



HAL
open science

Quelques applications des fonctions orthogonales en probabilité et statistiques

René Blacher

► **To cite this version:**

René Blacher. Quelques applications des fonctions orthogonales en probabilité et statistiques. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1990. Français. NNT: . tel-00338785

HAL Id: tel-00338785

<https://theses.hal.science/tel-00338785>

Submitted on 14 Nov 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.

711 18025

THESE

présentée par

BLACHER RENE

pour obtenir le titre de DOCTEUR

DE L'UNIVERSITE JOSEPH FOURIER-GRENOBLE 1

EN MATHEMATIQUES APPLIQUEES

**QUELQUES APPLICATIONS DES FONCTIONS ORTHOGONALES
EN PROBABILITE ET STATISTIQUE**

Le 26 OCTOBRE 1990

Président du jury : PHAM DINH TUAN

Examineurs : BOSQ DENIS
DAUXOIS JACQUES
VAN CUTSEM BERNARD

Thèse préparée au sein du laboratoire LMC

UNIVERSITE Joseph FOURIER (GRENOBLE I)

Président de l'Université :
M. NEMOZ Alain

Année Universitaire 1988 - 1989

MEMBRES DU CORPS ENSEIGNANT DE SCIENCES ET DE GEOGRAPHIE

PROFESSEURS DE 1ère Classe

ADIBA Michel	Informatique
ANTOINE Pierre	Géologie I.R.I.G.M.
ARNAUD Paul	Chimie Organique
ARVIEU Robert	Physique Nucléaire I.S.N.
AUBERT Guy	Physique C.N.R.S
AURIAULT Jean-Louis	Mécanique
AYANT Yves	Physique Approfondie
BARBIER Marie-Jeanne	Electrochimie
BARJON Robert	Physique Nucléaire ISN
BARNOUD Fernand	Biochimie Macromoléculaire Végétale
BARRA Jean-René	Statistiques-Mathématiques Appliquées
BECKER Pierre	Physique
BEGUIN Claude	Chimie Organique
BELORISKY Elie	Physique
BENZAKEN Claude	Mathématiques Pures
BERARD Pierre	Mathématiques Pures
BERNARD Alain	Mathématiques Pures
BERTRANDIAS Françoise	Mathématiques Pures
BERTRANDIAS Jean-Paul	Mathématiques Pures
BILLET Jean	Géographie
BOELHER Jean-Paul	Mécanique
BRAVARD Yves	Géographie
CARLIER Georges	Biologie Végétale
CASTAING Bernard	Physique
CAUQUIS Georges	Chimie Organique
CHARDON Michel	Géographie
CHIBON Pierre	Biologie Animale
COHEN ADDAD Jean-Pierre	Physique
COLIN DE VERDIERE Yves	Mathématiques Pures
CYROT Michel	Physique du Solide
DEBELMAS Jacques	Géologie Générale
DEGRANGE Charles	Zoologie
DEMAILLY Jean-Pierre	Mathématiques Pures
DENEUVILLE Alain	Physique
DEPORTES Charles	Chimie Minérale
DOLIQUE Jean-Michel	Physique des Plasmas
DOUCE Roland	Physiologie Végétale
DUCROS Pierre	Cristallographie
FINKE Gerde	Informatique
GAGNAIRE Didier	Chimie Physique
GAUTRON René	Chimie
GENIES Eugène	Chimie
GERMAIN Jean-Pierre	Mécanique,
GIDON Maurice	Géologie
GUITTON Jacques	Chimie
HICTER Pierre	Chimie
IDELMAN Simon	Physiologie Animale
JANIN Bernard	Géographie
JOLY Jean René	Mathématiques Pures

JOSELEAU Jean Paul	Biochimie
KAHANE André, détaché	Physique
KAHANE Josette	Physique
KRAKOWIAK Sacha	Mathématiques Appliquées
LAJZEROWICZ Jeanine	Physique
LAJZEROWICZ Joseph	Physique
LAURENT Pierre-Jean	Mathématiques Appliquées
LEBRETON Alain	Mathématiques Appliquées
DE LEIRIS Joël	Biologie
LHOMME Jean	Chimie
LLIBOUTRY Louis	Géophysique
LOISEAUX Jean-Marie	Sciences Nucléaires I.S.N.
LONGEQUEUE Nicole	Physique
LUNA Domingo	Mathématiques Pures
MACHE Régis	Physiologie Végétale
MASCLE Georges	Géologie
MAYNARD Roger	Physique du Solide
OMONT Alain	Astrophysique
OZENDA Paul	Botanique (Biologie Végétale)
PANNETIER Jean	Chimie
PAYAN Jean-Jacques	Mathématiques Pures
PEBAY-PEYROULA Jean-Claude	Physique
PERRIER Guy	Géophysique
PIERRE Jean Louis	Chimie Organique
RENARD Michel	Thermodynamique
RIEDTMANN Christine	Mathématiques
RINAUDO Marguerite	Chimie CERMAV
ROSSI André	Biologie
SAXOD Raymond	Biologie Animale
SENGEL Philippe	Biologie Animale
SERGERAERT Francis	Mathématiques Pures
SOUCHIER Bernard	Biologie
SOUTIF Michel	Physique
STUTZ Pierre	Mécanique
TRILLING Laurent	Mathématiques Appliquées
VAN CUTSEM Bernard	Mathématiques Appliquées
VIALON Pierre	Géologie

PROFESSEURS de 2^{ème} Classe

ARMAND Gilbert	Géographie
ATTANE Pierre	Mécanique
BARET Paul	Chimie
BERTIN José	Mathématiques
BLANCHI J.Pierre	STAPS
BLOCK Marc	Biologie
BLUM Jacques	Mathématiques Appliquées
BOITET Christian	Mathématiques Appliquées
BORNAREL Jean	Physique
BORRIONE Dominique	Automatique informatique
BOUVET Jean	Biologie
BROSSARD Jean	Mathématiques
BRUANDET J.François	Physique
BRUGAL Gérard	Biologie
BRUN Gilbert	Biologie
CASTAING Bernard	Physique
CERFF Rudiger	Biologie
CHIARAMELLA Yves	Mathématiques Appliquées
CHOLLET Jean Pierre	Mécanique
COLOMBEAU Jean François	Mathématiques (ENSL)
COURT Jean	Chimie
CUNIN Pierre Yves	Informatique
DAVID Jean	Géographie

DHOUAILLY Danielle
 DUFRESNOY Alain
 GASPARD François
 GIDON Maurice
 GIGNOUX Claude
 GILLARD Roland
 GIORNI Alain
 GONZALEZ SPRINBERG Gérardo
 GUIGO Maryse
 GUMUCHAIN Hervé
 HACQUES Gérard
 HERBIN Jacky
 HERAULT Jeanny
 HERINO Roland
 JARDON Pierre
 KERCKHOVE Claude
 MANDARON Paul
 MARTINEZ Francis
 MOREL Alain
 NEMOZ Alain
 NGUYEN HUY Xuong
 OUDET Bruno
 PAUTOU Guy
 PECHER Arnaud
 PELMONT Jean
 PELLETIER Guy
 PERRIN Claude
 PIBOULE Michel
 RAYNAUD Hervé
 REGNARD Jean René
 RICHARD Jean-Marc
 RIEDTMANN Christine
 ROBERT Danielle
 ROBERT Gilles
 ROBERT Jean-Bernard
 SARROT-REYNAULD Jean
 SAYETAT Françoise
 SERVE Denis
 STOECKEL Frédéric
 SCHOLL Pierre-Claude
 SUBRA Robert
 VALLADE Marcel
 VIDAL Michel
 VINCENT Gilbert
 VIVIAN Robert
 VOTTERO Philippe

Biologie
 Mathématiques Pures
 Physique
 Géologie
 Sciences Nucléaires
 Mathématiques Pures
 Sciences Nucléaires
 Mathématiques Pures
 Géographie
 Géographie
 Mathématiques Appliquées
 Géographie
 Physique
 Physique
 Chimie
 Géologie
 Biologie
 Mathématiques Appliquées
 Géographie
 Thermodynamique CNRS - CRTBT
 Informatique
 Mathématiques Appliquées
 Biologie
 Géologie
 Biochimie
 Astrophysique
 Sciences Nucléaires I.S.N.
 Géologie
 Mathématiques Appliquées
 Physique
 Physique
 Mathématiques Pures
 Chimie
 Mathématiques Pures
 Chimie Physique
 Géologie
 Physique
 Chimie
 Physique
 Mathématiques Appliquées
 Chimie
 Physique
 Chimie Organique
 Physique
 Géographie
 Chimie

MEMBRES DU CORPS ENSEIGNANT DE L' IUT 1

PROFESSEURS de 1^{ère} Classe

BUISSON Roger	Physique IUT 1
CHEHIKIAN Alain	E.E.A. I.U.T.1
DODU Jacques	Mécanique Appliquée IUT 1
NEGRE Robert	Génie Civil IUT 1
NOUGARET Marcel	Automatique IUT 1
PERARD Jacques	EEA. IUT 1

PROFESSEURS de 2^{ème} classe

BEE Marc	Physique IUT 1
BOUTHINON Michel	EEA. IUT 1
CHAMBON René	Génie Mécanique IUT 1
CHENAVAS Jean	Physique IUT 1

CHILO Jean	Physique IUT 1
CHOUTEAU Gérard	Physique IUT 1
CONTE René	Physique IUT 1
FOSTER Panayotis	Chimie IUT 1
GOSSE Jean-Pierre	EEA.IUT 1
GROS Yves	Physique IUT 1
HAMAR Roger	Chimie IUT 1
KUHN Gérard, (Détaché)	Physique IUT 1
LEVIEL Jean Louis	Physique IUT 1
MAZUER Jean	Physique IUT 1
MICHOULIER Jean	Physique IUT 1
MONLLOR Christian	EEA.IUT 1
PERRAUD Robert	Chimie IUT 1
PIERRE Gérard	Chimie IUT 1
TERRIEZ Jean-Michel	Génie Mécanique IUT 1
TOUZAIN Philippe	Chimie IUT 1
TURGEMAN Sylvain	Génie civil
VINCENDON Marc	Chimie IUT 1
ZIGONE Michel	Physique IUT 1

PROFESSEURS DE PHARMACIE

AGNIUS-DELORD Claudine	Physique	Faculté La Tronche
ALARY Josette	Chimie Analytique	Faculté La Tronche
BERIEL Hélène	Physiologie et Pharmacologie	Faculté La Tronche
CUSSAC Max	Chimie Therapeutique	Faculté La Tronche
DEMENGE Pierre	Pharmacodynamie	Faculté La Tronche
FAVIER Alain	Biochimie	C.H.R.G.
JEANNIN Charles	Pharmacie Galénique	Faculté Meylan
LATURAZE Jean	Biochimie	Faculté La Tronche
LUU DUC Cuong	Chimie Générale	Faculté La Tronche
MARIOTTE Anne-Marie	Pharmacognosie	Faculté La Tronche
MARZIN Daniel	Toxicologie	Faculté Meylan
RENAUDET Jacqueline	Bactériologie	Faculté La Tronche
ROCHAT Jacques	Hygiène et Hydrologie	Faculté La Tronche
SEIGLE-MURANDI Françoise	Botanique et Cryptogamie	Faculté Meylan
VERAIN Alice	Pharmacie Galénique	Faculté Meylan

MEMBRES DU CORPS ENSEIGNANT DE MEDECINE

PROFESSEURS CLASSE EXEPTIONNELLE ET 1ère CLASSE

AMBLARD Pierre	Dermatologie	C.H.R.G.
AMBROISE-THOMAS Pierre	Parasitologie	C.H.R.G.
BEAUDOING André	Pédiatrie-Puericulture	C.H.R.G.
BEZEZ Henri	Orthopédie-Traumatologie	Hopital SUD
BONNET Jean-Louis	Ophthalmologie	C.H.R.G.
BOUCHET Yves	Anatomie	Faculté La Merci
BUTEL Jean	Chirurgie Générale et Digestive	C.H.R.G.
CHAMBAZ Edmond	Orthopédie-Traumatologie	C.H.R.G.
CHAMPETIER Jean	Biochimie	C.H.R.G.
CHARACHON Robert	Anatomie-Topographique et Appliquée	C.H.R.G.
COLOMB Maurice	O.R.L.	C.H.R.G.
COUDERC Pierre	Immunologie	Hopital sud
DELORMAS Pierre	Anatomie-Pathologique	C.H.R.G.
DENIS Bernard	Pneumophtisiologie	C.H.R.G.
GAVEND Michel	Cardiologie	C.H.R.G.
	Pharmacologie	Faculté La Merci

HOLLARD Daniel	Hématologie	C.H.R.G.
LATREILLE René	Chirurgie Thoracique et Cardiovasculaire	C.H.R.G.
LE NOC Pierre	Bactériologie-Virologie	C.H.R.G.
MALINAS Yves	Gynécologie et Obstétrique	C.H.R.G.
MALLION Jean-Michel	Médecine du Travail	C.H.R.G.
MICOUD Max	Clinique Médicale et Maladies Infectieuses	C.H.R.G.
MOURIQUAND Claude	Histologie	Faculté La Merci
PARAMELLE Bernard	Pneumologie	C.H.R.G.
PERRET Jean	Neurologie	C.H.R.G.
RACHAIL Michel	Hépto-Gastro-Entérologie	C.H.R.G.
DE ROUGEMONT Jacques	Neurochirurgie	C.H.R.G.
SARRAZIN Roger	Clinique Chirurgicale	C.H.R.G.
STIEGLITZ Paul	Anesthésiologie	C.H.R.G.
TANCHE Maurice	Physiologie	Faculté La Merci
VIGNAIS Pierre	Biochimie	Faculté La Merci

PROFESSEURS 2ème CLASSE

BACHELOT Yvan	Endocrinologie	C.H.R.G.
BARGE Michel	Neurochirurgie	C.H.R.G.
BENABID Alim Louis	Biophysique	Faculté La Merci
BENSA Jean-Claude	Immunologie	Hopital Sud
BERNARD Pierre	Gynécologie-Obstétrique	C.H.R.G.
BESSARD Germain	Pharmacologie	ABIDJAN
BOLLA Michel	Radiothérapie	C.H.R.G.
BOST Michel	Pédiatrie	C.H.R.G.
BOUCHARLAT Jacques	Psychiatrie Adultes	Hopital Sud
BRAMBILLA Christian	Pneumologie	C.H.R.G.
CHIROUSSEL Jean-Paul	Anatomie-Neurochirurgie	C.H.R.G.
COMET Michel	Biophysique	Faculté La Merci
CONTAMIN Charles	Chirurgie Thoracique et Cardiovasculaire	C.H.R.G.
CORDONNIER Daniel	Néphrologie	C.H.R.G.
COULOMB Max	Radiologie	C.H.R.G.
CROUZET Guy	Radiologie	C.H.R.G.
DEBRU Jean-Luc	Médecine Interne et Toxicologie	C.H.R.G.
DEMONGEOT Jacques	Biostatistiques et Informatique Médicale	Faculté La Merci
DUPRE Alain	Chirurgie Générale	C.H.R.G.
DYON Jean-François	Chirurgie Infantile	C.H.R.G.
ETERRADOSSI Jacqueline	Physiologie	Faculté La Merci
FAURE Claude	Anatomie et Organogénèse	C.H.R.G.
FAURE Gilbert	Urologie	C.H.R.G.
FOURNET Jacques	Hépto-Gastro-Entérologie	C.H.R.G.
FRANCO Alain	Médecine Interne	C.H.R.G.
GIRARDET Pierre	Anesthésiologie	C.H.R.G.
GUIDICELLI Henri	Chirurgie Générale et Vasculaire	C.H.R.G.
GUIGNIER Michel	Thérapeutique et Réanimation Médicale	C.H.R.G.
HADJIAN Arthur	Biochimie	Faculté La Merci
HALIMI Serge	Endocrinologie et Maladies Métaboliques	C.H.R.G.
HOSTEIN Jean	Hépto-Gastro-Entérologie	C.H.R.G.
HUGONOT Robert	Médecine Interne	C.H.R.G.
JALBERT Pierre	Histologie-Cytogénétique	C.H.R.G.
JUNIEN-LAVILLAULOY Claude	O.R.L.	C.H.R.G.
KOLODIE Lucien	Hématologie Biologique	C.H.R.G.
LETOUBLON Christian	Chirurgie Générale	C.H.R.G.
MACHECOURT Jacques	Cardiologie et Maladies Vasculaires	C.H.R.G.
MAGNIN Robert	Hygiène	C.H.R.G.
MASSOT Christian	Médecine Interne	C.H.R.G.

MOUILLON Michel
PELLAT Jacques
PHELIP Xavier
RACINET Claude
RAMBAUD Pierre
RAPHAEL Bernard
SCHAERER René
SEIGNEURIN Jean-Marie
SELE Bernard
SOTTO Jean-Jacques
STOEBNER Pierre
VROUSOS Constantin

Ophthalmologie
Neurologie
Rhumatologie
Gynécologie-Obstétrique
Pédiatrie
Stomatologie
Cancérologie
Bactériologie-Virologie
Cytogénétique
Hématologie
Anatomie Pathologique
Radiothérapie

C.H.R.G.
C.H.R.G.
C.H.R.G.
Hopital Sud
C.H.R.G.
C.H.R.G.
C.H.R.G.
Faculté La Merci
Faculté La Merci
C.H.R.G.
C.H.R.G.
C.H.R.G.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

46 avenue Felix Viallet
38031 GRENOBLE cedex

Tél. : 76.57.45.00

Année universitaire 1989

Président de l'Institut :
Monsieur Georges LESPINARD

Professeurs des Universités

BARIBAUD Michel	ENSERG	JAUSSAUD Pierre	ENSIEG
BARRAUD Alain	ENSIEG	JOST Rémy	ENSPG
BAUDELET Bernard	ENSPG	JOUBERT Jean-Claude	ENSPG
BEAUFILS Jean-Pierre	INPG	JOURDAIN Geneviève	ENSIEG
BLIMAN Samuel	ENSERG	LACOUME Jean-Louis	ENSIEG
BOIS Philippe	ENSHMG	LADET Pierre	ENSIEG
BONNETAIN Lucien	ENSEEG	LESIEUR Marcel	ENSHMG
BONNET Guy	ENSPG	LESPINARD Georges	ENSHMG
BRISSONNEAU Pierre	ENSIEG	LONGEQUEUE Jean-Pierre	ENSPG
BRUNET Yves	IUFA	LORET Benjamin	ENSHMG
CAILLERIE Denis	ENSHMG	LOUCHET François	ENSEEG
CAVAIGNAC Jean-François	ENSPG	LUCAZEAU Guy	ENSEEG
CHARTIER Germain	ENSPG	MASSE Philippe	ENSIEG
CHENEVIER Pierre	ENSERG	MASSELOT Christian	ENSIEG
CHERADAME Hervé	UFR PGP	MAZARE Guy	ENSIMAG
CHERUY Arlette	ENSIEG	MOHR Roger	ENSIMAG
CHOVET Alain	ENSERG	MOREAU René	ENSHMG
COHEN Joseph	ENSERG	MORET Roger	ENSIEG
COLINET Catherine	ENSEEG	MOSSIERE Jacques	ENSIMAG
CORNUT Bruno	ENSIEG	OBLED Charles	ENSHMG
COULOMB Jean-Louis	ENSIEG	OZIL Patrick	ENSEEG
COUMES André	ENSERG	PA ULEAU Yves	ENSEEG
CROWLEY James	ENSIMAG	PERRET Robert	ENSIEG
DARVE Félix	ENSHMG	PIAU Jean-Michel	ENSHMG
DELLA-DORA Jean	ENSIMAG	PIC Etienne	ENSERG
DEPEY Maurice	ENSERG	PLATEAU Brigitte	ENSIMAG
DEPORTES Jacques	ENSPG	POUPOT Christian	ENSERG
DEROO Daniel	ENSEEG	RAMEAU Jean-Jacques	ENSEEG
DESRE Pierre	ENSEEG	REINISCH Raymond	ENSPG
DOLMAZON Jean-Marc	ENSERG	RENAUD Maurice	UFR PGP
DURAND Francis	ENSEEG	ROBERT André	UFR PGP
DURAND Jean-Louis	ENSPG	ROBERT François	ENSIMAG
FAUTRELLE Yves	ENSHMG	SABONNADIERE Jean-Claude	ENSIEG
FOGGIA Albert	ENSIEG	SAUCIER Gabrièle	ENSIMAG
FONLUPT Jean	ENSIMAG	SCHLENKER Claire	ENSPG
FOULARD Claude	ENSIEG	SCHLENKER Michel	ENSPG
GANDINI Alessandro	UFR PGP	SERMET Pierre	ENSERG
GAUBERT Claude	ENSPG	SILVY Jacques	UFR PGP
GENTIL Pierre	ENSERG	SIRIEYS Pierre	ENSHMG
GENTIL Sylviane	ENSIEG	SOHM Jean-Claude	ENSEEG
GREVEN Hélène	IUFA	SOLER Jean-Louis	ENSIMAG
GUEGUEN Claude	ENSIEG	SOUQUET Jean-Louis	ENSEEG
GUERIN Bernard	ENSERG	TROMPETTE Philippe	ENSHMG
GUYOT Pierre	ENSEEG	VINCENT Henri	ENSPG
IVANES Marcel	ENSIEG	ZADWORNY François	ENSERG

Personnes ayant obtenu le diplôme d'HABILITATION A DIRIGER DES RECHERCHES

BECKER Monique
BINDER Zdenek
CHASSERY Jean-Marc
CHOLLET Jean-Pierre
COEY John
COLINET Catherine
COMMAULT Christian
CORNUJOLS Gérard
COULOMB Jean- Louis
COURNIL M.
DALARD Francis
DANES Florin
DEROO Daniel
DIARD Jean-Paul
DION Jean-Michel
DUGARD Luc
DURAND Madeleine
DURAND Robert
GALERIE Alain
GAUTHIER Jean-Paul
GENTIL Sylviane

GHIBAUDE Gérard
HAMAR Sylvaine
HAMAR Roger
LACHENAL D.
LADET Pierre
LATOMBE Claudine
LE HUY H.
LE GORREC Bernard
MADAR Roland
MEUNIER G.
MULLER Jean
NGUYEN TRONG Bernadette
NIEZ J.J.
PASTUREL Alain
PLA Fernand
ROGNON J.P.
ROUGER Jean
TCHUENTE Maurice
VINCENT Henri
YAVARI A.R.

Chercheurs du C.N.R.S

DIRECTEURS DE RECHERCHE CLASSE 0

LANDEAU	Ioan
NAYROLLES	Bernard

Directeurs de recherche 1ère Classe

ANSARA Ibrahim
CARRE René
FRUCHART Robert
HOPFINGER Emile

JORRAND Philippe
KRAKOWIAK Sacha
LEPROVOST Christian
VACHAUD Georges
VERJUS Jean-Pierre

Directeurs de recherche 2ème Classe

ALEMANY Antoine
ALLIBERT Colette
ALLIBERT Michel
ARMAND Michel
AUDIER Marc
BERNARD Claude
BINDER Gilbert
BONNET Roland
BORNARD Guy
CAILLET Marcel
CALMET Jacques
CHATILLON Chritiant
CLERMONT Jean-Robert
COURTOIS Bernard
DAVID René
DION Jean-Michel
DRIOLE Jean
DURAND Robert
ESCUDIER Pierre
EUSTATHOPOULOS Nicolas
GARNIER Marcel
GUELIN Pierre

JOUD Jean-Charles
KAMARINOS Georges
KLEITZ Michel
KOFMAN Walter
LEJEUNE Gérard
MADAR Roland
MERMET Jean
MICHEL Jean-Marie
MEUNIER Jacques
PEUZIN Jean-Claude
PIAU Monique
RENOUARD Dominique
SENATEUR Jean-Pierre
SIFAKIS Joseph
SIMON Jean-Paul
SUERY Michel
TEODOSIU Christian
VAUCLIN Michel
VENNEREAU Pierre
WACK Bernard
YONNET Jean-Paul

**Personnalités agréées à titre permanent à diriger
des travaux de recherche
(décision du conseil scientifique)**

E.N.S.E.E.G

HAMMOU Abdelkader
MARTIN-GARIN Régina
SARRAZIN Pierre
SIMON Jean-Paul

E.N.S.E.R.G

BOREL Joseph

E.N.S.I.E.G

DESCHIZEAUX Pierre
GLANGEAUD François
PERARD Jacques
REINISCH Raymond

E.N.S.H.M.G

ROWE Alain

E.N.S.I.M.A.G

COURTIN Jacques

C.E.N.G

CADET Jean
COEURE Philippe
DELHAYE Jean-Marc
DUPUY Michel
JOUVE Hubert
NICOLAU Yvan
NIFENECKER Hervé
PERROUD Paul
PEUZIN Jean-Claude
TAIEB Maurice
VINCENDON Marc

Laboratoires extérieurs :

C.N.E.T

DEVINE Rodericq
GERBER Roland
MERCCKEL Gérard
PAULEAU Yves

Situation particulière

PROFESSEURS D'UNIVERSITE

DETACHEMENT

ENSIMAG	LATOMBE	J..Claude	Détachement	21/10/1989
ENSHMG	PIERRARD	J.Marie	Détachement	30/04/1989
ENSIMAG	VEILLON	Gérard	Détachement	30/09/1990
ENSIMAG	VERJUS	J.Pierre	Détachement	30/09/1989
ENSPG	BLOCH	Daniel	Recteur à c/	21/12/1988

SURNOMBRE

INPG	CHIAVERINA	Jean	30/09/1989
ENSHMG	BOUVARD	Maurice	30/09/1991
ENSEEG	PARIAUD	J.Charles	30/09/1991

Je tiens d'abord à exprimer ma reconnaissance à Monsieur Pham Dinh Tuan pour l'honneur qu'il me fait en acceptant la présidence de ce jury, et je profite de cette occasion pour le remercier des conseils dont il m'a fait profiter depuis mon arrivée à l'équipe statistique, conseils qui se sont toujours avérés profitables.

Je remercie également Monsieur Denis Bosq d'avoir accepté d'effectuer le déplacement de Paris, et ce, d'autant plus que ce sont ses travaux sur le Test Hilbertien qui m'ont poussé à effectuer mes propres recherches sur les coefficients de corrélation d'ordre supérieur.

Je remercie aussi Monsieur Jacques Dauxois d'être venu de Toulouse pour participer à ce jury. Je tiens aussi à le remercier pour ses conseils sur la corrélation canonique et j'espère que cette collaboration continuera.

Enfin, je remercie Monsieur Bernard Van Cutsem avec qui je travaille depuis tant d'années pour avoir accepté de diriger cette thèse et aussi pour m'avoir appris la recherche scientifique.

D'autre part, je ne veux pas oublier les personnes qui m'ont permis d'entrer au CNRS, et d'abord Monsieur Jean Della Dora, directeur du Laboratoire LMC. Je veux également remercier Madame Jacqueline Bonnifet, administrateur Délégué du CNRS pour la région Rhone-Alpes-Auvergne, ainsi que Madame Strasser de Paris. Je tiens à leur exprimer ma profonde gratitude.

Il en est de même pour les professeurs Jacques Demongeot et Jacques Gasqui sur l'amitié de qui j'ai toujours pu compter. Je les en remercie ici.

