Géographie
 Chimie
 Physiologie animale
 Biochimie
 Mathématiques appliquées
 Géologie

MM.	KRAKOWIACK Sacha	Mathématiques appliquées
	KUHN Gérard	Physique IUT 1
	KUPKA Yvon	Mathématiques pures
	LUNA Domingo	Mathématiques pures
	MACHE Régis	Physiologie végétale
	MARECHAL Jean	Mécanique
	MICHOULIER Jean	Physique IUT 1
Mme	MINIER Colette	Physique IUT 1
MM.	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique IUT 1
	OUDET Bruno	Mathématiques appliquées
	PEFFEN René	Métallurgie IUT 1
	PELMONT Jean	Biochimie
	PERRAUD Robert	Chimie IUT 1
	PERRIAUX Jean-Jacques	Géologie minéralogie
	PERRIN Claude	Sciences nucléaires
	PFISTER Jean-Claude	Physique du solide
	PIERRE Jean-Louis	Chimie organique
Mlle	PIERY Yvette	Physiologie animale
MM.	RAYNAUD Hervé	Mathématiques appliquées
	RICHARD Lucien	Biologie végétale
	ROBERT Gilles	Mathématiques pures
	ROBERT Jean-Bernard	Chimie physique
	ROSSI André	Physiologie végétale
	SAKAROVITCH Michel	Mathématiques appliquées
	SARROT REYNAUD Jean	Géologie
	SAXOD Raymond	Biologie animale
Mme	SOUTIF Jeanne	Physique
MM.	STUTZ Pierre	Mécanique
	VIALON Pierre	Géologie
	VIDAL Michel	Chimie organique
	VIVIAN Robert	Géographie

CHARGES D'ENSEIGNEMENT PHARMACIE

MM.	ROCHAS Jacques	Hygiène et hydrologie
	DEMENGE Pierre	Pharmacodynamie

PROFESSEURS SANS CHAIRE (médecine)

M.	BARGE Michel	Neuro-chirurgie
----	--------------	-----------------

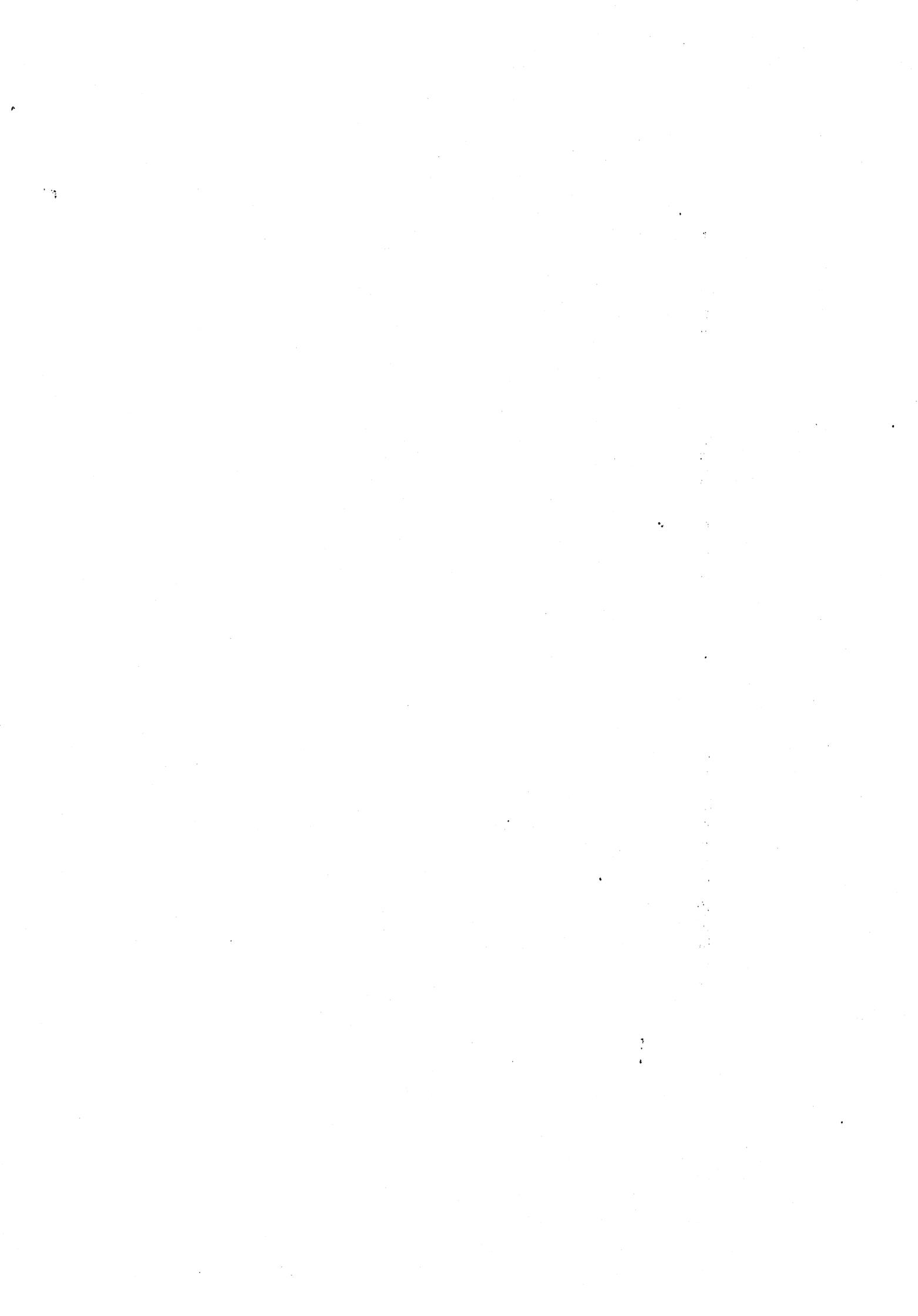
MM.	BOST Michel	Pédiatrie
	BOUCHARLAT Jacques	Psychiatrie
	CHAMBAZ Edmond	Biochimie (hormonologie)
	CHAMPETIER Jean	Anatomie
	COLOMB Maurice	Biochimie
	COULOMB Max	Radiologie
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	GROULADE Joseph	Biochimie A
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Gérontologie
	JALBERT Pierre	Histologie
	MAGNIN Robert	Hygiène
	PHELIP Xavier	Rhumatologie
	REYMOND Jean-Charles	Chirurgie générale
	STIEGLITZ Paul	Anesthésiologie
	VROUSOS Constantin	Radiothérapie

MAITRES DE CONFERENCES AGREGES (médecine)

MM.	BACHELOT Yvan	Endocrinologie
	BENABID Alim Louis	Médecine et chirurgie
	BERNARD Pierre	Gynécologie obstétrique
	CONTAMIN Charles	Chirurgie thoracique
	CORDONNIER Daniel	Néphrologie
	CROUZET Guy	Radiologie
	DEBRU Jean-Luc	Médecine interne
	DYON Jean-François	Chirurgie infantile
	FAURE Claude	Anatomie et organogénèse
	FAURE Gilbert	Urologie
	FLOYRAC Roger	Biophysique
	FOURNET Jacques	Hépatogastro-entérologie
	GAUTIER Robert	Chirurgie générale
	GIRARDET Pierre	Anesthésiologie
	GUIDICELLI Henri	Chirurgie générale
	GUIGNIER Michel	Thérapeutique (réanimation)
	JUNIEN-LAVILLAULOY Claude	Clinique O.R.L.
	KOLODIE Lucien	Hématologie biologique
	MALLION Jean-Michel	Médecine du travail
	MASSOT Christian	Médecine interne
	MOUILLON Michel	Ophtalmologie

MM. PARAMELLE Bernard
RACINET Claude
RAMBAUD Pierre
RAPHAEL Bernard
SCHAEFER René
SEIGNEURIN Jean-Marie
SOTTO Jean-Jacques
STOEBNER Pierre

Pneumologie
Gynécologie-Obstétrique
Pédiatrie
Stomatologie
Cancérologie
Bactériologie-virologie
Hématologie
Anatomie-pathologique



Je suis reconnaissant à Monsieur BARRA de l'honneur qu'il me fait en acceptant la présidence de ce jury .

Je remercie particulièrement Monsieur VAN CUTSEM d'avoir accepté de diriger mes travaux de recherche , ainsi que de la patience avec laquelle il a bien voulu lire les nombreux manuscrits que je lui ai remis .

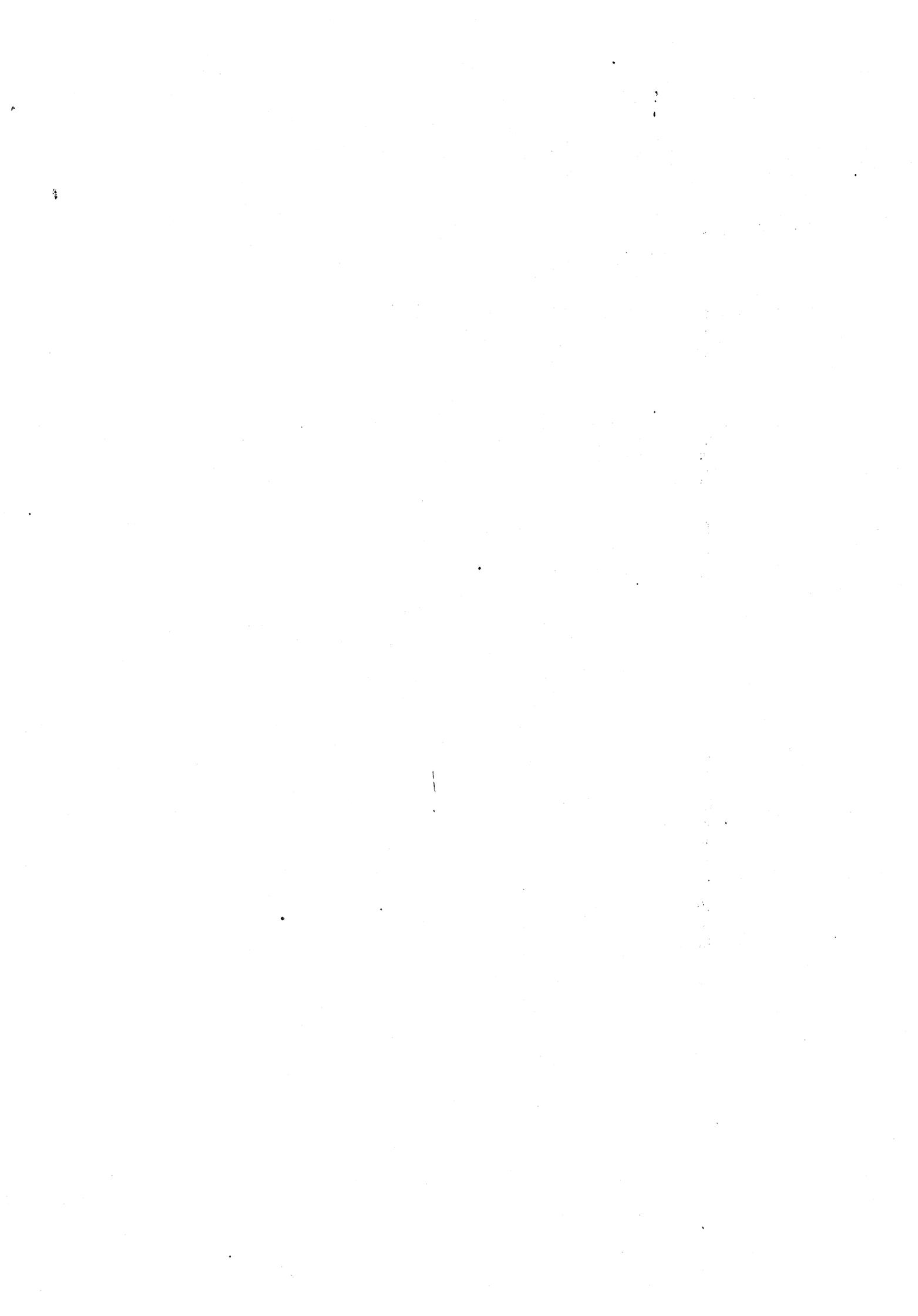
Je remercie également Monsieur LE BRETON de l'intérêt qu'il porte à mes travaux en acceptant d'être membre du jury .

L'équipe de statistique de Grenoble m'a toujours procuré l'aide et les conseils nécessaires , je tiens à le rappeler .

Je tiens aussi à remercier tous ceux qui m'ont aidé dans la réalisation matérielle de cette thèse , ainsi que des rapports de recherche qui l'ont précédée , et particulièrement les secrétaires de la Tour Maths et le service de reprographie de l'I.R.M.A. .

Je ne veux pas oublier l'aide que m'ont fournie Michèle et Hugues en relisant et corrigeant certaines parties de mes travaux .

Je tiens enfin à remercier Marie , tant pour ses conseils mathématiques que les autres , et sans qui cette thèse n'aurait pas eu lieu . Aussi je la Lui dédie .



SOMMAIRE

	Page
INTRODUCTION	1
NOTATIONS	4
<u>PARTIE 1 : INDICATEURS DE DEPENDANCE</u>	6
<u>PARTIE 2 : ETUDE DE CAS PARTICULIERS</u>	14
1 : PROPRIETES DES BASES ORTHONORMALES	15
2 : ASPECT PROBABILISTE	19
2-1 Cas des polynomes de Legendre	20
2-2 Cas des fonctions trigonométriques	24
2-3 Cas des fonctions de Walsh	28
2-4 Démonstrations	33
3 : COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR	39
3-1 Etude générale	40
1 Notations et hypothèses	40
2 Propriétés générales	41
3 Définition. et propriétés des coefficients de corrélation d'ordre supérieur	50
4 Coefficients de corrélation empiriques d'ordre supérieur	60
3-2 Cas de $(F_1(X), F_2(Y))$	63
3-3 Cas particuliers	67

4	: ASPECT STATISTIQUE	71
4-1	Test hilbertien	71
1	Cas de noyau fixe	71
2	Cas de noyau variant avec la taille de l'échantillon	73
3	Test hilbertien	74
4	Convergence du test	76
4-2	Estimateurs	77
4-3	Test d'indépendance	82
1	Test réel d'indépendance	82
2	Cas des polynomes de Legendre	89
3	Cas des fonctions trigonométriques	92
4	Cas des fonctions de Walsh	94
5	Application	98
6	Choix de l'ordre du développement	99
7	Comparaison des différentes séries	101
8	Conclusions	102
5	: TEST D'INDEPENDANCE PAR LE KHI-DEUX	103
5-1	Statistique hilbertienne et statistique du khi-deux	105
5-2	Cas de lois marginales connues	111
5-3	Cas de lois marginales inconnues	115
5-4	Conclusions	121

INTRODUCTION

Le test du khi-deux de Pearson est un des tests statistiques les plus connus , et très tôt s'est posé le problème de le rattacher à une classe de tests plus générale . Déjà Neymann en 1937 [14] montrait le lien existant avec le test des coefficients du développement de la densité de probabilité en série de polynomes de Legendre . Cette idée a été reprise par Barton en 1956 [16] , puis par Nadaraja [15] .

Enfin Bosq a complété ces résultats en 1978 par une étude de la classe des tests hilbertiens , le test des coefficients du développement de la densité de probabilité en série de fonctions orthogonales en étant un cas particulier .

Or, comme le faisait déjà remarquer Neyman , un des problèmes principaux de ce type de tests est celui du choix de l'ordre du développement de façon à avoir une approximation satisfaisante de la densité de probabilité .

Pour résoudre ce problème , nous avons alors eu besoin d'une mesure de dépendance et nous avons eu recours aux indicateurs D_h^N fournis par les distances χ^h classiques de la fonction de répartition du vecteur aléatoire à la fonction de répartition de l'indépendance . Les principales propriétés de ces indicateurs ont été étudiées par Renyi ([7], [8]) , Schweizer et Wolf [12] . Nous les avons complétées dans [2] et [3] de façon à pouvoir faire le lien avec les développements en série . Nous rappellerons ces propriétés dans une première partie .

Dans une deuxième partie , nous avons étudié le cas particulier de vecteurs aléatoires de lois marginales uniformes à valeurs dans $[0,1]^2$.

Comme bases orthogonales , nous avons choisi les fonctions produits des éléments des bases orthogonales classiques : polynomes de Legendre , fonctions trigonométriques ou fonctions indicatrices d'ensemble .

Nous avons déjà étudié , sous l'aspect probabiliste , les principales propriétés du lien entre série et dépendance dans [3] , nous complétons cette étude ici et nous étudions également l'aspect statistique à partir des résultats de Bosq .

Comme on pouvait l'espérer , on trouve alors que ce sont les termes de plus bas degré de la série qui sont les plus importants pour la dépendance mesurée par les coefficients D_h^N , ce qui fournit une solution au problème soulevé par Neyman .

L'avantage essentiel des bases que nous avons employées est de faire apparaître les lois marginales de façon simple .

Or , ce résultat est facilement généralisable au cas d'un vecteur aléatoire à valeurs dans \mathbb{R}^2 , de lois marginales quelconques et n'admettant pas forcément de densité de probabilité .

Nous obtenons ainsi des suites de coefficients mesurant chacun une dépendance particulière . En particulier nous pouvons définir des coefficients de corrélation d'ordre supérieur , complétant ρ dont on connaît les lacunes .

D'autre part , cette possibilité de faire apparaître les lois marginales de façon simple nous permettra aussi d'affiner le test d'indépendance par le khi-deux . En effet , le test classique - basé sur la statistique $\chi_{X,Y}^2$ - dans le cas de lois marginales connues , teste en fait des coefficients déjà connus .

Nous proposerons donc un autre test d'indépendance par le khi-deux qui apparaîtra plus naturel que le test classique , et dont la statistique $\chi_{X,Y}^2 - \chi_X^2 - \chi_Y^2$, sous l'hypothèse d'indépendance , suit asymptotiquement une loi du khi-deux ayant le même nombre de degrés de liberté que les lois marginales soient connues ou inconnues .

Nous concluerons enfin sur les rapports entre la dépendance et les coefficients du développement en série de la densité de probabilité dans une dernière partie où nous étudierons un exemple concret : les congruences linéaires considérées comme générateurs de nombres pseudo-aléatoires ce qui nous permettra de préciser un peu nos résultats de [4] ;

NOTATIONS

Nous noterons donc par (Ω, \mathcal{A}, P) un espace probabilisé et par V_E , l'ensemble des applications mesurables de (Ω, \mathcal{A}, P) dans un espace mesurable (E, \mathcal{E}) donné.

Si $E \subset \mathbb{R}^2$, alors $V_E^{m \otimes m'}$ désignera l'ensemble des vecteurs aléatoires $(X, Y) \in V_E$, de lois marginales m et m' .

Si $E = [0, 1]^2$, alors nous noterons V_E par V , et $V_E^{m \otimes m}$ par V^U lorsque m est la mesure de Lebesgue.

Soit $p \in \mathbb{N}^*$ et $\vec{X} \in V_{\mathbb{R}^p}$, nous noterons par F sa fonction de répartition, et par f sa densité de probabilité éventuelle.

De plus, si $(X, Y) \in V_{\mathbb{R}^2}$, nous noterons par F_1 et F_2 les fonctions de répartition de X et Y et par f_1 et f_2 leurs densités éventuelles.

Nous garderons les notations usuelles : $E(X)$, $\text{cov}(X, Y)$, $\sigma^2(X)$ et $\rho(X, Y)$ pour désigner respectivement l'espérance, la covariance, la variance et le coefficient de corrélation.

D'autre part, si H_i , $i = 1, 2, \dots, p$, est une suite d'applications de \mathbb{R} dans \mathbb{R} , nous noterons par

$$H_1 \otimes H_2 \otimes \dots \otimes H_p,$$

l'application de \mathbb{R}^p dans \mathbb{R} définie par :

$$H_1 \otimes H_2 \otimes \dots \otimes H_p (y_1, y_2, \dots, y_p) = H_1(y_1) \cdot H_2(y_2) \cdot \dots \cdot H_p(y_p).$$

Nous noterons encore : \mathbb{N}^{*2} l'ensemble $\mathbb{N}^{*2} = (\mathbb{N}^*)^2$;

U l'application de \mathbb{R}^p dans \mathbb{R} définie par

$$U(y_1, y_2, \dots, y_p) = y_1 \cdot y_2 \cdot \dots \cdot y_p \quad (U(x, y) = x \cdot y \text{ si } p=2);$$

et $\mathcal{C}(F)$ l'ensemble des applications continues de l'espace topologique F dans \mathbb{R} .

Nous garderons les notations classiques : \mathcal{B}_p , $p \in \mathbb{N}^*$, désignera la tribu des boréliens sur \mathbb{R}^p , ou sur un de ses sous-ensembles E ,

et nous simplifierons $\mathcal{L}^2(E, \mathcal{B}_p, m)$ en $\mathcal{L}^2(E, m)$, et en $\mathcal{L}^2(E)$ si m est la mesure borélienne.

Enfin, les signes $\langle \cdot, \cdot \rangle$ et $\| \cdot \|$ désigneront le produit scalaire et la norme usuels sur les espaces $\mathcal{L}^2(E, \mathcal{E}, Q)$ ou (E, \mathcal{E}, Q) est un espace mesuré.

Ces signes seront précisés par $\langle \cdot, \cdot \rangle_{\mathcal{L}^2(E, \mathcal{E}, Q)}$ et $\| \cdot \|_{\mathcal{L}^2(E, \mathcal{E}, Q)}$ lorsque ce sera nécessaire.

PARTIE 1 : INDICATEURS DE DEPENDANCE

Pour pouvoir juger du rôle des différents coefficients des développements en série de la densité de probabilité par rapport à la dépendance, il nous faudra avoir une mesure de celle-ci. Nous aurons alors recours aux coefficients D_h^N , $h \in \mathbb{N}^* \cup \{+\infty\}$, qui sont en fait les distances \mathcal{L}^h classiques de la fonction de répartition à la fonction de répartition de l'indépendance. Nous allons maintenant rappeler les définitions et les principales propriétés de ceux-ci. Pour les démonstrations, nous renvoyons à [2], [3], [12].

Quant au problème du choix de ces coefficients-ci plutôt que d'autres, nous le traiterons en partie 3-1-2.

1-1 : Notations : Nous supposons donc que $(X,Y) \in V$, et nous noterons par δ_h , $h \in \mathbb{N}^* \cup \{+\infty\}$, les distances normées \mathcal{L}^h classiques sur les fonctions boréliennes réelles de $[0,1]^2$:

$$\delta_h(g,t) = \|t-g\|_{\mathcal{L}^h([0,1]^2)},$$

avec la convention $\delta_\infty(g,t) = \delta(g,t) = \sup_{(x,y) \in [0,1]^2} |g(x,y) - t(x,y)|$.

Dans le cas où l'on possède un vecteur aléatoire à valeurs dans $[0,1]^P$, la généralisation naturelle de ces notations est immédiate et c'est elle que nous adopterons.

Nous choisirons comme mesures de l'uniformité les distances

$$\delta_h(F,U) \quad \text{et} \quad \delta_h\left(f, 1_{[0,1]^2}\right), \quad h \in \mathbb{N}^* \cup \{+\infty\},$$

U représentant l'application $U(x,y) = x.y$.

Dans le cas où les lois marginales sont uniformes - et l'on sait que l'on peut toujours se ramener à cette hypothèse si F_1 et F_2 sont continues [2] - ces mesures de l'uniformité nous fournissent alors toute une suite de coefficients de dépendance que nous noterons :

$$D_h(X,Y) = \delta_h(F,U), \quad D(X,Y) = \delta(F,U),$$

$$d_h(X,Y) = \delta_h\left(f, 1_{[0,1]^2}\right), \quad d(X,Y) = \delta\left(f, 1_{[0,1]^2}\right),$$

pour $h \in \mathbb{N}^* \cup \{+\infty\}$.

Nous noterons par \mathcal{P} l'ensemble des partitions boréliennes finies de $[0,1]^2$: $\{B_s\}, s \in S$, vérifiant $\mu(B_s) \neq 0$ - μ mesure de Lebesgue. Soit $\{B_s\} \in \mathcal{P}$, nous noterons alors par $f_{\{B_s\}}$, la fonction constante et égale à

$$\frac{P\{(X,Y) \in B_s\}}{\mu(B_s)} \quad \text{sur chaque } B_s.$$

Théorème 1-2 : Soit $(X, Y) \in V^U$,

alors on a les inégalités suivantes :

$$D_h(X, Y) < \left[\frac{(h!)^2}{(h+1)(2h+1)!} \right]^{\frac{1}{h}} \text{ pour } h \in \mathbb{N}^* \text{ , et } D(X, Y) < \frac{1}{4} .$$

Lorsque $h \in \mathbb{N}^*$, l'égalité est réalisée si et seulement si $X = Y$ p.s. ou $X = 1 - Y$ p.s. . Lorsque $h = +\infty$ l'égalité est réalisée si $X = Y$ p.s. ou $X = 1 - Y$ p.s.

Nous avons alors normalisé les coefficients D_h en posant :

$$D_h^N(X, Y) = \left[\frac{(h+1)(2h+1)!}{(h!)^2} \right]^{\frac{1}{h}} D_h(X, Y) \text{ pour } h \in \mathbb{N}^* ,$$

$$\text{et } D^N(X, Y) = D_\infty^N(X, Y) = 4 D(X, Y) .$$

Théorème 1-3 : Avec les notations et hypothèses précédentes alors

$$\frac{2}{3} (D(X, Y))^3 < D_1(X, Y) < D_2(X, Y) < D_3(X, Y) < \dots < D(X, Y) .$$

Théorème 1-4 : Soit $\{ I_{i,j} \}, (i, j) \in J \times J', \{ I_{i,j} \} \in \mathcal{P}$,

vérifiant : $I_{i,j} = I_i^1 \times I_j^2$, où I_i^1 et I_j^2

sont deux intervalles de $[0,1]$, de longueur L_i^1 et L_j^2 et soient

$$L_1 = \sup_{i \in J} (L_i^1) \quad \text{et} \quad L_2 = \sup_{j \in J'} (L_j^2).$$

Soit $(X,Y) \in \mathcal{V}^U$ et F sa fonction de répartition.

Soit G la fonction primitive :

$$G(x,y) = \int_0^x \int_0^y f_{\{I_{i,j}\}}(u,v) \cdot du \cdot dv.$$

$$\text{alors} \quad \delta(F,G) < \frac{1}{4} (L_1 + L_2).$$

1-5 : Axiomes de Renyi : Renyi a introduit ces axiomes en 1959, [6], [7]. Ceux-ci forment un ensemble raisonnable d'exigences pour des coefficients de dépendance.

Soient deux variables aléatoires X et Y de fonctions de répartition continues, alors le coefficient de dépendance $R(X,Y)$ vérifie les axiomes de Renyi si :

- 1°) $R(X,Y)$ est défini pour tout X et tout Y .
- 2°) $R(X,Y) = R(Y,X)$.
- 3°) $0 < R(X,Y) < 1$
- 4°) $R(X,Y) = 0$ si et seulement si X et Y sont indépendantes.

- 5°) $R(X,Y) = 1$ si et seulement si X est une fonction strictement monotone de Y , et Y est une fonction strictement monotone de X .
- 6°) Si f et g sont strictement monotones sur les ensembles de définition de X et Y , alors
 $R(f(X),g(Y)) = R(X,Y)$.
- 7°) Si la distribution de (X,Y) suit une loi normale avec $\rho(X,Y)$ pour coefficient de corrélation alors $R(X,Y)$ est une fonction strictement croissante de $\rho(X,Y)$.
- 8°) Si (X,Y) et (X_n, Y_n) , $n \in \mathbb{N}^*$, sont des variables aléatoires de fonctions de répartition F et F_n , et si la suite F_n converge faiblement vers F alors

$$\lim_{n \rightarrow \infty} R(X_n, Y_n) = R(X, Y).$$

Théorème 1-6 Les coefficients D_h^N , $h \in \mathbb{N}^*$, vérifient les axiomes de Renyi (cf [12]).

Théorème 1-7 : Soient $(X_i, Y_i) \in V^U$, $i=1,2$, de fonctions de répartition F_i et soient r et $q \in \mathbb{N}^* \cup \{+\infty\}$, alors

$$\left| D_q(X_1, Y_1) - D_q(X_2, Y_2) \right| < \delta_q(F_1, F_2),$$

$$\text{et } \delta_q(F_1, F_2) < \delta_r(F_1, F_2) \text{ si } r > q,$$

$$\text{et } \delta_q(F_1, F_2) < \left[\frac{3}{2} \delta_r(F_1, F_2) \right]^{\frac{1}{3}}, \text{ si } q > r.$$

Théorème 1-8 : Soit (X_n, Y_n) , $n \in \mathbb{N}$, une suite d'éléments de V^U , de fonctions de répartition F_n , et soit $h \in \mathbb{N}^* \cup \{+\infty\}$, si l'on suppose de plus que (X_0, Y_0) suit la loi uniforme sur $[0,1]^2$, alors on a l'équivalence :

$$\left((X_n, Y_n) \xrightarrow{h} (X_0, Y_0) \right) \iff \left(\lim_{n \rightarrow \infty} (D_h(X_n, Y_n)) = 0 \right).$$

Théorème 1-9 : Notons par b l'application de

$V^U \times [0,1]^2$ dans \mathbb{R} définie par :

$$b(X, Y, B_1, B_2) = P((X, Y) \in B_1 \times B_2) - (P(X \in B_1)) \times (P(Y \in B_2))$$

et par t la restriction de b à $V^U \times \mathcal{J}^2$

où \mathcal{J} est l'ensemble des intervalles inclus dans $[0,1]$.

Soient $\epsilon \in \mathbb{R}_+^*$, et (X, Y) et $(X_0, Y_0) \in V^U$.

alors les implications suivantes sont vérifiées :

$$1) |t(X, Y, \dots) - t(X_0, Y_0, \dots)| < \epsilon \implies \delta(F, F_0) < \epsilon$$

$$2) \delta(F, F_0) < \epsilon \implies |t(X, Y, \dots) - t(X_0, Y_0, \dots)| < 4\epsilon$$

$$3) \delta_1(f, f_0) < \epsilon \implies |b(X, Y, \dots) - b(X_0, Y_0, \dots)| < \epsilon.$$

Corollaire 1-10 : Soient h et $h' \in \mathbb{N}^* \cup \{+\infty\}$, on a également

- 4) $|t(X, Y, \dots)| < \varepsilon \implies D_h(X, Y) < \varepsilon$,
- 5) $D^N(X, Y) < \varepsilon \implies |t(X, Y, \dots)| < \varepsilon$,
- 6) $D_h(X, Y) < \varepsilon \implies |t(X, Y, \dots)| < 4 \left[\frac{3}{2} \varepsilon \right]^{\frac{1}{3}}$,
- 7) $d_h(X, Y) < \varepsilon \implies |b(X, Y, \dots)| < \varepsilon$,
- 8) $d_h(X, Y) < \varepsilon \implies D_{h'}(X, Y) < \varepsilon$.
-

Corollaire 1-11 : Soient $\{I_j\}$ et $\{I'_{j'}\}$, j et $j' \in \mathbb{N}$,
deux suites d'intervalles disjoints de $[0, 1]$.

Posons , pour tout couple $(k, k') \in \mathbb{N}^{*2}$:

$$B_k = \bigcup_{j=1}^k I_j \quad \text{et} \quad B'_{k'} = \bigcup_{j'=1}^{k'} I'_{j'} ,$$

alors

- 9) $D^N(X, Y) < \varepsilon \implies |b(X, Y, B_k, B'_{k'})| < k \cdot k' \cdot \varepsilon$,
- 10) $D_h(X, Y) < \varepsilon \implies |b(X, Y, B_k, B'_{k'})| < 4 k \cdot k' \left[\frac{3}{2} \varepsilon \right]^{\frac{1}{3}}$.
-

Exemple 1-12 : Nous rappelons que nous avons construits (1-16, [3]) , l'exemple d'une suite de V^U : (X_n, Y_n) , $n \in \mathbb{N}$, telle que pour tout entier n , il existe une bijection f_n de $[0,1]$ sur lui-même vérifiant $Y_n = f_n(X_n)$ et

$$D^N (X_n, Y_n) < \frac{2}{2^n} \quad \text{et} \quad d_h(X_n, Y_n) = +\infty \quad \text{pour } h > 2 .$$

Nous avons alors fait remarquer que par suite des arrondis de nos calculs , il existe un entier q tel que pour tout borélien $(B_1 \times B_2)$ de $[0,1]^2$, on trouvera l'égalité :

$$P ((X_q, Y_q) \in B_1 \times B_2) = (P(X_q \in B_1)) \times (P(Y_q \in B_2)) .$$

et donc que , normalement on estimera statistiquement que

$$D_h^N(X_q, Y_q) = d_h(X_q, Y_q) = 0 .$$

et que les deux variables X_q et Y_q sont indépendantes .

Nous nous servirons de cet exemple lors de nos conclusions .

PARTIE 2 : ETUDE DE CAS PARTICULIERS

Pour étudier concrètement les rapports entre la dépendance et les développements en série de la densité de probabilité , nous allons d'abord nous limiter au cas de vecteurs aléatoires à valeurs dans $[0,1]^2$, de lois marginales uniformes , et aux développements en série de produits de polynômes de Legendre , de fonctions trigonométriques et de fonctions de Walsh . Pour cela , nous rappellerons d'abord quelques propriétés des bases orthonormales , puis nous étudierons l'aspect probabiliste , puis statistique de ces développements . Nous généraliserons aussi les résultats obtenus en ce qui concerne les coefficients de corrélation aux coefficients de corrélation d'ordre supérieur , et les tests du khi-deux comme tests d'indépendance .

1 : PROPRIETES DES BASES ORTHONORMALES

Cette partie est en fait la clef de cette étude , car elle permet de voir comment développer f en série en faisant apparaître de façon simple les lois marginales .

Pour les démonstrations (excepté 2-5 et 2-6) nous renvoyons à [3] .

Soit μ' (resp μ'') , une probabilité sur les boréliens de \mathbb{R} , et soit $\{G'_i\}$ (resp $\{G''_i\}$) , $i \in \mathbb{N}$, une base orthonormale de $\mathcal{L}^2(\mathbb{R} , \mu')$ (resp $\mathcal{L}^2(\mathbb{R} , \mu'')$) .

Théorème 2-1 : Avec les notations et hypothèses précédentes , la famille $\{G'_i \otimes G''_j\}$, $(i,j) \in \mathbb{N}^2$, est une base orthonormale de $\mathcal{L}^2(\mathbb{R}^2 , \mu' \otimes \mu'')$.

Théorème 2-2 : Gardons les notations et les hypothèses du théorème 2-1 et soit $f \in \mathcal{L}^2(\mathbb{R}^2 , \mu' \otimes \mu'')$. Nous pouvons donc écrire

$$f = \sum_{(i,j) \in \mathbb{N}^2} \tau_{i,j} \cdot G'_i \otimes G''_j \quad \text{dans} \quad \mathcal{L}^2(\mathbb{R}^2 , \mu' \otimes \mu'') .$$

Posons alors : $\Gamma_i'(x) = \int_{-\infty}^x G_i'(t') \cdot \mu'(dt')$ pour $i \in \mathbb{N}$,

et $\Gamma_j''(y) = \int_{-\infty}^y G_j''(t'') \cdot \mu''(dt'')$ pour $j \in \mathbb{N}$,

et $F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(t',t'') \cdot \mu'(dt') \cdot \mu''(dt'')$,

$$\text{alors } F = \sum_{(i,j) \in \mathbb{N}^2} \tau_{i,j} \cdot \Gamma_i' \otimes \Gamma_j''$$

dans l'ensemble des fonctions réelles bornées sur \mathbb{R}^2 ,
muni de la topologie de la convergence uniforme.

Théorème 2-3 : Avec les notations et hypothèses du théorème 2-2,
et en supposant de plus que f est une densité de probabilité par rapport
à $\mu' \otimes \mu''$, et que $G_0 \equiv 1_{\mathbb{R}}$, alors on a les égalités :

$$\tau_{0,0} = 1, \text{ et}$$

$$\int_{\mathbb{R}} f(x,y) \cdot \mu''(dy) = f_1(x) = 1 + \sum_{i=1}^{\infty} \tau_{i,0} \cdot G_i'(x) \text{ dans } \mathcal{L}^2(\mathbb{R}, \mu')$$

$$\text{et } \int_{\mathbb{R}} f(x,y) \cdot \mu'(dx) = f_2(y) = 1 + \sum_{j=1}^{\infty} \tau_{0,j} \cdot G_j''(y) \text{ dans } \mathcal{L}^2(\mathbb{R}, \mu'')$$

Théorème 2-4 : Avec les notations et hypothèses du théorème 2-3 ,
et en supposant de plus que f est une densité de probabilité d'un
vecteur aléatoire à valeurs dans \mathbb{R}^2 , dont les lois marginales sont
 μ' et μ'' ,

alors , pour tout $i \in \mathbb{N}^*$, $\tau_{i,0} = \tau_{0,i} = 0$, et donc

$$f = 1 + \sum_{(i,j) \in \mathbb{N}^* \times \mathbb{N}^*} \tau_{i,j} \cdot G_i' \otimes G_j'' \quad \text{dans } \mathcal{L}^2(\mathbb{R}^2, \mu' \otimes \mu'') .$$

Théorème 2-5 : Soit $(X,Y) \in \mathcal{V}_{\mathbb{R}^2}^{\mu' \otimes \mu''}$, de loi Q , et

si nous posons pour tout $(i,j) \in \mathbb{N}^* \times \mathbb{N}^*$: $\tau_{i,j} = E(G_i'(X) \cdot G_j''(Y))$,

alors , pour tout $i \in \mathbb{N}^*$, $\tau_{i,0} = \tau_{0,i} = 0$.

En effet ,

$$\tau_{i,0} = \iint_{\mathbb{R}^2} G_i'(x) \cdot Q(dx, dy) = \int_{\mathbb{R}} G_i'(x) \cdot \mu'(dx) = 0 .$$

Remarque : On définit ainsi une généralisation des coefficients
 $\tau_{i,j}$ précédents .

Théorème 2-6 : Avec les notations et hypothèses du théorème 2-5 ,
pour tout couple $(i,j) \in \mathbb{N}^{*2}$, on a l'inégalité :

$$|\tau_{i,j}| < 1 .$$

Il suffit de remarquer que $\tau_{i,j}$ est le coefficient de
corrélation de $G'_i(X)$ et $G''_j(Y)$.

2 : ASPECT PROBABILISTE

Nous allons maintenant énoncer des propriétés qui permettent de donner un lien entre l'ordre du développement en série et la dépendance mesurée par les coefficients D_h^N .

Nous nous limitons ici au cas des séries de produits de polynômes de Legendre , de fonctions trigonométriques et de fonctions de Walsh .

La démonstration des résultats sera donné dans le quatrième paragraphe .

1 : Cas de polynomes de Legendre

Un des premiers développements en série d'une densité de probabilité qu'il vient à l'idée d'étudier est le développement en polynomes : dans le cas de lois marginales uniformes , il s'agira des polynomes de Legendre . Une des caractéristiques de ce développement , étant que le coefficient du polynome de plus bas degré de la série n'est autre que le coefficient de corrélation , nous suggérerons de remplacer celui-ci , dont on connaît les lacunes , par " une suite de coefficients de corrélation " .

Nous supposons donc que $(X,Y) \in \mathcal{V}^U$.

Définition 3-1 : Nous appellerons donc polynomes de Legendre sur $[0,1]$, les polynomes $L_i \in \mathbb{R}[X]$, $i \in \mathbb{N}$, vérifiant

$$\deg(L_i) = i \text{ pour tout } i \in \mathbb{N} ,$$

$$\text{et } \langle L_i, L_j \rangle_{\mathcal{L}^2([0,1]^2)} = 1 , \text{ si } i = j , \text{ et } 0 \text{ sinon .}$$

Proposition 3-2 : Pour tout $n \in \mathbb{N}$, on a l'égalité suivante :

$$L_n(x) = \frac{\sqrt{2n+1}}{n!} \cdot \sum_{i=0}^n \left(\frac{C_n^i \cdot (-1)^{n-i} \cdot ((n+i)!)}{i!} \right) \cdot x^i .$$

En particulier $L_0 \equiv 1$, $L_1(x) = 2\sqrt{3}(x - \frac{1}{2})$, $L_2(x) = \sqrt{5}(6x^2 - 6x + 1)$,

$$\text{et } L_3(x) = \sqrt{7}(20x^3 - 30x^2 + 12x - 1).$$

Théorème 3-3 : Si (X, Y) admet une densité de probabilité $f \in \mathcal{L}^2([0,1]^2)$, alors il existe une suite double réelle

$\{\theta_{i,j}\}$, $(i,j) \in \mathbb{N}^{*2}$, telle que

$$f = 1 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \theta_{i,j} \cdot L_i \otimes L_j \quad \text{dans } \mathcal{L}^2([0,1]^2),$$

$$\text{et } \theta_{1,1} = \rho(X, Y).$$

Nous généraliserons alors au cas où il n'y a pas de densité de probabilité en posant

$$\theta_{i,j} = E(L_i(X) \cdot L_j(Y)).$$

Théorème 3-4 : Avec les notations et hypothèses précédentes,

et si nous posons, pour tout $(i,j) \in \mathbb{Z}^2$:

$$\overline{\theta}_{i,j} = \begin{cases} \theta_{i,j} & , \text{ si } i > 1 \text{ et } j > 1 \\ 0 & \text{ sinon } , \end{cases}$$

$$\beta_{i,j} = \frac{\frac{\overline{\theta}_{i+1,j+1}}{\sqrt{(2i+3)(2j+3)}} + \frac{\overline{\theta}_{i-1,j-1}}{\sqrt{(2i-1)(2j-1)}} - \frac{\overline{\theta}_{i+1,j-1}}{\sqrt{(2i+3)(2j-1)}} - \frac{\overline{\theta}_{i-1,j+1}}{\sqrt{(2i-1)(2j+3)}}}{4 \sqrt{(2i+1)(2j+1)}} ,$$

alors

$$F - U = \sum_{(i,j) \in \mathbb{N}^2} \beta_{i,j} \cdot L_i \otimes L_j \quad \text{dans } \mathcal{L}^2([0,1]^2) .$$

Théorème 3-5 : Soit $\{M_{i,j}\}$, $(i,j) \in \mathbb{N}^2$, la suite double

définie par : $M_{0,0} = \frac{1}{12}$, $M_{1,1} = \frac{1}{60}$, $M_{0,1} = M_{1,0} = \sqrt{\frac{1}{720}}$,

$M_{0,s} = M_{s,0} = \frac{1}{s} \sqrt{\frac{1}{192}}$, et $M_{1,s} = M_{s,1} = \frac{1}{s} \sqrt{\frac{1}{960}}$ pour $s \in \mathbb{N}$ et $s > 2$,

$M_{i,j} = M_{j,i} = \frac{1}{16 \cdot ij}$, pour $(i,j) \in \mathbb{N}^2$, $i > 2$ et $j > 2$.

Alors on a les inégalités suivantes :

$$\forall (s,t) \in \mathbb{N}^2 , |\beta_{s,t}| < M_{s,t} .$$

3-6 : Cas où les $\theta_{i,j}$, $i \leq q$ et $j \leq q$, sont nuls

Soit \mathcal{P} une partie de \mathbb{N}^2 , vérifiant les axiomes suivants

1°) $\text{card}(\mathcal{P}) < +\infty$.

2°) Si $(i,j) \in \mathcal{P}$, alors pour tout $(i',j') \in \mathbb{N}^2$, tel que $0 < i' < i$ et $0 < j' < j$, $(i',j') \in \mathcal{P}$.

3°) Si nous notons \mathcal{P}' , l'ensemble

$$\mathcal{P}' = \{ (i,j) \in \mathbb{N}^2 \mid ((i+1), (j+1)) \in \mathcal{P} \},$$

on a l'inclusion : $[0,1]^2 \cap \mathbb{N}^2 \subseteq \mathcal{P}'$.

Nous supposons donc que pour tout $(i,j) \in \mathcal{P}$, $\theta_{i,j} = 0$.

alors

$$D_2^N(X,Y) < \left[90 \sum_{(i,j) \notin \mathcal{P}'} M_{i,j}^2 \right]^{\frac{1}{2}} = \epsilon(|\mathcal{P}|),$$

$$\text{où } |\mathcal{P}| = \inf_{(i,j) \notin \mathcal{P}} (\min(i,j)),$$

et où ϵ est une fonction vérifiant $\lim_{n \rightarrow \infty} \epsilon(n) = 0$.

L'application au cas particulier où $\mathcal{P}' = [0,q] \cap \mathbb{N}^2$ est immédiate.

2 : Cas des fonctions trigonométriques

Nous trouvons dans ce cas des résultats analogues aux précédents .
Nous supposons toujours que $(X,Y) \in \mathcal{V}^U$.

3-7 : Nous noterons C_i et S_i , les fonctions définies par :

$$C_i(x) = \sqrt{2} \cos(2\pi i x) \quad \text{et} \quad S_i(x) = \sqrt{2} \sin(2\pi i x) , \quad \text{pour } i \in \mathbb{N} .$$

Théorème 3-8 : Si (X,Y) admet une densité de probabilité $f \in \mathcal{L}^2([0,1]^2)$, alors il existe quatre suites doubles réelles :

$$\{a_{i,j}\} , \{b_{i,j}\} , \{c_{i,j}\} \text{ et } \{d_{i,j}\} , \quad (i,j) \in \mathbb{N}^{*2} ,$$

telles que :

$$f = 1 + \sum_{\substack{i=1 \\ j=1}}^{\substack{j=\infty \\ i=\infty}} (a_{i,j} C_i \otimes C_j + b_{i,j} C_i \otimes S_j + c_{i,j} S_i \otimes C_j + d_{i,j} S_i \otimes S_j)$$

$$\text{dans } \mathcal{L}^2([0,1]^2) .$$

Comme précédemment nous généraliserons ces notations en posant :

$$a_{i,j} = E(C_i(X) \cdot C_j(Y)) , \quad b_{i,j} = E(C_i(X) \cdot S_j(Y)) ,$$

$$c_{i,j} = E(S_i(X) \cdot C_j(Y)) , \quad d_{i,j} = E(S_i(X) \cdot S_j(Y)) .$$

et

$$f_q = 1 + \sum_{\substack{1 < i < q \\ 1 < j < q}} (a_{i,j} \cdot C_i \otimes C_j + b_{i,j} \cdot C_i \otimes S_j + c_{i,j} \cdot S_i \otimes C_j + d_{i,j} \cdot S_i \otimes S_j) ,$$

pour tout $q \in \mathbb{N}^*$,

et F_q , la primitive de f_q : $F_q(x,y) = \int_0^x \int_0^y f_q(u,v) \cdot du \cdot dv$.

Lemme 3-9 : En conservant les notations précédentes , et , si

$\{\gamma_{p,p'}\}$ désigne l'une des quatre suites doubles suivantes :

$$\{b_{p',p}\} , \text{ ou } \{c_{p,p'}\} , \text{ ou } \{d_{p,p'}\} \text{ ou } \{d_{p',p}\} ,$$

et si ϵ_q^r désigne le nombre $\epsilon_q^r = \sum_{s=q}^r \frac{1}{s^2}$,

alors on a les inégalités suivantes :

$$\left[\sum_{p=q}^r \frac{\gamma_{p,p'}}{p} \right]^2 < \epsilon_q^r , \text{ et } \left[\sum_{i=q}^r \sum_{j=q'}^{r'} \frac{d_{i,j}}{i \cdot j} \right]^2 < \epsilon_q^r \cdot \epsilon_{q'}^{r'} .$$

Proposition 3-10 : Si nous notons par :

$$\sum_{(i,j) \in \mathbb{N}^2} A_{i,j} \cdot C_i \otimes C_j + \sum_{\substack{j \in \mathbb{N}^* \\ i \in \mathbb{N}}} B_{i,j} \cdot C_i \otimes S_j +$$

$$\sum_{\substack{i \in \mathbb{N}^* \\ j \in \mathbb{N}}} C_{i,j} \cdot S_i \otimes C_j + \sum_{(i,j) \in \mathbb{N}^{*2}} D_{i,j} \cdot S_i \otimes S_j ,$$

le développement de $F - U$ dans $\mathcal{L}^2([0,1]^2)$,

alors on a les égalités suivantes :

pour tout $(i,j) \in \mathbb{N}^{*2}$:

$$A_{i,j} = \frac{d_{i,j}}{4\pi ij} , \quad B_{i,j} = \frac{-c_{i,j}}{4\pi ij} , \quad C_{i,j} = \frac{-b_{i,j}}{4\pi ij} , \quad D_{i,j} = \frac{a_{i,j}}{4\pi ij}$$

pour tout $q \in \mathbb{N}^*$:

$$A_{q,0} = - \left(\sum_{t \in \mathbb{N}^*} \frac{d_{q,t}}{2\sqrt{2} \pi qt} \right) , \quad A_{0,q} = - \left(\sum_{t \in \mathbb{N}^*} \frac{d_{t,q}}{2\sqrt{2} \pi qt} \right)$$

$$B_{0,q} = \sum_{t \in \mathbb{N}^*} \frac{c_{t,q}}{2\sqrt{2} \pi qt} , \quad C_{q,0} = \sum_{t \in \mathbb{N}^*} \frac{b_{q,t}}{2\sqrt{2} \pi qt} ;$$

et enfin $A_{0,0} = \sum_{(t,t') \in \mathbb{N}^{*2}} \frac{d_{t,t'}}{2\pi tt'}$.

Théorème 3-11 : En conservant les notations précédentes ,

on a l'inégalité :

$$\| F - F_q \|_{L^2([0,1]^2)}^2 < \left(\frac{5}{\pi^4} \right) \times \frac{1}{q} .$$

Théorème 3-12 : En conservant les notations et hypothèses du théorème précédent , et si l'on suppose de plus que :

$$a_{i,j} = b_{i,j} = c_{i,j} = d_{i,j} = 0 , \text{ pour } i < q \text{ et } j < q ,$$

alors on a l'inégalité :

$$D_2^N(X,Y) < \left(\frac{3 \sqrt{50}}{\pi^2} \right) \sqrt{\frac{1}{q}} \approx \frac{2,15}{\sqrt{q}} .$$

3 : Cas des fonctions de Walsh

Ces fonctions sont créées à partir des fonctions indicatrices des intervalles de $[0,1]$. Elles permettent d'obtenir des majorations des coefficients D_h^N meilleures que dans les cas précédents. De plus, on a également une majoration lorsque l'on ne suppose plus que les lois marginales sont uniformes, et ce, dans le cas de $[0,1]^p$, $p \in \mathbb{N}^*$, quelconque.

3-13 : Notations : Soient $p \in \mathbb{N}^*$ et $X \in \mathcal{V}_{[0,1]^p}$,

Nous noterons par M^* et M_k^* , $k \in \mathbb{N}$, les sous ensembles de \mathbb{N}^2 définis par :

$$M^* = \{ (i,j) \in \mathbb{N}^2 \mid 0 < j < 2^i \}$$

$$M_k^* = \{ (i,j) \in M^* \mid i < k \}$$

et par M et M_k , les ensembles :

$$M = M^* \cup \{(-\infty, 0)\}, \text{ et } M_k = M_k^* \cup \{(-\infty, 0)\}.$$

Nous rappelons que nous notons, pour tout $(i,j) \in M^*$,

$$I_{2^i, j}, \text{ l'intervalle } I_{2^i, j} = \left[\frac{j-1}{2^i}, \frac{j}{2^i} \right].$$

Alors , pour tout $(i,j) \in M^*$, nous définirons les fonctions réelles

$$B_{2^i,j} \text{ , par : } B_{2^i,j}(x) = \begin{cases} 1 \text{ , si } x \in I_{2^{i+1},2j-1} \\ -1 \text{ , si } x \in I_{2^{i+1},2j} \\ 0 \text{ , sinon .} \end{cases}$$

Nous définirons aussi : $B_{0,0}(x) \equiv 1$,

et par $A_{2^i,j}$, $(i,j) \in M$, les fonctions $B_{2^i,j}$ normalisées :

$$A_{2^i,j} = 2^{-\frac{i}{2}} \cdot B_{2^i,j} \text{ , si } (i,j) \in M^* \text{ , et } A_{0,0} = B_{0,0} .$$

Soit $J \in M^p$: $J = ((i_1,j_1) , (i_2,j_2) , \dots , (i_p,j_p))$,

nous noterons alors par A_J l'application de \mathbb{R}^p dans \mathbb{R} définie par :

$$A_J = A_{2^{i_1},j_1} \otimes A_{2^{i_2},j_2} \otimes \dots \otimes A_{2^{i_p},j_p} .$$

On peut alors énoncer la

Proposition 3-14 : La famille $\{ A_J \}$, $J \in M^p$, est une base orthonormale de $\mathcal{L}^2([0,1]^p)$.

Corollaire 3-15 : Si X admet une densité de probabilité $f \in \mathcal{L}^2([0,1]^p)$, alors il existe une suite $\{\alpha_J\}$, $J \in M^p$,

$$f = \sum_{J \in M^p} \alpha_J \cdot A_J \quad \text{dans } \mathcal{L}^2([0,1]^p) ;$$

et de plus $\alpha_{((-\infty,0), (-\infty,0), \dots, (-\infty,0))} = 1$.

Comme précédemment nous généraliserons ces notations en posant

$$\alpha_J = E(A_J(X)) ,$$

$$f_q = \sum_{J \in (M_q)^p} \alpha_J \cdot A_J ,$$

et $F_q(x_1, x_2, \dots, x_p) = \int_0^{x_1} \int_0^{x_2} \dots \int_0^{x_p} f_q \cdot dt$.