Je veux également remercier les membres de l'équipe de statistique de Grenoble pour l'aide qu'ils m'ont toujours apportée, en particulier lors de la rédaction et de la traduction de mes différents articles.

Je remercie enfin le service de reprographie de l'IMAG, et en particulier Daniel Iglesias pour l'accueil qu'il m'a toujours réservé.

SOMMAIRE

INTRODUCTION	p 0-5
DOCUMENT I : COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR	
Résumé	p I-1
I-I Introduction et Notations	p I-2
I-II Propriétés élémentaires des coefficients de corrélation d'ordre supérieur	p I-4
I-III Corrélation canonique	p I-7
I-III-A Coefficients de corrélation canonique	p I-7
I-III-B Opérateurs de projections orthogonales et corrélation maximale	p I-9
I-III-C Représentation matricielle	p I-12
I-III-D Coefficients de corrélation canonique généralisé.	p I-13
I-IV Caractérisation des divers type de dépendance fonctionnelle	p I-19
I-V Coefficients de corrélation d'ordre supérieur et probabilités conditionnelles	p I-22
I-VI Coefficients de corrélation d'ordre supérieur empiriques	p I-25
I-VII Test Hilbertien d'indépendance	p I-28
I-VIII Exemples	p I-38
I-VIII-A Coefficients de corrélation polynômiale	p I-38
I-VIII-B Coefficients de corrélation de Spearman	p I-41
I-VIII-C Coefficients de corrélation rectangulaire	p I-42
I-VIII-D Coefficients de corrélation canonique généralisés	p I-47
Bibliographie	p I-48

DOCUMENT II

A NEW FORM FOR THE CHI SQUARED INDEPENDENCE TEST

	Summary	p II-1
	Introduction	p II-1
II-I	Hilbertian Test and restricted chi-square test	p II-5
II-II	Chi-squared statistics and Hilbertian statistics	p II-7
II-III	Test of independence : the case of the known marginal distributions	p II-10
II-IV	Independence test : the case of the unknown marginal distributions	p II-13
II-V	Conclusions	p II-16
	References	p II-17

DOCUMENT III

TEST HILBERTIEN D'INDEPENDANCE

	Résumé	p III-1
	Introduction	p III-2
III-I	Loi asymptotique des coefficients de corrélation d'ordre supérieur empiriques	p III-5
III-II	Test Hilbertien d'indépendance	p III-12
	Bibliographie	p III-15

DOCUMENT IV

LOI DE LA SOMME DE N VECTEURS ALEATOIRES ET LOI DE LEURS FORMES QUADRATIQUES

Résumé	p IV 1
Sommaire	p IV 2
Introduction	p IV 3
Notations	p IV 5
I : Quelques propriétés des polynômes d'Hermite et de Laguerre.	p IV 6
II : Quelques propriétés des Transformées de Fourier.	p IV 10
III : Application au calcul d'une densité.	p IV 13
IV : Loi de la somme de n vecteurs aléatoires et lois de leurs formes quadratiques	p IV 15
IV-I Développement de la densité en polynômes orthogonaux	p IV 15
IV-II Développement de la densité en série de Mac Laurin	p IV 21
IV-III Densité sous forme d'intégrale.	p IV 23
V : Application au cas Gaussien	p IV 24
V-I Quelques propriétés des lois normales.	p IV 24
V-II Variances négatives dans les lois normales	p IV 28
V-III Formes quadratiques de vecteurs Gaussiens	p IV 30
Bibliographie	p IV 36

INTRODUCTION

INTRODUCTION

C'est en étudiant la dépendance entre deux variables aléatoires X et Y que l'on a rencontré pour la première fois une suite de coefficients de dépendance $\{\rho_{i,j}\}$ dont le premier terme $\rho_{1,1}$, n'était autre que le coefficient de corrélation linéaire. Il était alors facile de voir que ces coefficients complétaient et comblaient les lacunes bien connues de ce dernier en ce qui concerne la mesure de la dépendance. En fait, ces $\rho_{i,j}$ se comportaient naturellement comme des coefficients de corrélation d'ordre supérieur. Aussi, c'est cette appellation qu'on leur a donné dans un premier article publié en 1984 dans la revue "Statistique et Analyse des Données".

Bien sûr, il semblait clair que l'on pouvait obtenir à partir des $\rho_{i,j}$ une étude fine de la dépendance, et donc des applications pratiques intéressantes. Mais, il s'est avéré très vite que, c'étaient plutôt les fonctions orthogonales qui possèdent des propriétés probabilistes et statistiques intéressantes (on définit les $\rho_{i,j}$ par : $\rho_{i,j} = \mathbb{E}[P_i(X)Q_j(Y)]$, $\{P_i\}$ et $\{Q_j\}$ représentant les familles de polynômes orthogonaux associées à X et Y). En effet, on a vu aussi que les fonctions orthogonales permettaient d'introduire des variances d'ordre i mesurant la concentration d'une probabilité autour de i points distincts connus, avec un poids associé connu également. Ce résultat a été exposé au premier symposium international sur les polynômes orthogonaux de Bar Le Duc, et devrait être concrétisé prochainement par un article.

Enfin, on s'est aperçu également que la forme très particulière des Transformées de Fourier des polynômes d'Hermite permettait d'obtenir facilement la loi de la somme de n variables aléatoires.

Tout ceci a conduit à penser que les fonctions orthogonales sont des outils probabilistes particulièrement puissants. Aussi, dans cette thèse, on va s'intéresser à deux aspects des fonctions orthogonales considérées comme instruments purement probabilistes : les coefficients de corrélation d'ordre supérieur et le test Hilbertien d'indépendance d'une part, et l'étude de la loi de la somme de n vecteurs aléatoires et de leurs formes quadratiques d'autre part.

L'étude des coefficients de corrélation d'ordre supérieur constitue, bien sûr, la suite de l'article de 1984, et est destiné à être publié en article dans la revue STATISTICS. La version présentée dans cette thèse est la seconde version de cet article corrigée selon les indications des rapporteurs, et légèrement modifiée. En effet, vu sa longueur, il semblerait qu'il soit destiné à être publié en deux articles distincts.

Dans la version présentée dans cette thèse et en ce qui concerne l'aspect probabiliste, on généralise la définition des $\rho_{i,j}$ au cas vectoriel et à d'autres fonctions orthogonales que les polynômes orthogonaux, on étudie le lien entre les $\rho_{i,j}$ et les coefficients de corrélation canonique, on montre qu'ils peuvent détecter toute dépendance fonctionnelle du type $f(X)=g(Y)$, et enfin, qu'ils permettent d'explicitier les définitions et propriétés élémentaires des probabilités conditionnelles.

En ce qui concerne l'aspect statistique, on obtient la loi asymptotique des coefficients de corrélation d'ordre supérieur empiriques. Ce dernier résultat est à rattacher à l'étude du test d'indépendance Hilbertien que nous effectuerons ici de façon détaillée dans trois articles distincts : d'abord dans la fin de cet article sur les coefficients de corrélation d'ordre supérieur, ensuite, dans l'article "A new form for the chi squared test of independence" paru dans STATISTICS en 1988, et enfin, dans l'article "Test Hilbertien d'indépendance" proposé au Journal Canadien de Statistique, et corrigé selon les indications des rapporteurs.

Signalons, qu'en ce qui concerne le cas de lois marginales connues, ce test repose directement sur les résultats de BOSQ. Maintenant, s'il faut estimer les paramètres, on aura recours à la loi asymptotique des fonctions orthogonales empiriques dont la simplicité géométrique mérite d'être soulignée.

Quant à la loi de la somme de n vecteurs aléatoires et la loi des formes quadratiques associées, c'est une étude que l'on vient de terminer. Mais, il a semblé bon de l'inclure dans cette thèse car on voit là encore plus qu'ailleurs, que les fonctions orthogonales sont des instruments probabilistes particulièrement puissants.

Maintenant, il est bien clair que les fonctions orthogonales ont encore beaucoup d'autres propriétés qu'il n'était guère possible de réunir ici. En ce qui concerne nos recherches propres, à savoir l'étude des fonctions orthogonales en tant qu'instruments purement probabilistes ou statistiques, signalons : les variances d'ordre supérieur, la loi asymptotique des estimateurs des moindres carrés en régression linéaire, la loi asymptotique des probabilités et densités conditionnelles empiriques, des résultats sur le Théorème de la Limite Centrale et la Loi Faible des Grands Nombres.

DOCUMENT I

COEFFICIENTS DE
CORRELATION
D'ORDRE SUPERIEUR

COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR

Mots-clefs : Coefficient de corrélation, corrélation canonique, dépendance fonctionnelle, probabilité conditionnelle, fonctions orthogonales, test d'indépendance du chi-deux, test Hilbertien.

Résumé : On poursuit l'étude des coefficients de corrélation d'ordre supérieur qui complètent les coefficients de corrélation classiques en mesurant des dépendances de plus en plus fines. On montre qu'ils peuvent détecter toute dépendance fonctionnelle du type $f(X)=g(Y)$, qu'ils permettent d'expliciter les théorèmes classiques des probabilités conditionnelles, et de généraliser la théorie des coefficients de corrélation canonique. On obtient la loi asymptotique des coefficients de corrélation d'ordre supérieur empiriques. On en déduit une classe de tests Hilbertiens d'indépendance avec estimation de paramètres qui contient le test correspondant du chi-deux.

Summary : We continue the study of correlation coefficients of superior order. These coefficients complete the classical correlation coefficients by measuring some dependences more and more sharp. We prove that they can detect all dependence $f(X)=g(Y)$, and we explicit the classical theorems of conditionnal probability with these coefficients. Moreover we generalise the canonical correlation coefficients. By another way, we obtain the asymptotic distribution of empirical correlation coefficients. With theses results, we build the Hilbertian Test of Independence with estimation of parameters. Then, the Chi-squared Independence Test with estimation of parameters is a particular case of this test.

Indice de classification AMS : 60A05, 62F03, 62J02, 62J05.

I - INTRODUCTION ET NOTATIONS

On poursuit ici l'étude, commencée en 1984, des coefficients de corrélation d'ordre supérieur, les $\rho_{i,j}$, $(i,j) \in \mathbb{N}^* \times \mathbb{N}^*$. Cette appellation donnée aux $\rho_{i,j}$ avait été suggérée par le fait que $\rho_{1,1}$ n'est autre que le coefficient de corrélation linéaire, et n'est en fait que le premier terme d'une suite de coefficients de corrélation qui mesurent des dépendances de plus en plus fines, à savoir des dépendances polynômiales. Mais, on peut de la même façon, obtenir d'autres coefficients de corrélation d'ordre supérieur, qui mesurent d'autres types de dépendance. Par exemple, on aura également des coefficients de Spearman d'ordre supérieur.

En fait, les $\rho_{i,j}$ ont déjà été étudiés par Lancaster dans le cadre plus particulier de la corrélation canonique. D'autre part, la matrice des $\rho_{i,j}$ est la matrice d'un opérateur de projection orthogonale étudié par Csaki et Fischer dans le cadre de leurs recherches sur la corrélation maximale. La réunion de ces différents résultats permettra de généraliser de façon simple la définition des coefficients de corrélation canonique, ainsi que les théorèmes associés. On se retrouve ainsi dans le cadre de l'étude de Dauxois et Pousse sur la corrélation canonique généralisée.

On verra alors comment appliquer ces résultats dans le cas de l'existence d'une relation fonctionnelle du type $f(X)=g(Y)$ entre deux vecteurs aléatoires X et Y . En particulier, on trouvera une condition simple sur les $\rho_{i,j}$ pour détecter les dépendances du type $Y=f(X)$.

De plus, on exprimera les théorèmes élémentaires des probabilités conditionnelles au moyen des $\rho_{i,j}$, ce qui permettra d'explicitier l'espérance et la probabilité conditionnelle.

D'autre part, on généralisera la définition des coefficients de corrélation d'ordre supérieur empiriques, et on obtiendra leurs lois asymptotiques. On en déduira le Test d'indépendance Hilbertien avec estimation de paramètres, ce qui généralise ainsi le test Hilbertien de Bosq. Un cas particulier de cette classe de tests sera, bien sûr, le test d'indépendance du chi-deux dans le cas de lois marginales inconnues.

On introduit maintenant les notations et hypothèses que nous utiliserons tout au long de cet article.

1-1 Notation des vecteurs aléatoires : Dans cet article, (X,Y) désignera un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans \mathbb{R}^{p+q} . Nous supposons que $X = (X_1, X_2, \dots, X_p)$ est à valeurs dans \mathbb{R}^p , et $Y = (Y_1, Y_2, \dots, Y_q)$ à valeurs dans \mathbb{R}^q . Nous noterons par \mathbb{Q} , μ' et μ'' , les probabilités images de (X,Y) , X et Y , et par $F_{X,Y}$, F_X et F_Y , leurs fonctions de répartition.

1-2 Notation des familles orthonormales : Nous notons par $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ deux familles orthonormales associées à, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$, où $\mathcal{N} = \{0, 1, \dots, h\}$ ou $\mathcal{N} = \mathbb{N}$, et $\mathcal{M} = \{0, 1, \dots, k\}$ ou $\mathcal{M} = \mathbb{N}$, et vérifiant $P_0 = Q_0 \equiv 1$.

D'autre part, nous posons $\mathcal{N}^* = \mathcal{N} \setminus \{0\}$ et $\mathcal{M}^* = \mathcal{M} \setminus \{0\}$, et nous supposons $\mathcal{N}^* \neq \emptyset$ et $\mathcal{M}^* \neq \emptyset$.

De plus, dans les chapitres VI et VII, nous supposons $\mathcal{N} = \{0, 1, \dots, h\}$ et $\mathcal{M} = \{0, 1, \dots, k\}$.

Rappelons que, comme $L^2(\mathbb{R}^p, \mu')$ est séparable (Weidmann, p 20 ex-6, p 27 ex-11), toute famille orthonormale de cet espace est dénombrable (Weidmann, 2° prop p 47). En fait, \mathcal{N} représentera un ensemble fini, essentiellement dans deux cas : d'abord, lorsque μ' est concentrée en un nombre fini de points, ce qui sera le cas pour les mesures empiriques, ensuite, lorsque les P_i représentent les polynômes orthonormaux d'une probabilité μ' dont tous les moments ne sont pas définis.

1-3 - Notations classiques : On note par $\mathbb{E}(\cdot)$, $\sigma(\cdot)$ et $\rho(\cdot, \cdot)$, l'espérance, l'écart-type, et le coefficient de corrélation linéaire, et par $\langle \cdot, \cdot \rangle$ et $\| \cdot \|$, le produit scalaire et la norme dans les espaces L^2 .

De plus, soient f une application de \mathbb{R}^p dans \mathbb{R} , et g une application de \mathbb{R}^q dans \mathbb{R} , alors on note par $f \otimes g$, l'application de \mathbb{R}^{p+q} dans \mathbb{R} définie par : $(f \otimes g)(x, y) = f(x).g(y)$ pour tout $(x, y) \in \mathbb{R}^{p+q}$.

D'autre part, soit \mathbf{B} un Borélien de \mathbb{R}^p , alors, on note par $1_{\mathbf{B}}$ la fonction indicatrice de \mathbf{B} .

Soit $\{a_{i,j}\}_{(i,j) \in \mathcal{N} \times \mathcal{M}}$ une suite double réelle. Alors on note par $\{\{a_{i,j}\}\}_{(i,j) \in \mathcal{N} \times \mathcal{M}}$ la

$$\text{matrice, finie ou infinie : } \{\{a_{i,j}\}\}_{(i,j) \in \mathcal{N} \times \mathcal{M}} = \begin{bmatrix} a_{1,1}, a_{1,2}, a_{1,3}, \dots \\ a_{2,1}, a_{2,2}, a_{2,3}, \dots \\ a_{3,1}, a_{3,2}, a_{3,3}, \dots \\ \dots \end{bmatrix},$$

et on généralisera cette notation à d'autres ensembles d'indices que $\mathcal{N} \times \mathcal{M}$.

Enfin, $\delta_{i,j}$ représentera le symbole de Kronecker.

1-4 - Notation des sommes infinies : Les signes \sum opérant sur des éléments de L^2 représenteront toujours la limite dans l'espace normé L^2 considéré. De plus, les signes \sum opérant sur des réels représenteront la limite classique dans \mathbb{R} . Enfin, dans \mathbb{R}^2 , on définira $\sum_{(i,j) \in \mathbb{N}^2} a_{i,j} = \lim_{m \rightarrow \infty} \left\{ \sum_{i \leq m \text{ et } j \leq m} a_{i,j} \right\}$, et on généralisera de façon naturelle cette limite à des suites double définies sur d'autres ensembles d'indices que \mathbb{N}^2 .

II - PROPRIETES ELEMENTAIRES DES COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR

Dans ce paragraphe, on généralise les définitions et propriétés élémentaires des coefficients de corrélation introduits dans Lancaster 1958 et 1969 ou Blacher 1984-a .

On commence par donner leur définition.

Définition 2-1 : Soit $(i,j) \in \mathcal{N} * \times \mathcal{M} *$. On appelle coefficient de corrélation d'ordre (i,j) associé aux familles $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$, le nombre réel $\mathbb{E} [P_i(X)Q_j(Y)]$, que l'on note par $\rho_{i,j}(X,Y)$, ou $\rho_{i,j}$:

$$\rho_{i,j} = \rho_{i,j}(X,Y) = \mathbb{E} [P_i(X)Q_j(Y)] .$$

Pour simplifier les notations, on adopte la convention suivante.

Notations 2-2 : Nous généralisons les notations de 2-1 en posant pour tout $(i,j) \in \mathcal{N} \times \mathcal{M}$. $\rho_{i,j} = \mathbb{E} [P_i(X)Q_j(Y)]$, i.e. $\rho_{i,0} = \rho_{0,i} = 0$ si $i \neq 0$ et 1 sinon.

Par exemple, dans le cas où X et Y sont des variables aléatoires, et où les familles $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont les familles de polynômes orthonormaux, les $\rho_{i,j}$ sont les coefficients de corrélation polynômiale.

$$\text{En particulier, } P_1(x) = \frac{x - \mathbb{E}(X)}{\sigma(X)} \text{ et } Q_1(y) = \frac{y - \mathbb{E}(Y)}{\sigma(Y)} ,$$

et $\rho_{1,1}$ est donc le coefficient de corrélation linéaire. De même $\rho_{1,2}$, $\rho_{2,1}$ et $\rho_{2,2}$ mesurent des dépendances quadratiques, $\rho_{1,3}$, $\rho_{2,3}$, $\rho_{3,1}$, $\rho_{3,2}$ et $\rho_{3,3}$ mesurent des dépendances cubiques etc... (cf Blacher 1984) .

Rappelons aussi que le coefficient de corrélation d'ordre (i,j) est le coefficient de corrélation linéaire de $P_i(X)$ et $Q_j(Y)$.

Maintenant, la proposition suivante montre bien que le coefficient de corrélation linéaire n'est que le premier terme de toute une suite de coefficients de corrélation.

Théorème 2-3 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$. Alors, en tout $(x,y) \in \mathbb{R}^{p+q}$, $x=(x_1, x_1, \dots, x_p)$, $y=(y_1, y_1, \dots, y_q)$, on peut écrire la fonction de répartition $F_{X,Y}$, sous la forme suivante :

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) + \sum_{(i,j) \in \mathcal{N} \times \mathcal{M}^*} \rho_{i,j} \left(\int_{u=-\infty}^{u=x} P_i(u) \cdot \mu'(du) \right) \left(\int_{v=-\infty}^{v=y} Q_j(v) \cdot \mu''(dv) \right),$$

$$\text{où } \int_{u=-\infty}^{u=x} P_i(u) \cdot \mu'(du) = \int_{u_1=-\infty}^{u_1=x_1} \int_{u_2=-\infty}^{u_2=x_2} \dots \int_{u_p=-\infty}^{u_p=x_p} P_i(u_1, u_2, \dots, u_p) \cdot \mu'(du_1, du_2, \dots, du_p).$$

Pour démontrer ce théorème, on a besoin du résultat suivant.

Lemme 2-4: Soient $f \in L^2(\mathbb{R}^p, \mu')$ et $g \in L^2(\mathbb{R}^q, \mu'')$:

$$f = \sum_{i \in \mathcal{N}} \alpha_i P_i \text{ et } g = \sum_{j \in \mathcal{M}} \beta_j Q_j \quad . \quad \text{Alors, } \iint f \otimes g \cdot d\mathbf{0} = \sum_{(i,j) \in \mathcal{N} \times \mathcal{M}} \alpha_i \beta_j \rho_{i,j} \quad .$$

Il suffit de remarquer que l'on peut également écrire f et g sous la forme:

$$f = \sum_{i \in \mathcal{N}} \alpha_i P_i \otimes Q_0 \text{ et } g = \sum_{j \in \mathcal{M}} \beta_j P_0 \otimes Q_j \text{ dans } L^2(\mathbb{R}^{p+q}, \mathbf{0}),$$

et d'appliquer la continuité du produit scalaire.

Pour démontrer 2-3, il ne reste plus qu'à choisir f comme étant la fonction indicatrice de $]-\infty, x]$ et g celle de $]-\infty, y]$.

Dans le cas où (X, Y) admet une densité, ce théorème se traduit de la façon suivante.

Théorème 2-5 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$

Alors, (X, Y) admet une densité de probabilité f par rapport à $\mu' \otimes \mu''$, $f \in L^2(\mathbb{R}^{p+q}, \mu' \otimes \mu'')$

$$\text{si et seulement si } \sum_{(i,j) \in \mathcal{N} \times \mathcal{M}^*} \rho_{i,j}^2 < +\infty \quad .$$

De plus, sous cette hypothèse,

$$f = 1 + \sum_{(i,j) \in \mathcal{N} \times \mathcal{M}^*} \rho_{i,j} P_i \otimes Q_j \text{ dans } L^2(\mathbb{R}^{p+q}, \mu' \otimes \mu'') \quad .$$

Ces résultats sont classiques et sont des corollaires des propriétés élémentaires des bases orthonormales. Remarquons qu'on peut aussi les déduire de 2-3.

On rappelle maintenant le rapport entre les coefficients de corrélation d'ordre supérieur et l'indépendance. D'abord, comme $\rho_{i,j}$ est le coefficient de corrélation linéaire de $P_i(X)$ et $Q_j(Y)$, on a le théorème suivant.

Théorème 2-6 : Si X et Y sont indépendants, alors, pour tout $(i,j) \in \mathcal{N}^* \times \mathcal{M}^*$, $\rho_{i,j}(X,Y) = 0$.

D'autre part, les coefficients de corrélation d'ordre supérieur possèdent, eux, la fameuse réciproque que ne vérifie pas le coefficient de corrélation linéaire.

Théorème 2-7 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$

Alors, si pour tout $(i,j) \in \mathcal{N}^* \times \mathcal{M}^*$, $\rho_{i,j}(X,Y) = 0$, X et Y sont indépendants.

Il suffit d'appliquer 2-3.

Maintenant, les coefficients de corrélation d'ordre supérieur permettent aussi de calculer la corrélation de $f(X)$ et $g(Y)$.

Théorème 2-8 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$

Soient $f \in L^2(\mathbb{R}^p, \mu')$ et $g \in L^2(\mathbb{R}^q, \mu'')$: $f = \sum_{i \in \mathcal{N}} \alpha_i P_i$ et $g = \sum_{j \in \mathcal{M}} \beta_j Q_j$.

Alors, $\rho(f(X), g(Y))$, le coefficient de corrélation linéaire de $f(X)$ et $g(Y)$ est égal à :

$$\rho(f(X), g(Y)) = \frac{\sum_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*} \alpha_i \beta_j \rho_{i,j}}{\sqrt{\sum_{i \in \mathcal{N}^*} \alpha_i^2} \sqrt{\sum_{j \in \mathcal{M}^*} \beta_j^2}}.$$

Ce théorème est un corollaire évident de 2-4.

Il sera utile de rappeler une condition simple avec laquelle les hypothèses de 2-5 sont toujours vérifiées.

Théorème 2-9 : Nous supposons que X est concentré en un nombre fini de points. Alors (X,Y) admet toujours une densité de probabilité f par rapport à $\mu' \otimes \mu''$: $f \in L^2(\mathbb{R}^{p+q}, \mu' \otimes \mu'')$.

En effet, on sait (cf 16-10, p 261 de Choquet) qu'il existe toujours deux bases orthonormales, $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$, de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$.

D'autre part, d'après l'hypothèse, $\text{card}(\mathcal{N}) = \dim[L^2(\mathbb{R}^p, \mu')] < +\infty$.

Donc le lemme 3-17, dont il est clair que la démonstration est indépendante de celle-ci, montre que

$$\sum_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*} (\rho_{i,j})^2 \leq \text{card}(\mathcal{N}^*) < +\infty.$$

Il suffit alors d'appliquer 2-5 pour terminer la démonstration.

III : CORRELATION CANONIQUE

Un cas particulier des coefficients de corrélation d'ordre supérieur sera le cas des coefficients de corrélation canonique. En effet, les résultats de Lancaster montrent que les fonctions canoniques forment des familles orthogonales des espaces L^2 . D'autre part, Csaki et Fischer avaient constaté que les opérateurs de projection orthogonale peuvent servir à étudier la corrélation maximale. En fait, il est facile de s'apercevoir que ces divers résultats sont liés : la matrice de ces opérateurs n'est autre que la matrice des coefficients de corrélation d'ordre supérieur, et les vecteurs propres associés généralisent les coefficients de corrélation canonique.

III-A : COEFFICIENTS DE CORRELATION CANONIQUE

Nous allons maintenant rappeler les résultats de Lancaster sur les coefficients de corrélation canonique. Pour introduire ceux-ci, nous aurons besoin de la corrélation maximale dont on rappelle maintenant la définition.

Définition 3-1 : La corrélation maximale est définie par $\rho^*(X, Y) = \sup[\rho(f(X), g(Y))]$ lorsque f parcourt $L^2(\mathbb{R}^p, \mu')$ et g , $L^2(\mathbb{R}^q, \mu'')$.

On va d'abord se limiter au cas $p=q=1$ et $\mathcal{N}=\mathcal{M}=\mathbb{N}$ de façon à ne pas compliquer les notations.

Rappelons maintenant qu'il n'existe pas forcément deux fonctions f et g telles que $\rho^*(X, Y) = \rho(f(X), g(Y))$ (cf Renyi p 445). Cependant, sous l'hypothèse de distributions ϕ^2 -bornées, ce résultat est vrai.

Définition 3-2 : On suppose $p=q=1$. Lorsque $F_{X, Y}$ admet une densité de probabilité f par rapport à $\mu' \otimes \mu''$, $f \in L^2(\mathbb{R}^2, \mu' \otimes \mu'')$, on dit que $F_{X, Y}$ est ϕ^2 -bornée, et on définit ϕ^2 par $\phi^2 + 1 = \iint f^2 d(\mu' \otimes \mu'')$. Dans, le cas contraire, on pose $\phi^2 = +\infty$.

Donc, dans ce cas, on peut introduire les coefficients de corrélation canonique.