Théorème 3-16 : En conservant les notations précédentes ,

alors , pour tout $q \in \mathbb{N}$, on a l'inégalité :

$$\delta_1(F, F_q) < \frac{1}{2^{q+1}} \left[\sum_{t=0}^{p-1} \left[1 - \frac{1}{2^{q+1}} \right]^t \right].$$

En particulier , si pour tout $J \in (M_q)^p$, $\alpha_J = 0$,

excepté $\alpha_{(-\infty, 0), \dots, (-\infty, 0)}$ toujours égal à 1 , alors

$$\delta_1(F, U) < \frac{1}{2^{q+1}} \left[\sum_{t=0}^{p-1} \left[1 - \frac{1}{2^{q+1}} \right]^t \right].$$

Remarque 3-17 : On peut se demander quelle est la qualité de cette majoration . Pour l'éclairer un peu , nous prendrons l'exemple de la mesure de Dirac en $(0, 0, \dots, 0)$: i.e. $F = 1$ sur $[0, 1]^p$, et l'on trouvera facilement la minoration :

$$\delta_1(F, F_q) > \frac{1}{2} \left[\frac{p}{2^{q+1}} \left[1 - \frac{1}{2^{q+1}} \right]^{p-1} \right].$$

Dans le cas à deux variables nous noterons

$$\alpha_{((i, j), (i', j'))} \quad \text{par} \quad \alpha_{\substack{2^i, j \\ 2^{i'}, j'}}.$$

On peut alors énoncer le

Théorème 3-18 : En conservant les notations précédentes , et en supposant de plus que $(X,Y) \in V^U$, alors on a l'inégalité :

$$\delta(F, F_q) < \frac{1}{4 \cdot 2^q} .$$

En particulier , si pour tout $((i,j), (i',j')) \in (M_q^*)^2$, $\alpha_{\substack{2^i, j \\ 2^{i'}, j'}} = 0$

on obtient : $D^N(X,Y) < \frac{1}{2^q} .$

4 : Démonstrations

Tous les résultats ont été démontrés en [3] , dans le cas où l'on suppose l'existence d'une densité de probabilité $f \in \mathcal{L}^2$.
Nous donnons donc maintenant le complément de cette démonstration .

Lenne 3-19 : Le lemme 3-9 est vérifié même si l'on ne suppose plus l'existence d'une densité de probabilité .

Il suffit de se reporter à la démonstration de 3-9 dans [3] .
On voit alors aisément que comme

$$\int_{x=0}^{x=1} Q(dx , dy) = dy ,$$

cette démonstration reste identique dans le cas où l'on ne suppose plus l'existence de f .

Proposition 3-20 : Soit $(X,Y) \in V^U$, et soit $n \in \mathbb{N}^*$.
Notons par f_n ,

la fonction f sur $\{ B_i^n \}$ (cf 1-4) où $\{ B_i^n \}$ est la
partition de $[0,1]^2$ en 4^{n+1} carrés (cf 3-13)

$$C_{j,j'}^{4^n} = I_{2^{n+1},j} \times I_{2^{n+1},j'} ,$$

et par F_n la fonction primitive $\int_0^x \int_0^y f_n(u,v).du.dv$,

alors $\lim_{n \rightarrow \infty} (\sup_{(x,y) \in [0,1]^2} | F_n(x,y) - F(x,y) |) = 0$.

C'est le théorème 1-4 .

Lemme 3-21 : Les théorèmes 3-4 et 3-10 sont vérifiés même si l'on ne suppose pas l'existence d'une densité de probabilité .

Démonstration : Nous noterons par Q_n pour tout $n \in \mathbb{N}$ (resp Q) , les probabilités sur $[0,1]^2$ associées aux fonctions F_n définies par 3-20 (resp F) .

De plus

$\theta_{i,j}^n, a_{i,j}^n, b_{i,j}^n, c_{i,j}^n, d_{i,j}^n, \beta_{i,j}^n, A_{i,j}^n, B_{i,j}^n, C_{i,j}^n, D_{i,j}^n$,

seront les notations naturelles des

$\theta_{i,j}, a_{i,j}, b_{i,j}, c_{i,j}, d_{i,j}, \beta_{i,j}, A_{i,j}, B_{i,j}, C_{i,j}, D_{i,j}$,

lorsque l'on remplace F par F_n .

Avec ces notations , il est clair que Q_n converge étroitement vers Q , et donc que

$$\iint L_i \otimes L_j dQ_n \xrightarrow{n \rightarrow \infty} \iint L_i \otimes L_j dQ ,$$

$$\text{i.e. } \theta_{i,j}^n \xrightarrow{n \rightarrow \infty} \theta_{i,j} .$$

Et de même

$$\begin{aligned} a_{i,j}^n &\xrightarrow[n \rightarrow \infty]{} a_{i,j} \quad , & b_{i,j}^n &\xrightarrow[n \rightarrow \infty]{} b_{i,j} \quad , \\ c_{i,j}^n &\xrightarrow[n \rightarrow \infty]{} c_{i,j} \quad , & d_{i,j}^n &\xrightarrow[n \rightarrow \infty]{} d_{i,j} \quad . \end{aligned}$$

Comme d'autre part , d'après 3-20 ,

$$\sup_{(x,y) \in [0,1]^2} |F_n(x,y) - F(x,y)| \xrightarrow[n \rightarrow \infty]{} 0 \quad ,$$

et que $\beta_{i,j}^n = \iint (F_n(x,y) - xy) \cdot L_i(x) \cdot L_j(y) \cdot dx \cdot dy \quad ,$

$$\text{alors } \beta_{i,j}^n \xrightarrow[n \rightarrow \infty]{} \beta_{i,j} \quad .$$

On démontre de même que $A_{i,j}^n$ (resp $B_{i,j}^n$, $C_{i,j}^n$, $D_{i,j}^n$)

tend vers $A_{i,j}$ (resp $B_{i,j}$, $C_{i,j}$, $D_{i,j}$) .

Or Q_n admet la densité de probabilité $f_n \in \mathcal{L}^2([0,1]^2)$.
Alors , vu l'égalité de 3-4 pour les $\beta_{i,j}^n$, il est alors clair que
3-4 est aussi vérifiée pour $\beta_{i,j}$.

En ce qui concerne le développement en série de fonctions trigonométriques, la démonstration est identique pour les termes $A_{i,j}$, $B_{i,j}$, $C_{i,j}$ et $D_{i,j}$ qui sont fonctions d'un nombre fini de termes $a_{i',j'}$, $b_{i',j'}$, $c_{i',j'}$ et $d_{i',j'}$.

Nous allons maintenant démontrer que la proposition est vraie pour les termes $A_{i,0}$, par exemple.

$$\text{On sait que } A_{i,0}^n = - \left(\sum_{j=1}^{\infty} \frac{d_{i,j}^n}{\sqrt{2} \cdot ij} \right) \quad (3-10),$$

$$\text{que } A_{i,0}^n \xrightarrow{n \rightarrow \infty} A_{i,0}, \text{ et que } d_{i,j}^n \xrightarrow{n \rightarrow \infty} d_{i,j}.$$

Soit donc $\varepsilon > 0$, et soit $m \in \mathbb{N}^*$, ne dépendant pas de n (m sera défini grâce à 3-9 et 3-20) tel que

$$\sum_{j=m+1}^{\infty} \frac{d_{i,j}^n}{\sqrt{2} \cdot ij} < \frac{\varepsilon}{8},$$

$$\text{et que } \sum_{j=m+1}^{\infty} \frac{d_{i,j}}{\sqrt{2} \cdot ij} < \frac{\varepsilon}{8}.$$

Choisissons alors $n \in \mathbb{N}$ tel que

$$\left| \frac{d_{i,j}^n}{\sqrt{2} \cdot ij} - \frac{d_{i,j}}{\sqrt{2} \cdot ij} \right| < \frac{\varepsilon}{4m}, \text{ lorsque } j < m \text{ (} i \text{ est fixé) ,}$$

$$\text{et } |A_{i,0}^n - A_{i,0}| < \frac{\varepsilon}{2}.$$

$$\text{Alors } \left| \left(\sum_{j=1}^{\infty} \frac{d_{i,j}^n}{\sqrt{2} \cdot ij} \right) - \left(\sum_{j=1}^{\infty} \frac{d_{i,j}}{\sqrt{2} \cdot ij} \right) \right| < \frac{\epsilon}{2} ,$$

$$\text{et donc } \left| A_{i,0} - \sum_{j=1}^{\infty} \frac{d_{i,j}}{\sqrt{2} \cdot ij} \right| < \epsilon .$$

Comme cette inégalité est vraie pour tout $\epsilon > 0$,

$$\text{alors } A_{i,0} = \sum_{j=1}^{\infty} \frac{d_{i,j}}{\sqrt{2} \cdot ij} . \quad \text{C.Q.F.D.}$$

Lemme 3-22 : Les théorèmes 3-5 , 3-11 , 3-12 , 3-16 et 3-18 , ainsi que l'inégalité de 3-6 , sont vérifiés même si l'on ne suppose pas l'existence d'une densité de probabilité .

Démonstration : Il est facile de voir en se reportant aux démonstrations de 3-16 et 3-18 dans [2] et [3] , que les inégalités de ces théorèmes ont été démontrées sans supposer l'existence d'une densité de probabilité .

Dans le cas des polynomes de Legendre , en conservant les notations de la démonstration précédente , on voit clairement que l'inégalité

$$|\beta_{i,j}^n| < M_{i,j} \quad \text{impliquera} \quad |\beta_{i,j}| < M_{i,j} .$$

et que grâce à 3-21 , on obtiendra 3-6 de la même façon que lorsque l'on supposait l'existence de f .

Enfin , pour ce qui concerne les séries de fonctions
trigonométriques , en se reportant à la démonstration de 3-11 et 3-
12 dans [3] , on vérifie sans difficultés majeures grâce aux
limites des $a_{i,j}^n$, $A_{i,j}^n$, etc , que ces propriétés sont
toujours vérifiées .

3 : COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR

Nous avons vu lors de l'étude du cas des polynomes de Legendre (3-3) que si $(X,Y) \in V^U$, $\theta_{1,1} = \rho(X,Y)$, le coefficient de corrélation, et nous nous sommes alors posé la question de savoir si l'on ne pouvait pas considérer la suite $\{ \theta_{i,j} \}$, $(i,j) \in \mathbb{N}^{*2}$, comme une suite de coefficients de corrélation complétant $\rho = \theta_{1,1}$, dont on connaît les insuffisances.

Nous allons maintenant procéder à cette étude dans le cas général et définir les coefficients de corrélation d'ordre supérieur où $\rho(X,Y)$ apparaîtra naturellement comme le premier d'entre eux.

Nous examinerons ensuite le cas du vecteur aléatoire $(F_1(X), F_2(Y))$ qui est de lois marginales uniformes si F_1 et F_2 sont continues.

Et nous étudierons enfin dans un troisième temps quelques propriétés de cas particuliers et notamment le cas gaussien.

3-1 : Etude générale

1 : Notations et Hypothèses

4-1 : Dans cette partie (X, Y) représentera un vecteur aléatoire à valeurs dans \mathbb{R}^2 , de loi Q , et de lois marginales μ' et μ'' .

On supposera qu'il existe $T'_i \in \mathbb{N}^*$, $i = 1, 2$, tels que $T'_i > 1$ et tels que

$$\forall n \in \mathbb{N}, n < T'_1 : \int |x|^{2n} \cdot \mu'(dx) < +\infty,$$

$$\text{et } \forall n \in \mathbb{N}, n < T'_2 : \int |y|^{2n} \cdot \mu''(dy) < +\infty.$$

Nous noterons par n_1 et n_2 la dimension des espaces respectivement : $\mathcal{L}^2(\mathbb{R}, \mu')$ et $\mathcal{L}^2(\mathbb{R}, \mu'')$, que nous supposerons toujours supérieure ou égale à 2.

Et nous noterons par $T_i = \min(T'_i, n_i)$, pour $i = 1, 2$.

Dans ces conditions \mathbb{N}' , \mathbb{N}'' , \mathbb{N}'^* , \mathbb{N}''^* , $\underline{\mathbb{N}}^2$ et $\underline{\mathbb{N}}^{*2}$ désigneront les ensembles suivants :

$$N' = \{ n \in N \mid n < T_1 \} , \quad N'' = \{ n \in N \mid n < T_2 \} ,$$

$$N'^{*} = N' \setminus \{0\} , \quad N''^{*} = N'' \setminus \{0\} ,$$

$$\underline{N}^2 = N' \times N'' \quad \text{et} \quad \underline{N}^{*2} = N'^{*} \times N''^{*} .$$

2 : Propriétés générales

Nous rappelons et démontrons maintenant quelques propriétés qui servent de fondement à la suite de notre étude .

Lemme 4-2 : Soit $p \in \mathbb{N}^*$, et soit m une probabilité sur (R, \mathcal{B}_1) , alors les deux affirmations suivantes sont équivalentes .

1) : $\dim(\mathcal{L}^2(R, m)) = p .$

2) : Il existe $\mathcal{X} \subset R$ tel que $\text{card}(\mathcal{X}) = p$:

$$\mathcal{X} = (x_1, x_2, \dots, x_p) \text{ vérifiant}$$
$$m(\{x_i\}) = \lambda_i > 0 , \text{ pour } i = 1, 2, \dots, p \text{ et } \sum_{i=1}^p \lambda_i = 1 .$$

Démonstration : Supposons que 2) soit vérifiée .

Alors $\{ 1_{\{x_i\}} \}$, $i = 1, 2, \dots, p$,

est une famille linéairement indépendante de $\mathcal{L}^2(\mathbb{R}, m)$.

Soit donc $f \in \mathcal{L}^2(\mathbb{R}, m)$, alors

$$f = \sum_{i=1}^p f(x_i) \cdot 1_{\{x_i\}} \quad \text{dans } \mathcal{L}^2(\mathbb{R}, m) ,$$

et donc $\{ 1_{\{x_i\}} \}$, $i = 1, 2, \dots, p$, est une base de $\mathcal{L}^2(\mathbb{R}, m)$.

Supposons que $\dim(\mathcal{L}^2(\mathbb{R}, m)) = p$.

Soit q le nombre maximal de boréliens B'_i disjoints tels que $m(B'_i) > 0$.

Alors il est clair que $q \leq p$.

Soit donc $\{ B'_i \}$, $i = 1, 2, \dots, q$, une famille de boréliens disjoints vérifiant $m(B'_i) > 0$; on voit aisément que les B'_i sont des atomes .

Alors il existe un compact K tel que
 $m(B'_i \cap \bigcap K) = 0$ pour $i = 1, 2, \dots, q$.

En effet, il existe $n_i \in \mathbb{N}$ tel que $m(B'_i \cap [-n_i, +n_i]) > 0$,

sinon $m(B'_i) = \lim_{n \rightarrow \infty} m(B'_i \cap [-n, +n]) = 0$.

Donc comme B'_i est un atome, $m(B'_i \cap \bigcap [-n_i, +n_i]) = 0$.

Il suffit alors de prendre $K = \bigcap_{i=1}^q [-n_i, +n_i]$.

Nous noterons alors par $B_i = B'_i \cap K$, et nous définirons les suites $\{I_n^i\}$, $n \in \mathbb{N}$, d'intervalles emboîtés par :

$$I_0^i = K,$$

$$I_n^i = I_{n+1}^i \cup I'_{n+1}^i, \text{ où } m(I_{n+1}^i \cap B_i) > 0 \text{ et } L(I_{n+1}^i) = L(I'_{n+1}^i) = \frac{L(I_n^i)}{2}$$

- $L(I)$ représentant ici la longueur de l'intervalle I .

Alors il est clair que $m(I_{n+1}^i \cap B_i) = 0$

et que $\lim_{n \rightarrow \infty} m(I_n^i \cap B_i) = m(B_i)$,

et comme $\bigcap_{n=0}^{\infty} (I_n^i \cap B_i) = \{x_i\}$,

on en conclut que m sera de la forme $m = \sum_{i=1}^q \lambda_i \cdot 1_{\{x_i\}}$;

et d'après la première partie de la démonstration $p = q$. C.Q.F.D.

Lemme 4-3 : Soit $t \in \mathbb{N}^*$, $t < T_1$, alors la famille $\{x^i\}$, $i = 0, 1, \dots, t$, $x^i \in \mathbb{R}[X]$, est une famille linéairement indépendante de $\mathcal{L}^2(\mathbb{R}, \mu')$.

Démonstration : Soit donc λ_i , $i = 0, 1, \dots, t$, $\lambda_i \in \mathbb{R}$,

tels que $\sum_{i=0}^t \lambda_i \cdot x^i = 0$ dans $\mathcal{L}^2(\mathbb{R}, \mu')$,

et soit $H \in \mathbb{R}[X]$, défini par $H(x) = \sum_{i=0}^t \lambda_i \cdot x^i$.

Notons par A_0 l'ensemble des zéros de H
et supposons $\text{card}(A_0) \ll t$.

Alors il existe $B' \in \mathcal{B}_1$ tel que

$B' \cap A_0 = \emptyset$ et $\mu'(B') > 0$.

En effet, supposons l'inverse.

Soit $B \in \mathcal{B}_1$ tel que $\mu'(B) > 0$ et $B \cap A_0 = \{x_0\}$,

et soit $B_n = B \cap \left[x_0 - \frac{1}{n}, x_0 + \frac{1}{n} \right]$,

alors $\mu' \left(B \cap \bigcap_{B} B_n \right) = 0$, car $\left(B \cap \bigcap_{B} B_n \right) \cap A_0 = \emptyset$,

et donc $\mu'(B) = \lim_{n \rightarrow \infty} \mu'(B_n) = \mu'(\{x_0\})$.

Finalement on peut écrire μ' sous la forme $\mu' = \sum_{i=1}^t a_i \cdot 1_{\{x_i\}}$,

$x_i \in A_0$ pour $i = 1, 2, \dots, t$;

et donc d'après 4-2 $\dim(\chi^2(\mathbb{R}, \mu')) < t$,

ce qui est une contradiction.

Soit donc $B' \in \mathcal{B}_1$ tel que $B' \cap A_0 = \emptyset$ et $\mu'(B') > 0$,

et soient $B_1 = B' \cap H^{-1}(R_+)$ et $B_2 = B' \cap H^{-1}(R_-)$,

alors $B_i \in \mathcal{B}_1$ pour $i = 1, 2$,

et $\mu'(B_1) > 0$, par exemple,

et donc $\int_{B_1} H(x) \cdot \mu'(dx) > 0$,
ce qui est en contradiction avec $H = 0$ dans $\mathcal{L}^2(\mathbb{R}, \mu')$.

On en conclut donc que $\text{card}(A_0) > t$.

Alors comme $\text{deg}(H) = t$ et que H s'annule en $t+1$ points

$H \equiv 0$ et donc $\lambda_i = 0$ pour $i = 0, 1, \dots, t$,

ce qui prouve bien que la famille $\{ X^i \}$, $i = 0, 1, \dots, t$,
est linéairement indépendante .

Lemme 4-4 : Soit m une probabilité sur $(\mathbb{R}, \mathcal{B}_1)$ et soit
 $T \in \mathbb{N}^* \cup \{+\infty\}$, vérifiant $\dim(\mathcal{L}^2(\mathbb{R}, m)) \gg T$, et tel que

$\forall t < T$, $t \in \mathbb{N}$, $\int |x^{2t}| \cdot m(dx) < +\infty$;

alors il existe une famille orthonormale de $\mathcal{L}^2(\mathbb{R}, m)$

$\{ \Pi_i^m \}$, $i \in \mathbb{N}$, $i < T$, $\Pi_i^m \in \mathbb{R}[X]$,

telle que $\text{deg}(\Pi_i^m) = i$;

cette famille est définie par récurrence par

$$\pi_0^m \equiv 1 ,$$

$$\pi_i^m(x) = \frac{x^i - \sum_{s=0}^{i-1} \langle x^i, \pi_s^m \rangle \cdot \pi_s^m(x)}{\sqrt{\langle x^i, x^i \rangle - \sum_{s=0}^{i-1} \langle x^i, \pi_s^m \rangle^2}}$$

Ce lemme se déduit immédiatement de 4-3 et du procédé d'orthogonalisation de Gram-Schmidt .

Proposition 4-5 : En notant alors par

$$\{P_i^{\prime}\} , i \in N' \quad \text{et} \quad \{P_j^{\prime\prime}\} , j \in N'' ,$$

les familles de polynomes orthonormaux définies en 4-4 de

$$\mathcal{L}^2(\mathbb{R}, \mu^{\prime}) \quad \text{et} \quad \mathcal{L}^2(\mathbb{R}, \mu^{\prime\prime}) ,$$

on peut dire que la famille $\{P_i^{\prime} \otimes P_j^{\prime\prime}\} , (i,j) \in \mathbb{N}^2 ,$

est une famille orthonormale de $\mathcal{L}^2(\mathbb{R}^2, \mu^{\prime} \otimes \mu^{\prime\prime})$.

Lemme 4-6 : Pour tout $q' \in \mathbb{N}$, notons par $\mathcal{M}_{q'}$, l'ensemble des probabilités m sur $(\mathbb{R}, \mathcal{B}_1)$ telles que

$$\dim(\mathcal{X}^2(\mathbb{R}, m)) > q' \quad \text{et} \quad \int |x|^{2q'} \cdot m(dx) < +\infty .$$

Alors pour tout $q \in \mathbb{N}$ et tout $m \in \mathcal{M}_q$ et tout $(i, j) \in \mathbb{N}^2$, $i \leq q$ et $j \leq 2q$, nous noterons Π_i^m le polynome associé à m selon 4-4 par

$$\Pi_i^m(x) = \sum_{t=0}^i a_{i,t}^m \cdot x^t ,$$

et le moment d'ordre j par $M_j^m = \int x^j \cdot m(dx)$.

Alors il existe une suite d'applications $\{g_{r,s}\}$, $(r,s) \in \mathbb{N}^2$, $s \leq r$ de \mathbb{R}^{2r} dans \mathbb{R} , telle que

$$\forall q \in \mathbb{N}, \forall m \in \mathcal{M}_q, \forall (i,t) \in \mathbb{N}^2 : t < i < q ,$$

$$\text{on ait} \quad a_{i,t}^m = g_{i,t}(M_0^m, M_1^m, M_2^m, \dots, M_{2i}^m) .$$

De plus $g_{i,t}$ est continue en tout point de la forme

$$(M_0^m, M_1^m, M_2^m, \dots, M_{2i}^m) .$$

La démonstration de ce lemme est presque évidente . Il suffit pour s'en convaincre d'écrire les premiers polynomes Π_i^m , $i = 0, 1, 2, \dots$:

elle découle directement de 4-4 .

La continuité provient de ce que les dénominateurs sont toujours différents de zéro, car ils représentent la distance de X^i au sous-espace engendré par les $\{X^{i'}\}$, $i' < i$, et que $1, X, \dots$ et X^i sont linéairement indépendants d'après 4-3.

En simplifiant $M_i^{i'}$ en M_i , on trouve en particulier :

$$P_0' \equiv 1, \quad P_1'(x) = \frac{x - M_1}{\sqrt{M_2 - M_1^2}} = \frac{x - \bar{X}}{\sigma_X}$$

$$P_2'(x) = \frac{x^2 - \frac{(M_3 - M_1 \cdot M_2) \cdot (x - M_1)}{M_2 - M_1^2} - M_2}{\sqrt{M_4 - \frac{(M_3 - M_1 \cdot M_2)^2}{M_2 - M_1^2} - M_2^2}}$$

3 : Définition et propriétés des coefficients de corrélation d'ordre supérieur

Nous allons voir que , comme on pouvait le souhaiter , l'indépendance de X et Y entraîne la nullité des coefficients de corrélation , et que sous certaines hypothèses supplémentaires simples, la réciproque est également vraie .

Définition 4-8 : On rappelle que l'on note $\{P_i^I\}$, $i \in \mathbb{N}^I$, et $\{P_j^{II}\}$, $j \in \mathbb{N}^{II}$, les familles orthonormales de polynomes de $\mathcal{L}^2(\mathbb{R}, \mu^I)$ et $\mathcal{L}^2(\mathbb{R}, \mu^{II})$, (4-5) ,

Soit $(i,j) \in \mathbb{N}^{*2}$, on appelle coefficient de corrélation d'ordre (i,j) de X et Y le nombre réel $\rho_{i,j} = E(P_i^I(X) \cdot P_j^{II}(Y))$.

Alors les propriétés des polynomes orthonormaux permettent d'énoncer immédiatement la

Proposition 4-9 : Le coefficient de corrélation d'ordre (i,j) est le coefficient de corrélation habituel de $P_i^I(X)$ et $P_j^{II}(Y)$.

La formule de P'_1 donnée précédemment permet aussi d'énoncer le

Théorème 4-10 : Le coefficient de corrélation d'ordre (1,1) est le coefficient de corrélation habituel : i.e. $\rho_{1,1} = \rho(X,Y)$.

Théorème 4-11 : Si (X,Y) admet une densité de probabilité par rapport à $\mu' \otimes \mu''$: $f \in \mathcal{L}^2(\mathbb{R}^2, \mu' \otimes \mu'')$, et si $\{P'_i\}$, $i \in \mathbb{N}'$ et $\{P''_j\}$, $j \in \mathbb{N}''$, sont des bases orthonormales respectivement de $\mathcal{L}^2(\mathbb{R}, \mu')$ et $\mathcal{L}^2(\mathbb{R}, \mu'')$, alors

$$f = 1 + \sum_{(i,j) \in \mathbb{N}^{*2}} \rho_{i,j} \cdot P'_i \otimes P''_j \quad \text{dans } \mathcal{L}^2(\mathbb{R}^2, \mu' \otimes \mu'') .$$

Il suffit de remarquer que

$$\mathbb{E}(P'_i(X) \cdot P''_j(Y)) = \iint P'_i(x) \cdot P''_j(y) \cdot f(x,y) \cdot \mu'(dx) \cdot \mu''(dy) ,$$

et d'appliquer 2-4 .

Théorème 4-12 : Si X et Y sont indépendantes , alors pour tout $(i,j) \in \mathbb{N}^{*2}$, $\rho_{i,j} = 0$.

Avec 4-9 c'est évident .

Le théorème suivant est un corollaire de 4-11 .

Théorème 4-13 : Avec les hypothèses de 4-11 et si
pour tout $(i,j) \in \mathbb{N}^{*2}$, $\rho_{i,j} = 0$,
alors X et Y sont indépendantes .

Nous allons voir que si $\dim(\mathcal{L}^2(\mathbb{R}, \mu')) < +\infty$, nous pouvons nous ramener à ce cas : c'est le

Théorème 4-14 : Supposons que $\dim(\mathcal{L}^2(\mathbb{R}, \mu')) = p < +\infty$
alors (X,Y) admet une densité de probabilité f par rapport à
 $\mu' \otimes \mu''$: $f \in \mathcal{L}^2(\mathbb{R}^2, \mu' \otimes \mu'')$.

Démonstration : Si $\dim(\mathcal{L}^2(\mathbb{R}, \mu')) = p$, alors il existe $\mathcal{X} \subset \mathbb{R}$: $\mathcal{X} = (x_1, x_2, \dots, x_p)$ tel que

$$\mu(\{x_i\}) = \lambda_i > 0 \quad \text{et} \quad \sum_{i=1}^p \lambda_i = 1 \quad (4-2) .$$

Et donc Q et $\mu' \otimes \mu''$ sont concentrées sur $\mathcal{X} \times \mathbb{R}$.

Soit $\eta \in \mathbb{R}_+^*$ et $B \in \mathcal{B}_2$ tels que $\mu' \otimes \mu''(B) < \eta$.

Définissons alors B_i et B'_i par

$$B_i = B \cap (\{x_i\} \times \mathbb{R}) = \{x_i\} \times B'_i .$$

Alors comme

$$\mu' \times \mu''(B) = \sum_{i=1}^p \mu' \times \mu''(B_i) = \sum_{i=1}^p \lambda_i \cdot \mu''(B'_i) ,$$

on peut écrire $\mu''(B_i^c) < \frac{\eta}{\inf_{i=1,2,\dots,p} (\lambda_i)}$,

et donc

$$Q(B) = \sum_{i=1}^p Q(B_i) < \sum_{i=1}^p \mu''(B_i^c) < \frac{p \cdot \eta}{\inf(\lambda_i)} .$$

Cette dernière inégalité prouve que Q est absolument continue par rapport à $\mu' \otimes \mu''$, et donc il existe f mesurable telle que

$$Q(B) = \int_B f \cdot d(\mu' \otimes \mu'') \text{ pour tout } B \in \mathcal{B}_2 .$$

Il reste alors à prouver que $f \in \mathcal{L}^2$.

$$\text{On sait que } \int f \cdot d\mu' = \sum_{i=1}^p \lambda_i \cdot f(x_i, y) = 1$$

(étant donné que Q est de loi marginale μ') .

$$\text{donc } f(x_i, y) < \frac{1}{\inf(\lambda_i)} .$$

et donc f est borné, et donc $f \in \mathcal{L}^2(\mathbb{R}^2, \mu' \otimes \mu'')$. C.Q.F.D.

Théorème 4-15 : soit K un compact de \mathbb{R} ,
et supposons que $\mu' \otimes \mu''$ soit concentrée sur K^2 ,

alors si pour tout $(i, j) \in \mathbb{N}^2$, $\rho_{i,j} = 0$,

X et Y sont indépendantes .

Démonstration : il est clair que dans ce cas $T_1' = T_2' = +\infty$.

Supposons T_1 fini (par exemple) .

Alors , d'après 4-3 , la famille

$$\{ X^i \} , \quad i = 0, 1, \dots, n_1 - 1 , \quad X^i \in \mathbb{R}[X] ,$$

est une famille linéairement indépendante de $\mathcal{L}^2(\mathbb{R}, \mu')$.

Il en est donc de même de la famille $\{ P_i^1 \} , i = 0, 1, \dots, n_1 - 1$.

Il est donc clair que cette famille est une base de $\mathcal{L}^2(\mathbb{R}, \mu')$.

Si T_2 est fini $\{ P_j^2 \} , j = 0, 1, \dots, n_2 - 1$,

sera également une base de $\mathcal{L}^2(\mathbb{R}, \mu'')$.

Sinon le théorème de Stone-Weierstrass permet d'obtenir le même résultat .

Il suffit alors d'appliquer 4-14 et 4-13 pour achever la démonstration .

Supposons alors $T_1 = T_2 = +\infty$.

Montrons d'abord que

$$\iint P_i^1 \otimes P_j^2 . dQ = 0 \quad \text{si } (i, j) \neq (0, 0) \quad \text{et } 1 \quad \text{sinon} .$$

Il reste à prouver le cas des $(i, 0)$ et des $(0, j)$.

Il se démontre en remarquant que

$$\iint P_0^1 \otimes P_j^2 . dQ = \iint 1 \otimes P_j^2 . dQ = \int P_j^2 . d\mu'' = 0 \quad \text{si } j \neq 0$$

et 1 sinon .

Soit $g \in \mathcal{C}(K^2)$.

Nous allons prouver que $\iint g \cdot dQ = \iint g \cdot d(\mu' \otimes \mu'')$.

D'après Stone-Weierstrass , on sait qu'il existe une suite

$$\{g_p\}, p \in \mathbb{N}, g_p \in \mathbb{R}[X, Y], \text{ vérifiant}$$
$$\lim_{p \rightarrow \infty} \left(\sup_{(x,y) \in K^2} |g_p(x,y) - g(x,y)| \right) = 0 .$$

$$\text{Ecrivons } g_p = \sum_{\substack{(i,j) \in \mathbb{N}^2 \\ i < s(p) \text{ et } j < s(p)}} \lambda_{i,j}^p \cdot P_i' \otimes P_j''$$

avec $s(i') \in \mathbb{N}^*$.

Alors $\iint_{K^2} g_p \cdot dQ = \lambda_{0,0}^p$, d'après ce qui précède

$$\text{et } \iint_{K^2} g_p \cdot d(\mu' \otimes \mu'') = \lambda_{0,0}^p .$$

Or la convergence uniforme de g_p vers g implique la convergence de $\lambda_{0,0}^p$ vers un $\lambda_{0,0} \in \mathbb{R}$, ainsi que les égalités

$$\iint \lim_{p \rightarrow \infty} g_p \cdot dQ = \lim_{p \rightarrow \infty} \iint g_p \cdot dQ \text{ et}$$

$$\iint \lim_{p \rightarrow \infty} g_p \cdot d(\mu' \otimes \mu'') = \lim_{p \rightarrow \infty} \iint g_p \cdot d(\mu' \otimes \mu'') ,$$

$$\text{et donc } \iint g \cdot dQ = \iint g \cdot d(\mu' \otimes \mu'') .$$

Soit alors R° un rectangle compact de K^2 .

On sait qu'il existe une suite de fonctions continues

$\{h_n\}$, $n \in \mathbb{N}$, $h_n < 1$, telle que
 $h_n = 1$ sur R° et $h_n = 0$ sur l'ensemble des points de \mathbb{R}^2 de distance à R° supérieure ou égale à $\frac{1}{n}$.

$$\text{Et donc } \lim_{n \rightarrow \infty} \iint h_n \cdot dQ = Q(R^\circ)$$

$$\text{et } \lim_{n \rightarrow \infty} \iint h_n \cdot d(\mu' \otimes \mu'') = \mu' \otimes \mu''(R^\circ),$$

et, vu le résultat ci-dessus, on trouve donc $\mu' \otimes \mu''(R^\circ) = Q(R^\circ)$.

On en conclut que $\mu' \otimes \mu'' = Q$. C.Q.F.D.

Remarque 4-16 : Citons deux cas où la condition

$$" \forall (i,j) \in \mathbb{N}^{*2} : \rho_{i,j} = 0 "$$

n'implique pas l'indépendance de X et Y :

1°) Le cas où $\dim(\mathcal{L}^2(\mathbb{R}, \mu')) > T_1$:
ainsi lorsque μ' suit une loi de Cauchy de densité $\frac{1}{\pi \cdot (1+x^2)}$.

2°) Le cas où $T_1 = +\infty$ et où la famille $\{P_i\}$, $i \in \mathbb{N}$, n'est pas une base de $\mathcal{L}^2(\mathbb{R}, \mu')$ (cf 17-5, [8]).

Théorème 4-17 : Soit $(i,j) \in \mathbb{N}^{*2}$, alors $|\rho_{i,j}| < 1$,

l'égalité ayant lieu si et seulement si

$$P_i'(X) = P_j''(Y) \text{ p.s.} \quad \text{ou} \quad P_i'(X) = -P_j''(Y) \text{ p.s.}$$

Avec 4-9, c'est évident.

4-18 : Remarquons enfin que la liberté de choix des $\rho_{i,j}$ est relativement large, et que (X,Y) dépend d'une infinité dénombrable de paramètres "presque" indépendants les uns des autres (dans les cas où $\mathbb{N} = \mathbb{N}'$ ou \mathbb{N}'' bien sur!).

En effet, toute suite $\{\rho_{i,j}\}, (i,j) \in \mathbb{N}^{*2}$, telle que

$$\sum_{(i,j) \in \mathbb{N}^{*2}} \rho_{i,j}^2 < +\infty,$$

$$\text{et que } g = 1 + \sum_{(i,j) \in \mathbb{N}^{*2}} \rho_{i,j} \cdot P_i' \otimes P_j'' > 0,$$

définit une fonction $f = g$, $f \in \mathcal{L}^2(\mathbb{R}, \mu' \otimes \mu'')$, qui est une densité par rapport à $\mu' \otimes \mu''$ d'une probabilité Q de lois marginales μ' et μ'' .

La proposition suivante permet de préciser un peu cette liberté de choix des $\rho_{i,j}$.

Proposition 4-19 : Soit $n \in \mathbb{N}^*$ tel que $2n < \inf(T_1, T_2)$, alors il existe un hypercube C de \mathbb{R}^{n^2} de mesure borélienne non nulle tel que

$$\forall r \in C : r = (r_{2i,2j}) , (i,j) \in (1,2,\dots,n)^2 ,$$

on ait

$$1 + \sum_{\substack{i=1 \\ j=1}}^{\substack{j=n \\ i=n}} r_{2i,2j} \cdot P'_{2i} \otimes P''_{2j} > 0 .$$

Démonstration : Il est facile de voir qu'il existe un compact K^2 , $K^2 \subset \mathbb{R}^2$ tel que

$$\forall (i,j) \in (1,2,\dots,n)^2 \text{ et } \forall (x,y) \in \mathbb{R}^2 \cap K^2$$

on ait $P'_{2i}(x) \cdot P''_{2j}(y) \geq 0$.

Soit alors $L \in \mathbb{R}_+^*$, défini par

$$\frac{1}{L} = n^2 \cdot \left(\sup_{\substack{(x,y) \in K^2 \\ (i,j) \in (1,2,\dots,n)^2}} |P'_{2i}(x) \cdot P''_{2j}(y)| \right) .$$

Alors si $r_{2i,2j} \ll L$, pour tout $(i,j) \in (1,2,\dots,n)^2$,

$$\sum_{\substack{i=1 \\ j=1}}^{\substack{j=n \\ i=n}} r_{2i,2j} \cdot P'_{2i}(x) \cdot P''_{2j}(y) < 1 \text{ pour tout } (x,y) \in K^2 .$$

On en conclut facilement que

$$\forall (x,y) \in \mathbb{R}^2 : \left(1 + \sum_{\substack{i=1 \\ j=1}}^{i=n \\ j=n} r_{2i,2j} \cdot P'_{2i} \otimes P''_{2j} \right) > 0 .$$

C.Q.F.D.

Remarque 4-20 : Nous venons d'affirmer que la liberté de choix des $\rho_{i,j}$ était relativement large, mais il est clair que le choix de certains coefficients peut en déterminer d'autres : ainsi si $\rho_{1,1} = \rho(X,Y) = 1$ (remarquons que dans ce cas, il n'y a pas de densité de probabilité) .

4 : Coefficient de corrélation empirique d'ordre (i,j)

Comme on pouvait le souhaiter, nous allons trouver des estimateurs $\rho_{i,j}^n$ convergents presque sûrement et tels que $\rho_{1,1}^n$ soit le coefficient de corrélation empirique habituel.

4-21 : Notations : Soit (X_p, Y_p) , $p = 1, 2, \dots, n$, un n-échantillon du vecteur aléatoire (X, Y) .

Nous noterons par n' et n'' les entiers :
 $n' = \text{card}(X_p \mid p = 1, 2, \dots, n)$ et $n'' = \text{card}(Y_p \mid p = 1, 2, \dots, n)$,

et par μ_n' et μ_n'' les probabilités empiriques associées :

$$\mu_n'(\{x\}) = \frac{\text{card}(X_p = x \mid p=1, 2, \dots, n)}{n}$$

Enfin soient $\{P_i^{n'}\}$, $i = 0, 1, \dots, n'-1$, et $\{P_j^{n''}\}$, $j = 0, 1, \dots, n''-1$, les polynômes orthonormaux associés aux mesures,

$$\mu_n' \text{ et } \mu_n'' \quad (4-4)$$

4-22 : Définition : Soit $(i, j) \in \mathbb{N}^{*2}$, tel que $i < n'$ et $j < n''$, nous appellerons coefficient de corrélation empirique d'ordre (i, j) - associé au n-échantillon (X_p, Y_p) , $p = 1, 2, \dots, n$ - le nombre réel $\rho_{i,j}^n$ défini par :

$$\rho_{i,j}^n = \frac{1}{n} \cdot \left(\sum_{p=1}^n P_i^{n'}(X_p) \cdot P_j^{n''}(Y_p) \right)$$

Théorème 4-23 : Soit $(i,j) \in \mathbb{N}^{*2}$,

alors $\rho_{i,j}^n$ converge presque sûrement vers $\rho_{i,j}$.

Démonstration : Rappelons d'abord que $P(n' < N_0) \xrightarrow{n \rightarrow \infty} 0$,
pour tout $N_0 \in \mathbb{N}$, $N_0 < n_1$.

Et notons

$$P_i'(x) = \sum_{s=0}^i A_{i,s} \cdot x^s, \quad P_i'^n(x) = \sum_{s=0}^i a_{i,s}^n \cdot x^s,$$

$$P_j''(y) = \sum_{t=0}^j B_{j,t} \cdot y^t, \quad P_j''^n(y) = \sum_{t=0}^j b_{j,t}^n \cdot y^t,$$

$$M_i' = \int x^i \cdot \mu'(dx), \quad m_i'^n = \int x^i \cdot \mu_n'(dx),$$

$$M_j'' = \int y^j \cdot \mu''(dy), \quad m_j''^n = \int y^j \cdot \mu_n''(dy),$$

alors d'après 4-6,

$$a_{i,s}^n = g_{i,s}(m_0'^n, m_1'^n, \dots, m_{2i}'^n),$$

$$A_{i,s} = g_{i,s}(M_0', M_1', \dots, M_{2i}')$$

Or $m_i'^n$ converge presque sûrement vers M_i' (Th-8-7, [9]), et donc, vu la continuité (4-6), $a_{i,s}^n$ converge presque sûrement vers $A_{i,s}$.

Comme de plus, on a :

$$\rho_{i,j}^n = \sum_{s=0}^i \sum_{t=0}^j a_{i,s}^n \cdot b_{j,t}^n \cdot \left(\sum_{p=1}^n \frac{(X_p)^s \cdot (Y_p)^t}{n} \right) ,$$

$$\text{et } \rho_{i,j} = \sum_{s=0}^i \sum_{t=0}^j A_{i,s} \cdot B_{j,t} \cdot E(X^s \cdot Y^t)$$

on trouve bien que $\rho_{i,j}^n$ converge presque surement vers $\rho_{i,j}$.

Théorème 4-24 : L'estimateur $\rho_{1,1}^n$ de $\rho_{1,1} = \rho(X,Y)$ est le coefficient de corrélation empirique habituel .

Théorème 4-25 : On a l'inégalité : $|\rho_{i,j}^n| < 1$.

Pour 4-22, il suffit d'écrire $\rho_{1,1}^n$. Pour 4-23, il suffit d'appliquer 4-17 à μ_n' et μ_n'' .

3 - 2 : Cas de (F₁(X) , F₂(Y))

5-1 Notations : Dans cette partie nous supposons F₁ et F₂ continues . Dans ces conditions , le vecteur aléatoire (F₁(X),F₂(Y)) est à valeurs dans [0,1]² et de lois marginales uniformes . Nous sommes alors ramenés au cas des polynomes de Legendre , et nous allons étudier les estimateurs des

$$\theta_{i,j} = E(L_i(F_1(X)) \cdot L_j(F_2(Y)))$$

définis pour tout (i,j) ∈ IN^{*2} (cf partie 2-1) .

Ce cas est plus pratique que celui des coefficients de corrélation d'ordre supérieur , car , d'une part , nous connaissons les polynomes orthonormaux , d'autre part , nous avons l'équivalence logique :

$$" \forall (i,j) \in \mathbb{N}^{*2} , \theta_{i,j} = 0 " \iff " X \text{ et } Y \text{ sont indépendantes } "$$

et surtout nous avons l'implication :

$$" \forall (i,j) \in \mathbb{N}^{*2} \cap [1,q]^2 , \theta_{i,j} = 0 " \implies " D_2(F_1(X) \cdot F_2(Y)) < \sqrt{\sum_{\substack{i>q \\ \text{ou } j>q}} M_{i,j}^2} = \epsilon(q) " .$$

Nous supposons toujours que (X_p, Y_p) , p = 1,2,...,n , est un n-échantillon de (X,Y) et nous noterons par F₁ⁿ et F₂ⁿ les fonctions de répartition empiriques marginales .

Comme F₁ et F₂ sont continues nous supposons de plus que X_p ≠ X_{p'} , si p ≠ p' et Y_t ≠ Y_{t'} , si t ≠ t' .

Nous noterons alors par (X_p°) (resp (Y_p°)) , $p = 1, 2, \dots, n$,
 l'échantillon (X_p) (resp (Y_p)) ordonné , et par σ
 la permutation de $(1, 2, \dots, n)$ telle que pour tout $p' \in (1, 2, \dots, n)$
 il existe $\epsilon p \in (1, 2, \dots, n)$: $(X_{p'}^\circ, Y_{\sigma(p')}^\circ) = (X_p, Y_p)$.

Alors , avec ces notations , on peut énoncer le

Théorème 5-2 : Si nous posons

$$\theta_{i,j}^n = \frac{1}{n} \cdot \left(\sum_{p=1}^n L_i\left(\frac{p}{n}\right) \cdot L_j\left(\frac{\sigma(p)}{n}\right) \right) ,$$

alors $\theta_{i,j}^n$ est un estimateur convergent presque sûrement de $\theta_{i,j}$.

Démonstration : On sait que $\frac{p}{n} = F_1^n(X_p^\circ)$.

Alors, comme d'après Glivenko-Cantelli

$$\left(\sup_{p=1, 2, \dots, n} \left| F_1^n(X_p^\circ) - F_1(X_p^\circ) \right| \right) \xrightarrow{\text{p.s.}} 0 ,$$

$$\text{on a } \left(\sup_{p=1, 2, \dots, n} \left| L_i\left(\frac{p}{n}\right) - L_i(F_1(X_p^\circ)) \right| \right) \xrightarrow{\text{p.s.}} 0 ,$$

et donc ,

$$\left[\sup_{p=1, 2, \dots, n} \left| L_i\left(\frac{p}{n}\right) \cdot L_j\left(\frac{\sigma(p)}{n}\right) - L_i(F_1(X_p^\circ)) \cdot L_j(F_2(Y_{\sigma(p)}^\circ)) \right| \right] \xrightarrow{\text{p.s.}} 0 .$$

Et finalement

$$\left[\sum_{p=1}^n \frac{L_i\left(\frac{p}{n}\right) \cdot L_j\left(\frac{\sigma(p)}{n}\right)}{n} - \sum_{p=1}^n \frac{L_i(F_1(X_p^\circ)) \cdot L_j(F_2(Y_{\sigma(p)}^\circ))}{n} \right] \xrightarrow{\text{p.s.}} 0 .$$

Il suffit alors d'écrire que

$$\sum_{p=1}^n \frac{L_i(F_1(X_p^\circ)) \cdot L_j(F_2(Y_{\sigma(p)}^\circ))}{n} = \sum_{q=1}^n \frac{L_i(F_1(X_q)) \cdot L_j(F_2(Y_q))}{n}$$

qui converge presque sûrement vers $\theta_{i,j}$ (7-12) pour achever notre démonstration .

Théorème 5-3 : Les deux propositions suivantes sont équivalentes :

- 1°) X et Y sont indépendantes
- 2°) $\forall (i,j) \in \mathbb{N}^{*2}$, $\theta_{i,j} = 0$.

Ce théorème est une conséquence directe de 4-15 et de la

Proposition 5-4 : On a l'équivalence logique

$F_1(X)$ et $F_2(Y)$ sont indépendantes \iff X et Y sont indépendantes

Démonstration : Il reste à prouver l'implication directe .

Soit donc $[a,b] \times [c,d]$ un rectangle de \mathbb{R}^2 tel que
 $[a,b] \subset \text{supp}(\mu')$ et $[c,d] \subset \text{supp}(\mu'')$.

Posons $x' = \inf(x \in \mathbb{R} \mid F_1(x) = F_1(a))$ et
 $x'' = \sup(x \in \mathbb{R} \mid F_1(x) = F_1(b))$ et soient y' et y''
 définis de la même façon.

Notons aussi $F_1([a,b]) = [\alpha,\beta]$ et $F_2([c,d]) = [\gamma,\delta]$.

Alors $P(X \in [x',x'']) = P(X \in [a,b]) = P(F_1(X) \in [\alpha,\beta])$,

et $P((X,Y) \in [x',x''] \times [y',y'']) = P((X,Y) \in [a,b] \times [c,d])$
 $= P((F_1(X), F_2(Y)) \in [\alpha,\beta] \times [\gamma,\delta])$.

On en conclut aisément que si $F_1(X)$ et $F_2(Y)$ sont indépendantes,

alors $P((X,Y) \in [a,b] \times [c,d]) = P(X \in [a,b]) \cdot P(Y \in [c,d])$,

et donc $Q([a,b] \times [c,d]) = (\mu'([a,b])) \cdot (\mu''([c,d]))$,

i.e. $Q = \mu' \otimes \mu''$,

et donc X et Y sont indépendantes. C.Q.F.D.

Théorème 5-4 : Soit $q \in \mathbb{N}^*$, alors si pour tout
 $(i,j) \in \mathbb{N}^* \cap [1,q]^2$, $\theta_{i,j} = 0$ on a l'inégalité :

$$D_2^N(F_1(X), F_2(Y)) < \sqrt{90 \cdot \left(\sum_{\substack{i>q \\ \text{ou } j>q}} M_{i,j}^2 \right)}$$

C'est l'inégalité 3-6.

3 - 3 : Cas particuliers

Nous allons citer quelques cas particuliers de série de polynomes orthogonaux , et nous nous intéresserons plus particulièrement au cas gaussien .

Comme ce chapitre est en fait une digression du sujet général de cette thèse , nous nous contenterons d'énoncer les théorèmes : ils se démontrent d'ailleurs sans difficultés majeures .

Rappelons d'abord les définitions classiques (17-4 [8] , [10]) :

Définitions 6-1 :

a) Si m_1 est la mesure sur $[-1,+1]$ de densité $p_1(t) = (1-t)^\alpha \cdot (1+t)^\beta$ où $\alpha > -1$ et $\beta > -1$, les polynomes orthogonaux associés à cette mesure s'appellent polynomes de Jacobi .

b) Si $\alpha = \beta = 0$, on retrouve les polynomes de Legendre .

c) Si $\alpha = \beta = -\frac{1}{2}$, on obtient les polynomes de Tschebyscheff .

d) Si m_2 est la mesure sur $[0,+\infty[$ admettant comme densité $p_2(t) = e^{-t}$, on obtient les polynomes de Laguerre .

e) Si m_3 est la mesure sur \mathbb{R} de densité $p_3(t) = e^{-t^2}$, on obtient les polynomes d'Hermite .

Les propriétés générales de ces polynomes sont détaillées dans de nombreux ouvrages (en particulier [10]) ; citons plus particulièrement (17-5 [8]) le

Théorème 6-2 : Les polynomes de Laguerre et les polynomes d'Hermite sont des bases orthogonales de , respectivement $\mathcal{L}^2(\mathbb{R}, m_2)$ et $\mathcal{L}^2(\mathbb{R}, m_3)$.

On obtient donc ainsi à homotéthisie , translation , et normalisation près , les polynomes orthonormaux associés à la loi bêta de première espèce , à la loi uniforme , à la loi exponentielle , et à la loi de Gauss .

Nous allons plus particulièrement étudier cette dernière en rappelant d'abord la

Proposition 6-3 : Les polynomes d'Hermite sont également définis pour tout entier naturel n par l'égalité

$$H_n(x) = (-1)^n \cdot e^{x^2} \cdot \left(\frac{d^n (e^{-x^2})}{d x^n} \right) .$$

Alors en adoptant les

Notations 6-4 : Nous noterons par μ la mesure borélienne de densité

$$\frac{d\mu}{dx} = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} ,$$

et par \overline{H}_n , $n \in \mathbb{N}$, les polynomes

$$\overline{H}_n(x) = \left(2^{-\frac{1}{2} \cdot n} \cdot (n!)^{-\frac{1}{2}} \right) \cdot H_n\left(\frac{x}{\sqrt{2}}\right) .$$

on peut énoncer le

Corollaire 6-5 : Les polynomes $(\overline{H_n})$, $n \in \mathbb{N}$, forment une base orthonormale de $\mathcal{L}^2(\mathbb{R}, \mathcal{L})$.

On peut alors prouver le

Lemme 6-6 : Soit $(X, Y) \in V_{\mathbb{R}^2}$, un vecteur aléatoire suivant une loi normale centrée réduite - i.e. admettant une densité de probabilité g par rapport à la mesure de Lebesgue :

$$g(x, y) = \frac{1}{2\pi} e^{-\left(\frac{x^2 - 2\rho xy + y^2}{2}\right)}$$

et soit f l'application $f(x, y) = 2\pi \cdot e^{\frac{x^2}{2}} \cdot e^{\frac{y^2}{2}} \cdot g(x, y)$,

alors, dans $\mathcal{L}^2(\mathbb{R}^2)$, on a l'égalité

$$g(x, y) = \frac{e^{-\left(\frac{x^2+y^2}{2}\right)}}{2\pi} \cdot \left(1 + \sum_{(i,j) \in \mathbb{N}^* \times \mathbb{N}^*} \rho_{i,j} \cdot \overline{H_i(x)} \cdot \overline{H_j(y)} \right),$$

et de plus, $\rho_{i,j} = \langle \overline{H_i} \otimes \overline{H_j}, f \rangle_{\mathcal{L}^2(\mathbb{R}^2, \mathcal{L} \otimes \mathcal{L})}$,

i.e. $\rho_{i,j} = g_{i,j}(\rho)$, avec $g_{1,1} = 1_{\mathbb{R}}$,
 et on a l'implication " $\rho = 0 \implies g_{i,j}(\rho) = 0$ ".

Enfin, si $\rho \neq 0$, il existe $(i, j) \in \mathbb{N}^* \times \mathbb{N}^*$, $(i, j) \neq (1, 1)$,
 tel que $g_{i,j}(\rho) \neq 0$.

Conclusion 6-7 : Cette propriété permet de rappeler à quel point l'hypothèse gaussienne est forte : dans ce cas , on a un seul paramètre , alors que si l'on se contente de supposer que les lois marginales sont gaussiennes , on en aura une infinité dénombrable presque indépendants les uns des autres .

Elle montre aussi que le développement de g en polynômes d'Hermite ne se réduit pas à $g = \left(1 + \rho \cdot \overline{H_1} \otimes \overline{H_1} \right) \cdot \frac{e^{-\frac{1}{2}(x^2+y^2)}}{2 \pi}$.

4 : ASPECT STATISTIQUE

Nous allons donc maintenant étudier les estimateurs des différents coefficients introduits précédemment, ainsi que les tests d'indépendance - et parfois d'uniformité - construits à partir des développements en série. Le test hilbertien étudié par BOSQ servira de fondement à cette partie. Nous rappellerons d'abord ses principales propriétés. Pour leurs démonstrations, nous renvoyons le lecteur à [1].

4-1 : Test Hilbertien

4-1-a) Lois limites dans le cas d'un noyau fixe de type fini

7-1 : Notations : Soit $\{Z_i\}$, $i \in \mathbb{N}^*$, une suite de variables aléatoires indépendantes et de même loi, définies sur le même espace probabilisé (Ω, \mathcal{A}, P) , et à valeurs dans un espace mesurable (E, \mathcal{E}) , muni d'une probabilité m .