Définition 3-3 : On suppose que $p=q=1$, $\mathcal{N}=\mathcal{M}=\mathbb{N}$, et que $F_{X, Y}$ est ϕ^2 -bornée. On définit par récurrence les fonctions canoniques $\{\xi^{(i)}\}_{i \in \mathbb{N}}$ et $\{\eta^{(i)}\}_{j \in \mathbb{N}}$, de la manière suivante.

On pose $\xi^{(0)} = \eta^{(0)} \equiv 1$ et $\rho_0 = 1$. Alors, la i -ème paire de fonctions canoniques sera définie comme une paire de fonctions $(\xi^{(i)}, \eta^{(i)})$ telle que $\rho(\xi^{(i)}(X), \eta^{(i)}(Y)) = \sup[\rho(f(X), g(Y))]$ lorsque f parcourt l'orthogonal de $\{\xi^{(0)}, \xi^{(1)}, \dots, \xi^{(i-1)}\}$ dans $L^2(\mathbb{R}, \mu')$, et g , l'orthogonal de $\{\eta^{(0)}, \eta^{(1)}, \dots, \eta^{(i-1)}\}$ dans $L^2(\mathbb{R}, \mu'')$. On impose de plus, $\|\xi^{(i)}\| = \|\eta^{(i)}\| = 1$. On définit alors le coefficient de corrélation canonique associé ρ_i par $\rho_i = \rho(\xi^{(i)}(X), \eta^{(i)}(Y))$, et on impose enfin $\rho_i \geq 0$.

On remarque que la suite $\{\rho_i\}$ est décroissante et que $\rho^*(X,Y) = \rho_1$. Rappelons aussi que les fonctions canoniques ne sont pas forcément définies de manière unique : ce sera le cas dès qu'il existe i tel que $\rho_i = \rho_{i+1}$.

D'autre part, sous les hypothèses de 2-5, $\phi^2 = \sum_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*} \rho_{i,j}^2$.

En fait, on peut énoncer le théorème suivant (th 3-3, p 95 Lancaster 1969).

Théorème 3-4 : On suppose que $p=q=1$, $\mathcal{N}=\mathcal{M}=\mathbb{N}$, et que $F_{X,Y}$ est ϕ^2 -bornée. Alors

$$\phi^2 = \sum_{i \in \mathbb{N}^*} \rho_i^2$$

Maintenant, les fonctions canoniques possèdent la propriété de bi-orthogonalité (th 3-2, p 94 de Lancaster 1969, et th 2-3, p 126 de Lancaster 1980).

Théorème 3-5 : On suppose que $p=q=1$, $\mathcal{N}=\mathcal{M}=\mathbb{N}$, et que $F_{X,Y}$ est ϕ^2 -bornée. Alors, les fonctions canoniques possèdent la propriété de bi-orthogonalité, i.e. $E(\xi^{(i)}(X)\eta^{(j)}(Y)) = \rho_i \delta_{i,j}$. De plus, il existe deux bases orthonormales $\{\xi_i\}_{i \in \mathbb{N}}$ et $\{\eta_j\}_{j \in \mathbb{N}}$, de, respectivement, $L^2(\mathbb{R}, \mu')$ et $L^2(\mathbb{R}, \mu'')$, ayant la propriété de bi-orthogonalité, et l'on peut imposer $\{\xi^{(i)}\}_{i \in \mathbb{N}} \subset \{\xi_i\}_{i \in \mathbb{N}}$ et $\{\eta^{(j)}\}_{j \in \mathbb{N}} \subset \{\eta_j\}_{j \in \mathbb{N}}$, et écrire f , la densité de $F_{X,Y}$ par rapport à $\mu' \otimes \mu''$ sous la forme :

$$f = 1 + \sum_{i \in \mathbb{N}^*} \rho_i \xi^{(i)} \otimes \eta^{(i)} \quad \text{dans } L^2(\mathbb{R}^2, \mu' \otimes \mu'')$$

III-B : OPERATEURS DE PROJECTION ORTHOGONALE ET CORRELATION MAXIMALE

Dans leurs travaux de généralisation de la corrélation maximale, Csaki et Fischer ont utilisé des opérateurs de projection orthogonale. En particulier, leur article de 1963 constitue une étude approfondie de ces opérateurs, considérés dans cette perspective. Or, il est clair, à la lecture de cet article, qu'il y a également un lien entre les coefficients de corrélation canonique et ces opérateurs. En fait, nous verrons que ce sont les valeurs propres de ces opérateurs qui constituent une généralisation naturelle des coefficients de corrélation canoniques.

Nous allons d'abord introduire les notations nécessaires, et rappeler quelques définitions et propriétés élémentaires énoncées par Csaki et Fischer.

Opérateurs de projection orthogonale 3-6 : On note par $\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, les sous-espaces de, respectivement, $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$, engendrés par $\{P_i\}_{i \in \mathcal{N}^*}$ et $\{Q_j\}_{j \in \mathcal{M}^*}$, et par $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$, les sous-espaces de $L^2(\mathbb{R}^{p+q}, \mathcal{Q})$ suivants :

$$\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q = \left\{ f \in L^2(\mathbb{R}^{p+q}, \mathcal{Q}) \mid \exists f^\circ \in \mathcal{U}(\mathbb{R}^p, \mu') : \forall (x, y) \in \mathbb{R}^{p+q}, f(x, y) = f^\circ(x) \right\},$$

$$\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'') = \left\{ g \in L^2(\mathbb{R}^{p+q}, \mathcal{Q}) \mid \exists g^\circ \in \mathcal{Q}(\mathbb{R}^q, \mu'') : \forall (x, y) \in \mathbb{R}^{p+q}, g(x, y) = g^\circ(y) \right\}.$$

Alors, les opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{U}}$ et $\Pi_{\mathcal{U}}^{\mathcal{Q}}$ seront les opérateurs de projection orthogonale, de, respectivement, $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ sur $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$, et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$ sur $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$. De plus, $\mathcal{Y}^{\mathcal{U}}$ et $\mathcal{Y}^{\mathcal{Q}}$ seront les opérateurs définis, sur, respectivement, $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$ par :

$$\mathcal{Y}^{\mathcal{U}} = \Pi_{\mathcal{U}}^{\mathcal{Q}} \circ \Pi_{\mathcal{Q}}^{\mathcal{U}} \quad \text{et} \quad \mathcal{Y}^{\mathcal{Q}} = \Pi_{\mathcal{Q}}^{\mathcal{U}} \circ \Pi_{\mathcal{U}}^{\mathcal{Q}}.$$

Dans le cas où $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$, on notera

$$\mathcal{Y}^X, \mathcal{Y}^Y, \Pi_Y^X \text{ et } \Pi_X^Y \text{ au lieu de } \mathcal{Y}^{\mathcal{U}}, \mathcal{Y}^{\mathcal{Q}}, \Pi_{\mathcal{Q}}^{\mathcal{U}} \text{ et } \Pi_{\mathcal{U}}^{\mathcal{Q}}.$$

Normes des opérateurs de projection orthogonale 3-7 On définit $\delta(\mathcal{U}, \mathcal{Q})$ et $S(\mathcal{U}, \mathcal{Q})$ par :

$$\delta(\mathcal{U}, \mathcal{Q}) = \inf_{f \in \mathcal{U}, \|f\|=1} \left\| f - \Pi_{\mathcal{Q}}^{\mathcal{U}}(f) \right\| \quad \text{et} \quad S(\mathcal{U}, \mathcal{Q}) = \sqrt{1 - \delta^2(\mathcal{U}, \mathcal{Q})}.$$

$$\text{On définit le réel } \|\|\| \Pi_{\mathcal{Q}}^{\mathcal{U}} \|\|\| \text{ par : } \|\|\| \Pi_{\mathcal{Q}}^{\mathcal{U}} \|\|\| = \sqrt{\sum_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*} \langle \Pi_{\mathcal{Q}}^{\mathcal{U}} P_i, Q_j \rangle^2}.$$

Définition 3-8 : On appelle opérateur adjoint de l'opérateur $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$, l'opérateur $\left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^*$ de $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$ dans $\mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$, tel que pour tout $f \in \mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$, et tout $g \in \mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$,

$$\langle \Pi_{\mathcal{Q}}^{\mathcal{Y}} f, g \rangle = \langle f, \left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^* g \rangle .$$

Proposition 3-9 : On a l'égalité : $\left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^* = \Pi_{\mathcal{Y}}^{\mathcal{Q}}$.

Référence: th 1-3 Csaki-Fischer .

Définition 3-10 : On appelle valeur propre des opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$ et $\left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^*$ associée aux vecteurs propres f et g , $f \in \mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $g \in \mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$, $f \neq 0$ ou $g \neq 0$, tout réel λ tel que $\Pi_{\mathcal{Q}}^{\mathcal{Y}} f = \lambda g$ et $\left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^* g = \lambda f$.

Avec ces notations, on peut généraliser la corrélation maximale (th 1-2, Csaki, Fischer).

Proposition 3-11 : On a l'égalité suivante : $\mathbf{S}(\mathcal{Y}, \mathcal{Q}) = \sup \rho(f(X), g(Y))$, lorsque f parcourt $\mathcal{Y}(\mathbb{R}^p, \mu')$ et g , $\mathcal{Q}(\mathbb{R}^q, \mu'')$.

Maintenant, on a une condition pour l'existence de fonctions f et g telles que $\mathbf{S}(\mathcal{Y}, \mathcal{Q}) = \rho(f(X), g(Y))$.

Proposition 3-12 : Pour que $\mathbf{S}(\mathcal{Y}, \mathcal{Q}) = \langle f, g \rangle$, avec $f \in \mathcal{Y}(\mathbb{R}^p, \mu')$, $g \in \mathcal{Q}(\mathbb{R}^q, \mu'')$, $\|f\| = \|g\| = 1$, il faut et il suffit que f et g forment une paire de vecteurs propres des opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$ et $\left(\Pi_{\mathcal{Q}}^{\mathcal{Y}}\right)^*$, associés à la valeur propre $\lambda = \mathbf{S}(\mathcal{Y}, \mathcal{Q})$.

Référence : Csaki-Fischer, théorème 1-5.

Rappelons que, s'il n'existe pas toujours f et g telles que $\langle f, g \rangle = \mathbf{S}(\mathcal{Y}, \mathcal{Q})$, c'est parce que la boule unité d'un espace de Hilbert de dimension infinie n'est pas compact (cf Schwartz, paragr 12, p 133). D'autre part, on sait que, sous l'hypothèse de distributions ϕ^2 -bornées, les opérateurs associés sont des opérateurs d'Hilbert-Schmidt. Or ces opérateurs sont complètement continus, et, par définition, un opérateur complètement continu transforme un ensemble faiblement compact d'un espace de Hilbert en un ensemble fortement compact (cf Akhiezer, Glazmann, paragr 27, p 56). C'est donc pour cette raison que les $(\xi^{(i)}, \eta^{(i)})$ définis en 3-3 existent toujours.

A ce sujet, on peut aussi rappeler que la double norme $\| \cdot \|$ introduite en 3-7 définit une généralisation du ϕ^2 de Pearson (Csaki-Fischer, th 1-7).

$$\text{Proposition 3-13 : On a l'égalité : } \|\Pi_{\mathcal{Q}}^{\mathcal{P}}\| = \sqrt{\sum_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*} \rho_{i,j}^2}.$$

D'autre part, il nous sera utile de remarquer qu'il y a identité entre les valeurs propres des opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{P}}$ et $(\Pi_{\mathcal{Q}}^{\mathcal{P}})^*$, et de l'opérateur $\mathcal{Y}^{\mathcal{P}}$.

Proposition 3-14 : Soit ν une valeur propre de $\mathcal{Y}^{\mathcal{P}}$. Alors $0 \leq \nu \leq 1$. D'autre part, si f est un vecteur propre associé à la valeur propre $\nu=0$, alors f est orthogonal à $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$.

De plus, soit $\lambda \in \mathbb{R}^*$. Alors, les trois assertions suivantes sont équivalentes.

- 1) λ ou $-\lambda$ est valeur propre de $\Pi_{\mathcal{Q}}^{\mathcal{P}}$ et $(\Pi_{\mathcal{Q}}^{\mathcal{P}})^*$, de vecteurs propres associés : (f, g) .
- 2) λ^2 est valeur propre de $\mathcal{Y}^{\mathcal{P}}$, de vecteur propre associé f , et $\Pi_{\mathcal{Q}}^{\mathcal{P}} f = \lambda g$.
- 3) λ^2 est valeur propre de $\mathcal{Y}^{\mathcal{Q}}$, de vecteur propre associé g , et $\Pi_{\mathcal{P}}^{\mathcal{Q}} g = \lambda f$.

Démonstration : Définissons γ , φ et ψ par $\gamma = \Pi_{\mathcal{Q}}^{\mathcal{P}} f$, $\varphi = \mathcal{Y}^{\mathcal{P}} f$ et $\psi = f - \varphi$. Alors, d'après les propriétés des projections orthogonales, $|\nu| \leq 1$, $\|\varphi\| \leq \|\gamma\| \leq \|f\|$ et $\|\psi\| \leq \|\gamma\| \leq \|f\|$. Comme $\varphi = \nu f$ et $\psi = (1-\nu)f$, on en déduit que $|1-\nu| \leq 1$, i.e. $\nu \geq 0$.

Supposons maintenant $\mathcal{Y}^{\mathcal{P}} f = 0$. Si $\Pi_{\mathcal{Q}}^{\mathcal{P}} f = 0$, f est orthogonal à $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$.

Si $\Pi_{\mathcal{Q}}^{\mathcal{P}} f = \gamma \neq 0$, alors, $\Pi_{\mathcal{P}}^{\mathcal{Q}} \gamma = 0$. Donc γ est orthogonal à $\mathcal{P}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$, et donc $\Pi_{\mathcal{Q}}^{\mathcal{P}} f = 0$.

Maintenant, supposons $\mathcal{Y}^{\mathcal{P}} f = \lambda^2 f$, et définissons g par $\lambda g = \Pi_{\mathcal{Q}}^{\mathcal{P}} f$.

Donc, $\lambda^2 f = \mathcal{Y}^{\mathcal{P}} f = \Pi_{\mathcal{P}}^{\mathcal{Q}} \circ \Pi_{\mathcal{Q}}^{\mathcal{P}} f = \Pi_{\mathcal{P}}^{\mathcal{Q}} \lambda g$, i.e. $\Pi_{\mathcal{P}}^{\mathcal{Q}} g = \lambda f$.

On a donc montré que 2) implique 1). On en déduit le reste de la proposition.

L'intérêt de se ramener à l'étude des valeurs propres de $\mathcal{Y}^{\mathcal{P}}$ réside dans le fait que cet opérateur est auto-adjoint. En effet, on peut énoncer le résultat suivant, corollaire évident de 3-9.

Proposition 3-15 : l'opérateur $\mathcal{Y}^{\mathcal{P}}$ est auto-adjoint.

III-C : REPRESENTATION MATRICIELLE

Nous allons voir maintenant que la matrice des coefficients de corrélation d'ordre supérieur n'est autre que la matrice de l'opérateur $\Pi_{\mathcal{Y}}^{\mathcal{Q}}$.

Comme cette matrice est éventuellement infinie, il faudra donc avoir recours à la théorie des matrices infinies. Or, les opérateurs de projection orthogonale sont des opérateurs bornés, et dans ce cas, cette théorie ne présente pas de difficultés majeures. Aussi les résultats suivants sont assez évidents : on pourra, par exemple, se référer aux livres de, Akhiezer paragraphe 26, Smirnov chapitre 5, et Weidmann chapitre 6-3.

Notations 3-16 : On note par \mathbf{R} la matrice de transposée : $\mathbf{R} = \{\{\rho_{i,j}\}\}_{(i,j) \in \mathcal{N}^* \times \mathcal{M}^*}$,
et par \mathbf{F} , la matrice symétrique : $\mathbf{F} = \left\{ \left\{ \sum_{i \in \mathcal{N}^*} \rho_{i,j} \rho_{i,j'} \right\} \right\}_{(j,j') \in \mathcal{M}^* \times \mathcal{M}^*}$.

Il faut remarquer que la matrice \mathbf{F} est bien définie. En effet, c'est une conséquence du lemme suivant.

Lemme 3-17 : Pour tout $(j,j') \in \mathcal{M}^* \times \mathcal{M}^*$, on a les inégalités suivantes

$$\sum_{i \in \mathcal{N}^*} \rho_{i,j}^2 \leq 1 \quad \text{et} \quad \left| \sum_{i \in \mathcal{N}^*} \rho_{i,j} \rho_{i,j'} \right| \leq 1.$$

En effet, comme $\Pi_{\mathcal{Y}}^{\mathcal{Q}}$ est une projection orthogonale, $\Pi_{\mathcal{Y}}^{\mathcal{Q}} Q_j = \sum_{i \in \mathcal{N}^*} \mathbb{E}[P_i(X) Q_j(Y)] P_i$,

$$\text{et} \quad \sum_{i \in \mathcal{N}^*} \rho_{i,j}^2 = \|\Pi_{\mathcal{Y}}^{\mathcal{Q}} Q_j\|^2 \leq \|Q_j\|^2 = 1,$$

ce qui prouve la première inégalité. On en déduit la seconde en utilisant l'inégalité de Schwartz.

Remarquons que les égalités précédentes montrent aussi que la matrice des $\rho_{i,j}$ est la matrice de la projection orthogonale définie par l'opérateur $\Pi_{\mathcal{Y}}^{\mathcal{Q}}$.

Aussi, on peut énoncer le théorème suivant.

Théorème 3-18 Les matrices, respectivement, \mathbf{R} , ${}^t\mathbf{R}$, \mathbf{F} fournissent une représentation des opérateurs, respectivement $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$, $\Pi_{\mathcal{Y}}^{\mathcal{Q}}$ et $\mathcal{Y}^{\mathcal{Q}}$, dans les bases, respectivement, $\{P_i\}_{i \in \mathcal{N}^*}$ et $\{Q_j\}_{j \in \mathcal{M}^*}$, $\{Q_j\}_{j \in \mathcal{M}^*}$ et $\{P_i\}_{i \in \mathcal{N}^*}$, $\{Q_j\}_{j \in \mathcal{M}^*}$ et $\{Q_j\}_{j \in \mathcal{M}^*}$.

Pour la matrice associée à $\mathcal{Y}^{\mathcal{Q}}$, il suffit d'utiliser les produits de matrices infinies qui sont la simple généralisation - dans le cas d'opérateurs bornés - des produits de matrices finies (ch 26, Akhiezer, p 218 Smirnov). De même, on a le théorème suivant.

Proposition 3-19 On a l'égalité $\mathbf{F} = \mathbf{R} {}^t\mathbf{R}$.

Donc, finalement, pour trouver les coefficients de corrélation canonique, il faudra chercher les valeurs propres des matrices \mathbf{R} ou \mathbf{F} .

III-C : COEFFICIENTS DE CORRELATION CANONIQUE GENERALISES

On va voir maintenant comment obtenir une généralisation naturelle des coefficients de corrélation canonique. En fait, nous allons voir que, lorsque le spectre de $\mathcal{Y}^{\otimes 2}$ est dénombrable, les valeurs propres des opérateurs de projection orthogonale fournissent une généralisation évidente des coefficients de corrélation canonique. Dans le cas contraire, on aura recours aux travaux de Dauxois et Pousse sur ce sujet, et on obtiendra ainsi une généralisation de la corrélation canonique la plus complète possible.

Enfin, comme dans le cas pratique, on se limitera à l'étude de matrices finies, on remarquera alors qu'en procédant de cette façon, on obtiendra bien une approximation de \mathbf{Q} .

Rappelons d'abord que, dans le cas où $F_{X,Y}$ n'est pas ϕ^2 -bornée, on n'est pas assuré de pouvoir définir des fonctions canoniques. Cependant, dans le cas où c'est possible, on a une première généralisation simple de 3-3.

Généralisation des notations 3-20 : On définit toujours par récurrence les fonctions canoniques de $\mathcal{Y}^{\mu}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, $\{\xi^{(i)}\}_{i \in \Gamma}$ et $\{\eta^{(i)}\}_{j \in \Gamma}$, de la manière suivante. On pose d'abord $\xi^{(0)} = \eta^{(0)} \equiv 1$.

Alors, la i -ème paire $(\xi^{(i)}, \eta^{(i)})$ sera définie, lorsque elle existe, comme une paire de fonctions $(\xi^{(i)}, \eta^{(i)})$ telle que $\rho(\xi^{(i)}(X), \eta^{(i)}(Y)) = \sup[\rho(f(X), g(Y))]$ lorsque f parcourt l'orthogonal de $\{\xi^{(1)}, \dots, \xi^{(i-1)}\}$ dans $\mathcal{Y}^{\mu}(\mathbb{R}^p, \mu')$, et g l'orthogonal de $\{\eta^{(1)}, \dots, \eta^{(i-1)}\}$ dans $\mathcal{Q}(\mathbb{R}^q, \mu'')$. On impose de plus $\|\xi^{(i)}\| = \|\eta^{(i)}\| = 1$, et on définit toujours le coefficient de corrélation canonique associé par $\rho_i = \rho(\xi^{(i)}(X), \eta^{(i)}(Y)) \geq 0$.

On note alors par $\{\xi_i\}_{i \in \mathfrak{N}}$ et $\{\eta_j\}_{j \in \Delta}$, deux bases orthonormales de, respectivement, $\mathcal{Y}^{\mu}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$ qui complètent $\{\xi^{(i)}\}_{i \in \Gamma}$ et $\{\eta^{(i)}\}_{j \in \Gamma}$, i.e. $\{\xi^{(i)}\}_{i \in \Gamma} \subset \{\xi_i\}_{i \in \mathfrak{N}}$ et $\{\eta^{(i)}\}_{j \in \Gamma} \subset \{\eta_j\}_{j \in \Delta}$ avec $\Gamma \subset \mathfrak{N}$ et $\Gamma \subset \Delta$.

Maintenant, les fonctions canoniques ainsi définies possèdent toujours la propriété de bi-orthogonalité.

Théorème 3-21 : On garde les notations de 3-20, et soient $i \in \mathfrak{N}$ et $j \in \Delta$. On suppose $i \in \Gamma$ ou $j \in \Gamma$. Alors, $\mathbb{E}(\xi_i(X)\eta_j(Y)) = \rho_i \delta_{i,j}$.

Il suffit de reprendre la démonstration de Lancaster sous les hypothèses de 3-20 (th 3-2, p 94, 1969, et th 2-1, p 124, 1980).

D'autre part, les couples de fonctions canoniques forment des couples de vecteurs propres.

Théorème 3-22 : On garde les notations de 3-20, et soit $i \in \Gamma$. Alors, le couple $\xi^{(i)}, \eta^{(i)}$ est un couple de vecteurs propres des opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{P}}$ et $\left(\Pi_{\mathcal{Q}}^{\mathcal{P}}\right)^*$, de valeur propre associée ρ_i .

En effet, on déduit de 3-21, que
$$\Pi_{\mathcal{Q}}^{\mathcal{P}} \xi^{(i)} = \rho_i \eta^{(i)} \quad \text{et} \quad \Pi_{\mathcal{P}}^{\mathcal{Q}} \eta^{(i)} = \rho_i \xi^{(i)}.$$

Malheureusement, sans l'hypothèse ϕ^2 -bornée, l'existence de bases bi-orthogonales n'est plus assurée a priori. Cependant, on peut énoncer le théorème suivant.

Théorème 3-23 : Il existe toujours deux bases orthonormales $\{\tau_i\}_{i \in \mathcal{E} \cup \mathcal{V} \cup \mathcal{K}}$ et $\{\zeta_j\}_{j \in \mathcal{E} \cup \mathcal{W} \cup \mathcal{L}}$, de, respectivement, $\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, vérifiant les assertions suivantes.

- 1) \mathcal{E}, \mathcal{V} et \mathcal{K} sont des ensembles dénombrables deux à deux disjoints et pouvant être vides, et il en est de même pour \mathcal{E}, \mathcal{W} et \mathcal{L} .
- 2) Si $i \in \mathcal{E}$, le couple τ_i, ζ_i est un couple de vecteurs propres de $\Pi_{\mathcal{Q}}^{\mathcal{P}}$ et $\left(\Pi_{\mathcal{Q}}^{\mathcal{P}}\right)^*$, de valeur propre associée $\lambda_i \neq 0$.
- 3) Si $i \in \mathcal{V}$, alors τ_i est orthogonal à $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$. De plus, si $j \in \mathcal{W}$, alors, ζ_j est orthogonal à $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$.
- 4) Il n'y a aucun vecteur propre de, respectivement, $\mathcal{U}^{\mathcal{P}}$ et $\mathcal{U}^{\mathcal{Q}}$ dans les espaces engendrés par, $\{\tau_i\}_{i \in \mathcal{K}}$ et $\{\zeta_j\}_{j \in \mathcal{L}}$.
- 5) Si $i \in \mathcal{E}$ ou $j \in \mathcal{E}$, alors, $\mathbb{E}(\tau_i(X)\zeta_j(Y)) = \lambda_i \delta_{ij}$.

Démonstration : Ce résultat se déduit aisément des propriétés des opérateurs auto-adjoints. En effet, d'après 3-15, $\mathcal{U}^{\mathcal{P}}$ est auto-adjoint. De plus, $\mathcal{U}^{\mathcal{P}}$ est borné. Donc les sous-espaces propres associés à des valeurs propres distinctes sont fermés et orthogonaux (p 81, et th 3, p 86, Akhiezer-Glazman).

De plus, on sait que, si on note par $\mathcal{U}_p^{\mathcal{P}}$ le sous espace de $L^2(\mathbb{R}^p, \mu')$ engendré par les vecteurs propres de $\mathcal{U}^{\mathcal{P}}$, on peut écrire $\mathcal{U}(\mathbb{R}^p, \mu')$ sous la forme : $\mathcal{U}(\mathbb{R}^p, \mu') = \mathcal{U}_p^{\mathcal{P}} \oplus \mathcal{U}_c^{\mathcal{P}}$, où $\mathcal{U}_c^{\mathcal{P}}$ représente un sous-espace de $\mathcal{U}(\mathbb{R}^p, \mu')$ n'admettant aucun vecteur propre de $\mathcal{U}^{\mathcal{P}}$ (cf p 357, Riesz et Nagy). On choisit alors $\{\tau_i\}$, une base orthonormale de vecteurs propres de $\mathcal{U}_p^{\mathcal{P}}$, et on définit ζ_i par $\lambda_i \zeta_i = \Pi_{\mathcal{Q}}^{\mathcal{P}} \tau_i$, lorsque τ_i est un vecteur propre de $\mathcal{U}^{\mathcal{P}}$ de valeur propre $\lambda_i \neq 0$.

On complète alors les familles $\{\tau_i\}$ et $\{\zeta_j\}$ de façon à obtenir des bases de, respectivement,

$\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, et de façon à ce qu'une sous-famille de $\{\zeta_j\}$ engendre aussi le sous-espace des vecteurs propres de $\mathcal{U}^{\otimes 2}$ de valeur propre nulle. Il est facile de voir, en appliquant 3-14, que les familles $\{\tau_i\}$ et $\{\zeta_j\}$ ainsi définies ont bien les propriétés annoncées.

Finalement, il est clair que les ρ_i sont des λ_i particuliers. Plus précisément, on peut énoncer le théorème suivant.