On cherche à tester l'hypothèse H'_0 : « La loi commune des Z_i est m » contre sa négation.

Pour cela on considère une famille orthonormale $\{e_j\}$, $j \in J$, $J \subset \mathbb{N}^p$, $0 \in J$, de $L^2(E, \mathcal{E}, m)$, vérifiant

$$e_0 \equiv 1 \quad \text{et} \quad \text{card}(J) = k < +\infty.$$

Nous noterons aussi par J^* l'ensemble : $J^* = J \setminus \{0\}$.

Posons $K(x,t) = \sum_{j \in J} e_j(x) \cdot e_j(t)$, pour $(x,t) \in E^2$;

K s'appelle le noyau de l'espace vectoriel engendré par les e_j , $j \in J$.

Considérons alors les statistiques $T_n(\cdot)$ et $S_n(\cdot)$:

$$T_n(\cdot) = \sum_{i=1}^n K(Z_i, \cdot) = n \cdot \left(1 + \sum_{j \in J^*} \widehat{a}_{j,n} \cdot e_j(\cdot) \right) ,$$

$$S_n(\cdot) = n^{-1/2} \cdot (T_n(\cdot) - n) ,$$

$$\text{où } \widehat{a}_{j,n} = \frac{1}{n} \sum_{i=1}^n e_j(Z_i) , \text{ pour tout } j \in J^* .$$

Alors , nous pouvons énoncer les

Proposition 7-2 : Si la loi commune des Z_i est m , alors $\| S_n \|^2 \xrightarrow{\mathcal{L}} Q$, où Q suit une loi du khi-deux à $k-1$ degrés de liberté .

Proposition 7-3 : Si la loi commune des Z_i est une probabilité m' telle qu'il existe $j \in J^*$ pour lequel e_j soit m' -intégrable , et d'intégrale non nulle , alors :

$$\| S_n \|^2 \xrightarrow{\text{p.s.}} +\infty .$$

4-1-b) Lois limites dans le cas d'un noyau de type fini qui varie
avec la taille de l'échantillon

7-4 : Dans ce cas nous considérons une suite

$$\{ e_{j,n} \}, \quad j \in J_n, \quad n \in \mathbb{N}^*$$

avec $J_n \subset \mathbb{N}^p$, $0 \in J_n$, $\text{card}(J_n) = k_n < +\infty$,

$$\text{et } \lim_{n \rightarrow \infty} k_n = +\infty.$$

qui est une suite de systèmes orthonormaux dans $\mathcal{L}^2(E, \mathcal{E}, m)$, telle
que $e_{0,n} \equiv 1$, pour tout $n \in \mathbb{N}^*$.

Soit alors la suite de noyaux

$$K_n(x,t) = \sum_{j \in J_n} e_{j,n}(x) \cdot e_{j,n}(t); \quad (x,t) \in E^2, \quad n \in \mathbb{N}^*.$$

On pose encore :

$$S_n = n^{-1/2} \cdot \left(\sum_{i=1}^n K_n(Z_i, \cdot) - n \right),$$

$$\widehat{a}_{j,n} = \frac{1}{n} \sum_{i=1}^n e_{j,n}(Z_i); \quad j \in J_n, \quad n \in \mathbb{N}^*.$$

On peut alors énoncer la

Proposition 7-5 : Soit $\{ Q_n \}$, $n \in \mathbb{N}^*$, une suite de variables aléatoires réelles qui suivent des lois du khi-deux dont les degrés de liberté respectifs valent $k_n - 1$, alors

$$\sup_{a \in \mathbb{R}_+} |P(\| S_n \|^2 < a) - P(Q_n < a)| < \frac{c_0 \cdot (k-1)^3}{\sqrt{n}} \int \sum_{j \in J_n^*} |e_{j,n}|^3 \cdot dm,$$

pourvu que m soit la loi des Z_i .

(c_0 désignant une constante universelle)

4-1-c) Construction du test

7-6 : Cas de noyau fixe : Dans ce cas, on pourra tester :

H'_0 : « La loi des Z_i est m » .

contre $H'_1(J^*)$: « La loi des Z_i est une loi m' telle que $\int e_j \cdot dm'$ existe et soit non nulle pour au moins un $j \in J^*$ » .

en utilisant la statistique

$$\| S_n \|^2 = n \cdot \left(\sum_{j \in J^*} \widehat{a_{j,n}}^2 \right) :$$

$$\text{sous } H'_0, \quad \|S_n\|^2 \xrightarrow{L} \chi_{k-1}^2,$$

$$\text{sous } H'_1(J^*), \quad \|S_n\|^2 \xrightarrow{p.s.} +\infty;$$

et l'on choisira le test défini par une zone de rejet de l'hypothèse

$$H'_0 \text{ de la forme : } n \left[\sum_{j \in J^*} \widehat{a}_{j,n}^2 \right] > w_n.$$

7-7 : Cas de noyau variant avec la taille de l'échantillon

Dans ce cas, les résultats de Bosq définissent alors une suite de tests qui permettent de tester

$$H'_0 : \ll \text{La loi des } Z_i \text{ est } m \gg.$$

en utilisant un test de la forme

$$n \cdot \left[\sum_{j \in J_n^*} \widehat{a}_{j,n}^2 \right] > w_n.$$

L'hypothèse $H'_1(J_n^*)$ sera, par exemple, l'ensemble des lois m' telles qu'il existe une suite $j(n)$ de \mathbb{N}^p pour laquelle

$$0 < \inf_n \left| \int e_{j(n)} \cdot n \cdot dm' \right| < +\infty.$$

Si $e_{j,n} = e_j$, $n \in \mathbb{N}^*$, $j \in J$, l'hypothèse $H'_1(J_n^*)$ sera l'ensemble des lois m' telles que pour au moins un $j \in J^*$,

$\int e_j \cdot dm'$ existe et soit non nulle. $H'_1(J_n^*)$ contient alors toutes les lois différentes de m qui admettent une densité appartenant à $\mathcal{L}^2(E, \mathcal{E}, m)$.

4-1-d) Convergence du test

Celle-ci est donnée par les deux propositions suivantes

Proposition 7-8 : Dans le cas de test utilisant un noyau fixe ,
et si l'on suppose que les e_j sont bornés

pour que le test 7-6 soit convergent , il faut et il suffit que

$$\lim_{n \rightarrow \infty} w_n = +\infty \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{w_n}{n} = 0 .$$

Proposition 7-9 : Dans le cas de test utilisant un noyau variant
avec la taille de l'échantillon , et si

$$e_{j,n} = e_j , \quad n \in \mathbb{N}^* , \quad \lim_{n \rightarrow \infty} \frac{k_n^4}{\sqrt{n}} = 0 , \quad \text{avec toujours} \quad \lim_{n \rightarrow \infty} k_n = \infty ,$$

et en supposant de plus que les e_j sont uniformément bornés ,

alors pour que le test converge , il faut et il suffit que

$$\lim_{n \rightarrow \infty} \frac{w_n - k_n}{\sqrt{k_n}} = +\infty \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{w_n}{n} = +\infty .$$

4-2 : Estimateurs

Nous allons voir que si nous posons $a_j = E(e_j(Z_1))$,

Théorème 7-10 : Les $\widehat{a}_{j,n}$ sont des estimateurs convergents presque sûrement des a_j .

En effet, il suffit d'appliquer le théorème de la loi forte des grands nombres.

Théorème 7-11 : Avec ces notations, on a l'égalité

$$E(\widehat{a}_{j,n}) = a_j.$$

De plus, si $e_j^2 \in \mathcal{L}^2(E, \mathcal{E}, m)$, on a aussi

$$\sigma^2(\widehat{a}_{j,n}) = \frac{1}{n} (E(e_j^2(Z_1)) - a_j^2).$$

En particulier, si $Z_i \rightsquigarrow m$ et si $j \neq 0$,

$$\text{alors } \sigma^2(\widehat{a}_{j,n}) = \frac{1}{n}.$$

Les deux premières égalités sont de simples calculs. Pour la dernière, il suffit de remarquer qu'alors $f = 1$, et donc $a_j = 0$ et $E(e_j^2(Z_1)) = |e_j|^2 = 1$.

Dans ces conditions , en rappelant 3-4,3-5,3-6,3-9,3-10, les théorèmes ci-dessous sont triviaux .

Théorème 7-12 : Soient $(X,Y) \in V^U$, et (X_i, Y_i) , $i = 1, 2, \dots, n$, un n-échantillon de (X,Y) , posons alors :

$$\widehat{\theta}_{i,j}^n = \frac{1}{n} \sum_{p=1}^n L_i(X_p) \cdot L_j(Y_p) ,$$

$$\widehat{\alpha}_{2^i, 2^{i,j}}^n = \frac{1}{n} \sum_{p=1}^n A_{2^i, j}(X_p) \cdot A_{2^{i,j}}(Y_p) ,$$

$$\widehat{a}_{i,j}^n = \frac{1}{n} \sum_{p=1}^n C_i(X_p) \cdot C_j(Y_p) , \quad \widehat{b}_{i,j}^n = \frac{1}{n} \sum_{p=1}^n C_i(X_p) \cdot S_j(Y_p)$$

$$\widehat{c}_{i,j}^n = \frac{1}{n} \sum_{p=1}^n S_i(X_p) \cdot C_j(Y_p) , \quad \widehat{d}_{i,j}^n = \frac{1}{n} \sum_{p=1}^n S_i(X_p) \cdot S_j(Y_p)$$

alors $\widehat{\theta}_{i,j}^n$, $\widehat{\alpha}_{2^i, 2^{i,j}}^n$, $\widehat{a}_{i,j}^n$, $\widehat{b}_{i,j}^n$, $\widehat{c}_{i,j}^n$, $\widehat{d}_{i,j}^n$,

sont des estimateurs convergents presque surement de , respectivement

$$\theta_{i,j} , \alpha_{2^i, 2^{i,j}} , a_{i,j} , b_{i,j} , c_{i,j} , d_{i,j} .$$

Corollaire 7-13 : Si nous notons par

$$\overline{\theta_{i,j}^n}, (i,j) \in \mathbb{Z}^2 ; \text{ et } \widehat{\beta_{i,j}^n}, (i,j) \in \mathbb{N}^2 ,$$

les variables aléatoires définies par :

$$\overline{\theta_{i,j}^n} = \widehat{\theta_{i,j}^n} \text{ si } (i,j) \in \mathbb{N}^{*2}, \text{ et } 0 \text{ sinon .}$$

$$\widehat{\beta_{i,j}^n} = \frac{\frac{\overline{\theta_{i+1,j+1}^n}}{\sqrt{(2i+3)(2j+3)}} + \frac{\overline{\theta_{i-1,j-1}^n}}{\sqrt{(2i-1)(2j-1)}} - \frac{\overline{\theta_{i+1,j-1}^n}}{\sqrt{(2i+3)(2j-1)}} - \frac{\overline{\theta_{i-1,j+1}^n}}{\sqrt{(2i-1)(2j+3)}}}{\sqrt{4(2i+1)(2j+1)}}$$

alors les $\widehat{\beta_{i,j}^n}$ sont des estimateurs convergents presque sûrement des $\beta_{i,j}$.

Corollaire 7-14 : Avec les notations de 3-10 ,
on a les convergences presque sûres suivantes :

Pour tout $(i,j) \in \mathbb{N}^{*2}$,

$$\frac{\widehat{a_{i,j}^n}}{4\pi ij} \xrightarrow{\text{p.s.}} D_{i,j} , \quad -\left(\frac{\widehat{b_{i,j}^n}}{4\pi ij}\right) \xrightarrow{\text{p.s.}} C_{i,j} ,$$

$$-\left(\frac{\widehat{c_{i,j}^n}}{4\pi ij}\right) \xrightarrow{\text{p.s.}} B_{i,j} , \quad \frac{\widehat{d_{i,j}^n}}{4\pi ij} \xrightarrow{\text{p.s.}} A_{i,j} ;$$

Pour tout $q \in \mathbb{N}^*$ et tout $s \in \mathbb{N}^*$:

$$\sum_{t=1}^s \frac{\widehat{b_{q,t}^n}}{2\sqrt{2} \pi q t} \xrightarrow{\text{p.s.}} C_{q,0} + \eta_{s,q}^1,$$

$$\sum_{t=1}^s \frac{\widehat{c_{t,q}^n}}{2\sqrt{2} \pi q t} \xrightarrow{\text{p.s.}} B_{0,q} + \eta_{s,q}^2,$$

$$\sum_{t=1}^s \frac{\widehat{-d_{q,t}^n}}{2\sqrt{2} \pi q t} \xrightarrow{\text{p.s.}} A_{q,0} + \eta_{s,q}^3,$$

$$\sum_{t=1}^s \frac{\widehat{-d_{t,q}^n}}{2\sqrt{2} \pi q t} \xrightarrow{\text{p.s.}} A_{0,q} + \eta_{s,q}^4;$$

$$\sum_{\substack{t=1 \\ t'=1}}^{\substack{t=s \\ t'=s}} \frac{\widehat{d_{t,t'}^n}}{2\pi t t'} \xrightarrow{\text{p.s.}} A_{0,0} + \eta_s^5;$$

où $\eta_{s,q}^i \in \mathbb{R}$ et $\eta_{s,q}^i < \frac{1}{2\pi q \sqrt{2s}}$ pour $i = 1, 2, 3, 4$,

et $\eta_s^5 \in \mathbb{R}$ et $\eta_s^5 < \frac{1}{2\pi s}$.

Corollaire 7-15 : Il existe une application ϵ de \mathbb{N} dans lui-même, vérifiant

$$\epsilon(p) < \sum_{i>p \text{ ou } j>p} M_{i,j}^2 .$$

telle que pour tout $q \in \mathbb{N}^*$, on ait la convergence presque sûre :

$$\sum_{\substack{j=q \\ i=q \\ i=0 \\ j=0}} \widehat{p}_{i,j}^n \xrightarrow{\text{p.s.}} (D_2(X,Y))^2 + \epsilon(q) .$$

4-3 : Tests d'indépendance

Nous allons maintenant nous servir des résultats de Bosq pour tester l'indépendance de deux variables aléatoires. Notre étude ci-dessous porte sur le cas où (X,Y) est à valeurs dans $[0,1]^2$ et de lois marginales uniformes. Mais nous généraliserons parfois au cas des tests d'uniformité sur $[0,1]^p$.

Nous supposerons donc dans cette partie que nous sommes sous ces hypothèses.

4-3-a) Test réel d'indépendance

Le problème principal qui se pose est en fait de savoir ce que l'on teste réellement si l'on utilise le test hilbertien pour tester l'indépendance.

Bosq conclut ses calculs en disant que ceux-ci permettent de tester

H'_0 : « La loi des Z_i est m » ,

contre $H'_1(J^*)$: « La loi des Z_i est une loi m'
telle que $\int e_j \cdot dm'$ existe
et soit non nulle pour au moins un $j \in J^*$. »

Mais il est clair que ce que l'on teste réellement c'est :

" $a_j = 0$, $j \in J^*$ " à partir de

$$n \cdot \left[\sum_{j \in J^*} \widehat{a_{j,n}}^2 \right],$$

et donc $H_0(J^*) : \ll a_j = 0 , \forall j \in J^* \gg$

contre $H_1(J^*) = H'_1(J^*) :$

$\ll \exists j \in J^* : a_j \neq 0 \text{ et } \int e_j \cdot dm' \text{ existe} \gg$.

Comme les fonctions de Legendre , de Walsh ou trigonométriques sont bornées sur $[0,1]$, on a toujours l'inégalité :

$$\left| \int e_j \cdot dm' \right| < +\infty ,$$

et donc nous testerons en fait $H_0(J^*)$

contre $H_1(J^*) : \ll \exists j \in J^* : a_j \neq 0 \gg$.

Soit $h \in \mathbb{N}^* U \{+\infty\}$,

nous noterons alors par $C_0(\varepsilon)$ et C_1 , les hypothèses :

$$C_0(\varepsilon) : \ll (X, Y) \in V^U : D_h^N(X, Y) < \varepsilon . \gg ,$$

$$\text{ou } C_0(\varepsilon) : \ll (X_1, X_2, \dots, X_p) \in V_{[0,1]^p} : \delta_1(F, U) < \varepsilon . \gg$$

suivant le cas

$$\text{et } C_1 : \ll (X, Y) \in V^U : D_h^N(X, Y) \neq 0 . \gg$$

$$\text{ou } C_1 : \ll (X_1, X_2, \dots, X_p) \in V_{[0,1]^p} : \delta_1(F, U) \neq 0 . \gg$$

Comme J représente ici une partie finie de \mathbb{N}^p , nous pouvons noter par q l'entier naturel tel que

$$[0, q]^p \subset J \quad \text{et} \quad [0, q+1]^p \not\subset J ,$$

et nous poserons alors $H_s(J^*) = H_s(q)$ $s=1, 2$.

7-16 : Application : Test réel d'indépendance ou d'uniformité
fourni par le test hilbertien .

Avec les notations précédentes , il est clair (3-6,3-12,3-16,3-18,1-3) que si l'on effectue ce test avec les polynomes de Legendre , les fonctions trigonométriques ou les fonctions de Walsh , on a les inclusions suivantes :

$$H_0(q) \subseteq C_0(\varepsilon(q)) \quad \text{et} \quad H_1(q) \subseteq C_1 .$$

$$\text{et que } C_1 \cap C_0(\varepsilon(q)) \neq \emptyset$$

- ε étant une suite réelle vérifiant $\lim_{q \rightarrow \infty} \varepsilon(q) = 0$,

et ne dépendant que du choix de la base orthonormale et du choix de l'indicateur de dépendance $D_h^N(X,Y)$, $h \in \mathbb{N}^* \cup \{+\infty\}$, (lors du test d'indépendance , bien sur !) .

Concrètement , ces inclusions signifient que l'acceptation de l'hypothèse $H_0(q)$ entraîne l'acceptation de l'hypothèse

$C_0(\varepsilon(q))$: $\ll X$ et Y sont indépendantes

(resp $(X_1, X_2, \dots, X_p) \rightsquigarrow U([0,1]^p)$)

à $\varepsilon(q)$ près , si l'on mesure la dépendance

(resp la distance à l'uniformité) par l'indicateur

$D_h^N(X,Y)$ (resp par la distance $\delta_1(F,U)$) \gg .

Et , bien sûr , l'acceptation de $H_1(q)$ implique le rejet de l'indépendance (resp de l'uniformité) .

De plus ces hypothèses ne sont pas exclusives l'une de l'autre .

7-17 : Remarques : Rappelons que résumer ce genre de test au test de

H_0 : « X et Y sont indépendantes »

contre H_1 : « X et Y ne sont pas indépendantes »

est un abus de langage et une inexactitude mathématique .

Les résultats ci-dessus permettent donc de préciser notre langage et de résumer de façon simple le test d'indépendance réellement effectué .

Quant à la question du choix du coefficient D_h pour mesurer la dépendance , nous y reviendrons dans la partie 3-1-2 .

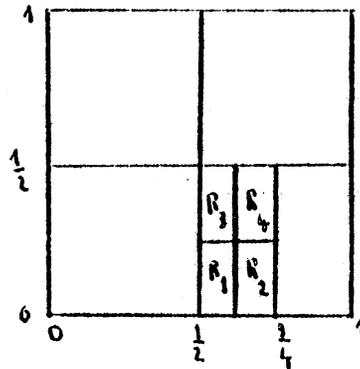
7-18 : Exemple concret : Nous nous plaçons ici dans le cas des fonctions de Walsh , et nous nous servirons du test hilbertien pour tester

W_0 : « $\alpha_{2,1}^{4,3} = 0$ » contre W_1 : « $\alpha_{2,1}^{4,3} \neq 0$ » .

Nous supposons donc disposer d'un n-échantillon (X_i, Y_i) , $i=1,2,\dots,n$.

Nous aurons donc recours à l'estimateur $\hat{\alpha}_{2,1}^{4,3}$ n.

Il est à peu près évident que celui-ci représente le nombre de points de l'échantillon contenus dans les rectangles R_1 et R_4 , diminué de celui des points contenus dans R_2 et R_3 , à une constante de normalisation près.



Supposons que ce soit l'indépendance de $(X,Y) \in V^U$, avec $Y = X$, que nous testions.

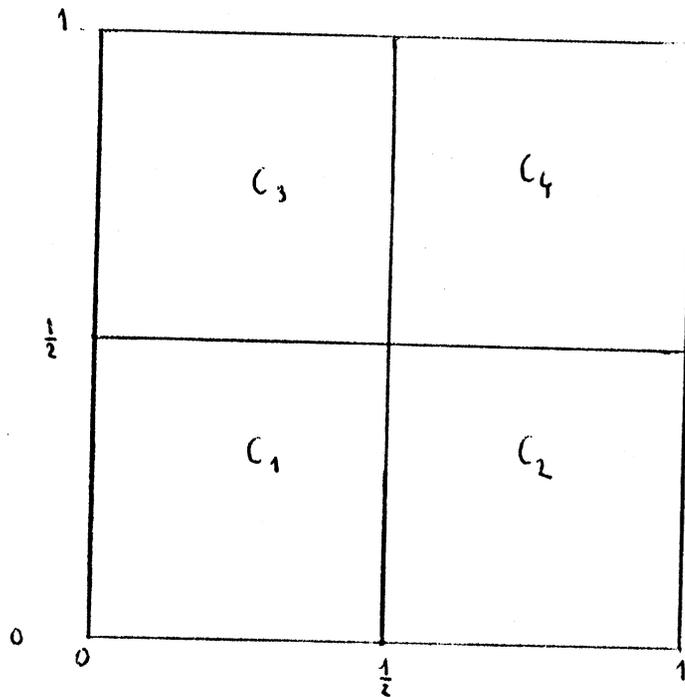
Il est clair que dans ce cas $\hat{\alpha}_{2,1}^{4,3} n = 0$.

Donc, si à partir de ce résultat, nous concluons autre chose que X et Y sont indépendantes à 1 près (i.e. le maximum), nous commettrions une grossière erreur.

Le résultat serait le même si l'on testait tous les coefficients

$$\alpha_{2^i, j'}^{2^i, j} \text{ vérifiant } j' < 2^{i-1} \text{ et } j > 2^{i-1},$$

$$\text{ou } j' > 2^{i-1} \text{ et } j < 2^{i-1},$$



en fait , les coefficients des fonctions $A_{2^i,j} \otimes A_{2^{i'},j'}$, de support inclus dans les carrés C_2 ou C_3 .

Si $Y = X$, tous ces coefficients sont nuls , mais il ne faut pas pour autant conclure à l'indépendance

Inversement , un de ces coefficients non nul signifiera , bien sur , qu'il n'y a pas indépendance .

Cet exemple montre , d'une part l'utilité de savoir ce que l'on teste (7-16,7-17) , et d'autre part la nécessité des calculs de majoration des coefficients D_h .

Il montre aussi que les bases orthonormales de Legendre , Fourier et Walsh sont des bases pratiques (cf remarque 6-5 [3]) : une base dont les premiers termes seraient les coefficients $\frac{2^i,j}{2^{i'},j'}$ ci-dessus ne serait guère utilisable .

4-3-b) : Cas des séries de polynomes

Nous nous plaçons toujours dans le cas où (X,Y) est à valeurs dans $[0,1]^2$.

1 : Test d'indépendance

Nous supposerons de plus, ici, que $(X,Y) \in V^U$.
Nous garderons les notations et hypothèses de 3-6 : \mathcal{P} et \mathcal{P}' représenteront donc des parties finies de \mathbb{N}^2 .

Alors, si l'on teste l'indépendance par le test hilbertien :

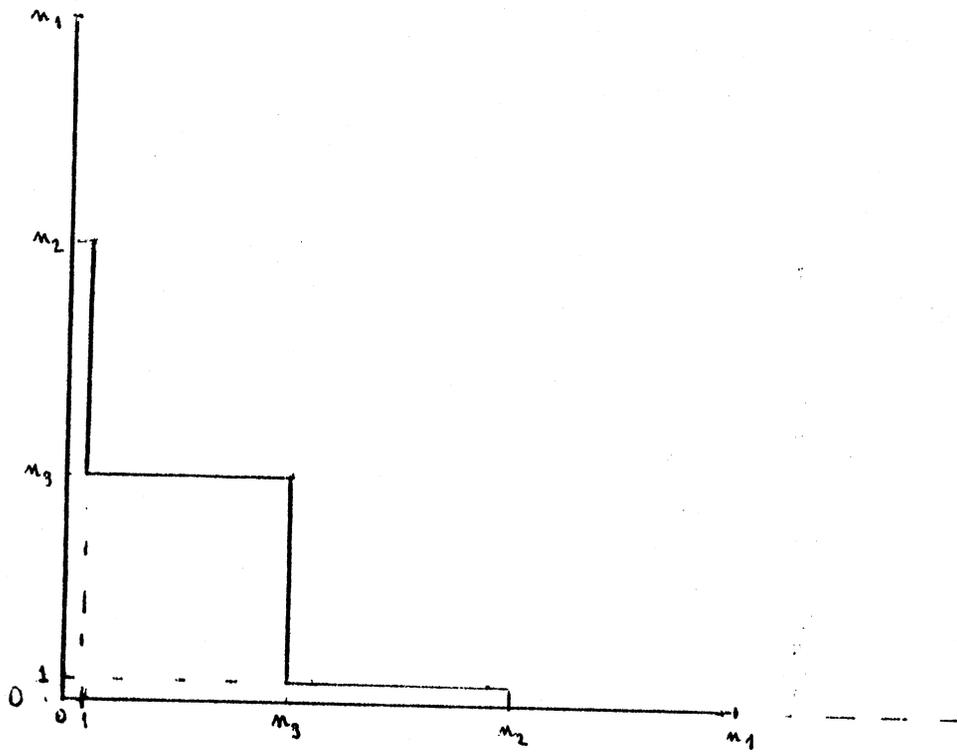
$$H_0(\mathcal{P}) : \langle\langle \theta_{i,j} = 0, \forall (i,j) \in \mathcal{P} \rangle\rangle$$

contre $H_1(\mathcal{P})$, sa négation, on pourra écrire le test en rappelant les inclusions suivantes : on teste donc

$$H_0(\mathcal{P}) \subset C_0(\epsilon(|\mathcal{P}|)) : \langle\langle (D_2^N(X,Y))^2 < 90. \left(\sum_{(i,j) \in \mathcal{P}} M_{i,j}^2 \right) \rangle\rangle$$

contre $H_1(\mathcal{P}) \subset C_1 : \langle\langle X \text{ et } Y \text{ ne sont pas indépendantes} \rangle\rangle$

Concrètement si P' est le sous-ensemble fermé de \mathbb{R}^2 défini sur le graphique ci-dessous, et si $\mathcal{P}' = P' \cap \mathbb{N}^2$, et si nous notons par $N_{f.o.}$ le nombre de fonctions orthogonales utilisées, alors on obtient le tableau suivant :



(n_1, n_2, n_3)	10, 5, 3	100, 25, 10	500, 150, 25	10 000, 2 500, 1 000
$D_2^N <$	0,545	0,250	0,143	0,025
$N_{f.o.}$	34	314	1 829	1 021 004

Nous remarquons bien sûr , le nombre élevé de fonctions orthogonales nécessaires pour obtenir de faibles majorations de $D_2^N(X,Y)$. Mais cela est assez normal : nous reviendrons sur cette question en partie 3-1-3 .

2 : Test du coefficient de corrélation

Dans le cas où nous nous contentons de tester

$$H_0(1) : \langle \langle \theta_{1,1} = 0 \rangle \rangle \quad \text{contre sa négation ,}$$

et vu que $\theta_{1,1} = \rho(X,Y)$ (3-3) , , on retrouve ainsi le test du coefficient de corrélation .

C'est un des tests d'indépendance les plus connus ; et l'on sait très bien que l'acceptation de l'hypothèse $H_0(1)$ ne signifie pas réellement l'acceptation de l'indépendance mais seulement l'acceptation de l'indépendance linéaire (nous avons donné un exemple où $\rho(X,Y) = 0$ et $D_2^N(X,Y) = 0,61$ (C-7, [2])). C'est d'ailleurs une des raisons pour lesquelles nous avons suggéré de remplacer $\rho(X,Y)$ par la suite des coefficients de corrélation d'ordre supérieur .

En tout cas , cet exemple illustre parfaitement notre propos : l'acceptation de $H_0(1)$ n'est que l'acceptation d'une indépendance particulière , et si l'on veut être plus assuré de l'indépendance réelle , il convient de procéder à des tests complémentaires .

3 : Test d'uniformité

Dans ce cas , nous ne possédons pas de majoration de $\delta_h(F,U)$, mais bien sur on pourra toujours tester

$H_0(E) : \langle \langle \theta_{i,j} = 0 , \forall (i,j) \in E \rangle \rangle$, contre sa négation ,

E étant ici un sous ensemble de $\mathbb{N}^2 \setminus \{0\}$ (et pas seulement de \mathbb{N}^{*2} : il faut aussi dans ce cas tester les lois marginales) .

Mais alors , nous ne pourrons plus assurer a priori que l'acceptation de $H_0(E)$ entraîne que la loi de (X,Y) soit assez proche de la loi uniforme , bien qu'intuitivement on puisse quand même l'espérer (cf graphes des L_i : 6-5, [3]) .

4-3-c) Cas des fonctions trigonométriques

On obtient ici exactement le même type de résultats que précédemment .

1 : Test d'indépendance

Si on ramène le test de l'indépendance au test de

$$H_0(q) : \ll a_{i,j} = b_{i,j} = c_{i,j} = d_{i,j} = 0 , \forall (i,j) \in [1,q]^2 \cap \mathbb{N}^2 \gg$$

contre sa négation , on pourra écrire en rappelant les inclusions suivantes (3-12) que l'on teste

$$H_0(q) \subset C_0(\varepsilon(q)) : \ll D_2^N(X,Y) < \frac{3\sqrt{50}}{\pi} \cdot \frac{1}{\sqrt{q}} \gg$$

contre $H_1(q) \subset C_1 : \ll X \text{ et } Y \text{ ne sont pas indépendantes} \gg$.

Concrètement , en notant toujours par $N_{f.o.}$ le nombre de fonctions orthogonales utilisées , on obtient le tableau suivant :

q	100	10 000
$D_2^N <$	0,215	0,0215
$N_{f.o.}$	40 000	400 000 000

On pourrait se dire que , étant donné le nombre élevé de fonctions orthogonales nécessaires pour tester une indépendance acceptable , l'utilisation des fonctions trigonométriques comme base orthogonale du test hibertien ne s'impose pas .

Aussi il convient de faire deux remarques .

D'abord , la majoration de $D_2(X,Y)$ obtenue dans ce cas est assez grossière et peut sans doute être affinée (cf 4-9, [3]) .

Ensuite , les fonctions trigonométriques sont celles qui sont le plus faciles à utiliser sur ordinateur , étant donné qu'elles y sont déjà programmées .

2 : Tests d'uniformité

Dans ce cas , les conclusions sont exactement les mêmes que pour les tests d'uniformité à partir des polynomes de Legendre (cf ci-avant) .

4-3-d) Cas des fonctions de Walsh

Ce cas est en fait très intéressant car , d'une part le test d'indépendance est lié directement au test du khi-deux , et d'autre part , il fournit des exemples faciles à visualiser et qui nous permettrons par la suite de mieux expliciter les rapports entre la dépendance et les développements en série , et enfin la majoration des indicateurs de dépendance est la meilleure que nous ayons obtenue et de plus nous avons également obtenu une majoration de la distance à l'uniformité .

1 : Test d'uniformité

Contrairement aux cas précédents , nous traitons d'abord ici le cas du test d'uniformité pour rappeler qu'alors la statistique hilbertienne et la statistique du khi-deux sont identiques . C'est le

Théorème 7-21 : Soit $X \in V$, et (X_i) , $i = 1, 2, \dots, n$, un n -échantillon de X sur $[0, 1]^p$, et soit χ_X^2 , la statistique du khi-deux associée à cet échantillon et à la partition de $[0, 1]^p$ muni de la loi uniforme en 4^{q+1} carrés $C_{j,j'}$, alors on a l'égalité :

$$\| S_n \|^2 = n \cdot \left(\sum_{\substack{J \in (M_q)^p \\ J \neq (-\infty, 0)^p}} (\hat{\alpha}_J)^2 \right) = \chi_X^2 .$$

Pour la démonstration , nous renvoyons à 8-8 .

On en déduit la

Conséquence 7-22 : Si l'on teste l'uniformité d'un vecteur aléatoire $X \in V_{[0,1]^p}$ par le test du khi-deux avec la partition de 7-21 - ou, ce qui revient au même par le test hilbertien avec un développement en fonctions de Walsh jusqu'à l'ordre q - on teste donc

$$H_0(q) \subset C_0(\epsilon(q)) : \ll \delta_1(F,U) < \frac{1}{2^{q+1}} \left(\sum_{t=0}^{p-1} \left(1 - \frac{1}{2^{q+1}} \right)^t \right) \gg$$

contre $H_1(q) \subset C_1 : \ll X \text{ ne suit pas la loi uniforme sur } [0,1]^p \gg$.

Concrètement on a le tableau suivant :

p	1	1	2	2	3	3	10	10
q	3	7	3	7	3	7	3	7
$\delta_1(F,U) <$	0,06	0,004	0,12	0,007	0,176	0,01	≈ 1	0,14
$N_{f.o.}$	15	255	255	65 535	4 095	16 777 215	16^{10}_{-1}	$2^{8 \times 10}_{-1}$

2 : Test d'indépendance

Ce cas est très intéressant car il va nous permettre de préciser le test d'indépendance fourni par le test du khi-deux lorsque l'on connaît les lois marginales .

En effet , si $(X,Y) \in V^U$, les coefficients $\alpha_{0,0}^{2^i,j}$ et $\alpha_{2^i,j}^{0,0}$ sont connus : ils sont nuls .
 Il n'y a donc plus besoin de les tester .

Alors , grâce à 7-21 , nous pouvons donner une forme simple à la statistique hilbertienne avec le

Théorème 7-23 : En conservant les notations précédentes et si $(X,Y) \in V^U$, alors on a l'égalité :

$$\| S_n \|^2 = \sum_{(J',J'') \in \binom{J',n}{q}} (M_q^*)^2 \alpha_{J''}^2 = \chi_{X,Y}^2 - \chi_X^2 - \chi_Y^2 ,$$

et de plus , $\chi_{X,Y}^2 - \chi_X^2 - \chi_Y^2$ suit asymptotiquement une loi du khi-deux à $(2^{q+1} - 1)^2$ degrés de liberté .

Pour la démonstration , nous renvoyons à 8-13 .

D'autre part , comme précédemment , on pourra écrire le test d'indépendance sous la forme

$$H_0(q) \subset C_0(\varepsilon(q)) : \langle\langle D^N(X,Y) < \frac{1}{2^q} \rangle\rangle$$

contre $H_1(q) \subset C_1 : \langle\langle X \text{ et } Y \text{ ne sont pas indépendantes} \rangle\rangle$.

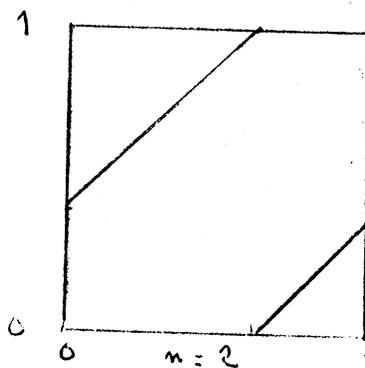
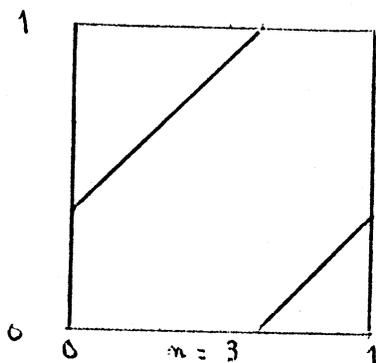
Concrètement on obtient le tableau suivant :

q	2	3	7
$D^N(X,Y) <$	1 4	1 8	1 128
$N_{f.o.}$	49	225	16 129

4-3-e) Application

Les résultats que nous avons acquis vont nous permettre de mieux comprendre que pour qu'un n -échantillon de $[0,1]^p$ puisse servir à tester l'uniformité, il faut que n soit assez grand.

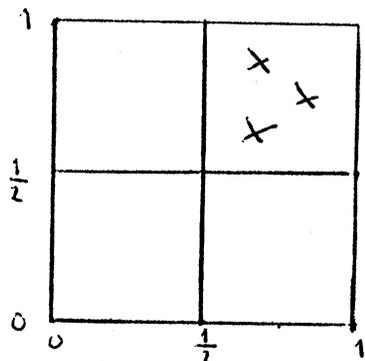
Ainsi, si $n \ll p$, il est bien évident que l'on ne peut pas parler d'échantillon significatif. Intuitivement, on sent bien qu'il faut au moins assurer $n \gg 2^p$. Par exemple un échantillon de taille 2 ou 3 sera insuffisant dans le cas de $[0,1]^2$ pour prouver une quelconque indépendance.



cas $p = 0$, $D_2^m(x, Y) \approx 0,61$ (C-7[R])

Par contre, on peut très bien obtenir avec de tels échantillons des présomptions sur l'hypothèse opposée.

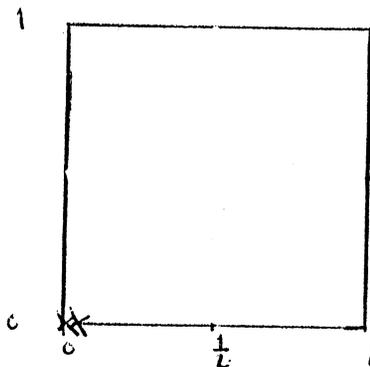
$n = 3$



$\alpha_{1,1}^{1,1} \neq 0, \alpha_{1,1}^{0,0} \neq 0$

$\alpha_{0,0}^{1,1} \neq 0$

$n = 2$



mesure de Dirac en $(0,0)$

Et il n'est pas difficile de voir sur ces exemples que ce que l'on peut tester en fait, ce sont les coefficients α_j , $J \in M^q$, et que ce qui peut apparaître à partir de tels échantillons, c'est qu'un de ces coefficients soit non nul.

Et donc la seule conclusion que l'on puisse obtenir; c'est qu'il n'y a pas indépendance.

Remarquons aussi que d'après les inégalités de 3-16 et 3-17 la taille minimum nécessaire de l'échantillon croît en fonction de p , ce qui veut dire qu'il est beaucoup plus difficile de tester l'uniformité d'une variable à valeurs dans $[0,1]^{12}$, par exemple, que dans $[0,1]^2$ (en fait, ce résultat est parfaitement naturel: si l'on testait l'indépendance, ce serait l'indépendance de 12 variables aléatoires ou de 2). Ainsi un échantillon de taille 4 096 = 2^{12} ne nous apportera que peu de conclusions sur l'uniformité dans $[0,1]^{12}$, alors qu'il sera bien convenable dans $[0,1]^2$.

4-3-f) conclusion : choix de l'ordre du développement

C'est bien sûr la question la plus importante qui se pose alors, et à laquelle nous allons répondre maintenant.

1°) Cas où l'on dispose d'échantillons de taille limitée

Dans ce cas, l'ordre du développement sera limité de facto, vu qu'à partir d'un n -échantillon, il n'est pas possible d'effectuer un test qui se base sur l'estimation de plus de n paramètres.

2°) Cas où l'on dispose d'échantillons de taille arbitraire

Pour résoudre le problème dans ce cas , il faut en fait savoir quelle indépendance on veut obtenir , et il est bien évident que ce choix, et donc le choix de l'ordre du développement dépend des besoins de l'utilisateur .

La partie 3-2 est un exemple de ce genre de problèmes . Mais on peut aussi prendre le cas de 1-12 : dans ce cas l'acceptation de l'indépendance que le test vérifie en principe doit elle être considérée comme une réponse aberrante ou non ? Il est clair que ce choix dépend essentiellement des besoins de l'utilisateur .

3°) Test d'indépendance réelle

Dans le cas où l'on veut l'indépendance réelle , il convient de choisir l'ordre du développement maximum en fonction de la taille de l'échantillon selon les résultats de Bosq .

Ainsi , par exemple pour tester l'indépendance par le test du khi-deux , il convient que les nombres $n.p_i$ - n : taille de l'échantillon , p_i probabilités des carrés utilisés - soient minimums (tout en étant suffisamment élevés pour que l'approximation de la loi multinomiale par la loi du khi-deux soit correcte) . Sinon , on risque de ne pas déceler certaines dépendances .

Ainsi , il sera insuffisant si l'on dispose d'un 10^3 -échantillon d'un vecteur aléatoire $(X,Y) \in V^U$ de diviser $[0,1]^2$ en 4^2 carrés égaux : il faudra plutôt opérer une partition de 10^2 carrés .

4-3-g) Comparaison des différentes séries

Nous résumons ici les résultats acquis

1°) La vitesse de convergence $\epsilon(q)$ des indicateurs de dépendance vers zéro, en fonction de l'ordre q du développement est bien meilleure pour les fonctions de Walsh, et ce pour la meilleure distance : $h = +\infty$, $\delta_h(f,g) = \sup |f - g|$.

De plus, dans ce cas, et dans ce cas seul, nous avons également obtenu une majoration de la distance à l'uniformité lorsque l'on ne fait plus d'hypothèses sur les lois marginales.

2°) Les fonctions trigonométriques étant déjà programmées sur les ordinateurs sont les plus facilement utilisables pour les tests de l'indépendance réelle.

3°) Les polynômes de Legendre sont évidemment moins maniables, néanmoins ils permettent d'étudier des dépendances particulières (dépendance linéaire, quadratique, etc) et nous ramènent au cas des coefficients de corrélation d'ordre supérieur.

4-3-h) Conclusion

Comme on sait que de toutes façons , à moins de posséder des données physiques concrètes - par exemple le lancement de deux dés - , on ne pourra jamais être sûr , non seulement de l'indépendance réelle de deux variables aléatoires X et Y , mais même qu'il n'existe pas de bijection entre elles , il est clair que tout test d'indépendance ne sera significatif que par rapport à son utilisation . Le choix de l'ordre du développement ne pourra donc dépendre que des besoins de l'utilisateur , et l'on devra toujours se ramener au test de H_0 : « X et Y peuvent être considérées comme indépendantes par l'utilisateur » contre H_1 sa négation .

5 : TEST DU KHI-DEUX D'INDEPENDANCE

Bosq a fait remarquer que , comme les fonctions indicatrices d'ensemble 1_I , I appartenant à une partition de \mathbb{R}^2 , sont des fonctions orthogonales , et bien qu'elles ne vérifient pas :

$\int 1_I \cdot dm = 0$, le test du khi-deux est contenu dans la classe des tests hilbertiens .

Nous allons alors introduire de nouvelles fonctions orthogonales A_j qui sont combinaisons linéaires de fonctions indicatrices d'ensemble et qui vérifient $\int A_j \cdot dm = 0$.

Dans le cas à deux variables , nous pourrons alors faire apparaître les densités marginales sous la forme

$$f_1 = 1 + \sum_{i=1}^{\infty} \tau_{i,0} \cdot A_{i,0} \quad (2-3) .$$

Alors , si nous testons l'indépendance à partir de lois marginales connues , nous n'aurons pas à tester les coefficients $\tau_{i,0}$, (car nous saurons déjà qu'ils sont nuls) , ce que l'on fait en fait avec le test du khi-deux classique .

Finalement , nous trouverons que la statistique hilbertienne sera : $\chi_{X,Y}^2 = \chi_X^2 + \chi_Y^2$, et qu'elle suivra asymptotiquement une loi du khi-deux à $(k-1)(k'-1)$ degrés de liberté au lieu de $kk'-1$ dans le cas classique .

Nous étudierons alors le test d'indépendance avec des lois marginales inconnues, et nous découvrirons que celui-ci est basé sur la statistique $\widehat{\chi^2_{X,Y}} - \widehat{\chi^2_X} - \widehat{\chi^2_Y}$, avec dans ce cas $\widehat{\chi^2_X} = \widehat{\chi^2_Y} = 0$, ce qui nous permettra de voir qu'en fait, dans tous les cas, la statistique naturelle du khi-deux pour tester l'indépendance sera de la forme $\overline{\chi^2_{X,Y}} - \overline{\chi^2_X} - \overline{\chi^2_Y}$, et que, sous l'hypothèse d'indépendance, elle suit toujours une loi du khi-deux à $(k-1)(k'-1)$ degrés de liberté.

Comme ce résultat est important, nous détaillerons nos démonstrations qui sont d'ailleurs assez simples.

1 : Statistique du khi-deux et statistique hilbertienne

Dans cette partie , nous allons démontrer que la statistique hilbertienne portant sur tous les coefficients $\tau_{i,j}$, est exactement la statistique du khi-deux . Alors , il sera simple , par la suite , de démontrer les propriétés annoncées ci-dessus .

8-1 : Notation des partitions de \mathbb{R}^p : Soit $p \in \mathbb{N}^*$, et soit $\{c(s)\}$, $s = 1, 2, \dots, p$, une suite finie de \mathbb{N}^* : $c(s) \in \mathbb{N}^*$.

Soient \mathbb{D}^s et \mathbb{D}^{*s} , $s = 1, 2, \dots, p$, les sous-ensembles de \mathbb{N}
 $\mathbb{D}^s = (1, 2, \dots, c(s))$ et $\mathbb{D}^{*s} = (2, 3, \dots, c(s))$.

et soient \mathbb{D} et \mathbb{D}^* les sous-ensembles de \mathbb{N}^p :
 $\mathbb{D} = \prod_{s=1}^p \mathbb{D}^s$, et $\mathbb{D}^* = \mathbb{D} \setminus \{ (1, 1, \dots, 1) \}$.

Soit alors $\{ \mathcal{P}^s \}$, $s = 1, 2, \dots, p$, une suite finie de partitions de \mathbb{R} en intervalles disjoints : $\mathcal{P}^s = \{ I_i^s \mid i \in \mathbb{D}^s \}$,

et soit \mathcal{J} la partition de \mathbb{R}^p :

$$\mathcal{J} = \left\{ \prod_{s=1}^p I_{i(s)}^s \mid s \in (1, 2, \dots, p) ; i(s) \in \mathbb{D}^s \right\} .$$

8-2 : Notations des mesures : Soit $\{\mu_i\}$, $i = 1, 2, \dots, p$, une suite de probabilité sur \mathbb{R} , et soit μ la probabilité produit sur \mathbb{R}^p : $\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_p$.

Soit alors X un vecteur aléatoire à valeurs dans \mathbb{R}^p , de loi Q et de lois marginales $\mu_1, \mu_2, \dots, \mu_p$.

Et soit (X_i) , $i = 1, 2, \dots, n$, un n -échantillon de X , nous noterons alors par μ_e la probabilité empirique :

$$\mu_e(\{V\}) = \frac{1}{n} \text{card}(i \mid X_i = V),$$

et par f_e , la densité empirique associée à \mathcal{J} :

$$\text{pour tout } I \in \mathcal{J} \text{ et tout } x \in I : f_e(x) = \frac{\mu_e(I)}{\mu(I)}.$$

Nous supposerons donc que pour tout $I \in \mathcal{J}$, on a $\mu(I) > 0$.

8-3 Notations des fonctions orthogonales : Soit alors

$\{A_i^s\}$, $s = 1, 2, \dots, p$, $i \in \mathbb{D}^s$, une suite finie de familles

orthonormales de, respectivement, $\mathcal{L}^2(\mathbb{R}, \mu_s)$, vérifiant $A_1^s \equiv 1$,

et engendrant le même sous-espace de $\mathcal{L}^2(\mathbb{R}, \mu_s)$ que les familles

d'indicatrices d'ensemble : $\{1_{I_i^s}\}$, $i \in \mathbb{D}^s$:

ces fonctions existent d'après le procédé d'orthogonalisation de Gram-Schmidt .

Nous noterons enfin par $\{ A_J \}$, $J \in \mathcal{D}$, la famille :

$$\{ A_{J_1}^1 \otimes A_{J_2}^2 \otimes \dots \otimes A_{J_p}^p \mid J_s \in \mathcal{D}^s \text{ pour tout } s \in (1, 2, \dots, p) \} .$$

Alors la propriété suivante est évidente .

Proposition 8-4 : La famille $\{ A_J \}$, $J \in \mathcal{D}$,
est une famille orthonormale de $\mathcal{L}^2(\mathbb{R}^p, \mu)$,
engendrant le même sous-espace que la famille $\{ 1_I \}$, $I \in \mathcal{J}$.

De plus , $A_{1,1}, \dots, 1 \equiv 1$.

8-5 Notations des coefficients des A_J

Nous noterons par α_J , $J \in \mathcal{D}$, le nombre $\alpha_J = E(A_J(X))$,

et par $\widehat{\alpha_J^n}$, l'estimateur de α_J (7-1),

$$\widehat{\alpha_J^n} = \frac{1}{n} \sum_{j=1}^n A_J(X_j) .$$

8-6 Notations des lois marginales

Nous noterons par \mathbb{D}_m^s , $s \in (1, 2, \dots, p)$, les sous-ensembles de \mathbb{D} :

$$: \mathbb{D}_m^s = \{ (i_1, i_2, \dots, i_p) \mid i_s \neq 1, i_j = 1 \text{ si } j \neq s \}$$

et par $\mathbb{D}_m = \bigcup_{s=1}^p \mathbb{D}_m^s$, et $\overline{\mathbb{D}} = \mathbb{D}^* \setminus \mathbb{D}_m$.

alors les coefficients des lois marginales sont les

$$\alpha_{J_0}, J_0 \in \mathbb{D}_m.$$

Proposition 8-7 Soit $J_0 \in \mathbb{D}_m$, alors $\alpha_{J_0} = 0$.

En effet, supposons, par exemple $J_0 = (i_1, 1, 1, \dots, 1)$, alors

$$\begin{aligned} \alpha_{J_0} &= \int \dots \int_{\mathbb{R}^p} A_{i_1, 1, 1, \dots, 1} \cdot dQ = \int \dots \int_{\mathbb{R}^p} A_{i_1} \otimes 1 \otimes 1 \otimes \dots \otimes 1 \cdot dQ \\ &= \int_{\mathbb{R}} A_{i_1} \cdot d\mu_1 = 0. \end{aligned}$$

Nous pouvons maintenant énoncer la proposition fondamentale de cette partie : celle-ci nous apprend que la statistique hilbertienne et la statistique du khi-deux sont identiques .

Théorème 8-8 : Gardons les notations précédentes et notons pour tout borélien B de \mathbb{R}^p , N_B l'entier $N_B = n \cdot \mu_e(B)$, alors, on a l'égalité :

$$n \cdot \left(\sum_{J \in \mathcal{D}}^* \widehat{\alpha}_J^n \right)^2 = \sum_{I \in \mathcal{J}} \frac{(N_I - n \cdot \mu(I))^2}{n \cdot \mu(I)}$$

Démonstration : On sait que l'on peut écrire $A_J = \sum_{I \in \mathcal{J}} \lambda_I \cdot 1_I$,

et donc $\frac{1}{n} \left(\sum_{i=1}^n A_J(x_i) \right) = \iint \dots \int A_J \cdot d\mu_e = \sum_{I \in \mathcal{J}} \lambda_I \cdot \mu_e(I) =$

$$\sum_{I \in \mathcal{J}} \lambda_I \cdot \left(\iint \dots \int_I \frac{\mu_e(I)}{\mu(I)} d\mu \right) = \langle f_e, A_J \rangle_{\mathcal{L}^2(\mathbb{R}^p, \mu)}$$

i.e. $\widehat{\alpha}_J^n = \langle f_e, A_J \rangle_{\mathcal{L}^2(\mathbb{R}^p, \mu)}$

Comme $\{A_J\}$, $J \in \mathcal{D}$, et $\{1_I\}$, $I \in \mathcal{J}$,

engendrent le même sous-espace, on peut écrire :

$$\sum_{J \in \mathcal{D}} \widehat{\alpha}_J^n \cdot A_J = f_e, \quad \text{i.e.} \quad \sum_{J \in \mathcal{D}}^* \widehat{\alpha}_J^n \cdot A_J = f_e^{-1}$$

et donc $n. \left(\sum_{J \in \mathcal{D}^*} \widehat{\alpha}_J^n \right)^2 = n. \left(\int \dots \int_{\mathbb{R}^p} (f_e - 1)^2 . d\mu \right) =$

$$n. \left(\sum_{I \in \mathcal{J}} \int \dots \int_I (f_e - 1)^2 . d\mu \right) = \sum_{I \in \mathcal{J}} \frac{(N_I - n \cdot \mu(I))^2}{n \cdot \mu(I)} .$$

2 : Test d'indépendance : cas de lois marginales connues

Etant donné que les coefficients α_{j_0} , $j_0 \in \mathbb{D}_m$, des lois marginales sont nuls, nous n'avons pas à les tester, et nous effectuerons donc le test hilbertien à partir de la statistique $n.(\sum \hat{\alpha}_j^2)$ ne portant pas sur ces coefficients, contrairement à ce que l'on fait avec le test classique du khi-deux.