Théorème 3-24 : On garde les notations de 3-23. Alors, on peut écrire \mathcal{E} sous la forme: $\mathcal{E} = \mathcal{R} \cup \mathcal{S}$, $\mathcal{R} \cap \mathcal{S} = \emptyset$, \mathcal{R} ou \mathcal{S} pouvant être vides, \mathcal{R} et \mathcal{S} vérifiant les assertions suivantes.

- 1) Il existe deux familles orthonormales $\{\xi^{(i)}\}_{i \in \Gamma}$ et $\{\eta^{(j)}\}_{j \in \Gamma}$ de $\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, construites selon 3-20, et telles que, pour tout $i \in \mathcal{R}$, il existe $k \in \Gamma$, telle que $\tau_i = \xi^{(k)}$, $\zeta_i = \eta^{(k)}$.
- 2) Si $i \in \mathcal{S}$, alors, de quelque façon que soient construites les familles $\{\xi^{(i)}\}_{i \in \Gamma}$ et $\{\eta^{(j)}\}_{j \in \Gamma}$ selon les hypothèses de 3-20, il ne peut exister $k \in \Gamma$, telle que $\tau_i = \xi^{(k)}$ et $\zeta_i = \eta^{(k)}$.
- 3) Soient $i \in \mathcal{R}$ et $i' \in \mathcal{S}$. Alors $\lambda_i > \lambda_{i'}$.

Les assertions 1) et 2) sont presque évidentes. Quant à l'assertion 3), par construction, il est clair que $\rho_k = \lambda_i \geq \lambda_{i'}$. Maintenant si $\lambda_i = \lambda_{i'}$, cela signifie que λ_i est valeur propre. Il suffit alors d'imposer à la famille $\{\xi^{(i)}\}_{i \in \Gamma}$ de contenir une base du sous-espace associé à cette valeur propre.

Il est donc clair que les familles $\{\tau_i\}$ et $\{\zeta_j\}$ généralisent de façon naturelle les définitions de 3-3 et 3-20. On peut donc appeler les λ_i coefficients de corrélation canonique généralisés. En fait, on rejoint ici les définitions de Dauxois et Pousse sur l'analyse canonique généralisée (déf 3, p 363). Aussi, on adopte les définitions suivantes.

Définitions 3-25 : On appelle les couples (τ_i, ζ_i) , $i \in \mathcal{E}$, couples de fonctions canoniques généralisées de $\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$, de coefficients de corrélation canonique généralisés associés λ_i .

De plus, lorsque $\mathcal{K} = \mathcal{L} = \emptyset$, on appelle le triplet $(\{\lambda_i\}_{i \in \mathcal{E}}, \{\tau_i\}_{i \in \mathcal{E} \cup \mathcal{V}}, \{\zeta_i\}_{i \in \mathcal{E} \cup \mathcal{W}})$ analyse canonique de $\mathcal{U}(\mathbb{R}^p, \mu')$ et $\mathcal{Q}(\mathbb{R}^q, \mu'')$.

Très précisément, Dauxois et Pousse ont appelé le triplet ci-dessus analyse canonique de type dénombrable. En effet, ils ont pu étendre leur définition au cas le plus général. Pour cela, ils ont eu besoin du théorème suivant (p 357-358).

Théorème 3-26 : Soit \mathbf{I} l'opérateur identité et soient $\Pi_{\mathcal{U}}$ et $\Pi_{\mathcal{Q}}$ les opérateurs de projection orthogonale sur $\mathcal{U}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$. Soit A l'opérateur défini par $(\Pi_{\mathcal{U}} + \Pi_{\mathcal{Q}} - \mathbf{I})^2$. Alors, il existe une mesure π dont le support est identique au spectre de A , et telle qu'il existe une représentation isométrique de $L = L^2(\mathbb{R}^{p+q}, \mathcal{Q})$ par une intégrale Hilbertienne :

$$\mathfrak{H} = \int^{\oplus} L(u) \cdot \pi(du) .$$

De plus, si l'on note par $\mathfrak{H}_{\mathcal{Y}_p}$ et $\mathfrak{H}_{\mathcal{Q}}$ les transformées des sous-espaces $\mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$, respectivement, et par $\mathcal{S}_{\mathcal{Y}_p}$ et $\mathcal{S}_{\mathcal{Q}}$, les spectres des opérateurs, respectivement, $\mathcal{Y}^{\mathcal{Y}_p}$ et $\mathcal{Y}^{\mathcal{Q}}$, alors on peut écrire $\mathfrak{H}_{\mathcal{Y}_p}$ et $\mathfrak{H}_{\mathcal{Q}}$ sous forme des intégrales Hilbertiennes suivantes :

$$\mathfrak{H}_{\mathcal{Y}_p} = \int^{\oplus}_{\mathcal{S}_{\mathcal{Y}_p}} L_{\mathcal{Y}_p}(u) \pi(du) \quad \text{et} \quad \mathfrak{H}_{\mathcal{Q}} = \int^{\oplus}_{\mathcal{S}_{\mathcal{Q}}} L_{\mathcal{Q}}(u) \pi(du) .$$

Enfin, $\mathcal{S}_{\mathcal{Y}_p} \cup \{0\} = \mathcal{S}_{\mathcal{Q}} \cup \{0\}$ (que l'on note \mathcal{S}).

Remarquons que la restriction de A à $\mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ est égale à $\mathcal{Y}^{\mathcal{Y}_p}$, et que A est transformée par l'isométrie en la multiplication par u : $\{g(u)\}_{u \in \mathbb{R}^q} \in \mathfrak{H} \mapsto \{ug(u)\}_{u \in \mathbb{R}^q} \in \mathfrak{H}$.

On déduit alors de 3-26 le théorème suivant.

Théorème 3-27 : Soit ν la restriction de π à \mathcal{S} . Alors $\mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$ sont isométriques à, respectivement, $\int^{\oplus}_{\mathcal{S}} L_{\mathcal{Y}_p}(u) \nu(du)$ et $\int^{\oplus}_{\mathcal{S}} L_{\mathcal{Q}}(u) \nu(du)$.

On peut maintenant définir l'analyse canonique la plus générale.

Définition 3-28 : On garde les notations de 3-26 et 3-27. On appelle alors analyse canonique de $\mathcal{Y}(\mathbb{R}^p, \mu') \otimes \mathbb{R}^q$ et $\mathbb{R}^p \otimes \mathcal{Q}(\mathbb{R}^q, \mu'')$, le quadruplet $(\mathcal{S}, \nu, \{L_{\mathcal{Y}_p}(u)\}_{u \in \mathcal{S}}, \{L_{\mathcal{Q}}(u)\}_{u \in \mathcal{S}})$.

En fait, avec ces théorèmes on effectue l'analyse spectrale de A. On se place donc dans le cas le plus général et on généralise ainsi la corrélation canonique même au cas où il n'existe pas de vecteurs propres.

Signalons aussi que, dans le cas où il existe une base de vecteurs propres, on peut montrer que la connaissance du quadruplet de 3-28 est équivalente à celle de $(\{\lambda_i\}, \{\tau_i\}, \{\zeta_j\})$. Comme exemple de cette situation, il y a, bien sûr, le cas des distributions ϕ^2 - bornées à valeurs dans \mathbb{R}^2 .

A ce sujet, on a déjà remarqué lors du théorème 3-13 que la norme $\| \cdot \|$ fournit une généralisation du ϕ^2 de Pearson. On complète maintenant ce théorème par la proposition suivante.

Théorème 3-29 : Si $\| \Pi_Y^X \| < +\infty$, alors $\mathcal{K}=\mathcal{L}=\mathcal{S}=\emptyset$. De plus, $\| \Pi_Y^X \|^2 = \sum_{i \in \Gamma} \rho_i^2$.

Démonstration : D'après 6-22, page 152 de Weidmann, Π_Y^X est un opérateur de Hilbert-Schmidt. Donc, d'après le théorème 7-10, page 175 de Weidmann, $\| \Pi_Y^X \|^2 = \sum_{i \in \mathcal{E}} \lambda_i^2$.

De plus, d'après le théorème 7-6, page 170 de Weidmann, $\mathcal{K}=\mathcal{L}=\emptyset$. Enfin, la suite $\{\lambda_i\}$ converge vers 0 . Il est alors facile d'en déduire qu'elle peut être identifiée à la suite $\{\rho_i\}$, et donc, que $\mathcal{S}=\emptyset$.

Enfin, des théorèmes, 2-3, 2-5, 3-13, 3-23, et 3-29 , on déduit la décomposition de $F_{X,Y}$ dans le cas où il existe une base de vecteurs propres

Théorème 3-30 : On garde les notations de 3-20 et 3-23 , et on suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$, et que $\mathcal{K}=\mathcal{L}=\emptyset$.

Alors, en tout point $(x,y) \in \mathbb{R}^{p+q}$,

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) + \sum_{i \in \mathcal{E}} \lambda_i \left(\int_{u=-\infty}^{u=x} \tau_i(u) \cdot \mu'(du) \right) \left(\int_{v=-\infty}^{v=y} \zeta_i(v) \cdot \mu''(dv) \right) .$$

De plus, si $\| \Pi_Y^X \| < +\infty$, alors (X,Y) admet une densité par rapport à $\mu' \otimes \mu''$:

$$f = 1 + \sum_{i \in \Gamma} \rho_i \xi^{(i)} \otimes \eta^{(i)} \text{ dans } L^2(\mathbb{R}^{p+q}, \mu' \otimes \mu'') .$$

Maintenant, l'étude pratique de la corrélation canonique se ramènera au calcul des vecteurs et valeurs propres des matrices symétriques finies $\left\{ \left\{ \sum_{i=1}^h \rho_{i,j} \rho_{i,j'} \right\} \right\}_{(j,j') \in \{1,2,\dots,k\}^2}$.

Aussi, il convient de remarquer que l'on a bien ainsi une approximation de la probabilité Q .

Théorème 3-31 : On suppose que $\mathcal{M}=\mathcal{N}=\mathbb{N}$ et que $\{P_i\}_{i \in \mathbb{N}}$ et $\{Q_j\}_{j \in \mathbb{N}}$ sont des bases de $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$. Soient h et $k \in \mathbb{N}^*$. On note alors par \mathcal{U}_h et \mathcal{Q}_k les sous-espaces de $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$ engendrés par $\{P_i\}_{i=1,2,\dots,h}$ et $\{Q_j\}_{j=1,2,\dots,k}$.

Alors, il existe deux bases orthonormales, de, respectivement, \mathcal{U}_h et \mathcal{Q}_k : $\{\xi_{h,k}^{(i)}\}_{i=1,2,\dots,h}$ et $\{\eta_{h,k}^{(j)}\}_{j=1,2,\dots,k}$ telles que $\mathbb{E} [\xi_{h,k}^{(i)}(X) \eta_{h,k}^{(j)}(Y)] = \rho_i \delta_{i,j}$ pour tout $1 \leq i \leq h$ et tout $1 \leq j \leq k$.

De plus, les couples $(\xi_{h,k}^{(i)}, \eta_{h,k}^{(i)})$ sont des couples de vecteurs propres des opérateurs $\Pi_{\mathcal{Q}_k}^{\mathcal{Y}_h}$ $\left(\Pi_{\mathcal{Q}_k}^{\mathcal{Y}_h}\right)^*$, de valeurs propres associées $\rho_i^{h,k}$.

Enfin, en tout point $(x,y) \in \mathbb{R}^{p+q}$,

$$F_X(x)F_Y(y) + \lim_{h \rightarrow \infty, k \rightarrow \infty} \left\{ \sum_{i=1}^{\min(h,k)} \rho_i^{h,k} \left(\int_{u=-\infty}^{u=x} \xi_{h,k}^{(i)}(u) \cdot \mu'(du) \right) \left(\int_{v=-\infty}^{v=y} \eta_{h,k}^{(i)}(v) \cdot \mu''(dv) \right) \right\}.$$

En reprenant les démonstrations de 2-3 et 2-4, on prouve facilement ce théorème en appliquant les propriétés élémentaires des opérateurs $\mathcal{Y}^{\mathcal{Y}}$ citées dans le paragraphe III-B.

IV : CARACTERISATION DES DIVERS TYPES DE DEPENDANCE FONCTIONNELLE

On va voir maintenant comment appliquer les théorèmes énoncés dans le chapitre précédent dans le cas de dépendance fonctionnelle du type $f(X) = g(Y)$. En particulier, on trouvera une condition simple sur les coefficients de corrélation d'ordre supérieur pour détecter les dépendances du type $Y=f(X)$.

On énonce d'abord le résultat le plus général sous deux formes distinctes.

Théorème 4-1: Soit λ^2 , une valeur propre de l'opérateur $\mathcal{Y}^{\mathcal{Q}}$. Alors $0 \leq \lambda^2 \leq 1$. De plus, pour que 1 soit valeur propre de $\mathcal{Y}^{\mathcal{Q}}$, il faut et il suffit qu'il existe une relation fonctionnelle du type $f(X) = g(Y)$ presque-sûrement, où $f \in \mathcal{Y}^{\mathcal{P}}(\mathbb{R}^p, \mu')$, $f \neq 0$, et $g \in \mathcal{Q}(\mathbb{R}^q, \mu'')$, $g \neq 0$. Dans ce cas, g est un vecteur propre associé à la valeur propre 1, et $f = \Pi_{\mathcal{Y}^{\mathcal{P}}}^{\mathcal{Q}} g$.

De plus, si X et Y sont indépendants, alors toutes les valeurs propres de $\mathcal{Y}^{\mathcal{Q}}$ sont nulles.

Théorème 4-2 : Soit λ une valeur propre des opérateurs $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$ et $(\Pi_{\mathcal{Q}}^{\mathcal{Y}})^*$. Alors, $-1 \leq \lambda \leq 1$.

De plus, pour que 1 (ou -1) soit valeur propre de $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$ et $(\Pi_{\mathcal{Q}}^{\mathcal{Y}})^*$, il faut et il suffit qu'il existe une relation fonctionnelle du type $f(X)=g(Y)$ presque sûrement, où $f \in \mathcal{Y}^{\mathcal{P}}(\mathbb{R}^p, \mu')$, $f \neq 0$, et $g \in \mathcal{Q}(\mathbb{R}^q, \mu'')$, $g \neq 0$. Dans ce cas, f et g sont des vecteurs propres associés à la valeur propre 1.

De plus, si X et Y sont indépendants, toutes les valeurs propres de $\Pi_{\mathcal{Q}}^{\mathcal{Y}}$ et $(\Pi_{\mathcal{Q}}^{\mathcal{Y}})^*$ sont nulles.

Ces théorèmes se déduisent facilement de 3-11, 3-12 et 3-14.

Maintenant, si l'on impose certaines contraintes à g , par exemple d'être un polynôme de degré donné, ces conditions se ramèneront à calculer les valeurs propres d'une matrice finie : il suffit, en effet, d'appliquer 3-18.

Théorème 4-3 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ représente une base de $L^2(\mathbb{R}^p, \mu')$ et que $\mathcal{M} = \{0, 1, 2, \dots, k\}$. Alors, pour qu'il existe une relation fonctionnelle $f(X) = g(Y)$ presque sûrement où $g \in \mathcal{Q}(\mathbb{R}^q, \mu'')$, $g \neq 0$, et $f \in L^2(\mathbb{R}^p, \mu')$, $f \neq 0$, il faut et il suffit que la matrice $k \times k$: $F = \left\{ \left\{ \sum_{i \in \mathcal{N}^*} e_{i,j} e_{i,j'} \right\} \right\}_{(j,j') \in \{1,2,\dots,k\}^2}$, admette 1 pour valeur propre.

Dans ce cas, si $(\alpha_1, \alpha_2, \dots, \alpha_k)$ est un vecteur propre associé à la valeur propre 1, on peut choisir f et g de la façon suivante : $g = \sum_{j=1}^k \alpha_j Q_j$ et $f = \sum_{i \in \mathcal{N}^*} \left(\sum_{j=1}^k \alpha_j e_{i,j} \right) P_i$.

De plus, si X et Y sont indépendants, alors toutes les valeurs propres de F sont nulles.

On peut remarquer que la matrice F est symétrique, et donc diagonalisable. D'autre part, on peut aussi obtenir une condition analogue sur la matrice R .

Maintenant, pour le cas plus particulier d'une dépendance du type $Y=f(X)$, on a la proposition suivante, qui est la généralisation du théorème 1 de Lancaster 1963-b.

Proposition 4-4 : On suppose que $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont des bases de $L^2(\mathbb{R}^p, \mu')$ et $L^2(\mathbb{R}^q, \mu'')$. Alors, pour qu'il existe une fonction mesurable f telle que $Y=f(X)$ presque sûrement, il faut et il suffit que $R'R = I$, où I représente la matrice identité.

Démonstration : Supposons qu'il existe f telle que $Y=f(X)$, et soit $g \in L^2(\mathbb{R}^q, \mu'')$.

Alors $g(Y) = g[f(X)]$, et $g \circ f \in L^2(\mathbb{R}^p, \mu')$,

i.e. $(x,y) \mapsto g(y)$ est égale à $(x,y) \mapsto g[f(x)]$ dans $L^2(\mathbb{R}^{p+q}, \mathbf{0})$.

Donc, $\prod_X^Y g = g \circ f$ et $\prod_Y^X g \circ f = g$, i.e. $\prod_Y^X \circ \prod_X^Y$ est l'opérateur identité.

Donc, la matrice associée $R'R$, vérifie $R'R = I$.

Réciproquement, supposons que $R'R = I$.

Alors, l'opérateur $\prod_Y^X \circ \prod_X^Y$ est l'opérateur identité.

Soit $g \in L^2(\mathbb{R}^q, \mu'')$. Alors, $\|g\| = \|\prod_Y^X \circ \prod_X^Y g\| \leq \|\prod_X^Y g\| \leq \|g\|$, i.e. $\|\prod_X^Y g\| = \|g\|$.

Donc, d'après les propriétés des projections orthogonales, $\prod_X^Y g = g$ dans $L^2(\mathbb{R}^{p+q}, \mathbf{0})$.

Donc, il existe $\varphi_g \in L^2(\mathbb{R}^p, \mu')$ telle que $g = \varphi_g$ dans $L^2(\mathbb{R}^{p+q}, \mathbf{0})$.

On en déduit que, dans le cas où $Y_1 \in L^2(\Omega, \mathcal{A}, \mathbf{P})$, il existe $f^1 \in L^2(\mathbb{R}^p, \mu')$ telle que $Y_1 = f^1(X)$ dans $L^2(\Omega, \mathcal{A}, \mathbf{P})$.

Dans le cas où $Y_1 \notin L^2(\Omega, \mathcal{A}, \mathbf{P})$, on sait qu'il existe une suite de partitions disjointes de Ω , $\{\mathcal{O}_m\}_{m \in \mathbb{N}}$, telle que, pour tout $m \in \mathbb{N}$, $Y_1^m \in L^2(\Omega, \mathcal{A}, \mathbf{P})$, Y_1^m représentant la fonction identité à Y_1 sur \mathcal{O}_m , et nulle ailleurs, i.e. $Y_1(\omega) = \sum_{m \in \mathbb{N}} Y_1^m(\omega)$ en tout $\omega \in \Omega$.

De plus, il existe une suite $\{f_m^1\}_{m \in \mathbb{N}}$, $f_m^1 \in L^2(\mathbb{R}^p, \mu')$, $Y_1^m = f_m^1(X)$ dans $L^2(\Omega, \mathcal{A}, \mathbf{P})$.

Donc, $Y_1(\omega) = \sum_{m \in \mathbb{N}} f_m^1[X(\omega)]$ en presque tout $\omega \in \Omega$.

On raisonne de même pour Y_2, Y_3, \dots

Maintenant, sous l'hypothèse L^2 , on a une condition simple sur les $\rho_{i,1}$ qui permet de déterminer l'existence des relations fonctionnelles du type $Y=f(X)$.

Théorème 4-5 On suppose $q=1$, $Y \in L^2(\Omega, \mathcal{A}, \mathbf{P})$ et $Q_1(y) = \frac{y - \mathbb{E}(Y)}{\sigma(Y)}$.

Alors, pour qu'il existe une relation fonctionnelle $Y=f(X)$ presque sûrement, $f \in \mathcal{U}(\mathbb{R}^p, \mu')$, il faut et il suffit que $\sum_{i \in \mathcal{N}^*} [\rho_{i,1}(X, Y)]^2 = 1$.

De plus, dans ce cas, $Y = \mathbb{E}(Y) + \sigma(Y) \left(\sum_{i \in \mathcal{N}^*} \rho_{i,1} P_i(X) \right)$ presque sûrement.

Il suffit d'appliquer le théorème 4-3 dans le cas où $k=1$.

Par exemple dans le cas où les familles $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ sont les familles de polynômes orthonormaux, l'existence d'une relation fonctionnelle $Y = aX^2 + bX + c$ est équivalente à la condition $(\rho_{1,1})^2 + (\rho_{2,1})^2 = 1$. Ce résultat signifie aussi que le couple $(\rho_{1,1}, \rho_{2,1})$ mesure une dépendance quadratique.

En fait, on pourrait également obtenir le théorème 4-5 directement en remarquant que $\sum_{i \in \mathcal{N}^*} \rho_{i,1}^2$ est le coefficient de corrélation linéaire de $Q_1(Y)$ et $\sum_{i \in \mathcal{N}^*} \rho_{i,1} P_i(X)$.

D'ailleurs, tous les résultats de ce chapitre constituent une généralisation de la normalité du coefficient de corrélation linéaire comme nous le rappelle plus particulièrement le théorème suivant.

Théorème 4-6 : Soit $(i, j) \in \mathcal{N}^* \times \mathcal{M}^*$. On a l'inégalité : $|\rho_{i,j}(X, Y)| \leq 1$, l'égalité ayant lieu si et seulement si $P_i(X) = Q_j(Y)$ ou $P_i(X) = -Q_j(Y)$ presque sûrement.

V : COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR ET PROBABILITES CONDITIONNELLES

Dans ce chapitre, nous allons donner la traduction des théorèmes classiques des probabilités conditionnelles en termes de coefficients de corrélation d'ordre supérieur, et faire ainsi apparaître explicitement la probabilité, l'espérance et la densité conditionnelles.

Pour effectuer cette étude, il nous faudra supposer que les familles $\{P_i\}$ et $\{Q_j\}$ sont des bases. Aussi, on introduit les hypothèses suivantes.

Hypothèses 5-1 : Dans ce chapitre, on suppose que $\{P_i\}_{i \in \mathcal{N}}$ est une base de $L^2(\mathbb{R}^p, \mu')$, et de plus, dans les théorèmes 5-4 et 5-6, on suppose que $\{Q_j\}_{j \in \mathcal{M}}$ est une base de $L^2(\mathbb{R}^q, \mu'')$.

On explicite d'abord l'espérance conditionnelle dans le cas où Y est une variable aléatoire.

Théorème 5-2 : On garde les hypothèses de 5-1, et on suppose de plus que $q=1$, $\mathbb{E}(Y^2) < +\infty$, et que $Q_1(y) = \frac{y - \mathbb{E}(Y)}{\sigma(Y)}$. Soit $\mathbb{E}(Y|X=x)$, une version régulière de l'espérance conditionnelle de

Y sachant $X=x$. Alors $\mathbb{E}(Y|X=x) \in L^2(\mathbb{R}^p, \mu')$, et on a l'égalité :

$$\mathbb{E}(Y|X=x) = \mathbb{E}(Y) + \sigma(Y) \sum_{i \in \mathcal{N}^*} \rho_{i,1} P_i(x) \quad \text{dans } L^2(\mathbb{R}^p, \mu').$$

Démonstration : Comme $\mathbb{E}(Y^2) < +\infty$, $\mathbb{E}(Y|X=x) \in L^2(\mathbb{R}^p, \mu')$ et est la projection orthogonale dans $L^2(\mathbb{R}^{p+1}, \mathcal{Q})$ de $(x_1, x_2, \dots, x_p, y) \mapsto y$ sur $L^2(\mathbb{R}^p, \mu') \otimes \mathbb{R}$, ce qui suffit à démontrer le théorème.

On déduit alors de ce théorème l'expression de la probabilité conditionnelle.

Théorème 5-3 : On garde les hypothèses de 5-1, et on note par $P(\cdot | X=x)$ la probabilité conditionnelle sachant $X=x$. Alors, pour tout $A \in \mathcal{A}$, on a l'égalité :

$$P(A|X=x) = P(A) + \sum_{i \in \mathcal{N}^*} \left(\int_A P_i(X) dP \right) P_i(x) \quad \text{dans } L^2(\mathbb{R}^p, \mu').$$

De plus, on a l'inégalité :
$$\sum_{i \in \mathcal{N}^*} \left(\int_A P_i(X) \cdot dP \right)^2 \leq P(A)[1-P(A)] .$$

On peut maintenant expliciter la probabilité conditionnelle de Y sachant $X=x$.

Théorème 5-4 : On garde les hypothèses de 5-1 et on note par μ''^x la distribution conditionnelle de Y sachant $X=x$. Alors pour tout borélien F de \mathbb{R}^q , $P(Y \in F | X=x) \in L^2(\mathbb{R}^p, \mu')$, et on a l'égalité

$$P(Y \in F | X=x) = \mu''(F) + \sum_{i \in \mathcal{N}^*} \left[\int_F \left(\sum_{j \in \mathcal{M}^*} \rho_{i,j} Q_j \right) d\mu'' \right] P_i(x) \quad \text{dans } L^2(\mathbb{R}^p, \mu') .$$

De plus, on a la majoration :
$$\sum_{i \in \mathcal{N}^*} \left[\int_F \left(\sum_{j \in \mathcal{M}^*} \rho_{i,j} Q_j \right) d\mu'' \right]^2 \leq \mu''(F)[1-\mu''(F)] .$$

Démonstration : On peut écrire l'égalité
$$\int_{Y \in F} P_i(X) \cdot dP = \int_F \mathbb{E}(P_i(X) | Y=y) \cdot \mu''(dy) ,$$

et il ne reste plus qu'à utiliser 5-2 , 5-3 et l'orthonormalité des P_i .

Supposons maintenant que (X,Y) admette une densité de probabilité f par rapport à $\mu' \otimes \mu''$, et que X et Y admettent respectivement d' et d'' pour densité par rapport à la mesure de Lebesgue. Alors (X,Y) admet $f(x,y)d'(x)d''(y)$ pour densité par rapport à la mesure de Lebesgue. Donc, g^x , la densité conditionnelle de Y par rapport à la mesure de Lebesgue, sachant $X=x$, vérifie l'égalité $g^x(y) = f(x,y) \cdot d''(y)$. Donc, la densité conditionnelle de Y par rapport à μ'' sachant $X=x$, f^x , vérifie l'égalité $f^x(y) = f(x,y)$. C'est la généralisation de ce résultat qu'exprime le théorème suivant.

Théorème 5-5 : On suppose que (X,Y) admet une densité de probabilité f par rapport à $\mu' \otimes \mu''$. Alors, μ''^x admet une densité $f^x_{\mu''}$ par rapport à μ'' , μ' -presque sûrement, définie par $f^x_{\mu''}(y) = f(x,y)$ pour μ' -presque tout x .

On démontre ce théorème exactement comme le théorème classique de l'existence d'une densité conditionnelle par rapport à la mesure de Lebesgue.