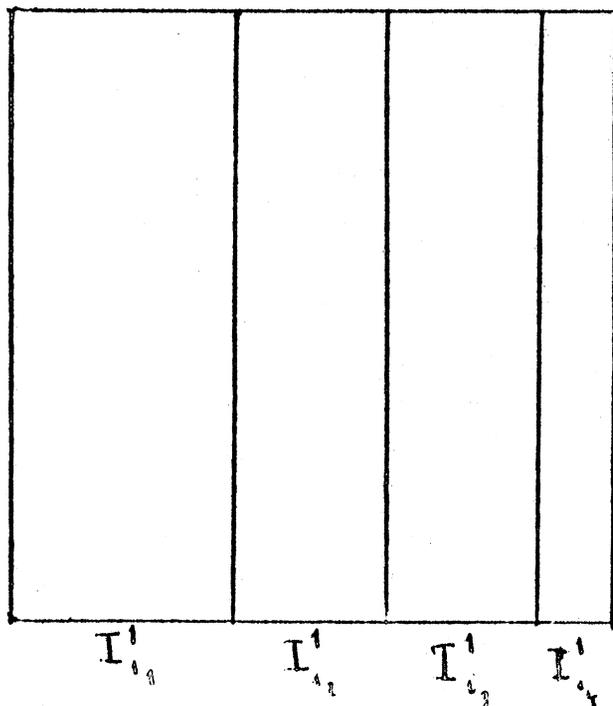
8-9 Notation du prolongement des partitions de \mathbb{R} à \mathbb{R}^p : Soit

$s \in (1, 2, \dots, p)$, nous noterons alors par \mathcal{J}_s la partition de \mathbb{R}^p

$$\mathcal{J}_s = \{ E_{i_1}^1 \times E_{i_2}^2 \times \dots \times E_{i_p}^p \mid i_s \in \mathbb{D}^s ; E_{i_s}^s = I_{i_s}^s ; E_{i_j}^j = \mathbb{R} \text{ si } j \neq s \}$$

et par $\mathcal{J}_0 : \mathcal{J}_0 = \mathcal{J}$.

cas $p=2 : (x, y) \in V : \mathcal{J}_1$



8-10 Notation des statistiques du khi-deux : Soit $s \in (0,1,\dots,p)$

nous noterons alors la statistique du khi-deux sur respectivement (\mathbb{R}^p, μ) , ou (\mathbb{R}, μ_i) , $i = 1,2,\dots,p$, prolongé à \mathbb{R}^p , par

$$\text{respectivement } \chi_i^2 = \sum_{I \in \mathcal{J}_i} \frac{(N_I - n \cdot \mu(I))^2}{n \cdot \mu(I)}, \quad i = 0,1,\dots,p.$$

on est alors en mesure d'énoncer le

Théorème 8-11 : Avec les notations précédentes, et en notant par $\|S_n\|^2$ la statistique hilbertienne définie par 7-1 :

$$\|S_n\|^2 = n \cdot \left(\sum_{J \in \mathcal{Q}} \widehat{\alpha_J^n} \right)^2,$$

$$\text{on a l'égalité : } \|S_n\|^2 = \chi_0^2 - \sum_{s=1}^p \chi_s^2.$$

c'est une conséquence directe de 8-8.

Corollaire 8-12 : Sous l'hypothèse d'indépendance, la statistique $\chi_0^2 - \sum_{s=1}^p \chi_s^2$ suit asymptotiquement une loi du khi-deux à

$$\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \quad \text{degrés de liberté.}$$

C'est le résultat général de Bosq

Corollaire 8-13 : Si $p = 2$, et si l'on note $k = c(1)$ et $k' = c(2)$, alors : $\| S_n \|^2 = \chi_0^2 - \chi_1^2 - \chi_2^2$, et suit asymptotiquement , sous l'hypothèse d'indépendance , une loi du khi-deux à $(k-1)(k'-1)$ degrés de liberté ,

8-14 : test d'indépendance : Il est clair que si l'on teste l'indépendance grâce au test hilbertien , dans le cas de lois marginales connues , on testera l'hypothèse

$H_0(\overline{\mathbb{D}}) : \ll \forall J \in \overline{\mathbb{D}} : \alpha_J = 0 \gg$ contre sa négation , grâce à la statistique $\| S_n \|^2$ de 8-11 et aux théorèmes 7-2 et 7-3 .

8-15 Commentaire : Il est clair aussi que dans ce cas , on ne teste pas les coefficients α_{J_0} , $J_0 \in \mathbb{D}_m$, car d'après 8-7 , nous savons déjà qu'ils sont nuls .

Par contre , dans le test classique du khi-deux , d'après 8-8 , on teste de nouveau ces coefficients que l'on connaît déjà .

8-16 Commentaire On peut aussi formuler les affirmations précédentes en faisant remarquer que , comme le test classique du khi-deux sert à tester l'hypothèse $H'_0(\mathbb{D}^*) : \ll \forall J \in \mathbb{D}^* : \alpha_J = 0 \gg$, il est utilisé en fait comme test de $C'_0 \ll \llcorner$ La loi de X est $\mu \gg$ contre sa négation .

Et comme l'on sait que $\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_p$, on en déduit l'indépendance. Il est alors clair que dans ce cas, on oublie certaines données : à savoir la connaissance des lois marginales.

Et, par contre, le test hilbertien permet quant à lui de tester l'hypothèse C_0 : « Les variables sont indépendantes » contre sa négation.

8-17 : Comparons maintenant les deux tests d'indépendance : le test classique du khi-deux et le test hilbertien d'indépendance.

En fait, étant donné 8-8, nous sommes en présence de deux tests hilbertiens :

d'une part, le test classique du khi-deux qui est donc absolument identique au test hilbertien de $H_0(\mathbb{D}^*)$ contre sa négation,

d'autre part, le test hilbertien d'indépendance qui est donc le test de $H_0(\overline{\mathbb{D}})$ contre sa négation.

Si l'on se rappelle que les α_j sont des paramètres ("presque") indépendants (4-18 et partie 3-1-3), on ne peut que dire que le test d'indépendance le plus naturel est le second, ce qui ne veut d'ailleurs pas dire que ce soit le plus puissant, bien qu'intuitivement on puisse espérer un résultat de cet ordre.

On remarque aussi bien sur que, sous l'hypothèse d'indépendance, la loi de $\|S_n\|^2$ est asymptotiquement une loi du khi-deux ayant le même nombre de degrés de liberté que la statistique du khi-deux dans le cas du test d'indépendance avec estimation de paramètres. Aussi allons nous maintenant étudier ce cas.

3 : Test d'indépendance : cas de lois marginales inconnues

Bosq n'a pas étudié le test hilbertien avec estimation de paramètres, cependant, comme on se rend compte très facilement que dans ce cas $\widehat{\chi}_X^2 = \widehat{\chi}_Y^2 = 0$, (si $p=2$), on en déduira aisément que le test d'indépendance le plus naturel est toujours donné par une statistique de la forme $\overline{\chi_{X,Y}^2} - \overline{\chi_X^2} - \overline{\chi_Y^2}$ (et non $\overline{\chi_{X,Y}^2}$) qui suit toujours asymptotiquement, sous l'hypothèse d'indépendance, une loi du khi-deux à $(k-1)(k'-1)$ degrés de liberté.

8-18 Notations Soit $s \in (1, 2, \dots, p)$,

nous noterons alors par $\widehat{\mu}_s$ la mesure empirique marginale, admettant une densité de probabilité constante sur chaque intervalle

$$I_i^s, i \in \mathbb{O}^s, \text{ et égale à } \widehat{\mu}_s(I_i^s) = \mu_e(I_i^s),$$

que nous supposerons strictement positif.

Nous noterons aussi par

$\widehat{\mu}$ la mesure produit $\widehat{\mu} = \widehat{\mu}_1 \otimes \widehat{\mu}_2 \otimes \dots \otimes \widehat{\mu}_p$,
 par $[A_j]$, $J \in \mathbb{O}$, les fonctions orthonormales associées,
 par $\widehat{\alpha}_j$, les coefficients de ces fonctions, et par $\widehat{\alpha}_j^n$ leurs estimateurs.

8-19 Rappel : Le test du khi-deux d'indépendance dans le cas de lois marginales inconnues s'opère à partir de la statistique du khi-deux avec estimation de paramètres :

$$\hat{\chi}_0^2 = \sum_{I \in J} \frac{(N_I - n \cdot \hat{\mu}(I))^2}{n \cdot \hat{\mu}(I)}$$

qui suit asymptotiquement, sous l'hypothèse d'indépendance, une loi du khi-deux à

$$\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right) \text{ degrés de liberté .}$$

Proposition 8-20 : Les statistiques du khi-deux des lois marginales :

$$\hat{\chi}_s^2 = \sum_{I \in J_s} \frac{(N_I - n \cdot \hat{\mu}(I))^2}{n \cdot \hat{\mu}(I)}, \quad s \neq 0, \quad \text{sont identiquement nulles .}$$

La démonstration est triviale .

Proposition 8-21 : On a les égalités

$$\hat{\chi}_0^2 = n \cdot \left(\sum_{J \in \mathcal{O}^*} \hat{\alpha}_J^n \right)^2 \quad \text{et} \quad \hat{\chi}_s^2 = n \cdot \left(\sum_{J_0 \in \mathcal{O}_m^s} \hat{\alpha}_{J_0}^n \right)^2 = 0 .$$

C'est exactement la même démonstration que 8-8 en remplaçant μ par $\hat{\mu}$.

Proposition 8-22 : Quel que soit $J_0 \in \mathbb{D}_m$, on a $\widehat{\alpha}_{J_0}^n = 0$.

C'est une conséquence directe de 8-20 et 8-21.

Proposition 8-23 : En notant $\|\widehat{S}_n\|^2$ la statistique hilbertienne : $\|\widehat{S}_n\|^2 = n \cdot \left(\sum_{J \in \mathbb{D}} \widehat{\alpha}_J^2 \right)$,

alors, on a l'égalité $\|\widehat{S}_n\|^2 = \widehat{\chi}_0^2 - \sum_{s=1}^p \widehat{\chi}_s^2 = \widehat{\chi}_0^2$,

et $\|\widehat{S}_n\|^2$ suit asymptotiquement,

sous l'hypothèse d'indépendance, une loi du khi-deux à

$\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right)$ degrés de liberté.

C'est une autre forme des théorèmes précédents

Nous allons maintenant démontrer que les $\widehat{\alpha}_J^n$ sont des estimateurs convergents presque sûrement des α_J .

Il nous faudra cependant imposer certaines contraintes évidentes : à savoir que les A_J et les \widehat{A}_J soient construits par la même méthode.

Aussi, démontrerons nous ce résultat dans un cas particulier, mais il est évident que le procédé reste général.

8-24 Définition d'une suite de familles orthonormales

Nous définirons la suite $\{ B_j^S \}$, $j \in \mathbb{D}^S$,
 pour tout $s \in (1, 2, \dots, p)$, $B_j^S \in \mathcal{L}^2(\mathbb{R}, \mu_s)$, par

$$B_1^S \equiv 1, \text{ et}$$

$$B_j^S(x) = \begin{cases} \dots\dots\dots 0 & \text{si } x \in \bigcup_{t=1}^{j-1} I_t^S, \\ \dots\dots \sum_{t=j+1}^{c(s)} \mu_s(I_t^S) & \text{si } x \in I_j^S, \\ \dots\dots\dots -(\mu_s(I_j^S)) & \text{si } x \in \bigcup_{t=j+1}^{c(s)} I_t^S, \end{cases}$$

si $j \in \mathbb{D}^{*s}$;

et nous noterons désormais par

A_j^S les fonctions B_j^S normalisées dans $\mathcal{L}^2(\mathbb{R}, \mu_s)$:

$$A_j^S = \frac{B_j^S}{\sqrt{\left(\sum_{t=j+1}^{c(s)} \mu_s(I_t^S) \right) \left(\mu_s(I_j^S) \right) \left(\sum_{t=j}^{c(s)} \mu_s(I_t^S) \right)}}$$

Nous définirons alors A_j , $j \in \mathbb{D}$, comme en 8-3, et de même \widehat{A}_j en remplaçant μ_s par $\widehat{\mu}_s$ dans les définitions ci-dessus.

Il est clair que ces familles vérifient 8-4.

Proposition 8-25 : Notons les décompositions de A_J et \widehat{A}_J par rapport à la base $\{1_I\}, I \in J$, par

$$A_J = \sum_{I \in J} \lambda_I \cdot 1_I \quad \text{et} \quad \widehat{A}_J = \sum_{I \in J} \widehat{\lambda}_I \cdot 1_I,$$

alors $\widehat{\lambda}_I$ converge presque sûrement vers λ_I .

C'est presque évident .

Proposition 8-26 : Pour tout $J \in \mathcal{D}^*$, l'estimateur $\widehat{\alpha}_J^n$ converge presque sûrement vers α_J .

C'est une conséquence simple de 8-26 .

8-27 : Test d'indépendance : On sait que si l'on teste l'indépendance par le test hilbertien, on teste l'hypothèse

$$H_0 : \ll \forall J \in \overline{\mathcal{D}} : \alpha_J = 0 \gg \quad \text{contre sa négation .}$$

Il est clair que le test classique du khi-deux avec estimation de paramètres teste ces coefficients et eux seuls, et que dans ce cas la statistique hilbertienne d'indépendance et la statistique du khi-deux sont identiques :

$$\| \widehat{S}_n \|^2 = \widehat{\chi}_0^2 .$$

8-28 : Il est clair également que cette égalité provient des égalités $\widehat{\chi}_s^2 = 0$ lorsque $s \in (1, 2, \dots, p)$.

8-29 : Il est clair aussi que dans ce cas on a encore l'égalité

$$\| \widehat{S}_n \|^2 = \widehat{\chi}_0^2 - \sum_{s=1}^p \widehat{\chi}_s^2 ,$$

et que $\| \widehat{S}_n \|^2$ suit asymptotiquement une loi du khi-deux à

$$\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \text{ degrés de liberté .}$$

8-30 : il est enfin tout aussi clair que dans le cas de l'estimation de paramètres , on ne peut tester l'hypothèse

$$C'_0 : \ll \text{La loi de } X \text{ est } \mu \gg .$$

4 : Conclusion

On déduit facilement des résultats précédents que

1°) Que l'on connaisse les lois marginales μ_s , $s = 1, 2, \dots, p$, ou que l'on n'en connaisse aucune, le test d'indépendance le plus naturel repose toujours sur une statistique de la forme

$$\overline{\chi_0^2} = \sum_{s=1}^p \overline{\chi_s^2} \quad (\overline{\chi_{X,Y}^2} - \overline{\chi_X^2} - \overline{\chi_Y^2} \text{ si } p=2) - \text{ et non } \overline{\chi_0^2} -$$

qui, sous l'hypothèse d'indépendance, suit toujours asymptotiquement une loi du khi-deux à

$$\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \prod_{s=1}^p (c(s) - 1) \right) \text{ degrés de liberté}$$

$$((k-1)(k'-1) \text{ degrés de liberté si } p=2) .$$

2°) La statistique $\overline{\chi_0^2}$ n'est pas la statistique la plus naturelle à utiliser pour le test d'indépendance car elle conduit à tester des coefficients déjà connus. Cette statistique est en fait la statistique associée au test de l'hypothèse

$$C'_0 : \ll \text{La loi de } X \text{ est } \mu \gg \text{ contre sa négation}$$

3°) Les résultats obtenus ci-dessus l'ont été, soit sous l'hypothèse où toutes les lois marginales sont connues, soit sous l'hypothèse où toutes les lois marginales sont inconnues. Mais, étant donné leur identité, on peut espérer qu'ils resteront vérifiés si q lois marginales sont connues et $p-q$ inconnues.

4°) Il faut remarquer aussi que les résultats précédents font apparaître de façon naturelle les différences de degrés de liberté entre les différentes statistiques utilisées

Les degrés de liberté sont en fait le nombre de paramètres réellement testés, l'estimation de paramètres se traduisant ici par l'annulation des coefficients $\hat{\alpha}_{j_0}^n$, $j_0 \in \mathbb{D}_m$, ce qui revient en fait à ne plus tester les α_{j_0} correspondants.

On peut se demander si cette méthode ne peut se généraliser de façon à fournir un nouveau mode d'étude de l'effet de l'estimation des paramètres dans les lois du khi-deux. La question est ouverte.

PARTIE 3 : CONCLUSIONS GENERALES SUR LES RAPPORTS
ENTRE LA DEPENDANCE ET LES COEFFICIENTS DES SERIES

Les résultats acquis vont nous permettre maintenant de tirer les conclusions d'ordre général sur les rapports entre séries et dépendance .

Dans une première partie , nous rappellerons quelques considérations générales sur la dépendance , ensuite nous concluerons sur les coefficients D_h^N , enfin , nous rappellerons nos résultats sur les séries et nous en tirerons les conséquences .

Nous développerons ces résultats dans une deuxième partie en étudiant le cas concret des congruences linéaires générateurs de nombres pseudo-aléatoires .

Enfin nous rappellerons nos conclusions générales dans une troisième et dernière partie .

1 : DIVERS MODES DE MESURE DE LA DEPENDANCE

1 : Considérations sur la dépendance

Le principal problème de la dépendance est celui de sa mesure : l'indépendance, elle, possède une définition mathématique précise, mais il est beaucoup plus difficile de savoir comment étalonner les écarts à l'indépendance et éventuellement les comparer. Jusqu'à ce jour, on a créé beaucoup de mesures de dépendance : aucune ne s'est avérée parfaite.

Dans cette étude, nous avons eu recours aux coefficients D_h^N , mais on peut tout aussi bien se servir de plusieurs coefficients au lieu d'un seul : les $\rho_{i,j}$, par exemple.

Et comme on ne peut les prendre tous, on se contentera de prendre les premiers d'entre eux que l'on estime mesurer les écarts à l'indépendance les plus importants : ce que l'on fait souvent en se contentant d'étudier $\rho_{1,1} = \rho(X,Y)$ et la dépendance linéaire.

Mais il est bien clair que c'est là un critère subjectif et, qu'en fait, on se contente de mesurer la dépendance qui nous intéresse en tant qu'observateur.

Reprenons le cas de l'exemple 1-12, où, à cause des erreurs d'arrondis des calculs, nous trouvons pour tout borélien $B_1 \times B_2$ de $[0,1]^2$:

$$P((X_q, Y_q) \in B_1 \times B_2) = (P(X_q \in B_1))(P(Y_q \in B_2)) ,$$

bien qu'il existe une bijection entre X_q et Y_q .

Alors , comment faut-il mesurer cette dépendance ?

Par 1 (dans le cas de coefficients normalisés) ou par un ϵ très petit ?

En fait , il est bien évident que ce choix dépend des besoins de l'utilisateur : mathématiquement , nous devrions - vu la bijection - la mesurer par 1 , mais l'observateur qui , lui , constatera la vérification de l'égalité d'indépendance ci-dessus , pourra très bien se dire que la dépendance qui l'intéresse lui n'existe pas , ou à peu près pas , et il aimerait la mesurer par ϵ .

En conclusion , on voit qu'il ne peut y avoir une mesure de dépendance , et c'est pourquoi toutes sont à priori justifiables .

2 : Les indicateurs D_h^N

Nous allons maintenant donner les raisons pour lesquelles nous avons eu recours à ces coefficients pour mesurer la dépendance .

a) Propriétés générales

Rappelons d'abord qu'ils ont de bonnes propriétés mathématiques : nous les avons énoncées dans la partie 1 . En particulier , ils vérifient les axiomes de Renyi (1-5,1-6) ; de plus , ils sont faciles à estimer ($D, [2]$) ; enfin ils sont comparables aux indices classiques (1-9,1-10,1-11) .

Ajoutons qu'ils ont un bon comportement par rapport à la taille des échantillons . Nous voulons dire par là que si un échantillon est de taille suffisante , l'estimation que nous pourrons faire de D_h^N sera à priori satisfaisante .

Ainsi , dans le cas de l'exemple 1-12 , bien qu'il y ait bijection

entre X_q et Y_q on aura $D_h^N(X_q, Y_q) \approx 0$,

et c'est l'estimation que nous obtiendrons en principe ,

alors que $d_h(X_q, Y_q) = +\infty$,

et que nous estimerons - vu l'égalité d'indépendance - $d_h(X_q, Y_q) = 0$.

On peut traduire ce phénomène en disant que les indicateurs D_h^N correspondent à un aspect pratique et quasi-intuitif de la dépendance : l'envie de dire que lorsque des échantillons de taille élevée vérifient l'indépendance , la dépendance devrait , elle , être proche de 0 . Il est évident que ce n'est pas l'aspect probabiliste , et que c'est donc un choix que nous faisons en préférant cet aspect pratique .

b) Rapport avec les séries

Posons nous maintenant la question de savoir si le fait de prendre ces indicateurs pour étudier la dépendance associée aux coefficients du développement en série de la densité de probabilité était une décision judicieuse .

Bien entendu , il s'agit encore ici d'un choix . Celui-ci se justifie , d'une part , par les résultats obtenus (les coefficients d_h ne possèdent pas de telles propriétés : cf 1-12 , par exemple) . D'autre part , nous venons de voir qu'ils répondent à une certaine idée de la dépendance . Cela se traduira dans leurs rapports avec les séries à ce qu'ils seront surtout sensibles , aux coefficients $\theta_{i,j}$, par exemple , tels que $i+j$ soit assez petit , et donc , en particulier , à $\theta_{1,1} = \rho(X,Y)$, le coefficient de corrélation , ce qui est bien conforme aux pratiques et aux espérances habituelles .

Concluons en disant que les résultats obtenus justifient a posteriori l'utilisation de ces indicateurs pour mesurer la dépendance induite par les coefficients des séries .

3 : Mesure de la dépendance par les coefficients des séries

Si , en général , on mesure la dépendance par un seul nombre , rien n'empêche de la mesurer par plusieurs ou par une infinité . Ainsi , par exemple , par les $\rho_{i,j}$, $(i,j) \in \underline{\mathbb{N}}^{*2}$.

a) Propriétés générales

En fait , chacun de ces coefficients mesurera une dépendance particulière : $\rho_{1,1}$, la dépendance linéaire ,
 $\rho_{1,2}$, $\rho_{2,1}$, $\rho_{2,2}$, des dépendances quadratiques , etc ...

De plus , ces dépendances seront "à peu près indépendantes" les unes des autres : la seule contrainte qui pèsera sur elles sera en fait $f \geq 0$ - ou F croissante (cf 4-18 ou ci-après paragraphe d) . Cette indépendance se traduira lors du test hilbertien en degrés de liberté .

Enfin ces coefficients sont normalisés (2-6,4-17) .

b) Choix de la série

Comme , bien sûr , concrètement , on ne pourra prendre qu'un nombre fini de termes , il convient de se rappeler qu'il faut être prudent pour le choix de ceux-ci : les premiers termes de certaines séries pourraient laisser passer certaines dépendances que l'on estime grossières .

Ainsi , la série ayant pour premiers termes les fonctions orthogonales $A_{2^i, j} \otimes A_{2^{i'}, j'}$ de 7-18 ne percevra pas la dépendance linéaire de X et Y , lorsque $X = Y$, si on se limite à étudier ces premiers termes .

c) Nombre de fonctions orthogonales nécessaires

Posons nous maintenant la question de savoir si le grand nombre de fonctions orthogonales nécessaires pour obtenir que les coefficients D_h^N soient assez petits (cf partie 2-4-3 : 2,3,4) est justifié .

Il semble bien qu'il faille répondre par oui .

Le cas des fonctions de Walsh - que nous allons rappeler ci-après - qui est en fait celui des fonctions indicatrices de carrés de $[0,1]^2$, et avec lequel nous avons obtenu les meilleures majorations , montre bien que sinon des dépendances simples nous échapperont .

D'autre part , le fait que la dépendance de p variables aléatoires - ou l'uniformité dans $[0,1]^p$ - demande pour être testée l'introduction de partitions (et d'échantillons) de tailles de plus en plus élevées lorsque p croît abonde dans ce sens (cf partie 2-4-3-5)

Rappelons aussi que , sauf dans le cas des fonctions trigonométriques , les majorations que nous avons données des indicateurs $D_h^N(X, Y)$ ou de $\delta_1(F, U)$ ne sont pas trop grossières (3-16, 3-17, 6-2, [3]) . En particulier la majoration de $D^N(X, Y)$ pour les fonctions de Walsh est presque exacte (cf Pg 17, [2]) .

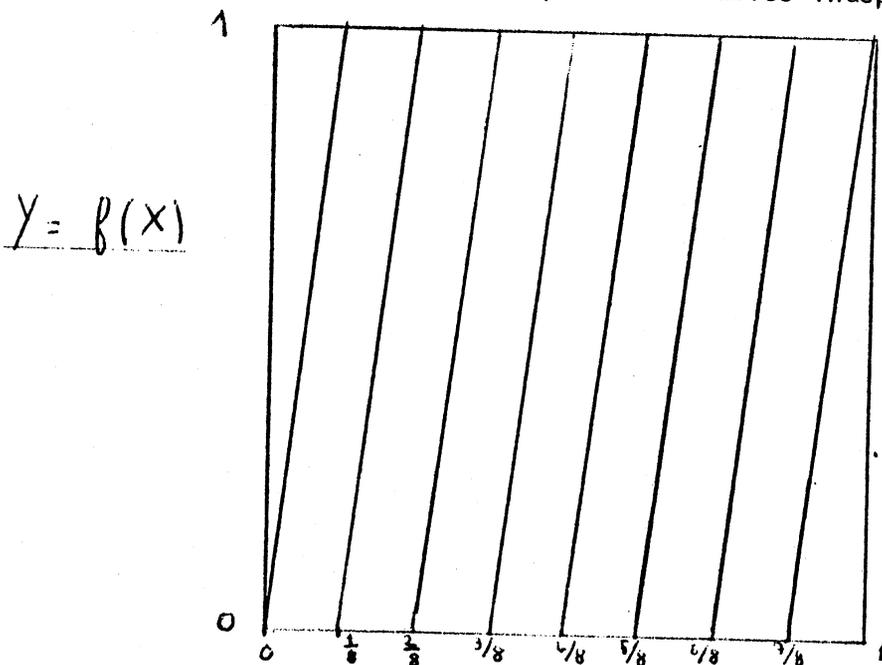
Finalement , il nous paraît normal , si l'on veut tester une indépendance de qualité de faire intervenir un grand nombre de fonctions orthogonales et des échantillons de tailles élevées .

Ce point est à souligner car il nous est apparu au cours de nos recherches que très fréquemment , lors du test du khi-deux par exemple , on ne prenait pas le nombre de classes nécessaire pour s'assurer de l'indépendance voulue (cf partie 3-2-3) .

Rappelons que l'ordre du développement , et donc le nombre de classes , doivent être choisis en fonction des besoins des utilisateurs ; nous donnons un exemple concret en partie 3-2 .

Ainsi , si l'on se contente d'une "indépendance" : " $D^N(X,Y) < \frac{1}{8}$ " , il faudra un découpage de $[0,1]^2$ en 256 carrés égaux .

Remarquons que dans ce cas , on considère que $(X,Y) \in V^U$, défini par le graphique ci-dessous est un couple de variables indépendantes .