Finalement la densité f est une fonction tout à fait remarquable du vecteur aléatoire (X,Y) . En effet, on vient de voir que f est à la fois la densité conditionnelle de Y sachant $X=x$, et la densité

conditionnelle de X sachant $Y=y$. De plus, il est clair que f définit complètement la dépendance. En particulier, X et Y sont indépendants si et seulement si $f \equiv 1$, et d'autre part, $\|f\|^2 - 1$ est égal au ϕ^2 de Pearson. Aussi, pour toutes ces raisons, on peut appeler f densité de dépendance de (X, Y) ou densité conditionnelle de (X, Y) .

Nous revenons maintenant au cadre particulier de notre étude, i.e. au cas où $f \in L^2$.

Théorème 5-6 : On garde les notations et hypothèses de 5-1 et 5-5, et on suppose de plus que $f \in L^2(\mathbb{R}^{p+q}, \mu' \otimes \mu'')$. Alors $f_{\mu''}^x \in L^2(\mathbb{R}^q, \mu'')$ μ' -presque sûrement. De plus, pour tout $x \in \mathbb{R}^p$, tel que $f_{\mu''}^x$ soit définie et appartienne à $L^2(\mathbb{R}^q, \mu'')$, on peut écrire

$$f_{\mu''}^x(y) = 1 + \sum_{j \in \mathcal{M}^*} \beta_j^x Q_j(y) \quad \text{dans } L^2(\mathbb{R}^q, \mu''),$$

où, pour tout $j \in \mathcal{M}^*$, β_j^x considérée comme fonction de x appartient à $L^2(\mathbb{R}^p, \mu')$ et vérifie l'égalité

$$\beta_j^x = \sum_{i \in \mathcal{N}^*} \rho_{i,j} P_i(x) \quad \text{dans } L^2(\mathbb{R}^p, \mu').$$

Démonstration : On sait que $\int \int [f(x,y)]^2 \mu''(dy) \mu'(dx) = \|f\|^2 < +\infty$. On en déduit que $f_{\mu''}^x \in L^2(\mathbb{R}^q, \mu'')$ μ' -presque sûrement.

On notera alors par $\underline{\mathbb{R}}^p$ le sous-ensemble des x appartenant à \mathbb{R}^p tel que $f_{\mu''}^x$ soit une densité et que $f_{\mu''}^x \in L^2(\mathbb{R}^q, \mu'')$.

Alors, $\mu'(\mathbb{R}^p \setminus \underline{\mathbb{R}}^p) = 0$, et de plus, pour tout $x \in \underline{\mathbb{R}}^p$, il existe une suite $\{\beta_j^x\}_{j \in \mathcal{M}}$ telle que

$$f_{\mu''}^x = \sum_{j \in \mathcal{M}} \beta_j^x Q_j \quad \text{dans } L^2(\mathbb{R}^q, \mu'').$$

On définira alors β_j^x sur $\mathbb{R}^p \setminus \underline{\mathbb{R}}^p$ en posant $\beta_j^x = 0$ si $j \neq 0$ et $\beta_0^x = 1$ sinon.

Alors, pour tout $j \in \mathcal{M}$, $\beta_j^x \in L^2(\mathbb{R}^p, \mu')$.

$$\begin{aligned} \text{En effet, } & \int [\beta_j^x]^2 \mu'(dx) = \int \left[\int [f(x,y) Q_j(y)] \mu''(dy) \right]^2 \mu'(dx) \\ & \leq \int \left[\int [f(x,y)]^2 \mu''(dy) \right] \left[\int [Q_j(y)]^2 \mu''(dy) \right] \mu'(dx) \quad (\text{d'après l'inégalité de Schwartz}) \\ & = \int \left[\int [f(x,y)]^2 \mu''(dy) \right] \mu'(dx) = \|f\|^2 < +\infty. \end{aligned}$$

Donc, on peut écrire β_j^x dans $L^2(\mathbb{R}^p, \mu')$ sous forme d'un développement en fonctions orthogonales.

Calculons alors $\langle \beta_j^x, P_i \rangle$. On a $\int [\beta_j^x P_i(x)] \mu'(dx) = \int \left[\int [f(x,y) Q_j(y)] \mu''(dy) \right] P_i(x) \mu'(dx)$
 $= \rho_{i,j}$. Donc, $\beta_j^x = \sum_{i \in \mathcal{N}^*} \rho_{i,j} P_i(x)$ dans $L^2(\mathbb{R}^p, \mu')$ si $j \geq 1$, et $\beta_0^x = 1$.

VI-COEFFICIENTS DE CORRELATION EMPIRIQUES D'ORDRE SUPERIEUR

On a déjà vu (Blacher 1984) que les polynômes orthogonaux possèdent des estimateurs naturels: les polynômes orthogonaux empiriques, et on en avait déduit des coefficients de corrélation d'ordre supérieur empiriques. En fait, pour obtenir ceux-ci, il suffit de remplacer les probabilités par les probabilités empiriques. On va maintenant généraliser cette définition et obtenir les coefficients de corrélation d'ordre supérieur empiriques les plus généraux possibles.

En fait, ces estimateurs correspondent au cas où l'on ne connaît pas les lois marginales. Dans le cas contraire, les résultats de Bosq fournissent d'autres estimateurs qui forment ainsi une deuxième famille de coefficients de corrélation d'ordre supérieur empiriques.

On rappelle que dans ce chapitre on suppose $\text{card}(\mathcal{N}^) = h < +\infty$ et $\text{card}(\mathcal{M}^*) = k < +\infty$.*

On va commencer par introduire les fonctions orthogonales empiriques. Comme celles-ci doivent être construites par la même méthode que leurs homologues non empiriques, on va devoir revenir à la définition des fonctions orthogonales par le procédé de Gram-Schmidt.

6-1 - Construction des fonctions orthogonales : Soit (Z, S) un vecteur aléatoire à valeurs dans \mathbb{R}^{h+k+2} tel que $Z = (Z_0, Z_1, \dots, Z_h)$ soit fonction de $X : Z = \varphi(X)$ et $S = (S_0, S_1, \dots, S_k)$, fonction de $Y : S = \gamma(Y)$, et tel que $Z_0 = S_0 \equiv 1$.

Soient C_s les fonctions définies par $C_s(X) = Z_s$ pour $s \in \{0, 1, \dots, h\}$. On suppose que, pour tout $s \in \{0, 1, \dots, h\}$, $C_s(Z) \in L^2(\Omega, \mathcal{A}, \mathbf{P})$. On suppose aussi que Z est de rang $h+1$, i.e. les C_s sont linéairement indépendantes.

Soit alors $\{\bar{P}_i\}_{i \in \{0, 1, 2, \dots, h\}}$ la famille suivante, définie par le procédé d'orthogonalisation de Gram-Schmidt : $\bar{P}_0 = C_0 \equiv 1$,

$$\text{et pour } i > 0, \quad \bar{P}_i = C_i - \sum_{s=0}^{i-1} \frac{\langle C_i, \bar{P}_s \rangle}{\|\bar{P}_s\|^2} \bar{P}_s.$$

On suppose alors que la famille orthonormale $\left\{ \frac{\bar{P}_s}{\|\bar{P}_s\|} \right\}_{s=0, 1, \dots, h}$ est identique à la famille

$\{P_i\}_{i \in \{0, 1, \dots, h\}}$ définie en 1-2.

On suppose que la condition correspondante pour $\{Q_j\}_{j=0, 1, \dots, k}$ est aussi vérifiée.

Par exemple, dans le cas où X est une variable aléatoire, et où $Z = (1, X, X^2, \dots, X^h)$, la famille $\{P_i\}_{i=0, 1, \dots, h}$ sera la famille des polynômes orthonormaux de degré inférieur à h .

6-2 Notation des mesures empiriques : Soit $\{X_j, Y_j\}_{j \in \mathbb{N}^*}$ une suite de vecteurs aléatoires définis sur (Ω, \mathcal{A}, P) , à valeurs dans \mathbb{R}^{p+q} , indépendants et de même loi que (X, Y) , et soit $Z_j = \varphi(X_j)$ et $S_j = \gamma(Y_j)$.

Soit alors $n \in \mathbb{N}^*$. Nous noterons par O_n, μ'_n et μ''_n , les mesures empiriques associées :

$$\mu'_n(B') = \frac{\text{card}\{j \leq n \mid X_j \in B'\}}{n}, \quad \mu''_n(B'') = \frac{\text{card}\{j \leq n \mid Y_j \in B''\}}{n},$$

$$O_n(B' \times B'') = \frac{\text{card}\{j \leq n \mid (X_j, Y_j) \in B' \times B''\}}{n},$$

pour tous boréliens B' et B'' de \mathbb{R}^p et \mathbb{R}^q .

On note alors par $\langle \cdot, \cdot \rangle_n$ et $\|\cdot\|_n$, le produit scalaire et la norme empiriques.

Remarquons que $\{Z_j, S_j\}_{j \in \mathbb{N}^*}$ est une suite de vecteurs aléatoires indépendants et de même loi que (Z, S) .

On peut maintenant définir les fonctions orthogonales empiriques.

6-3 Fonctions orthogonales empiriques : Comme Z est de rang $h+1$, on peut montrer facilement que la probabilité que les $C_s, s=0,1,\dots,h$, soient linéairement dépendants dans $L^2(\mathbb{R}^p, \mu'_n)$ tend vers 0 si n tend vers l'infini. Aussi, nous supposons qu'il n'en est rien.

Soit alors $\{\bar{P}^n_i\}_{i=0,1,2,\dots,h}$ les familles orthogonales suivantes: $\bar{P}^n_0 = C_0 \equiv 1$,

$$\text{et pour } i > 0, \quad \bar{P}^n_i = C_i - \sum_{s=0}^{i-1} \frac{\langle C_i, \bar{P}^n_s \rangle_n}{\|\bar{P}^n_s\|_n^2} \bar{P}^n_s,$$

et $\{P^n_i\}_{i=0,1,2,\dots,h}$ la famille orthonormale associée.

On définit alors de la même façon la famille orthonormale associée à μ''_n : $\{Q^n_j\}_{j=0,1,2,\dots,k}$.

Par exemple, si $Z = (1, X, X^2, \dots, X^h)$, la famille $\{P^n_i\}_{i=0,1,2,\dots,h}$ sera la famille des polynômes orthonormaux empiriques de degré inférieur à h . Il suffit, pour l'obtenir de remplacer les moments, par les moments empiriques (Blacher 1984).

En particulier, $\bar{P}^n_1(x) = x - \mathbb{E}(X)$, $\bar{P}^n_1(x) = x - \mathbb{E}_n(X)$,

$$P_1(x) = \frac{x - \mathbb{E}(X)}{\sigma(X)} \quad \text{et} \quad P^n_1(x) = \frac{x - \mathbb{E}_n(X)}{\sigma_n(X)},$$

où $\mathbb{E}_n(X)$ et $\sigma_n(X)$ représentent l'espérance et l'écart type empiriques.

On peut maintenant définir les coefficients de corrélation empiriques d'ordre supérieur.

6-4 Coefficients de corrélation empiriques d'ordre supérieur : Soit $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$. On appelle coefficient de corrélation empirique d'ordre (i,j) le nombre réel $\rho_{i,j}^n$ défini par :

$$\rho_{i,j}^n = \frac{\sum_{r=1}^n P_i^n(X_r) Q_j^n(Y_r)}{n}$$

Il est facile de voir que le coefficient de corrélation polynômiale empirique d'ordre $(1,1)$ n'est autre que le coefficient de corrélation linéaire empirique classique.

En fait, ces coefficients de corrélation empiriques d'ordre supérieur correspondent aux cas où l'on ne connaît pas les lois marginales. Lorsque celles-ci sont connues, on peut également définir d'autres estimateurs des $\rho_{i,j}$ que l'on pourra appeler, si besoin est, coefficients de corrélation empiriques d'ordre supérieur de deuxième type.

Notations 6-5 : Soit $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$. Alors, on note $\rho_{i,j}^{*n}$, le nombre réel défini par :

$$\rho_{i,j}^{*n} = \frac{\sum_{r=1}^n P_i(X_r) Q_j(Y_r)}{n}$$

Alors, ces coefficients sont des estimateurs des $\rho_{i,j}$.

Théorème 6-6 : Soit $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$. Alors $\rho_{i,j}^n$ et $\rho_{i,j}^{*n}$ convergent presque sûrement vers $\rho_{i,j}$.

Pour $\rho_{i,j}^n$, on utilise la même méthode qu'en 3-3 de Blacher 1984-a. Pour $\rho_{i,j}^{*n}$, il suffit d'appliquer la loi Forte de Grands Nombres.

Remarquons que, d'après 2-9, \mathbf{O}_n admet toujours une densité appartenant à L^2 par rapport à $\mu'_n \otimes \mu''_n$.

VII-TEST HILBERTIEN D'INDEPENDANCE

On va maintenant étudier le test Hilbertien d'indépendance qui généralise le test du chi-deux d'indépendance. En fait, on aura deux types de tests, suivant que les lois marginales sont connues ou inconnues.

Dans le cas de lois marginales connues, on retrouve la classe des tests Hilbertiens déjà étudiée par Bosq.

Dans le cas de lois marginales inconnues, on aura un test Hilbertien avec estimation de paramètres. Pour construire celui-ci, on devra étudier la loi asymptotique des fonctions orthogonales empiriques que l'on obtiendra sous une forme géométrique simple. On en déduira la loi des coefficients de corrélation d'ordre supérieur empirique et le test d'indépendance associé.

On terminera ce chapitre en démontrant que ces tests sont convergents.

On garde, dans ce chapitre, les notations et hypothèses du chapitre VI.

On commence par rappeler les bases du Test Hilbertien (cf Bosq).

Théorème 7-1: On suppose que, pour tout $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$, $\mathbb{E} [P_i(X)^2 Q_j(Y)^2] < +\infty$.

Alors, le vecteur aléatoire à valeurs dans \mathbb{R}^{hk} , $\sqrt{n} \{e_{i,j}^{*n} - e_{i,j}\}_{(i,j) \in \{1,\dots,h\} \times \{1,\dots,k\}}$,

suit asymptotiquement la loi Gaussienne centrée sur \mathbb{R}^{hk} de matrice de variances-covariances $\left\{ \left\{ \mathbb{E} [P_i(X)P_i(X)Q_j(Y)Q_j(Y)] \right\} \right\}_{((i,j),(i',j')) \in \{1,\dots,h\} \times \{1,\dots,k\}^2}$.

Ce résultat est une simple conséquence du Théorème de la Limite Centrale. On en déduit le théorème suivant.

Théorème 7-2 : On garde les hypothèses de 7-1, et on note par $\|S_n^*\|^2$ la variable aléatoire:

$$\|S_n^*\|^2 = n \left\{ \sum_{i=1}^h \sum_{j=1}^k (e_{i,j}^{*n})^2 \right\}.$$

Alors, si X et Y sont indépendantes, $\|S_n^*\|^2$ suit asymptotiquement une loi du chi-deux à hk degrés de liberté.

D'autre part, s'il existe $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$, tel que $e_{i,j} \neq 0$, alors $\|S_n^*\|^2$ tend presque-sûrement vers $+\infty$.

Ce théorème fonde le test Hilbertien d'indépendance. Rappelons que le test Hilbertien est un test des coefficients des fonctions orthogonales et est lui-même un cas particulier de la w-divergence (cf Bosq et Kagan).

Dans le cas de lois marginales inconnues, on va obtenir des résultats analogues. Mais, pour simplifier la présentation et les démonstrations, on aura besoin de définir une relation d'équivalence.

Relation d'équivalence 7-3 : Soient $\{U_n\}_{n \in \mathbb{N}^+}$ et $\{V_n\}_{n \in \mathbb{N}^+}$ deux suites de variables aléatoires définies sur (Ω, \mathcal{F}, P) . Alors, on emploiera la notation $U_n \approx V_n$ pour signifier que $n^{1/2}(U_n - V_n)$ converge en probabilité vers 0.

Maintenant, soient $\{g_n\}_{n \in \mathbb{N}^*}$ et $\{g'_n\}_{n \in \mathbb{N}^*}$ deux suites de fonctions aléatoires telles que $g_n = \sum_{i=0}^h \sum_{j=0}^k a_{i,j}^n P_i \otimes Q_j$ et $g'_n = \sum_{i=0}^h \sum_{j=0}^k b_{i,j}^n P_i \otimes Q_j$, où $a_{i,j}^n$ et $b_{i,j}^n$ représentent des variables aléatoires. Alors, la notation $g_n \approx g'_n$ signifiera que, pour tout $(i,j) \in \{0,1,\dots,h\} \times \{0,1,\dots,k\}$, $n^{1/2}(a_{i,j}^n - b_{i,j}^n)$ converge en probabilité vers 0.

On va maintenant énoncer les résultats qui fondent le test Hilbertien avec estimation de paramètres, les démonstrations étant effectuées en fin de paragraphe.

En fait, si on peut obtenir la loi asymptotique de $\rho_{i,j}^n$, c'est grâce à la simplicité géométrique de la loi des fonctions orthogonales empiriques.

Théorème 7-4 : On suppose que, pour tout $s \in \{1,2,\dots,h\}$, $\mathbb{E}[P_s(X)^4] < +\infty$. Alors, pour tout $i \in \{0,1,\dots,h\}$, on peut écrire P_i^n sous la forme :

$$P_i^n = P_i + \sum_{s=0}^i \varepsilon_{i,s} P_s$$

où, pour tout $i \in \{0,1,\dots,h\}$, $\varepsilon_{i,s} \approx - \int P_i P_s \cdot d\mu'_n$ si $s < i$ et $\varepsilon_{i,i} \approx \frac{1 - \int P_i^2 \cdot d\mu'_n}{2}$.

Pour énoncer sous une forme relativement simple la loi asymptotique de $\rho_{i,j}^n$, on introduit les notations suivantes.

Notations 7-5 : Soit $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$. On note par $\Pi[Q_j|0,1,\dots,i]$ et $\Pi[P_i|0,1,\dots,j]$, les projections orthogonales, dans $L^2(\mathbb{R}^{p+q}, \mathbf{0})$, de, respectivement, $Q_j(y)$ et $P_i(x)$ sur les sous-espaces engendrés, respectivement, par $P_0(x), P_1(x), \dots, P_i(x)$ et $Q_0(y), Q_1(y), \dots, Q_j(y)$. De plus, on note par $\odot[Q_j|P_i]$ et $\odot[P_i|Q_j]$, les vecteurs de projection orthogonale dans $L^2(\mathbb{R}^{p+q}, \mathbf{Q})$ de, respectivement, $Q_j(y)$ sur $P_i(x)$, et, $P_i(x)$ sur $Q_j(y)$, c'est à dire $\odot[Q_j|P_i](x,y) = Q_j(y) - \rho_{i,j} P_i(x)$.

On définit alors la fonction $C_{i,j} \in L^2(\mathbb{R}^{p+q}, \mathbf{Q})$ par :

$$C_{i,j}(x,y) = \frac{Q_j(y) \odot [P_i|Q_j](x,y) + P_i(x) \odot [Q_j|P_i](x,y)}{2} - Q_j(y) \Pi[P_i|0,1,\dots,j-1](y) - P_i(x) \Pi[Q_j|0,1,\dots,i-1](x)$$

En fait la loi des coefficients de corrélation d'ordre supérieur empiriques sera un corollaire du théorème suivant.

Théorème 7-6 : On suppose que, pour tout $s \in \{1, 2, \dots, h\}$, $\mathbb{E}[P_s(X)^4] < +\infty$, et que pour tout $t \in \{1, 2, \dots, k\}$, $\mathbb{E}[Q_t(Y)^4] < +\infty$. Soit $(i, j) \in \{1, 2, \dots, h\} \times \{1, 2, \dots, k\}$.

$$\text{Alors, } \rho_{i,j}^n - \rho_{i,j} \approx \frac{\sum_{j=1}^n C_{i,j}(X, Y, j)}{n}.$$

On déduit de ce résultat la loi asymptotique de $\rho_{i,j}^n$.

Théorème 7-7 : On garde les hypothèses de 7-6.

Alors, le vecteur aléatoire à valeurs dans \mathbb{R}^{hk} , $\sqrt{n} \{ \rho_{i,j}^n - \rho_{i,j} \}_{(i,j) \in \{1, \dots, h\} \times \{1, \dots, k\}}$,

suit asymptotiquement la loi Gaussienne centrée sur \mathbb{R}^{hk} de matrice de variances-covariances

$$\left(\mathbb{E}[C_{i,j}(X, Y) C_{i',j'}(X, Y)] \right)_{((i,j), (i',j')) \in (\{1, \dots, h\} \times \{1, \dots, k\})^2}.$$

On constate maintenant que, sous l'hypothèse d'indépendance, la loi limite de $\rho_{i,j}^n$ devient très simple.

Théorème 7-8 : On garde les hypothèses de 7-6, et on suppose de plus, que X et Y sont indépendantes.

Alors, le vecteur aléatoire $\sqrt{n} \{ \rho_{i,j}^n \}_{(i,j) \in \{1, \dots, h\} \times \{1, \dots, k\}}$ suit asymptotiquement la loi normale

centrée sur \mathbb{R}^{hk} , de matrice de variances-covariances, la matrice identité.

Pour démontrer ce théorème, il suffit d'utiliser les diverses relations d'orthogonalité existantes sous l'hypothèse d'indépendance.

Ce résultat permet de construire le test Hilbertien d'indépendance avec estimation de paramètres. En effet, on déduit de 7-6 et 7-8 le théorème suivant.

Théorème 7-9 : On garde les hypothèses de 7-6, et on note par $\|S_n\|^2$ la variable aléatoire:

$$\|S_n\|^2 = n \left\{ \sum_{i=1}^h \sum_{j=1}^k (\rho_{i,j}^n)^2 \right\}.$$

Alors, si X et Y sont indépendantes, $\|S_n\|^2$ suit asymptotiquement une loi du chi-deux à hk degrés de liberté.

D'autre part, s'il existe $(i, j) \in \{1, 2, \dots, h\} \times \{1, 2, \dots, k\}$, tel que $\rho_{i,j} \neq 0$, alors $\|S_n\|^2$ tend presque-sûrement vers $+\infty$.

Avec ces résultats, on retrouve la loi du coefficient de corrélation linéaire empirique (cf Serfling p 126).

Théorème 7-10 : Soit (X, Y) un vecteur aléatoire à valeurs dans \mathbb{R}^2 tel que $\mathbb{E}(X^4) < +\infty$, $\mathbb{E}(Y^4) < +\infty$, et que, ni X, ni Y, ne soient constantes. Soit ρ le coefficient de corrélation linéaire de X et Y, et ρ^n le coefficient de corrélation linéaire empirique associé.

$$\text{Soient, de plus, } P(x) = \frac{x - \mathbb{E}(X)}{\sigma(X)} \text{ et } Q(y) = \frac{y - \mathbb{E}(Y)}{\sigma(Y)}.$$

Alors, la variable aléatoire $\sqrt{n}(\rho^n - \rho)$ suit asymptotiquement la loi normale centrée de variance $\left(1 + \frac{\rho^2}{2}\right) E\{P(X)^2 Q(Y)^2\} + \rho^2 \frac{E\{P(X)^4\} + E\{Q(Y)^4\}}{4} - \rho E\{P(X)Q(Y)^3\} - \rho E\{P(X)^3 Q(Y)\}$.

De plus, lorsque X et Y sont indépendantes, $\sqrt{n}\rho^n$ suit asymptotiquement la loi normale centrée réduite.

Maintenant, nous allons voir que les tests Hilbertiens d'indépendance sont des tests convergents. Rappelons que l'on dit qu'un test est convergent si son niveau converge vers 0 et sa puissance vers 1 pour toute loi ν de l'alternative.

Notations 7-11 : On suppose que μ' et μ'' sont des probabilités *données* et que Q est une probabilité *quelconque* de lois marginales μ' et μ'' . On définit alors les hypothèses H_0 et $H_1(h,k)$ par $H_0 : \langle\langle X \text{ et } Y \text{ sont indépendants} \rangle\rangle$ et $H_1(h,k) : \langle\langle \text{Il existe un } \rho_{i,j}, 1 \leq i \leq h \text{ et } 1 \leq j \leq k, \text{ tel que } \rho_{i,j} \neq 0 \rangle\rangle$.

Soit $\|S_n^*\|^2 > w_n^*$ le test de H_0 contre $H_1(h,k)$. On note par α_n^* son niveau et, pour toute loi ν de $H_1(h,k)$, on note par $p_n^*(\nu)$ la puissance de ce test. De même, on note par α_n et $p_n(\nu)$, le niveau et la puissance du test $\|S_n\|^2 > w_n$.

Remarquons bien que l'on ne suppose pas que μ' et μ'' sont connues. En fait, on va étudier la puissance asymptotique des tests d'indépendance Hilbertiens lorsque les lois marginales restent fixes.

Alors, sous leurs hypothèses de définition, on peut montrer que ces tests sont bien convergents.

Théorème 7-12 : On garde les hypothèses de 7-1. Alors, pour que le test $\|S_n^*\|^2 > w_n^*$ soit convergent, il faut et il suffit que w_n^* tende vers l'infini et w_n^*/n vers 0

Théorème 7-13 : On garde les hypothèses de 7-6. Alors, pour que le test $\|S_n\|^2 > w_n$ soit convergent, il faut et il suffit que w_n tende vers l'infini et w_n/n vers 0.

On peut affiner ces théorèmes en montrant que l'on a une convergence uniforme dès que $E[P_s(X)^4] < +\infty$ et $E[Q_t(Y)^4] < +\infty$.

Théorème 7-14 : On suppose que les hypothèses du théorème 7-6 sont vérifiées. Soit $H \subset H_1(h,k)$ telle que $\delta(H) = \inf_{\nu \in H} \{\rho_{i_\nu, j_\nu}^2\} > 0$, où $|\rho_{i_\nu, j_\nu}| = \max_{1 \leq i \leq h, 1 \leq j \leq k} (|\rho_{i,j}|)$.

Alors, si le test $\|S_n^*\|^2 > w_n^*$ (resp. $\|S_n\|^2 > w_n$) est convergent, la convergence de la puissance est uniforme sur H .

Le théorème suivant donne une idée plus précise de la puissance asymptotique de ces tests.

Théorème 7-15 : On suppose que les hypothèses du théorème 7-6 sont vérifiées. Soit $\{v_n\} \subset H_1(h,k)$ telle que $(\rho_{i_n j_n})^2 = (\rho_{i_{v_n} j_{v_n}})^2 \geq m^2 \frac{w_n^*}{n}$ avec $m > 1$. Alors, si w_n^* converge vers $+\infty$, α_n^* converge vers 0 et $p_n^*(v_n)$ vers 1.

De même, si w_n converge vers $+\infty$ et $(\rho_{i_n j_n})^2 \geq m^2 w_n/n$, alors α_n converge vers 0 et $p_n(v_n)$ vers 1.

Démonstration du théorème 7-4

En gardant les notations de 6-1 et 6-3, il est clair que l'on peut aussi écrire \bar{P}_i^n sous la forme :

$$\bar{P}_i^n = \bar{P}_i + \sum_{s=0}^{i-1} \bar{\epsilon}_{i,s} P_s .$$

Aussi on d'abord démontre par récurrence sur i que, pour tout $i, 1 \leq i \leq h$, et pour tout $s, 0 \leq s < i$, il existe une suite $\{U^{i,s}_j\}_{j \in \mathbb{N}^*}$ de variables aléatoires indépendantes, centrées, de même loi, vérifiant $E[(U^{i,s}_j)^2] < +\infty$, telle que, pour tout $n \in \mathbb{N}^*$,

$$\bar{\epsilon}_{i,s} \approx \frac{1}{n} \sum_{j=1}^n U^{i,s}_j ,$$

et on démontrera en même temps qu'il en est de même pour les $\epsilon_{i,s}, 0 \leq i \leq h$ et $0 \leq s < i$.

Le cas $i=0$ est évident: $P_0^n = P_0 \equiv 1$, et donc $\epsilon_{0,0} = 0$.

On suppose désormais que la récurrence est vraie pour tout $s < i, i \leq h$.

En employant les théorèmes classiques de convergence des variables aléatoires (Serfling, p 19), et en tenant compte du fait que $\epsilon_{s,t} \epsilon_{s',t'} \approx 0$, on trouve facilement les équivalences suivantes :

$$\begin{aligned} \bar{P}_i^n &= C_i - \sum_{s=0}^{i-1} \langle C_i, P_s^n \rangle_n P_s^n \\ \approx \bar{P}_i - \sum_{s=0}^{i-1} [\langle C_i, P_s \rangle_n - \langle C_i, P_s \rangle] P_s - \sum_{s=0}^{i-1} \sum_{t=0}^s \epsilon_{s,t} \langle C_i, P_t \rangle P_s - \sum_{s=0}^{i-1} \sum_{t=0}^s \epsilon_{s,t} \langle C_i, P_s \rangle P_t \end{aligned}$$

que l'on écrit sous la forme $\bar{P}_i + \sum_{s=0}^{i-1} \bar{\epsilon}_{i,s} P_s$.

Il est alors facile de constater que les $\bar{\epsilon}_{i,s}$ sont bien de la forme annoncée.

On démontre alors la récurrence pour P_i^n .

Pour cela, on remarque d'abord que l'on a l'équivalence suivante :

$$\int \bar{P}_i^n \bar{P}_i^n . d\mu'_n \approx \int \bar{P}_i \bar{P}_i . d\mu'_n + \int \left(\sum_{s=0}^{i-1} \varepsilon_{i,s} P_s \right) \bar{P}_i . d\mu'_n + \int \bar{P}_i \left(\sum_{t=0}^{i-1} \varepsilon_{i,t} P_t \right) . d\mu'_n .$$

Comme $\int \bar{P}_i \bar{P}_s . d\mu' = 0$ si $s \neq i$ et $\|\bar{P}_i\|^2$ sinon, d'après la Loi Forte des Grands Nombres, $\int \bar{P}_i \bar{P}_s . d\mu'_n$ converge presque sûrement vers 0 si $s \neq i$, et $\|\bar{P}_i\|^2$ sinon.

On en déduit $\|\bar{P}_i^n\|^2 \approx \int \bar{P}_i \bar{P}_i . d\mu'_n$.

Or, on trouve facilement l'équivalence : $P_i^n \approx P_i - \frac{\|\bar{P}_i^n\|^2 - \|\bar{P}_i\|^2}{2\|\bar{P}_i\|^2} P_i + \frac{1}{\|\bar{P}_i\|} \sum_{s=0}^{i-1} \varepsilon_{i,s} P_s .$

On en déduit la récurrence.

En employant alors la même méthode pour $\int P_i^n P_s^n . d\mu'_n$ que pour $\int \bar{P}_i^n \bar{P}_i^n . d\mu'_n$, on obtient les équivalences : $\int P_i^n P_s^n . d\mu'_n \approx \int P_i P_s . d\mu'_n + \varepsilon_{i,s} \int P_s P_s . d\mu'_n \approx \int P_i P_s . d\mu'_n + \varepsilon_{i,s}$, ce qui nous donne bien le résultat annoncé, vu que $\int P_i^n P_s^n . d\mu'_n = 0$ si $s < i$.

Enfin, on démontre de la même façon le résultat pour $\varepsilon_{i,i}$.

Démonstration du théorème 7-6 :

D'après 7-4, on peut écrire : $P_i^n = P_i + \sum_{s=0}^i \varepsilon'_{i,s} P_s$ et $Q_j^n = Q_j + \sum_{t=0}^j \varepsilon''_{j,t} Q_t$.

On en déduit les équivalences suivantes :

$$\begin{aligned} & e^n_{i,j} - e_{i,j} \approx \\ & \int P_i(x) Q_j(y) . Q_n(dx, dy) - e_{i,j} + \sum_{s=0}^i \varepsilon'_{i,s} \int P_s(x) Q_j(y) . Q_n(dx, dy) + \sum_{t=0}^j \varepsilon''_{j,t} \int P_i(x) Q_t(y) . Q_n(dx, dy) \\ & \approx \\ & \int P_i(x) Q_j(y) . Q_n(dx, dy) - e_{i,j} - e_{i,j} \frac{\int P_i(x) P_i(x) . \mu'_n(dx) - 1}{2} - e_{i,j} \frac{\int Q_j(y) Q_j(y) . \mu''_n(dy) - 1}{2} \\ & - \sum_{s=0}^{i-1} [E [P_s(X) Q_j(Y)]] \int P_i(x) P_s(x) . \mu'_n(dx) - \sum_{t=0}^{j-1} [E [P_i(X) Q_t(Y)]] \int Q_j(y) Q_t(y) . \mu''_n(dy) \end{aligned}$$

$$=$$

$$\frac{\int P_i(x)[Q_j(y) - \rho_{ij}P_i(x)].Q_n(dx,dy)}{2} + \frac{\int Q_j(y)[P_i(x) - \rho_{ij}Q_j(y)].Q_n(dx,dy)}{2}$$

$$- \int \left[\sum_{s=0}^{i-1} \mathbb{E}[P_s(X)Q_j(Y)]P_s(x) \right] P_i(x) \cdot \mu'_n(dx) - \int \left[\sum_{t=0}^{j-1} \mathbb{E}[P_i(X)Q_t(Y)]Q_t(y) \right] Q_j(y) \cdot \mu''_n(dy) .$$

Démonstration du théorème 7-7 : Les diverses orthogonalités existantes montrent que
 $\mathbb{E}(\mathcal{O}[Q_j|P_i](X,Y).P_i(X)) = \mathbb{E}(\Pi[Q_j|0,1,\dots,i-1](X).P_i(X)) = \mathbb{E}(\mathcal{O}[P_i|Q_j](X,Y).Q_j(X))$
 $= \mathbb{E}(\Pi[P_i|0,1,\dots,j-1](Y).Q_j(Y)) = 0 .$

De plus, on déduit de l'inégalité de Holder et de l'hypothèse que $\mathbb{E}[C_{i,j}(X,Y)^2] < +\infty .$

Il ne reste plus qu'à appliquer 7-6 et le Théorème de la Limite Centrale pour obtenir 7-7.

Démonstration de 7-12 et 7-13 : On démontre de la même façon ces deux théorèmes. Par exemple, on va démontrer 7-13. En fait, on emploie à peu près la même méthode que dans Bosq 1978-a.

D'abord, il est clair que α_n converge vers 0 si et seulement si w_n converge vers $+\infty .$

On suppose maintenant que w_n/n tend vers 0.

On considère une loi ν telle que $\rho_{i,j} = \int P_i(x)Q_j(y).\nu(dx,dy) \neq 0 .$

Alors, il existe $N \in \mathbb{N}^*$ tel que, pour tout $n \geq N$, $\frac{w_n}{n} \leq \frac{\rho_{i,j}^2}{m^2}$, où $m > 1 .$

Donc, si $\sum_{i=1}^h \sum_{j=1}^k (\rho_{i,j}^n)^2 \leq \frac{w_n}{n}$, alors $|\rho_{i,j}^n| \leq \frac{|\rho_{i,j}|}{m}$ et $|\rho_{i,j}^n - \rho_{i,j}| \geq |\rho_{i,j}| \frac{m-1}{m}$ pour $n \geq N$.

Donc, $1 - p_n(\nu) \leq \mathbb{P}\left\{ |\rho_{i,j}^n - \rho_{i,j}| \geq |\rho_{i,j}| \frac{m-1}{m} \right\}$. Donc, comme $\sqrt{n}(\rho_{i,j}^n - \rho_{i,j})$ suit asymptotiquement une loi normale centrée, on en déduit que $p_n(\nu)$ tend vers 1 .

Maintenant, pour démontrer la réciproque, on aura besoin du lemme suivant.

Lemme 7-16 : Pour tout $\eta > 0$, il existe $\nu \in H_1(h, k)$ tel que

$$0 < \sum_{i=1}^h \sum_{j=1}^k \left(\int P_i(x) Q_j(y) \cdot \nu(dx, dy) \right)^2 < \eta .$$

Démonstration : Comme $\mathcal{N}^* \neq \emptyset$, on peut partager \mathbb{R}^p en deux sous-ensembles \mathfrak{L} et \mathfrak{Y} de mesures non nulles. On définit \bar{A}_1 par : $\bar{A}_1(x) = \mu'(\mathfrak{Y})$ sur \mathfrak{L} , et $\bar{A}_1(x) = -\mu'(\mathfrak{L})$ sur \mathfrak{Y} . Il est clair que $\int \bar{A}_1 \cdot d\mu' = 0$.

On définit alors \bar{R}_1 de même façon sur \mathbb{R}^q et on pose $f(x, y) = 1 + \bar{\alpha}_{1,1} \bar{A}_1(x) \bar{R}_1(y)$.

Il suffit alors de choisir $\bar{\alpha}_{1,1}$ assez petit pour que $\sup |f(x, y) - 1|^2 < \eta$ et $f(x, y) > 0$. En effet, la mesure ν admettant f pour densité par rapport à $\mu' \otimes \mu''$ répond aux conditions du lemme, vu que

$$\sum_{i=1}^h \sum_{j=1}^k e_{i,j}^2 \leq \|f - 1\|_{\mu' \otimes \mu''}^2 .$$

On démontre maintenant la réciproque. Pour cela, on suppose que le test est convergent et que w_n/n ne tend pas vers 0.

Donc, on peut trouver un $\eta > 0$ et une partie infinie \mathbb{N}_1 de \mathbb{N} telle que $w_n/n > \eta$ pour tout $n \in \mathbb{N}_1$. Soit alors ν définie par le lemme précédent.

Comme le test est convergent, $\mathbf{P} \left\{ \sum_{i=1}^h \sum_{j=1}^k (e_{i,j}^n)^2 \leq \eta \right\}$ tend vers 0 lorsque $n \in \mathbb{N}_1$.

Or, $\sum_{i=1}^h \sum_{j=1}^k (e_{i,j}^n)^2$ converge presque sûrement vers $\sum_{i=1}^h \sum_{j=1}^k e_{i,j}^2 < \eta$.

Donc, $\mathbf{P} \left\{ \sum_{i=1}^h \sum_{j=1}^k \left((e_{i,j}^n)^2 - e_{i,j}^2 \right) \leq \eta - \sum_{i=1}^h \sum_{j=1}^k e_{i,j}^2 \right\}$ converge vers 1, ce qui est contradictoire.

Démonstration de 7-14 : Pour n assez grand, on aura $w_n^*/n < \delta(H)/m^2$, où $m > 1$.

Donc, en utilisant les inégalités de Bienaymé-Tschébéchev et de Schwartz, on voit que

$$1 - p_n^*(\nu) \leq \mathbf{P} \left\{ \left| e_{i_\nu, j_\nu}^{*n} - e_{i_\nu, j_\nu} \right| \geq |e_{i_\nu, j_\nu}| \frac{m-1}{m} \right\} \leq \frac{\sigma^2 \left(e_{i_\nu, j_\nu}^{*n} - e_{i_\nu, j_\nu} \right)}{\left(|e_{i_\nu, j_\nu}| \frac{m-1}{m} \right)^2}$$

$$\leq \frac{1}{n} \frac{\mathbb{E}\{P_{i_v}(X)^2 Q_{j_v}(Y)^2\}}{\delta(H)} \left(\frac{m}{m-1}\right)^2 \leq \frac{1}{n} \frac{\sqrt{\mathbb{E}\{P_{i_v}(X)^4\} \mathbb{E}\{Q_{j_v}(Y)^4\}}}{\delta(H)} \left(\frac{m}{m-1}\right)^2.$$

On démontre maintenant la convergence du test $\|S_n\|^2 > w_n$.

En gardant les notations de la démonstration de 7-6, on voit facilement que

$$\left\{ \left| e_{i_v j_v}^n - e_{i_v j_v} \right| \geq |e_{i_v j_v}| \frac{m-1}{m} \right\}$$

c

$$\left\{ \left| e_{i_v j_v}^{*n} - e_{i_v j_v} \right| \geq |e_{i_v j_v}| \frac{m-1}{4m} \right\}$$

$$\bigcup_{t=0}^{j_v} \left\{ \left| \varepsilon''_{j_v, t} \int P_{i_v}(x) Q_t(y) \cdot Q_n(dx, dy) \right| \geq |e_{i_v j_v}| \frac{m-1}{4(k+1)m} \right\}$$

$$\bigcup_{s=0}^{i_v} \left\{ \left| \varepsilon'_{i_v, s} \int P_s(x) Q_{j_v}(y) \cdot Q_n(dx, dy) \right| \geq |e_{i_v j_v}| \frac{m-1}{4(h+1)m} \right\}$$

$$\bigcup_{s=0}^{i_v} \bigcup_{t=0}^{j_v} \left\{ \left| \varepsilon'_{i_v, s} \varepsilon''_{j_v, t} \int P_s(x) Q_t(y) \cdot Q_n(dx, dy) \right| \geq |e_{i_v j_v}| \frac{m-1}{4(h+1)(k+1)m} \right\}.$$

Or, par exemple, $\left\{ \left| \varepsilon''_{j_v, t} e_{i_v, t}^{*n} \right| \geq |e_{i_v j_v}| \frac{m-1}{4(k+1)m} \right\}$

$$c \left\{ \left| \varepsilon''_{j_v, t} e_{i_v, t} \right| \geq |e_{i_v j_v}| \frac{m-1}{8(k+1)m} \right\} \cup \left\{ \left| \varepsilon''_{j_v, t} (e_{i_v, t}^{*n} - e_{i_v, t}) \right| \geq |e_{i_v j_v}| \frac{m-1}{8(k+1)m} \right\}$$

c

$$\left\{ |\varepsilon''_{j_v,t}| \geq \frac{|e_{i_v,j_v}|}{|e_{i_v,t}|} \frac{m-1}{8(k+1)m} \right\} \cup \left\{ |e^*_{i_v,t} - e_{i_v,t}| \geq |e_{i_v,j_v}| \frac{m-1}{8(k+1)m} \right\}$$

$$\cup \left\{ \left\{ |e^*_{i_v,t} - e_{i_v,t}| \leq |e_{i_v,j_v}| \frac{m-1}{8(k+1)m} \right\} \cap \left\{ |\varepsilon''_{j_v,t}| \geq \frac{|e_{i_v,j_v}|}{|e^*_{i_v,t} - e_{i_v,t}|} \frac{m-1}{8(k+1)m} \right\} \right\}$$

c

$$\left\{ |\varepsilon''_{j_v,t}| \geq \frac{m-1}{8(k+1)m} \right\} \cup \left\{ |e^*_{i_v,t} - e_{i_v,t}| \geq |e_{i_v,j_v}| \frac{m-1}{8(k+1)m} \right\}$$

$$\cup \left\{ |\varepsilon''_{j_v,t}| \geq \frac{|e_{i_v,j_v}|}{|e_{i_v,j_v}| \frac{m-1}{8(k+1)m}} \frac{m-1}{8(k+1)m} \right\}$$

$$c \left\{ |\varepsilon''_{j_v,t}| \geq \frac{m-1}{8(k+1)m} \right\} \cup \left\{ |e^*_{i_v,t} - e_{i_v,t}| \geq |e_{i_v,j_v}| \frac{m-1}{8(k+1)m} \right\} .$$

On montre, comme précédemment, que $P\left\{|e^*_{s,t} - e_{s,t}| \geq K|e_{i_v,j_v}|\right\} \leq \frac{\sqrt{E[P_s(X)^4]E[Q_t(Y)^4]}}{nK^2\delta(H)}$.

De plus, la loi de $\varepsilon''_{j,t}$ est uniquement fonction de $\{Y_j\}_{j=1,2,\dots,n}$ et ne dépend donc que de Y et de n. On en déduit le théorème.

Démonstration de 7-15 : En raisonnant comme précédemment, on trouvera que

$$1 - p_n^*(v_n) \leq P\left\{|e^*_{i_n,j_n} - e_{i_n,j_n}| \geq |e_{i_n,j_n}| \frac{m-1}{m}\right\} \leq \frac{\sqrt{E[P_{i_n}(X)^4]E[Q_{j_n}(Y)^4]}}{(m-1)^2 w_n^*} .$$

Pour démontrer la convergence du test $\|S^n\|^2 > w_n$, on emploie alors la même méthode que pour 7-14.

VIII EXEMPLES

VIII-A : COEFFICIENTS DE CORRELATION POLYNOMIALE

Dans le cas où (X, Y) est un vecteur aléatoire à valeurs dans \mathbb{R}^2 , et où $\{P_i\}_{i \in \mathcal{N}}$ et $\{Q_j\}_{j \in \mathcal{M}}$ représentent les familles de polynômes orthonormaux associées à X et Y , les $\rho_{i,j}$ mesurent des dépendances polynômiales. En particulier, $\rho_{1,1}$ est le coefficient de corrélation linéaire, et $\rho_{1,1}^n$, le coefficient de corrélation linéaire empirique (cf Blacher 1984 a et b).

D'autre part \mathcal{N} sera un ensemble fini lorsque X sera concentrée en un nombre fini de points ou lorsque tous les moments ne sont pas définis. Par exemple, dans le cas de la loi de Cauchy, la famille $\{P_i\}_{i \in \mathcal{N}}$ sera réduite à $\{P_0\}$. Rappelons aussi que, même lorsque tous les moments existent, $\{P_i\}_{i \in \mathbb{N}}$ ne sera pas forcément une base de $L^2(\mathbb{R}, \mu')$ (cf Natanson p 149-150). Par contre cela sera vrai, dès que X est concentrée sur un compact ou pour les lois Normales et Gammas (cf Natanson, p 154 et 158).

Considérons, par exemple le cas des lois Gaussiennes. On sait que la famille de polynômes orthonormaux associée à la loi Normale centrée réduite, $N(0,1)$, est la famille des polynômes d'Hermite qui sont définis par

$$H_i(x) = \frac{(-1)^i}{\sqrt{i!}} \left[e^{\frac{x^2}{2}} \frac{d^i \left(e^{-\frac{x^2}{2}} \right)}{dx^i} \right] \quad (\text{cf Chihara p 145 et suivantes}).$$

Alors, dans le cas où (X, Y) est un vecteur aléatoire Gaussien à valeurs dans \mathbb{R}^2 , de lois marginales centrées réduites, et de coefficient de corrélation ρ , on trouve que $\rho_{i,j} = E[H_i(X)H_j(Y)] = \rho^i$ si $i=j$ et 0 sinon, c'est à dire que les familles $\{H_i(x)\}_{i \in \mathbb{N}^*}$ et $\{H_j(y)\}_{j \in \mathbb{N}^*}$ sont les familles de fonctions canoniques de coefficients de corrélation canonique associés ρ^i (cf Lancaster 1969, p 197 ou Gupta p 792).

En particulier, la densité de la loi binormale, par rapport à la mesure de Lebesgue est

$$\left(1 + \sum_{i=1}^{\infty} \rho^i H_i(x)H_i(y) \right) e^{-\frac{x^2+y^2}{2}} \frac{1}{2\pi},$$

et, la densité de la loi binormale, par rapport à la mesure $N(0,1) \otimes N(0,1)$:

$$1 + \sum_{i=1}^{\infty} \rho^i H_i(x)H_i(y).$$

Ce résultat permet de rappeler à quel point l'hypothèse Gaussienne est forte. En effet, dans ce cas, la dépendance est fonction d'un seul paramètre, ρ . et les autres $\rho_{i,j}$ sont alors entièrement déterminés par ρ . Par contre, si l'on suppose uniquement que les lois marginales sont

Gaussiennes, la densité par rapport à la mesure de Lebesgue sera

$$\left(1 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{i,j} H_i(x) H_j(y) \right) e^{-\frac{x^2+y^2}{2}} \frac{1}{2\pi},$$

où les $\rho_{i,j}$ sont des paramètres "quasi-indépendants" les uns des autres (cf Blacher 1984-a), et la dépendance sera donc fonction d'un infinié dénombrable de paramètres.

Alors, en posant $N(x) = \int_{u=-\infty}^{u=x} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$, on peut écrire la fonction de répartition sous la forme :

$$\begin{aligned} F_{X,Y}(x,y) &= N(x)N(y) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{i,j} \left(\int_{u=-\infty}^{u=x} H_i(u) \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du \right) \left(\int_{v=-\infty}^{v=y} H_j(v) \frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}} dv \right) \\ &= N(x)N(y) + \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{\rho_{i,j}}{\sqrt{ij}} H_{i-1}(x) H_{j-1}(y) \right) e^{-\frac{x^2+y^2}{2}} \frac{1}{2\pi}. \end{aligned}$$

Dans le cas où (X,Y) est Gaussien, on retrouve le développement de $F_{X,Y}$ en fonctions tetrachoriques, valable également si $\rho=1$, i.e. $X=Y$.

Maintenant, pour les vecteurs aléatoires (X_1, X_2, \dots, X_p) , $p > 1$, il n'y a plus d'orthogonalisation canonique, et il faudra donc orthogonaliser les monômes $x_1^{i_1} x_2^{i_2} \dots x_p^{i_p}$ selon ce qui sera le plus pratique. Maintenant, lorsque il y a indépendance, on pourra employer des bases du type $P_{i_1}^1 \otimes P_{i_2}^2 \otimes \dots \otimes P_{i_p}^p$.

Par exemple, dans le cas d'un vecteur aléatoire Gaussien non dégénéré centré, $((X'_1, X'_2, X'_3), T)$, si l'on veut étudier la dépendance entre T et (X'_1, X'_2, X'_3) , on peut transformer linéairement (X'_1, X'_2, X'_3) en un vecteur (U, V, W) tel que (U, T) , V et W soient indépendants. On dit alors que $((U, V, W), T)$ est distribué canoniquement. Dans ce cas, la corrélation maximale $\rho^*((U, V, W), T) = \rho_{U,T}$, le coefficient de corrélation linéaire de U et T (Csaki-Fischer, p 39 à 43). De plus, la famille $\{H_i(u)H_j(v)H_k(w)\}_{(i,j,k) \in \mathbb{N}^3}$ est une base orthonormale de $L^2(\mathbb{R}^3, \mu_{X'_1, X'_2, X'_3})$, où $\mu_{X'_1, X'_2, X'_3}$ représente la probabilité image de (X'_1, X'_2, X'_3) , et où u, v et w sont des variables réelles combinaisons linéaires de x'_1, x'_2 et x'_3 .

Dans ce cas, et en supposant de plus que U, V, W et T sont centrées réduites, la densité de la loi quadrinormale de (U, V, W, T) s'écrit

$$\left(\sum_{i=0}^{\infty} \rho_{U,T}^i H_i(u) H_i(t) \right) e^{-\frac{u^2+v^2+w^2+t^2}{2}} \frac{1}{\sqrt{(2\pi)^4}}.$$

$$\text{Donc, } \mathbb{E} |H_i(U)H_j(V)H_k(W)H_l(T)| = \rho_{U,T}^i \delta_{i,l} \delta_{j,0} \delta_{k,0} ,$$

c'est à dire que les couples $H_i(u), H_i(t)$, sont les couples de fonctions canoniques généralisées de coefficients de corrélation canonique associés : $(\rho_{U,T})^i$. De plus, avec les notations de 3-23, $\mathcal{E} = \mathbb{N}_x \times \{0\}_x \times \{0\}$, $\mathcal{K} = \mathcal{L} = \mathcal{V} = \emptyset$, et $\mathcal{W} = \{(i,j,k) \in \mathbb{N}^3, j \neq 0 \text{ ou } k \neq 0\}$. C'est à dire que la famille $\{H_j(t)\}_{j \in \mathbb{N}}$ est bien une base de $L^2(\mathbb{R}, \mu_T)$ (μ_T représentant la loi de T), mais que, par contre, la famille $\{H_i(u)\}_{i \in \mathbb{N}}$ doit être complétée par la famille $\{H_i(u)H_j(v)H_k(w)\}_{(i,j,k) \in \mathcal{W}}$ pour obtenir une base de $L^2(\mathbb{R}^3, \mu_{X'_1, X'_2, X'_3})$.

Maintenant, on peut se contenter de transformer linéairement (X'_1, X'_2, X'_3) en un vecteur (X, Y, Z) tel que X, Y et Z soient indépendantes. Dans ce cas, la famille $\{H_i(x)H_j(y)H_k(z)\}_{(i,j,k) \in \mathbb{N}^3}$ sera aussi une base orthonormale de $L^2(\mathbb{R}^3, \mu_{X'_1, X'_2, X'_3})$.

Maintenant, pour savoir s'il existe une relation fonctionnelle du type $T = f(X'_1, X'_2, X'_3)$, d'après le théorème 4-5 , il faut calculer

$$\sum_{l=1}^{\infty} \{ \mathbb{E} |P_l(X'_1, X'_2, X'_3) Q_l(T)| \}^2 , \text{ i.e. } \sum_{(i,j,k) \neq (0,0,0)} \{ \mathbb{E} |H_i(X)H_j(Y)H_k(Z) H_l(T)| \}^2 .$$

Or, on sait que si X, Y, Z et T sont centrées réduites, la densité de la loi quadrinormale de (X, Y, Z, T) s'écrit

$$\left(\sum_{i,j,k=0}^{\infty} \frac{\sqrt{(i+j+k)!}}{\sqrt{i! j! k!}} \rho_{X,T}^i \rho_{Y,T}^j \rho_{Z,T}^k H_i(x)H_j(y)H_k(z)H_{i+j+k}(t) \right) e^{-\frac{x^2+y^2+z^2+t^2}{2}} \frac{1}{\sqrt{(2\pi)^4}} \quad (\text{Gupta, p797}) .$$

Or, pour que $i+j+k=1$, il faut que l'un de ces trois nombres soit égal à 1 , et les autres nuls. Donc, dans ce cas, $\mathbb{E} |H_i(X)H_j(Y)H_k(Z) H_l(T)| = \rho_{X,T}$ ou $\rho_{Y,T}$ ou $\rho_{Z,T}$, suivant le cas.

Autrement, $\mathbb{E} |H_i(X)H_j(Y)H_k(Z) H_l(T)| = 0$.

$$\text{Donc, } \sum_{(i,j,k) \neq (0,0,0)} \{ \mathbb{E} |H_i(X)H_j(Y)H_k(Z) H_l(T)| \}^2 = \rho_{X,T}^2 + \rho_{Y,T}^2 + \rho_{Z,T}^2 .$$

D'autre part, $\rho_{X,T}H_1(X) + \rho_{Y,T}H_1(Y) + \rho_{Z,T}H_1(Z)$ est la projection orthogonale de $H_1(T)$.