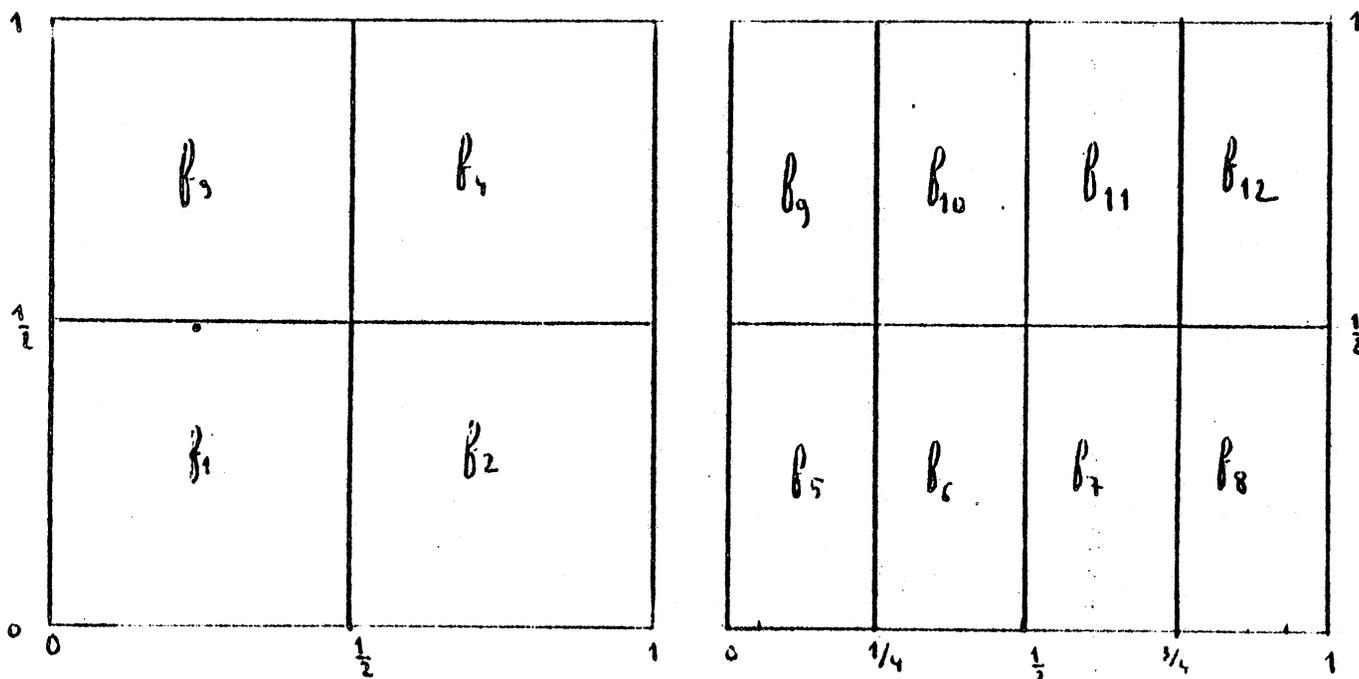


Par contre , si l'on estime que " $D^N(X,Y) < 2^{-7} = \frac{1}{128} = 0,0078$ " est à peine suffisant comme critère d'indépendance , il faudra recourir à une partition de $[0,1]^2$ en au moins 2^{16} carrés ($2^{16} = 65.536$!) .

d) Exemple : développement en fonctions de Walsh

Nous allons donc voir maintenant les différentes sortes de dépendance mesurée par les fonctions de Walsh .

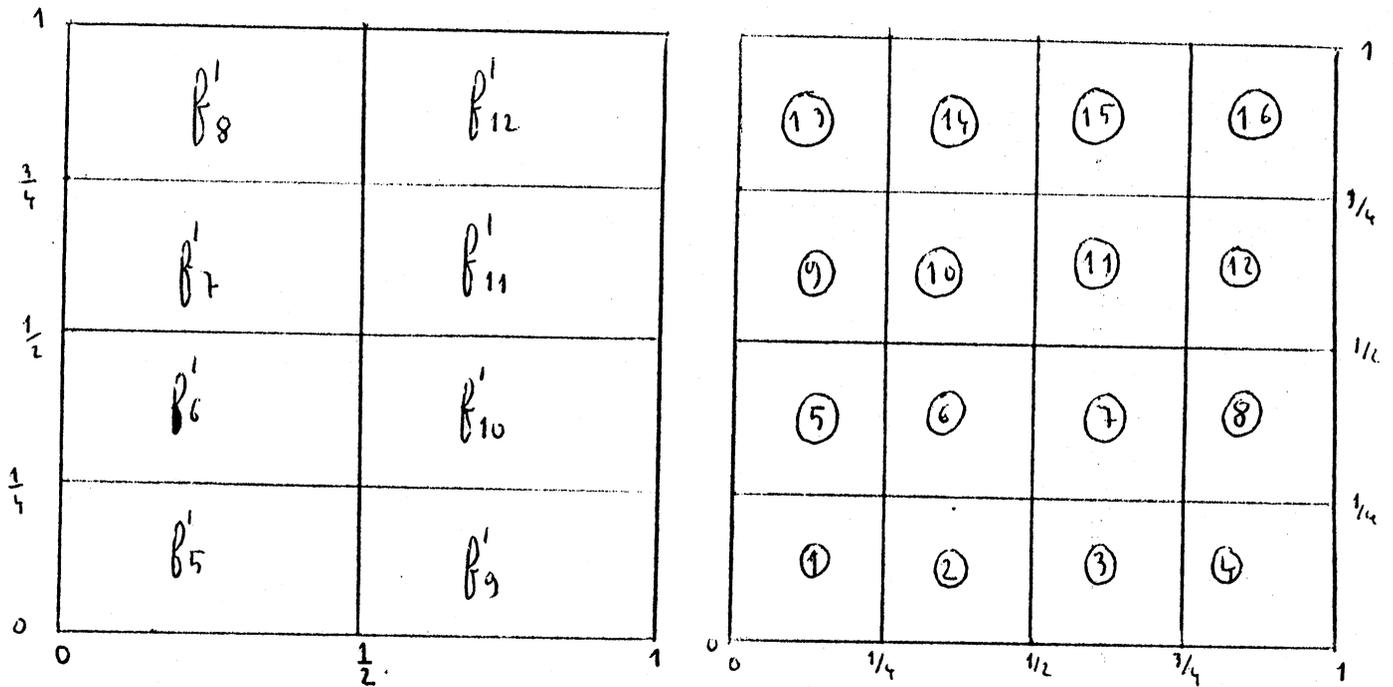
Dans les graphiques ci-dessous les f_i ou les \textcircled{i} mesurent les densités des rectangles R_i . Les K_i désigneront des constantes dont les définitions exactes sont donnés par les formules de [3] .



$$K_1 \cdot \alpha_{1,1}^{1,1} = ((f_1 + f_4) - (f_2 + f_3)) \quad ,$$

$$K_2 \cdot \alpha_{1,1}^{2,1} = ((f_5 + f_{10}) - (f_6 + f_9)) \quad , \quad K_2 \cdot \alpha_{1,1}^{2,2} = ((f_7 + f_{12}) - (f_8 + f_{11}))$$

$$K_2 \cdot \alpha_{2,1}^{1,1} = ((f'_5 + f'_{10}) - (f'_6 + f'_9)) \quad , \quad K_2 \cdot \alpha_{2,2}^{1,1} = ((f'_7 + f'_{12}) - (f'_{11} + f'_8))$$



$$K_3 \cdot \alpha_{2,1}^{2,1} = ((1) + (6)) - ((2) + (5)) \quad , \quad K_3 \cdot \alpha_{2,2}^{2,1} = ((13) + (10)) - ((9) + (14))$$

$$K_3 \cdot \alpha_{2,1}^{2,2} = ((3) + (8)) - ((7) + (4)) \quad , \quad K_3 \cdot \alpha_{2,2}^{2,2} = ((11) + (16)) - ((15) + (12))$$

Il est donc facile de voir quelle dépendance est mesurée par les coefficients $\alpha_{2^i, j'}^{2^i, j}$.

Par exemple, $\alpha_{2,1}^{2,1}$ mesure la différence - à une constante de normalisation près - de la densité des carrés 1 et 6 et des carrés 2 et 5, différence qui est nulle lorsque il y a indépendance.

Il est clair que ce sont des dépendances que l'on estime intuitivement de plus en plus fines, et nous rappelons que nous considérons que ces séries sont bien ordonnées (6-5, [3]).

Rappelons aussi que ces mesures de dépendance sont "à peu près indépendantes" les une des autres : dans le sens où le fait que $\alpha_{1,2}^{2,2} = 0,5$, par exemple, n'influera en rien sur la valeur des autres coefficients : la seule contrainte qu'ils devront respecter étant $f \geq 0$ - ou F croissante (cf 4-18) .

Remarquons que la connaissance des coefficients $\alpha_{2^i, 2^{j'}}^{2^i, 2^j}$, $i < l$ et $i' < l$, fournit la connaissance des densités approximées sur les carrés ci-dessus de largeur 0,25 (cf 5-10, [3]) et il a donc fallu 9 fonctions orthogonales (dans le cas où l'on sait que les lois marginales sont uniformes) pour obtenir l'inégalité $D^N(X,Y) < 0,5$.

Il est donc bien clair que, d'une part les coefficients $\alpha_{2^i, 2^{j'}}^{2^i, 2^j}$ nous fournissent des indications que l'on estime en général de plus en plus fines sur la dépendance, et que d'autre part, la connaissance de celle-ci demande l'introduction d'un nombre élevé, mais intuitivement normal de fonctions orthogonales .

2 : EXEMPLE CONCRET : TEST DES CONGRUENCES LINEAIRES
GENERATEURS DE NOMBRES PSEUDO-ALEATOIRES

Nous allons voir maintenant comment mettre en pratique nos résultats dans ce cas concret . Cela nous permettra de préciser un peu les résultats de [4] .

Nous rappelons d'abord que nous désirions obtenir à partir des congruences linéaires des simulations convenables de la loi uniforme sur $[0,1]$. Pour cela , nous avons étudié l'indépendance de p nombres simulés successifs , $p \in \mathbb{N}^*$, et ce , de façon exhaustive dans le cas $p = 2$ (i.e. l'uniformité dans $[0,1]^2$) .

1 : rappel des résultats acquis

1-1 Cas de $[0,1]^2$

Soit T une congruence linéaire inversible : $T(x) \equiv ax + b \pmod{m}$,
avec $a, b, m \in \mathbb{N}$,

et soit t l'application associée de $[0,1]$ dans lui-même vérifiant

$$t\left(\frac{k}{m}\right) = \frac{1}{m} \overline{T(k)} ,$$

pour tout $k \in (0,1,\dots,(m-1))$,

et où \overline{r} , $r \in \mathbb{Z}$,

désigne l'unique entier congru à r modulo m tel que $0 < \overline{r} < m$.

Soit $n \in \mathbb{N}^*$ et $(X, t(X), \dots, t^{n-1}(X))$,
un n -pseudo-échantillon simulé (où $X = \frac{k}{m}$, $k \in \mathbb{N}$) ,
et soit P la probabilité empirique associée .

Nous noterons alors par \mathcal{R} l'ensemble des rectangles de $[0,1]^2$ de
cotés parallèles aux axes .

Alors , pour tout $R \in \mathcal{R}$, nous noterons par S_R son aire ,
par N_R le nombre de points du pseudo-échantillon qu'il contient ,

et par i_R et G_R les nombres

$$i_R = \frac{N_R}{n} - S_R, \quad G_R = \frac{|N_R - n \cdot S_R|}{\sqrt{n \cdot S_R \cdot (1 - S_R)}}.$$

On remarque que $i_R = P((X, T(X)) \in R) - S_R$,

et, en particulier si $R = C_{j,j}^{4q}$ (Walsh), $i_R = f_{q-1}^{-1}$.

Alors, il existe deux applications B' et B'' de \mathbb{N}^3 dans \mathbb{R}_+ qui, à toute congruence linéaire T ($T \rightarrow (a, b, m)$), associent les réels $B'(T)$ et $B''(T)$ vérifiant

$$\forall R \in \mathcal{R} : B'(T) < S_R < B''(T) \implies G_R < 2,57,$$

$$\text{et }]R' : S_{R'} = s \in [0, B'(T) \cap]B''(T), 1], \text{ et } G_{R'} > 2,57.$$

ce qui revient à dire que sur tous les rectangles de \mathcal{R} de surface comprise entre $B'(T)$ et $B''(T)$, on ne détectera pas de rupture d'indépendance, si l'on prend comme référence le fait que, lorsque elle est associée à un vecteur aléatoire de loi uniforme sur $[0, 1]^2$, la statistique G_R est la valeur absolue d'une statistique approximativement gaussienne centrée réduite, 2,57 étant alors le seuil à 1% de la loi associée.

Et par contre, il existera des rectangles de surface inférieure à $B'(T)$ où cette indépendance n'est pas vérifiée.

Nous allons par la suite nous intéresser uniquement au cas d'écart à l'indépendance existant sur des rectangles R' : $S_{R'} < B'(T)$.

En effet , le cas $S_{R'} > B''(T)$, d'une part , se ramène à celui-ci par passage au complémentaire , et , d'autre part , ne se retrouve plus lorsque l'on passe à l'étude des développements en série du type de Walsh , par exemple : les carrés associés au développement en fonctions de Walsh à l'ordre n sont de surface $4^{-(n+1)}$, c'est à dire assez petite et bien inférieure à $B''(T)$.

1-2 Cas de $[0,1]^p$

Les résultats acquis dans ce cas ne sont pas aussi précis que les précédents ; de plus ils n'entrent pas directement dans le cadre de cette thèse , aussi nous contenterons nous de signaler qu'il existe des critères nécessaires (qui peuvent éventuellement se révéler aussi suffisants) pour assurer l'indépendance des pseudo-échantillons de taille p .

1-3 Avertissement

Il est bien entendu que les résultats cités précédemment ne sont qu'un résumé des résultats acquis , et que sur ce sujet de nombreuses études restent à faire . Mais cela n'a pas d'importance ici , où il s'agit uniquement de montrer un lien entre la dépendance et l'ordre des développements en série de la densité de probabilité .

2 : Applications

2-1 : Critère de choix de l'ordre du développement

Pour opérer celui-ci , rappelons -nous ce que désire un utilisateur de simulation : posséder une simulation convenable de la loi uniforme sur $[0,1]$ pour un pseudo-échantillon de taille donnée (celle dont il a besoin) ; et cela lui sera complètement indifférent que pour des échantillons de taille supérieure (ou pour des simulations d'autres lois) , on puisse trouver à partir de la congruence utilisée des ruptures d'indépendance spectaculaires .

Ceci étant rappelé , voyons maintenant la

2-2 : Première conséquence de nos résultats

Nous allons maintenant faire apparaître un lien entre les paramètres de T et la taille maximale des échantillons utilisables .

Supposons que nous possédions une congruence vérifiant les autres conditions d'indépendance (de p nombres simulés successifs , $p = 3,4,\dots,n$,) , alors le carré $[0,1]^2$ se remplira de façon à peu près régulière . Nous voulons dire par là que les points fournis par le générateur ne se concentreront pas d'abord dans certains sous-ensembles , et , qu'avant de remplir les rectangles R' tels que $S_{R'} < B'(T)$, ils rempliront de façon à peu près uniforme les rectangles R de surfaces supérieures à $B'(T)$.

On ne pourra donc commencer à déceler la dépendance existant dans certains rectangles R' que si ceux-ci contiennent un certain nombre de points .

Par exemple $n.S_{R'} \geq 16$, si l'on veut tester la répartition des points de $[0,1]^2$ grâce au test du khi-deux avec une partition de R' en deux sous-ensembles et la contrainte $np_i \geq 8$ pour assurer une bonne approximation de la loi multinomiale .

Et donc lorsque $n < \frac{16}{B^1(T)}$, on ne pourra pas , par définition , déceler cette dépendance .

Donnons un exemple concret ([4]) : le générateur T_L :

$$T_L(x) \equiv 3\,125x + 1 \pmod{2^{35}} .$$

Ce générateur a été utilisé par Laborde pour la construction des "nombres aux hasard" des "Tables statistiques et financières" [13] , et donc il semble normal de lui reconnaître a priori une certaine qualité .

Or , il est facile de voir qu'il existe des rectangles $R'' \in \mathcal{R}$, de hauteur 1 et de largeur $\frac{1}{a} = \frac{1}{3\,125}$, tels que les points $(X, T(X))$ fournis par le générateur soient concentrés sur la première diagonale .

Alors si l'on prend tous les m points qu'il est possible d'engendrer ainsi dans $[0,1]^2$, on trouvera plus de $1,1 \times 10^7$ points concentrés dans ces rectangles R'' , et donc uniquement sur cette diagonale , ce qui constituera une rupture d'indépendance énorme .

Or cette rupture d'indépendance ne devrait commencer à apparaître que lorsque $n > \frac{16}{S_R''}$, i.e. $n > 50\ 000$; et donc un utilisateur de pseudo-échantillon de taille $n < 50\ 000$, pourra très bien utiliser ce générateur, et il lui sera complètement indifférent de savoir que pour les échantillons de taille supérieure des ruptures d'indépendance nettes apparaissent.

Et donc concrètement, un utilisateur de pseudo-échantillon de taille n se contentera de vérifier que : $n.B'(T) < C^{te}$ ($C^{te} = 16$ dans le choix opéré ci-dessus).

Rappelons que nous ne connaissons pas les critères suffisants pour assurer l'indépendance de p nombres simulés successifs, ni dans le cas général, ni dans le cas du générateur T_L , et qu'il ne s'agit ici que de montrer le lien entre la surface des rectangles et la taille de l'échantillon et, accessoirement, de donner une méthode éventuelle pour assurer la qualité des n -échantillons.

Il est simple de faire le lien entre la surface des rectangles et l'ordre du développement en série.

Plaçons nous dans le cas des fonctions de Walsh. Il sera alors inutile, d'après les résultats précédents d'étudier les carrés

$$C_{j,j'}^{4^{q+1}} = I_{2^{q+1},j} \times I_{2^{q+1},j'} \quad (3-13)$$

de surface $4^{-(q+1)}$ tels que

$$\frac{1}{4^q} < \frac{16}{n} \quad (\text{i.e. } q > \frac{\log(\frac{n}{16})}{\log(4)})$$

Donc il suffira de choisir k , l'ordre du développement, tel que

$$k-1 < \frac{\log(\frac{n}{16})}{\log(4)}$$

et il sera donc inutile - en ce qui concerne les carrés $C_{j,j'}^{4^d}$ - de développer la série plus loin que cet ordre.

2-3 : Deuxième conséquence

Nous venons de voir dans le cas de T_L l'inutilité pour un utilisateur d'échantillon de taille inférieure à 50 000 de connaître la répartition des points engendrés par la simulation dans les rectangles R'' . Par contre, il est bien évident que si l'on désire des échantillons de taille 10^6 par exemple, et si l'on n'effectue pas les tests sur des rectangles assez petits, on laissera passer les dépendances citées plus haut, et on utilisera T_L pour engendrer de tels échantillons, ce qui sera une erreur.

Cela veut dire que si les conditions d'indépendance de p nombres simulés successifs sont assurées pour $p = 3, 4, \dots, n$, la condition $n.B'(T) \leq C^{te}$ (ainsi que $n.(1-B''(T)) \leq C^{te2}$, mais celle-ci se déduit de la première) sera vraisemblablement une condition nécessaire et suffisante pour assurer la qualité du n -pseudo-échantillon .

Pour l'ordre du développement en fonctions de Walsh , cela se traduira par le fait que l'on ne pourra pas prendre k tel que :

$$k < \frac{\log\left(\frac{n}{16}\right)}{\log(4)} .$$

Cela déterminerait k de manière unique si l'on ne faisait pas une

2-4 : Remarque importante

Nous sommes passés dans les paragraphes précédents d'une façon trop simpliste de la surface des rectangles à l'ordre des développements en fonctions de Walsh , car les résultats cités plus haut portent sur tous les rectangles $R \in \mathcal{R}$, et pas uniquement sur les carrés $C_{j,j'}^{4^d}$.

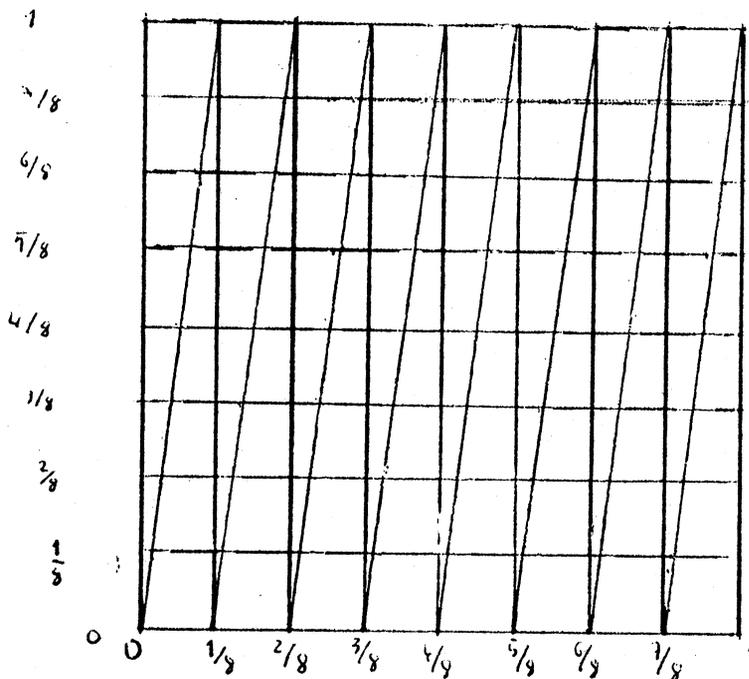
En fait , si nous ne possédions pas ces résultats , il faudrait effectuer des tests , soit sur tous les développements en série de fonctions indicatrices 1_R , $R \in \mathcal{R}$, et donc concrètement sur un nombre convenable de ceux-ci , soit pousser le développement en fonctions de Walsh jusqu'à un ordre plus élevé de façon à pouvoir assurer l'uniformité désirée (il faudrait vraisemblablement se servir de 3-18) .

Si nous n'agissions pas ainsi , il existerait certaines dépendances - i.e. certains rectangles $R \in \mathcal{R}$ - que nous n'étudierions pas et nous ne serions donc plus en mesure de garantir la qualité de la simulation pour les échantillons de taille désirée .

3 : Conséquences pratiques

Nous venons de voir comment calculer l'ordre du développement . Or , il nous faut bien faire remarquer que trop souvent on choisit cet ordre trop petit : par exemple , en prenant une partition trop grossière de $[0,1]^2$ pour y opérer le test du khi-deux , et l'on obtient alors des conclusions complètement erronnées .

Supposons que nous testions par exemple la simulation engendrée par $T(x) \equiv 2^3 x \pmod{1}$, par le développement en fonctions de Walsh jusqu'à l'ordre 2 - i.e. sur les carrés $C_{j,j}^{64}$, de largeur 0,125 .



Il est clair que dans ce cas , on aura $\alpha_{2^i, j}^{2^{i'}, j'} = 0$ pour i et i' inférieurs ou égaux à 2 , et que c'est la nullité de ces coefficients que l'on testera .

Supposons de plus que m soit grand , de l'ordre de 10^8 par exemple , que t , la période effective [1] , soit petite , de l'ordre de 400 par exemple , et que les points fournis par un pseudo-échantillon de taille t soient répartis de façon raisonnablement uniforme dans $[0,1]^2$.

Alors , avec des tests opérés sur des échantillons de taille 400 , on conclura que $\alpha_{2^i, j}^{2^{i'}, j'} = 0$, pour i et $i' \leq 2$, et vu les propriétés des périodes effectives ([4], [11]) , on trouvera que cette estimation est confirmée lorsque n croît , ainsi que les résultats des tests hilbertiens associés .

Si l'on ne prend pas de précautions , on conclura donc

- 1°) que cette simulation est bonne ,
- 2°) que la qualité des échantillons s'améliore lorsque leur taille croît :

exactement le contraire de la réalité !

Or cet exemple n'est pas si grossier que cela . Bien sûr , nous avons choisi a petit pour bien faire ressortir ce qui peut se produire si l'on effectue le test du khi-deux avec des partitions mal choisies . Mais , au cours de notre étude sur les simulations , il nous est arrivé parfois de rencontrer des congruences linéaires de période effective t petite , et que l'on testait par des tests du khi-deux ayant un nombre de classe du type précédent , et dont on tirait des conclusions de ce genre , qui sont en général aussi fausses l'une que l'autre .

Signalons que pour parer à la lacune des tests classiques sur les générateurs de nombres pseudo-aléatoires , de nouveaux tests ont été mis au point [5] . Mais ce n'est pas ici notre propos qui est simplement de montrer comment choisir le test hilbertien en fonction de l'utilisation désirée du générateur T .

3 : CONCLUSIONS

1°) Si l'on approxime la fonction de répartition d'un vecteur aléatoire (et donc sa loi) par la primitive d'une densité de probabilité développée en série jusqu'à un ordre k fini , le choix de cet ordre ne devra et ne pourra être effectué qu'avec le plus grand soin possible en fonction des besoins de l'utilisateur : dans le cas de tests de générateurs de nombres pseudo-aléatoires , il conviendra de lier cet ordre à la taille désirée du pseudo-échantillon .

2°) Cet ordre risque d'être assez élevé , et c'est normal .

3°) Si l'on veut tester l'indépendance réelle de deux variables aléatoires réelles X et Y , il faudra prendre cet ordre maximum , tout en se souvenant qu'à moins de posséder des données physique concrètes (ex : lancement de deux dés) , on ne pourra même jamais être sûr que X et Y ne sont pas en bijection .

BIBLIOGRAPHIE

- [1] BOSQ D. (1978)
Tests d'ajustement hilbertiens . Université de Lille 1 ; n° 125 .
- [2] BLACHER R. (1982)
Etude des coefficients de dépendance entre deux variables
aléatoires créés à partir des distances mathématiques classiques
Université de Grenoble 1 ; R.R. n° 289 .
- [3] BLACHER R. (1982)
Influence sur la dépendance des coefficients du développement en
série de la densité de probabilité . Université de Grenoble 1 ;
R.R. n° 343 .
- [4] BLACHER R. (1983)
Quelques propriétés des congruences linéaires considérées comme
générateurs de nombres pseudo-aléatoires .
Université de Grenoble 1 ; R.R. n° 345 .
- [5] BARRA J.R. (1971)
Contrôle statistique des suites de nombres "au hasard" .
Revue de statistique appliquée ; Vol XIX n° 3 .
- [6] RENYI A. (1959)
On measures of dependence .
Acta. Math. Acad. Sci. Ungar 10 , 441-451 .
- [7] RENYI A. (1970)
Probability theory . North Holland ; Amsterdam .
- [8] CHOQUET G. (1964)
Topologie . Masson .
- [9] FOURGEAU D. , FUCHS A. (1967)
Statistique . Dunod ; Paris .

- [10] NATANSON I. P. (1965)
Constructive function theory : Volume 2 , Approximation
in mean . Frederick Ungar Publishing C° ; New-York .
- [11] VAN CUTSEM B. (1980)
Etude de la période de la suite de nombres pseudo-aléatoires
engendrée par une congruence linéaire .
Bulletin de l'A.P.M. n° 324 .
- [12] SCHWEIZER B. , WOLF E. F. (1981)
On non parametric measures of dependence for random variables
Ann. Math. Statistic. 9 , 879-885 .
- [13] LABORDE J. (1971)
Tables statistiques et financières . Dunod
- [14] NEYMANN J (1937)
"Smooth test" for goodness of fit .
Skand. Aktuarietidskr. 20 , 149 .
- [15] NADARAJA E. A. (1976)
A quadratic measure of the deviation of a density estimator .
Theory of proba. and applications , 21 , p. 843-50 .
- [16] BARTON D. E. (1956)
Neyman's χ^2 test of goodness of fit when the null hypothesis
is composite . Skand. Aktuarietidskr . 39 , 216-245 .
-

Dernière page d'une thèse

VU

Grenoble, le 20 mars 1983

pour Le Président de la thèse

B. La Cuvier

Vu, et permis d'imprimer,

Grenoble, le 11-3-83

Le Président de l'Université Scientifique et Médicale



Le Président

M. TANCHE

Tanche