Donc, comme, il existe une combinaison linéaire $aX+bY+cZ$ telle que

$$\rho^*([X, Y, Z], T) = \rho(aX+bY+cZ, T) ,$$

$$\text{on en déduit facilement que } \rho^*([X, Y, Z], T) = \sqrt{\rho_{X,T}^2 + \rho_{Y,T}^2 + \rho_{Z,T}^2} .$$

VIII-B : COEFFICIENTS DE CORRELATION DE SPEARMAN

On considère maintenant un vecteur aléatoire (X, Y) à valeurs dans \mathbb{R}^2 , tel que les fonctions de répartition marginales F_X et F_Y soient continues. On sait que le coefficient de corrélation de Spearman est égal au coefficient de corrélation linéaire de $F_X(X)$ et $F_Y(Y)$ (cf V.3.4, p 287, Capéraà - Van Cutsem). En fait, $(F_X(X), F_Y(Y))$ est un vecteur aléatoire à valeurs dans $[0, 1]^2$ de lois marginales uniformes. On en vient donc à considérer les coefficients $\theta_{i,j} = E[L_i[F_X(X)] \cdot L_j[F_Y(Y)]]$, où les L_i représentent les polynômes orthonormaux associés à la loi uniforme sur $[0, 1]$, i.e. les polynômes de Legendre (cf Chihara p 143 et suivantes).

On pose $\tilde{\theta}_{i,j} = \theta_{i,j}$, si $i \geq 1$ et $j \geq 1$, et 0 sinon. Alors, d'après 3-7 de Blacher-1987, on sait que l'on peut écrire la fonction de répartition de (X, Y) sous la forme :

$$F_{X,Y}(x,y) = F_X(x) F_Y(y) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \beta_{i,j} L_i[F_X(x)] L_j[F_Y(y)] \quad \text{dans } L^2(\mathbb{R}^2, \mu' \otimes \mu) ,$$

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

De plus, on sait alors que, pour tout $(i,j) \in \mathbb{N}^2$, on a l'inégalité $(\beta_{i,j})^2 \leq M_i M_j$, où la suite $\{M_n\}_{n \in \mathbb{N}}$ est définie par :

$$M_0 = \frac{1}{12}, M_1 = \frac{1}{60} \text{ et } M_n = \frac{1}{2(2n+3)(2n-1)} \text{ pour } n \geq 2,$$

De cette inégalité, on peut déduire une majoration de la distance de $F_{X,Y}$ à $F^{(m)}$, la fonction de répartition de l'approximation de la probabilité $\mathbf{0}$ de (X, Y) par la mesure - pas forcément positive - admettant la densité :

$$g^{(m)}(x,y) = 1 + \sum_{i=1}^m \sum_{j=1}^m \theta_{i,j} L_i[F_X(x)] L_j[F_Y(y)] .$$

En particulier, si $\theta_{i,j} = 0$ pour $1 \leq i \leq m$ et $1 \leq j \leq m$, $\mathfrak{D}(X, Y) \leq 1,372 \sqrt{\log\left(\frac{2m+1}{2m-3}\right)}$,

$$\text{où } \mathfrak{D}(X, Y) = \sqrt{90 \left(\iint [F_{X,Y}(x,y) - F_X(x)F_Y(y)]^2 \mu'(dx) \mu''(dy) \right)} ,$$

est un indicateur de dépendance normalisé qui vérifie les axiomes de Renyi (cf aussi Schweizer et Wolf) .

VIII-C : COEFFICIENTS DE CORRELATION RECTANGULAIRE

Lorsque on veut effectuer le test d'indépendance du chi-deux entre deux variables aléatoires on décompose \mathbb{R}^2 en un nombre fini de rectangles $J_i \times A_j$, et on cherche à évaluer la dépendance en comparant $O(J_i \times A_j)$ et $\mu'(J_i) \cdot \mu''(A_j)$. Les coefficients de corrélation rectangulaire $\alpha_{i,j}$ seront les coefficients de corrélation correspondant à cette approche : les $\alpha_{i,j}$ seront des combinaisons linéaires des $O(J_i \times A_j)$ qui s'annuleront lorsque $O = \mu' \otimes \mu''$. Il n'est donc pas étonnant que ces coefficients permettent de considérer le test d'indépendance du chi-deux sous une forme plus géométrique comme nous allons le voir maintenant.

Pour construire ces coefficients, on utilise des fonctions orthogonales qui sont connues sous le nom de système de Haar. Ces fonctions sont définies de la façon suivante: on suppose que l'on a une suite de partitions emboîtées de \mathbb{R}^p en boréliens $\{J_{m,s}\}$, $m \in \mathbb{N}^*$ et $s=1,2,\dots,c(m)$, où $c(m)$ représente une suite entière strictement croissante et où $\mu'(J_{m,s}) > 0$. Par le procédé de Gram-Schmidt, on orthogonalise la famille des fonctions indicatrices $1_{J_{m,s}}$ des ensembles $J_{m,s}$ en imposant que la première fonction orthogonale A_0 soit identique à 1. Il est alors clair que les fonctions orthonormales A_i , $0 \leq i < c(m)$, ainsi définies sont des combinaisons linéaires des fonctions indicatrices des ensembles $J_{m,s}$, $1 \leq s \leq c(m)$, qui engendrent le même sous-espace de $L^2(\mathbb{R}^p, \mu')$ (cf Lancaster ou Blacher 1987, 1988).

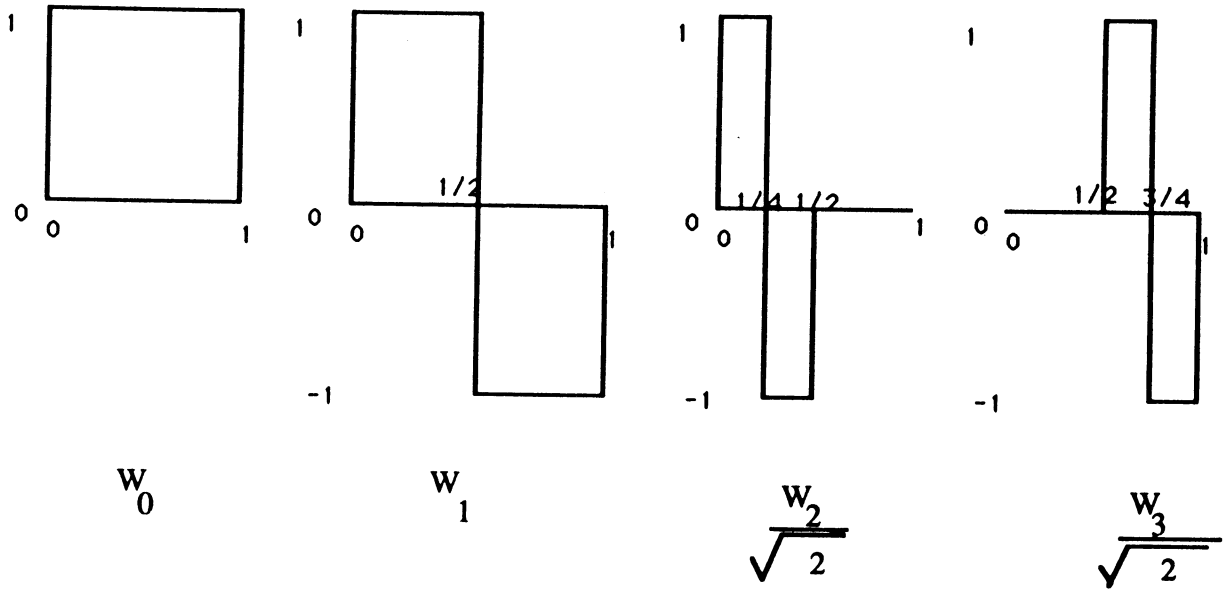
En choisissant convenablement la suite $\{J_{m,s}\}$, on peut facilement imposer que la famille $\{A_i\}_{i \in \mathcal{N}}$ soit une base orthonormale de $L^2(\mathbb{R}^p, \mu')$.

On définit de la même façon une base orthonormale $\{B_j\}_{j \in \mathcal{M}}$ de $L^2(\mathbb{R}^q, \mu'')$ sur une suite de partitions de \mathbb{R}^q en boréliens $\{A_{m,s}\}$.

Lorsque μ' représente la mesure Borélienne sur $[0,1]$, on peut alors employer les fonctions de Walsh ou le système spécial de Haar : $\{W_i\}_{i \in \mathbb{N}}$ (cf Beauchamps).

On le définit en posant $W_0 \equiv 1$ sur $[0,1]$,

$$\text{et, pour } i=2^k+m \text{ avec } m < 2^k, \quad W_i(x) = \begin{cases} \sqrt{2^k} & \text{si } x \in \left[\frac{2m}{2^{k+1}}, \frac{2m+1}{2^{k+1}} \right[, \\ -\sqrt{2^k} & \text{si } x \in \left[\frac{2m+1}{2^{k+1}}, \frac{2m+2}{2^{k+1}} \right[, \\ 0 & \text{sinon.} \end{cases}$$



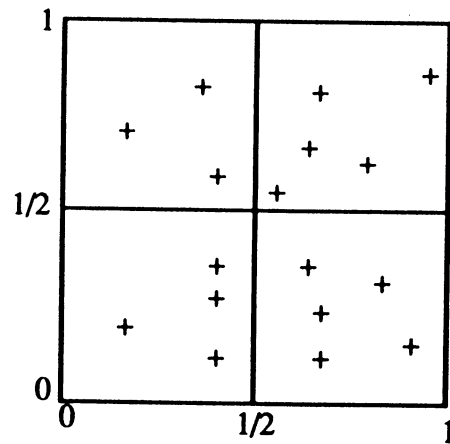
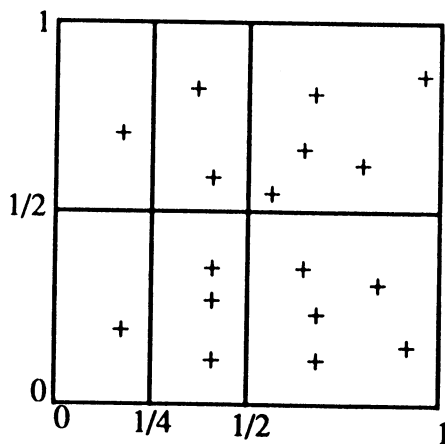
Alors, si (X,Y) est un vecteur aléatoire à valeurs dans $[0,1]^2$, de loi \mathbf{Q} , de lois marginales uniformes, les coefficients de corrélation rectangulaire d'ordre $(1,1)$ et $(2,1)$ sont définis par :

$$\alpha_{1,1} = \mathbb{E}[W_1(X)W_1(Y)] =$$

$$\mathbf{Q}\left\{\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]\right\} + \mathbf{Q}\left\{\left[\frac{1}{2}, 1\right] \times \left[\frac{1}{2}, 1\right]\right\} - \mathbf{Q}\left\{\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]\right\} - \mathbf{Q}\left\{\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]\right\} ,$$

$$\alpha_{2,1} = \mathbb{E}[W_2(X)W_1(Y)] =$$

$$\sqrt{2}\left(\mathbf{Q}\left\{\left[0, \frac{1}{4}\right] \times \left[0, \frac{1}{2}\right]\right\} + \mathbf{Q}\left\{\left[\frac{1}{4}, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]\right\} - \mathbf{Q}\left\{\left[\frac{1}{4}, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]\right\} - \mathbf{Q}\left\{\left[0, \frac{1}{4}\right] \times \left[\frac{1}{2}, 1\right]\right\}\right)$$



De même, pour un échantillon de taille n , $\{X_j, Y_j\}_{j=1,2,\dots,n}$ de (X,Y) ,

$n \cdot \alpha^{*n}_{1,1}$ est égal au

nombre de points de l'échantillon appartenant à $\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]$ et $\left[\frac{1}{2}, 1\right] \times \left[\frac{1}{2}, 1\right]$,
 diminué du nombre de points appartenant à $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]$ et $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$.

Nous obtiendrons alors une approximation de $\mathbf{0}$ en limitant l'ordre du développement :

$$f^{(k)} = 1 + \sum_{i=1}^{2^{k-1}} \sum_{j=1}^{2^{k-1}} \alpha_{i,j} W_i \otimes W_j .$$

En effet, la fonction $f^{(k)}$ est la densité de la probabilité constante sur chaque élément de la partition de $[0,1]^2$ en 4^k carrés égaux.

De plus, avec les notations de 5-2, on voit que l'espérance conditionnelle est égale à

$$\mathbb{E}(Y | X=x) = \lim_{k \rightarrow \infty} \left\{ \mathbb{E}(Y) + \sigma(Y) \sum_{i=1}^{2^{k-1}} \rho_{i,1} W_i(x) \right\} = \lim_{k \rightarrow \infty} \left\{ \sum_{i=0}^{2^k-1} \mathbb{E}[Y \cdot W_i(X)] W_i(x) \right\} .$$

$$\text{Or, } \sum_{i=0}^{2^k-1} \mathbb{E}[Y \cdot W_i(X)] W_i(x)$$

est la projection orthogonale dans $L^2([0,1]^2, \mathbf{0})$ de $(x,y) \mapsto y$ sur le sous-espace engendré

par les fonctions indicatrices d'ensemble $\mathbf{1}_{\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right]}(x)$, et est donc aussi égal à

$$\begin{aligned} & \sum_{m=0}^{2^k-1} \left(\int_{x=\frac{m}{2^k}}^{x=\frac{m+1}{2^k}} \int_{y=0}^{y=1} y \cdot \mathbf{0}(dx, dy) \right) \frac{\mathbf{1}_{\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right]}(x)}{2^{-k}} \\ &= \sum_{m=0}^{2^k-1} \mathbb{E}\left(Y | X \in \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right] \right) \mathbf{1}_{\left[\frac{m}{2^k}, \frac{m+1}{2^k}\right]}(x) . \end{aligned}$$

Donc, en définissant pour $k \in \mathbb{N}$ et $x \in [0,1[$, $m_x^k \in \mathbb{N}$, comme vérifiant $\frac{m_x^k}{2^k} \leq x < \frac{m_x^k+1}{2^k}$,

alors,
$$\mathbb{E}(Y | X=x) = \lim_{k \rightarrow \infty} \left\{ \mathbb{E} \left(Y | X \in \left[\frac{m_x^k}{2^k}, \frac{m_x^k+1}{2^k} \right] \right) \right\} \text{ dans } L^2([0,1]).$$

Ce résultat rappelle le fait que, sous certaines hypothèses, *en chaque point x*, $\mathbb{E}(Y|X=x)$ est la limite de l'espérance de Y, sachant que X appartient à un intervalle J contenant x lorsque la longueur de celui-ci tend vers 0.

En ce qui concerne la probabilité conditionnelle d'une variable aléatoire X de loi uniforme sur $[0,1]$, on obtient, de la même façon l'égalité suivante dans $L^2([0,1])$:

$$P(A | X=x) = \lim_{k \rightarrow \infty} \left\{ P \left(A | X \in \left[\frac{m_x^k}{2^k}, \frac{m_x^k+1}{2^k} \right] \right) \right\} .$$

Bien sûr, on obtient exactement le même type de résultats pour des systèmes définis sur des suites de partitions de \mathbb{R}^p et \mathbb{R}^q .

Dans ce cas, le test d'indépendance associé sera le test du chi-deux.

En effet, on suppose maintenant que (X, Y) est un vecteur aléatoire à valeurs dans \mathbb{R}^2 , et que $\{J_i\}_{i=1,2,\dots,c}$ et $\{I_j\}_{j=1,2,\dots,d}$ représentent 2 partitions finies de \mathbb{R} . Soient $\{A_i\}_{i=0,1,\dots,c-1}$ et $\{B_j\}_{j=0,1,\dots,d-1}$ (respectivement, $\{A_i^n\}_{i=0,1,\dots,c-1}$ et $\{B_j^n\}_{j=0,1,\dots,d-1}$), les familles orthonormales de $L^2(\mathbb{R}, \mu')$ et $L^2(\mathbb{R}, \mu'')$ (respectivement, $L^2(\mathbb{R}, \mu'_n)$ et $L^2(\mathbb{R}, \mu''_n)$) obtenues en orthogonalisant les fonctions indicatrices d'ensembles 1_{J_i} et 1_{I_j} et en imposant $A_0 = B_0 = A_0^n = B_0^n \equiv 1$.

On note alors par $n_{i,j}$ (resp. n_i, n_j) le nombre de points de l'échantillon appartenant à $J_i \times I_j$ (resp. $J_i \times \mathbb{R}, \mathbb{R} \times I_j$).

On pose $\alpha_{i,j} = \mathbb{E}[A_i(X)B_j(Y)]$, $\alpha_{i,j}^{*n} = \frac{\sum_{I=1}^n A_i(X_I)B_j(Y_I)}{n}$ et $\alpha_{i,j}^n = \frac{\sum_{I=1}^n A_i^n(X_I)B_j^n(Y_I)}{n}$.

On considère alors l'approximation $\mathbf{Q}^{(c,d)}$ de \mathbf{Q} admettant pour densité par rapport à $\mu' \otimes \mu''$:

$$f^{(c,d)}(x,y) = \sum_{i=1}^c \sum_{j=1}^d \frac{\mathbf{Q}(J_i \times I_j)}{\mu'(J_i)\mu''(I_j)} 1_{J_i}(x) 1_{I_j}(y) .$$

On estimera $f^{(c,d)}$ par les densités empiriques f_n^* et f_n :

$$f_n^*(x,y) = \sum_{i=1}^c \sum_{j=1}^d \frac{\frac{n_{i,j}}{n}}{\mu'(J_i)\mu''(I_j)} 1_{J_i}(x) 1_{I_j}(y) \text{ et } f_n(x,y) = \sum_{i=1}^c \sum_{j=1}^d \frac{\frac{n_{i,j}}{n}}{\frac{n_i}{n} \frac{n_j}{n}} 1_{J_i}(x) 1_{I_j}(y) .$$

Les propriétés élémentaires des bases orthonormales montrent alors que

$$f^{(c,d)}(x,y) = 1 + \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} \alpha_{i,j} A_i(x) B_j(y) \text{ et } f_n(x,y) = 1 + \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} \alpha_{i,j}^n A_i^n(x) B_j^n(y) .$$

On pose alors

$$\chi_{X,Y}^{*2} = \sum_{i=1}^c \sum_{j=1}^d \frac{(n_{i,j} - n \cdot \mu'(j_i) \mu''(\lambda_j))^2}{n \cdot \mu'(j_i) \mu''(\lambda_j)}, \quad \chi_{X,Y}^2 = \sum_{i=1}^c \sum_{j=1}^d \frac{\left(n_{i,j} - \frac{n_i \cdot n_j}{n}\right)^2}{\frac{n_i \cdot n_j}{n}},$$

$$\chi_X^{*2} = \sum_{i=1}^c \frac{(n_i - n \cdot \mu'(j_i))^2}{n \cdot \mu'(j_i)} \quad \text{et} \quad \chi_Y^{*2} = \sum_{j=1}^d \frac{(n_j - n \cdot \mu''(\lambda_j))^2}{n \cdot \mu''(\lambda_j)}.$$

Dans le cas où l'on sait que les lois marginales sont μ' et μ'' , l'emploi des coefficients de corrélation d'ordre supérieur a permis de prouver que

$$\chi_{X,Y}^{*2} - \chi_X^{*2} - \chi_Y^{*2} = n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\alpha_{i,j}^n)^2,$$

et suit asymptotiquement, sous l'hypothèse d'indépendance, une loi du chi-deux à $(c-1)(d-1)$ degrés de liberté. En fait, ce test est le test du chi-deux restreint, et il est donc plus puissant que le test associé à la statistique $\chi_{X,Y}^{*2}$ car ce dernier teste de nouveau que les lois marginales sont bien μ' et μ'' (Blacher 1988).

Dans le cas de lois marginales inconnues, on a les égalités suivantes :

$$\begin{aligned} \sum_{i=1}^c \sum_{j=1}^d \frac{\left(n_{i,j} - \frac{n_i \cdot n_j}{n}\right)^2}{\frac{n_i \cdot n_j}{n}} &= \sum_{i=1}^c \sum_{j=1}^d \left(\frac{n_i \cdot n_j}{n}\right) \left(\frac{n_{i,j}}{n_i \cdot n_j} - 1\right)^2 \\ &= n \sum_{i=1}^c \sum_{j=1}^d \iint_{J_i \times \lambda_j} (f_n(x,y) - 1)^2 \mu'_n(dx) \mu''_n(dy) \\ &= n \iint (f_n(x,y) - 1)^2 \mu'_n(dx) \mu''_n(dy) = n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\alpha_{i,j}^n)^2. \end{aligned}$$

Donc, la statistique du chi-deux avec estimation de paramètre est identique à la statistique Hilbertienne avec estimation de paramètres, i.e. le test du chi-deux d'indépendance avec estimation de paramètres est un cas particulier du test défini par le théorème 7-9.

VIII-D : COEFFICIENTS DE CORRELATION CANONIQUE GENERALISES

On reprend l'exemple donné par Lancaster en 1962 page 535. On considère un vecteur aléatoire (X, Y) à valeurs dans $[0, 1]^2$, de loi \mathbf{Q} et de lois marginales uniformes.

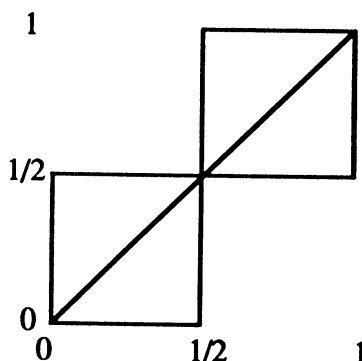
On suppose $\mathbf{Q}\left\{\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]\right\} = \mathbf{Q}\left\{\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]\right\} = 0$.

Soit Δ_1 la diagonale de $\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]$. On suppose que la restriction de \mathbf{Q} à Δ_1 est uniforme et que $\mathbf{Q}(\Delta_1) = \beta_1$.

De même, on suppose que \mathbf{Q} est uniforme sur $\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right] \setminus \Delta_1$,

et que $\mathbf{Q}\left(\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right] \setminus \Delta_1\right) = \frac{1}{2} - \beta_1$.

On suppose qu'il en est de même pour le carré $\left[\frac{1}{2}, 1\right] \times \left[\frac{1}{2}, 1\right]$, mais avec une probabilité $\beta_2 < \beta_1$.



Soit alors $\{P_i\}_{i \in \mathbb{N}}$ et $\{P^*_i\}_{i \in \mathbb{N}}$ deux bases orthonormales quelconques associées aux lois uniformes sur $[0, 1/2]$ et $[1/2, 1]$: $P_i(x) \equiv 0$ si $x \notin [0, 1/2]$, $P^*_i(x) \equiv 0$ si $x \notin [1/2, 1]$, avec $P_0 = P^*_0 \equiv 1$. De plus, on note toujours par W_i les fonctions de Walsh, pour $i=0, 1$.

Alors la famille $\{\tau^{(i)}\}_{i \in \mathbb{N}}$, définie par $\tau^{(i)} = W_i$, pour $i=0, 1$, $\tau^{(2i)} = P_i$, $\tau^{(2i+1)} = P^*_i$, pour $i \geq 1$, est une famille orthonormale associée à la loi uniforme sur $[0, 1]$.

Il est alors facile de voir que les fonctions $\{\tau^{(i)}\}_{i \in \mathbb{N}}$ possèdent la propriété de bi-orthogonalité : $\mathbb{E}(\tau^{(i)}(X)\tau^{(j)}(Y)) = \delta_{ij}$ si $i=0, 1$, $\beta_1 \delta_{ij}$ si i est pair et $i \geq 2$, et $\beta_2 \delta_{ij}$ si i est impair et $i \geq 3$.

Donc les paires $(\tau^{(i)}, \tau^{(i)})$ représentent les paires de fonctions canoniques généralisées, et les paires $(\tau^{(1)}, \tau^{(1)})$ et $(\tau^{(2i)}, \tau^{(2i)})$ représentent les paires de fonctions canoniques définies selon la façon classique introduite par Lancaster et généralisée en 3-20.

Enfin, il est facile de voir que $\{\tau^{(i)}\}_{i \in \mathbb{N}}$ est une base de $L^2([0, 1])$, et donc que l'on peut écrire la fonction de répartition de (X, Y) sous la forme de 3-30.

BIBLIOGRAPHIE

- AKHIEZER N.I. et GLAZMAN I. M. 1966 . Theory of Linear Operators in Hilbert Space, vol 1, Second Printing, Frederick Ungar Publishing Co, New York.
- BEAUCHAMPS K.G. 1975 . Walsh functions and their applications. Academic Press, London.
- BLACHER R. 1984-a . Coefficients de corrélation d'ordre supérieur. Statistique et analyse des données. Vol 9, n°2 , p 48-67 .
- BLACHER R. 1984-b . Coefficients de corrélation d'ordre (i,j) et variances d'ordre i. Symposium international de Bar le Duc, Polynômes orthogonaux et applications. Springer-Verlag n° 1171.
- BLACHER R. 1987 . Divers indices de dépendance fournis par les fonctions orthogonales. Statistique et analyse des données, vol 12, n° 1-2, p 1-27 .
- BLACHER R. 1988 . A New Form for The chi-Squared Test of Independence. Statistics 19-4, 519-536.
- BOSQ D. 1978-a . Test d'ajustement Hilbertien. Rapport interne de l'Université de Lille1, n° 125.
- BOSQ D. 1978-b. Test Hilbertien et test du khi-deux. C.R. Académie des Sciences de Paris, t 286, p 946-948.
- BOSQ D. 1983. Lois limites et efficacité asymptotique des tests Hilbertiens de diverses lois sous des hypothèses adjacentes. "Statistique et analyse des données", Volume 8, n° 1 , p 1-40 .
- CAPERAA P. VAN CUTSEM B. 1988 . Méthodes et Modèles en Statistique non Paramétrique. Presses de L'université Laval, Paris.
- CHIHARA T.S. 1978 . An introduction to orthogonal polynomials. Gordon and Breach, Science Publishers. New York, Paris.
- CHOQUET 1964 . Topologie. Masson
- CSAKI P. et FISCHER. J. 1963 . On the General Notion of Maximal Correlation. Magyar tud. Akad. Mat. Kutato int. Közl, 8, 27-51.
- DAUXOIS et J. POUSSE A. 1975. Une extension de l'analyse canonique. Quelques applications. Ann. Inst. Henri Poincaré, vol XI, n°4, p 355-379.
- GUPTA S.S. 1963 . Probability Integrals of Multivariate Normal and Multivariate t. Ann Math Statistics. 34, 792-828.
- KAGAN A.M. 1975. On a measure of divergency between two scalars products and its Statistical Application. Sankhya 37, 211-228.
- KOYAK R.A. 1987 . On measuring internal dependance in a set of random variables. The Annals of Statistics, vol 15, n° 3, 1215-1228.
- LANCASTER H.O. 1958 . The structure of bivariate distributions. Ann Math Statis. 29, 719-736. Correction 35, 1388.

- LANCASTER H.O. 1963-a . Correlation and canonical forms of bivariate distributions. Ann Math Statis. 34, 532-538.
- LANCASTER H.O. 1963-b . Correlation and complete dependence of random variables Ann Math Statis. 34, 1315-1321.
- LANCASTER H.O. 1969 . The chi-squared distribution. Wiley, New York.
- LANCASTER H.O. 1980 . Orthogonal models for contingency tables. Developments in Statistics, Academic Press, New-York.
- NATANSON I.P. 1965 . Constructive Function Theory, vol 2, Approximation in Mean. Frederick Ungar Publishing C°. New-York.
- RENYI A. 1959 On measures of dependence. Acta. Math. Acad. Sci. Ungar. 10. 441-451.
- RIESZ F. et NAGY B.SZ. 1955 . Leçons d'Analyse Fonctionnelle, 3° édition. Académie des Sciences de Hongrie, Budapest.
- SERFLING R. J. 1980 . Approximation theorems of mathematical statistics. Wiley. New-York.
- SCHWARTZ L. 1979 . Analyses Hilbertienne. Hermann, Paris.
- SCHWEIZER B. et WOLFF. E.F. 1981. On non parametric measures of dependence for random variables. The Annals of Statistics, vol 9 n°4, 879-885.
- SMIRNOV 1961 . Linear Algebra and Group Theory. McGraw-Hill Company, Inc, New York, the Maple Press Company, York.
- WEIDMANN J. 1980 . Linear Operators in Hilbert Space. Springer-Verlag, New York.

DOCUMENT II

A NEW FORM FOR
THE CHI SQUARED
INDEPENDENCE TEST

ARTICLE PUBLIE EN 1988 DANS *STATISTICS*

statistics 19 (1988) 4, 519–536

A New Form for the Chi-Squared Test of Independence

R. BLACHER

Université Scientifique et Médicale de Grenoble

Summary. We study in a new way the chi-squared tests of independence: tests with and without estimated marginal distributions, and the chi-squared restricted test for which we simplify the statistic test. We obtain thus a new form of the chi-squared tests of independence which simplify the presentation of them. Moreover our method enlightens the difference of these tests and produces examples of minimum chi-squared estimators, HILBERTIAN tests, and different chi-squared tests.

AMS 1980 subject classification: 62F05.

Key words: Chi-squared test, independence test, HILBERTIAN test.

Introduction

Using DENIS BOSQ's HILBERTIAN tests [6], and orthogonality in L^2 spaces, we want to differentiate chi-squared tests of independence.

Let us remain some basic principles of orthogonality and HILBERTIAN tests.

Let μ be a probability on \mathbb{R}^p . Two functions h and h' of $L^2(\mathbb{R}^p, \mu)$ are orthogonal in this space if $\int h \cdot h' \cdot d\mu = 0$. Let $\{A_i\}_{i \in \mathbb{N}}$ be a family of $L^2(\mathbb{R}^p, \mu)$. This family is an orthonormal basis of $L^2(\mathbb{R}^p, \mu)$ if it generates $L^2(\mathbb{R}^p, \mu)$ and if $\int A_i \cdot A_{i'} \cdot d\mu = 0$ if $i \neq i'$ and 1 if not.

For example, if U is LEBESGUE measure on $[0, 1]$, then $\sin(2\pi x)$ and $\cos(2\pi x)$ are orthogonal on $L^2([0, 1], U)$. If N is the standard normal distribution, the HERMITE polynomials $\{H_i(x)\}_{i \in \mathbb{N}}$ are an orthonormal basis of $L^2(\mathbb{R}, N)$. If N^2 is a normal distribution on \mathbb{R}^2 , with marginal laws N and with ρ for correlation coefficient, then the family $\{H_i(x)H_j(y)\}_{(i,j) \in \mathbb{N}^2}$ is an orthonormal basis of $L^2(\mathbb{R}^2, N^2)$ if $\rho = 0$, and is not orthogonal if not.

Let μ_1 and μ_2 be two probabilities on \mathbb{R} . We consider a random vector (X_1, X_2) with values in \mathbb{R}^2 such that there exists a probability density function f_{X_1, X_2} with respect to $\mu_1 \otimes \mu_2$.

Then,

$$P((X_1, X_2) \in B_1 \times B_2) = \iint_{B_1 \times B_2} f_{X_1, X_2} d\mu_1 d\mu_2, \quad (\text{I})$$

where P is the probability, and B_1, B_2 are two borelian subsets of \mathbb{R} , and, we can write

$$f_{X_1, X_2} = \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \alpha_{j_1, j_2} A_{j_1}^1 \otimes A_{j_2}^2 \quad \text{in } L^2(\mathbb{R}^2, \mu_1 \otimes \mu_2), \quad (\text{II})$$

where, respectively, $\{A_{j_1}^1\}_{j_1 \in N^*}$ and $\{A_{j_2}^2\}_{j_2 \in N^*}$ are orthonormal basis of $L^2(\mathbb{R}, \mu_1)$ and $L^2(\mathbb{R}, \mu_2)$, and where $A_{j_1}^1 \otimes A_{j_2}^2(x, y) = A_{j_1}^1(x) A_{j_2}^2(y)$.

Moreover, we set $A_1^1 \equiv A_1^2 \equiv 1$. Then

$$\alpha_{1,1} = 1 \quad \text{and} \quad A_1^1 \otimes A_1^2 \equiv 1. \tag{III}$$

At last, in the case where the distribution of (X_1, X_2) is $\mu_1 \otimes \mu_2$, we have $f_{X_1, X_2} \equiv 1$: cf (I). Then $\alpha_{j_1, j_2} = 0$ if $(j_1, j_2) \neq (1, 1)$: cf (II) and (III).

The Bosq's results show that, for all finite subset D^* of $(\mathbb{N}^*)^2 \setminus \{(1, 1)\}$, the HILBERTIAN test can test the hypothesis: $H_0(D^*) : \langle\langle \forall (j_1, j_2) \in D^*, \alpha_{j_1, j_2} = 0 \rangle\rangle$, and that we can denote the associated HILBERTIAN statistics by

$$n \left(\sum_{(j_1, j_2) \in D^*} \widehat{\alpha_{j_1, j_2}^n}^2 \right).$$

Thus, by an appropriate choice of D^* , this test is used to test the hypothesis

$$C'_0(\mu_1 \otimes \mu_2) : \langle\langle \text{The law of } (X_1, X_2) \text{ is } \mu_1 \otimes \mu_2 \rangle\rangle.$$

We recall these results in the first part.

Now, let f_{X_1} be the probability density function of X_1 with respect to μ_1 . We can write $f_{X_1} = 1 + \sum_{j_1=2}^{\infty} \alpha_{j_1, 1} A_{j_1}^1$ in $L^2(\mathbb{R}, \mu_1)$. We have the equality $P(X_1 \in B_1) = \int_{B_1} f_{X_1} d\mu_1$. Then, the HILBERTIAN test is also used to test

$$C'_0(\mu_1) : \langle\langle \text{The law of } X_1 \text{ is } \mu_1 \rangle\rangle,$$

by the test of

$$H_0(\{2, 3, \dots, c(1)\} \times \{1\}) : \langle\langle \forall j \in \{2, 3, \dots, c(1)\}, \alpha_{j, 1} = 0 \rangle\rangle.$$

Now, we can test $C'_0(\mu_1)$ using the chi-squared test.

Let $(X_1^p)_{p=1, 2, \dots, n}$ be an n -sample of X_1 , and let $\{I_i^1\}_{i=1, 2, \dots, c(1)}$ be a partition of \mathbb{R} on disjoint intervals. The classical chi squared test uses the χ^2 statistics:

$$\chi_{X_1}^2 = \sum_{i=1}^{c(1)} \frac{(n_i - n\mu_1(I_i^1))^2}{n\mu_1(I_i^1)},$$

where n_i is the number of X_1^p which belongs to I_i^1 .

Now, Bosq has remarked that the chi squared test is a HILBERTIAN test. Indeed, let $\{1_{I_i^1}\}_{i=1, 2, \dots, c(1)}$ the family of characteristic functions of the I_i^1 sets. Then.

we choose $\{A_{j_1}^1\}_{j_1=1, 2, \dots, c(1)}$ the orthonormal family such that each $A_{j_1}^1$ is a linear combination of $1_{I_s^1}$.

In particular, $A_1^1 = \sum_{s=1}^{c(1)} 1_{I_s^1}$.

We shall prove the equality:

$$n \left(\sum_{j_1=2}^{c(1)} \widehat{\alpha_{j_1, 1}^n}^2 \right) = \chi_{X_1}^2.$$

Then the HILBERTIAN statistics is equal at the classical chi squared statistics and then the χ^2 test is as HILBERTIAN one.

In the same way, let $\{I_i^2\}_{i=1,2,\dots,c(2)}$ be another partition of \mathbb{R} . We impose the same condition to the $\{A_{j_2}^2\}_{j_2=1,2,\dots,c(2)}$'s family, and we denote by $\chi_{X_2}^2$ and χ_{X_1, X_2}^2 the chi squared statistics on \mathbb{R} and \mathbb{R}^2 , respectively, associated with the $\{I_i^2\}_{i=1,2,\dots,c(2)}$ and $\{I_i^1 \times I_i^2\}_{i=1,2,\dots,c(1)}$ partitions. We shall obtain the equalities

$$\chi_{X_2}^2 = n \left(\sum_{j_2=2}^{c(2)} \widehat{\alpha_{1, j_2}^n}^2 \right),$$

and

$$\chi_{X_1, X_2}^2 = n \left(\sum_{\substack{1 \leq j_1 \leq c(1) \\ 1 \leq j_2 \leq c(2) \\ (j_1, j_2) \neq (1,1)}} \widehat{\alpha_{j_1, j_2}^n}^2 \right).$$

We shall prove these results in a second part.

If we know that the law of X_i is μ_i , for $i = 1, 2$, we have the equality $P(X_i \in B_i) = \int_{B_i} d\mu_i$. Then $\alpha_{j,1} = \alpha_{1,j} = 0$ for $j \neq 1$. Consequently, for testing the independence of X_1 and X_2 , we must test

$$H_0((2, 3, \dots, c(1)) \times (2, 3, \dots, c(2))) : \\ \langle\langle \forall (j_1, j_2) : j_1 = 2, 3, \dots, c(1), j_2 = 2, 3, \dots, c(2), \alpha_{j_1, j_2} = 0 \rangle\rangle.$$

The associated HILBERTIAN statistics is

$$\|S_n\|^2 = n \left(\sum_{j_1=2}^{c(1)} \sum_{j_2=2}^{c(2)} \widehat{\alpha_{j_1, j_2}^n}^2 \right).$$

Moreover, we have the equality

$$\|S_n\|^2 = \chi_{X_1, X_2}^2 - \chi_{X_1}^2 - \chi_{X_2}^2,$$

and, under the independence hypothesis, $\|S_n\|^2, \chi_{X_1}^2, \chi_{X_2}^2$ are asymptotically independent.

Now, we can also test the independence using the classical χ^2 -independence test.

In the case where we know the marginal laws, the statistics of this test is χ_{X_1, X_2}^2 if we suppose always that the law of X_i is μ_i .

And we have the equality

$$\chi_{X_1, X_2}^2 = n \left(\sum_{j_1=2}^{c(1)} \sum_{j_2=2}^{c(2)} \widehat{\alpha_{j_1, j_2}^n}^2 \right) + n \left(\sum_{j_1=2}^{c(1)} \widehat{\alpha_{j_1, 1}^n}^2 \right) + n \left(\sum_{j_2=2}^{c(2)} \widehat{\alpha_{1, j_2}^n}^2 \right).$$

Then, the classical chi-squared tests of independence test again that the $\alpha_{j_1, 1}$ and α_{1, j_2} vanish. Then the classical chi squared independence test tests also that the marginal distributions are μ_1 and μ_2 .

This means that the classical test tests again a known hypothesis.

The Bosq's results show that, under the independence hypothesis, the limit

distribution of the statistic

$$\chi^2_{X_1, X_2} - \chi^2_{X_1} - \chi^2_{X_2}$$

is a χ^2 law with $(c(1) - 1)(c(2) - 1)$ degrees of freedom.

These results recall those of restricted χ^2 test [8]. Indeed, we know that this test allows to test only some of the parameters. In the present case, if we want to test $H_0(\{2, 3, \dots, c(1)\} \times \{2, 3, \dots, c(2)\})$ by means of the restricted χ^2 test with an n -sample $(X_1^p, X_2^p)_{p=1,2,\dots,n}$ of (X_1, X_2) , the statistics of this test, $\overline{\chi^2_{X_1, X_2}}$ is equal to

$$\overline{\chi^2_{X_1, X_2}} = \sum_{\substack{1 \leq i \leq c(1) \\ 1 \leq i' \leq c(2)}} \frac{(n_{i,i'} - n \cdot \mu_1(I_i^1) \cdot \mu_2(I_{i'}^2))^2}{n \cdot \mu_1(I_i^1) \cdot \mu_2(I_{i'}^2)} - \sum_{\substack{1 \leq i \leq c(1) \\ 1 \leq i' \leq c(2)}} \frac{(n_{i,i'} - n \cdot \widehat{p}_{i,i'})^2}{n \cdot \mu_1(I_i^1) \cdot \mu_2(I_{i'}^2)}$$

where $n_{i,i'}$ is the number of (X_1^p, X_2^p) which belong to $I_i^1 \times I_{i'}^2$ and where $\{\widehat{p}_{i,i'}\}$ is the χ^2 minimum reduced estimator.

Then, we shall prove the equalities:

$$\overline{\chi^2_{X_1, X_2}} = \|S_n\|^2,$$

and

$$\chi^2_{X_1} + \chi^2_{X_2} = \sum_{\substack{1 \leq i \leq c(1) \\ 1 \leq i' \leq c(2)}} \frac{(n_{i,i'} - n \cdot \widehat{p}_{i,i'})^2}{n \cdot \mu_1(I_i^1) \cdot \mu_2(I_{i'}^2)}$$

Therefore, in this case, the restricted χ^2 test of independence and the HILBERTIAN test of independence are the same.

We shall prove these results in the third part.

In the fourth part, we shall study the case where the marginal laws are unknown. In this case, we must estimate the parameters.

Let $\widehat{\chi^2_{X_1, X_2}}$, $\widehat{\chi^2_{X_1}}$ and $\widehat{\chi^2_{X_2}}$ be the χ^2 statistics with estimation of parameters, and let $\|S_n\|^2 = \widehat{\chi^2_{X_1, X_2}} - \widehat{\chi^2_{X_1}} - \widehat{\chi^2_{X_2}}$.

We remark that $\widehat{\chi^2_{X_1}} = \widehat{\chi^2_{X_2}} = 0$. Therefore $\|S_n\|^2 = \widehat{\chi^2_{X_1, X_2}}$.

Then, the statistics of the two tests are the same. Therefore, the two tests are the same.

Finally, in the fifth part, we recall the conclusions. In particular, we remark that, under the independence hypothesis, the both statistics

$$\chi^2_{X_1, X_2} - \chi^2_{X_1} - \chi^2_{X_2} \quad \text{and} \quad \widehat{\chi^2_{X_1, X_2}} - \widehat{\chi^2_{X_1}} - \widehat{\chi^2_{X_2}}$$

have asymptotically a chi-squared distribution with $(c(1) - 1)(c(2) - 1)$ degrees of freedom.

1. HILBERTIAN test and restricted chi-square test

At first, we recall some definitions and properties of HILBERTIAN tests as in Bosq ([6], [7]).

1.1. Notation. Let $\{Z_i\}, i \in \mathbb{N}^*$, be a sequence of independent identically distributed random vectors, defined on the same probability space (Ω, \mathcal{A}, P) , and with values in a BOREL space $(\mathbb{R}^p, \mathcal{B}_p)$. Let m be a given probability measure on $(\mathbb{R}^p, \mathcal{B}_p)$.

We want to test the hypothesis $C'_0 : \langle\langle$ The common distribution of Z_i is $m \rangle\rangle$ against its negation.

Bosq considers an orthonormal family $\{e_J\}, J \in D$, of $L^2(\mathbb{R}^p, \mathcal{B}_p, m)$ such that $1 \in \{e_J\}$ and $2 \leq \text{card}(D) = k < +\infty$. Here $D \subset (\mathbb{N}^*)^p, \sup_{x \in \mathbb{R}^p} |e_J(x)| < +\infty, 1_p = (1, 1, \dots, 1) \in D$ and $e_{1_p} \equiv 1$.

Since (e_J) is not necessary a basis, then we have

$$\dim(L^2(\mathbb{R}^p, \mathcal{B}_p, m)) \geq k.$$

We set $D^* = D \setminus \{1_p\}$, and, for $n \in \mathbb{N}^*$ and $J \in D$, we denote by \widehat{a}_J^n the statistics:

$$\widehat{a}_J^n = \frac{1}{n} \left[\sum_{i=1}^n e_J(Z_i) \right]$$

which is an estimator converging almost surely to $a_J = E(e_J(Z_1))$ when

$$E(|e_J(Z_1)|) < +\infty.$$

Let $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ be respectively the norm and the scalar product of $L^2(\mathbb{R}^p, \mathcal{B}_p, m)$. Then, for $n \in \mathbb{N}^*$, we denote by $S_n(\cdot)$ the statistics:

$$S_n(\cdot) = \sqrt{n} \left(\sum_{J \in D^*} \widehat{a}_J^n e_J(\cdot) \right).$$

Consequently

$$\|S_n\|^2 = n \left(\sum_{J \in D^*} \widehat{a}_J^n^2 \right).$$

Therefore, if the law of Z_1 is m , then $E(e_J(Z_1) \cdot e_{J'}(Z_1)) = 0$ if $J \neq J'$ and 1 if not, and the Central Limit Theorem proves the following theorem.

Proposition 1.2. *If the law of Z_1 is m , then the sequence of random vectors $n^{1/2} \{\widehat{a}_J^n\}_{J \in D^*}$ has asymptotically a $(k-1)$ -variate normal limit distribution with mean zero and covariance structure the identity matrix.*

Then, we deduce the following propositions.

Proposition 1.3. *If the law of Z_1 is m , then the random variables $n^{1/2} \widehat{a}_J^n, J \in D^*$, are asymptotically independent.*

Proposition 1.4. *If the common distribution of Z_i is m , then $\|S_n\|^2 \xrightarrow{\text{Distribution}} T$, where T has a chi-squared distribution with $k - 1$ degrees of freedom.*

Proposition 1.5. *If the common distribution of Z_i is a probability m' such that there exists $J \in D^*$ for which e_J is m' -integrable, and with non-vanishing integral. then*

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

1.6. Construction of the test. These results allow us to define a test of

$$H_0(D^*) : \langle\langle \forall J \in D^*, a_J = 0 \rangle\rangle$$

against its negation $H_1(D^*)$. by using the statistics $\|S_n\|^2$: under $H_0(D^*)$, we have $\|S_n\|^2 \xrightarrow{\text{Distribution}} \tau_{k-1}^2$. and under $H_1(D^*)$, we have $\|S_n\|^2 \xrightarrow{\text{a.s.}} +\infty$.

Then, by use of the HILBERTIAN test, we can test the Hypothesis

$$C'_0 : \langle\langle \text{The law of } Z_1 \text{ is } m \rangle\rangle .$$

The above theorems constitute the foundation of HILBERTIAN tests. This very simple case was developed by many authors. Note that BOSQ's results allow to test an infinite number of parameters ([6], [7]). Moreover BOSQ proposes to use a sequence $\{(e_j^n)_{j \in N^*}\}_{n \in N^*}$ of orthogonal functions families in $L^2(\mathbb{R}^p, m)$. Other way, GERLEIN and PINCUS [10] have considered the more general situation where the HILBERTIAN test is a particular case of the w -divergence [14] between two scalar products. Other references are BICKELL and ROSENBLATT [2] and GREGORY ([11], [12]).

We know also that χ^2 tests are special HILBERTIAN tests by the choice of characteristic functions of set as orthogonal functions.

Notations 1.7. Let $n \in \mathbb{N}^*$ and $\{I_i\}_{i \in \{1, 2, \dots, c\}}$ be a partition of \mathbb{R}^p . For all $i \in \{1, 2, \dots, c\}$, we denote by p_i and π_i the probabilities: $p_i = P(Z_1 \in I_i)$ and $\pi_i = m(I_i)$. Then, let G denote the hypothesis, $G : \langle\langle \{p_i\} = \{\pi_i\} \rangle\rangle$. Moreover, let ν_i be the real number $\nu_i = (1/n) \text{card} (\{t \leq n \mid Z_t \in I_i\})$.

Of course, for testing C'_0 , we may use the classical chi squared test. However, frequently one does know the type of deviations to be expected if the hypothesis is not true, and in such cases the chi-squared test can be modified so as to increase its asymptotic power against the alternatives of interest by concentrating it on these alternatives.

Notation 1.8. Let $s \in \mathbb{N}^*$, $s \leq c$. Let \mathcal{E} be the affine subspace of \mathbb{R}^c have a parametric representation

$$p_i = \gamma_{i,0} + \sum_{r=1}^s \gamma_{i,r} \vartheta_r \quad i = 1, 2, \dots, c$$

and let

$$\pi_i = \gamma_{i,0} + \sum_{r=1}^s \gamma_{i,r} \vartheta_r^0 \quad i = 1, 2, \dots, c$$

where the $\gamma_{i,r}$ are real constants.

Let K be the restricted class of alternatives to G defined by

$$K : \langle \langle \{p_i\} \in \mathfrak{E}, \{p_i\} \neq \{\pi_i\} \rangle \rangle.$$

In that case, in order to test G vs K , we may use the restricted χ^2 test (p. 308 [15], see also [8]).

Definition 1.9. The restricted chi squared test of level α , rejects G in favour of K if

$$\sum_{i=1}^c \frac{n (v_i - \pi_i)^2}{\pi_i} - \sum_{i=1}^c \frac{n (v_i - \widehat{p}_i)^2}{\pi_i} = \sum_{i=1}^c \frac{n (\widehat{p}_i - \pi_i)^2}{\pi_i} > C$$

where the \widehat{p}_i minimize $\sum_{i=1}^c \frac{(v_i - q_i)^2}{\pi_i}$ subject $\{q_i\} \in \mathfrak{E}$. The constant $C > 0$ is determined by

$$\int_C^{+\infty} 1/[2\Gamma(n/2)] (z/2)^{s/2-1} e^{-z/2} dz = \alpha.$$

Remark 1.10. As a matter of fact the restricted χ^2 test tests $\langle \langle \vartheta_r = \vartheta_r^0, r = 1, 2, \dots, s \rangle \rangle$. Thus, for alternatives lying on \mathfrak{E} , the restricted χ^2 test is more powerful than the classical χ^2 test. This result follows from the study done by FIX, HODGES and LEHMANN (p. 98 [8]). Of course, if the alternative does not belong to \mathfrak{E} , this result is not necessarily true.

We recall that the restricted χ^2 test is defined under weaker hypothesis. Indeed, we can suppose \mathfrak{E} is a surface and $\{p_i\}$ is another estimator than the χ^2 minimum reduced estimator [8].

2. Chi-squared statistics and HILBERTIAN statistics

We now prove that the HILBERTIAN statistics involving all the coefficients α_{j_1, j_2} is exactly the chi-squared statistics.

2.1. Notations for the partitions of \mathbb{R}^p : If $p \in \mathbb{N}^*$, let $\{c(s)\}$, $s = 1, 2, \dots, p$, be a finite sequence of \mathbb{N}^* : $c(s) \in \mathbb{N}^*$. We set $\mathfrak{D}^s = (1, 2, \dots, c(s))$ and $\mathfrak{D}^{*s} = (2, 3, \dots, c(s))$.

We consider also the subsets of \mathbb{N}^p :

$$\mathfrak{D} = \prod_{s=1}^p \mathfrak{D}^s, \quad \text{and} \quad \mathfrak{D}^* = \mathfrak{D} \setminus \{(1, 1, \dots, 1)\}.$$

Let $\{\mathfrak{P}^s\}$, $s = 1, 2, \dots, p$, be a finite sequence of partitions of \mathbb{R} on disjointed intervals I_i^s : $\mathfrak{P}^s = \{I_i^s \mid i \in \mathfrak{D}^s\}$.

Moreover let \mathfrak{I} be the following partition of \mathbb{R}^p :

$$\mathfrak{I} = \left\{ \prod_{i=1}^p I_{i(s)}^s \mid s \in (1, 2, \dots, p) \Rightarrow i(s) \in \mathfrak{D}^s \text{ and } I_{i(s)}^s \in \mathfrak{P}^s \right\}.$$

2.2. Notations for measures. Let $\{\mu_i\}, i = 1, 2, \dots, p$, be a sequence of probability measures on \mathbb{R} , and let $\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_p$ be the product measure on \mathbb{R}^p . Let $(X_i), i \in \mathbb{N}$, be a sequence of independent identically distributed random vectors, defined on the same probability space $(\Omega, \mathfrak{A}, P)$, and with values in the BOREL space \mathbb{R}^p . Let $X = X_0$. We suppose the distribution of X is Q .

Let $n \in \mathbb{N}^*$. We denote by μ_e the empirical probability:

$$\mu_e(\{V\}) = \frac{1}{n} \text{card} (i \mid i \leq n, X_i \in V)$$

and by f_e the empirical density with respect to μ associated to \mathfrak{I} : for all $I \in \mathfrak{I}$ and $x \in I$, we have $f_e(x) = \frac{\mu_e(I)}{\mu(I)}$. We have therefore supposed that $\mu_s(I_{i(s)}^s) > 0$, for all $s \in (1, 2, \dots, p)$ and $I_{i(s)}^s \in \mathfrak{P}^s$.

2.3. Notations for orthogonal functions. Let, for $s = 1, 2, \dots, p$, $\{A_{j_s}^s\}, j_s \in \mathfrak{D}^s$, be an orthonormal family of $L^2(\mathbb{R}, \mu_s)$ such that $A_1^s \equiv 1$. We suppose that it spans the same subspace of $L^2(\mathbb{R}, \mu_s)$ than the family of characteristic functions

$$\{1_{I_{i(s)}^s}\}_{i(s) \in \mathfrak{D}^s} \text{ of sets } I_{i(s)}^s \in \mathfrak{P}^s.$$

Moreover for each $s \in (1, 2, \dots, p)$, $A_{j_s}^s$ is a linear combination of $1_{I_{i(s)}^s}$.

Since $\mu_s(I_{i(s)}^s) > 0$ for all $I_{i(s)}^s \in \mathfrak{P}^s$, we have, according to GRAM-SCHMIDT orthogonalisation process, that such functions always exist.

Finally, we denote by $\{A_J\}, J \in \mathfrak{D}$, the family:

$$\{A_{j_1}^1 \otimes A_{j_2}^2 \otimes \dots \otimes A_{j_p}^p \mid j_s \in \mathfrak{D}^s, \text{ for all } s \in (1, 2, \dots, p)\},$$

where

$$A_{j_1}^1 \otimes A_{j_2}^2 \otimes \dots \otimes A_{j_p}^p(x_1, x_2, \dots, x_p) = A_{j_1}^1(x_1) A_{j_2}^2(x_2) \dots A_{j_p}^p(x_p).$$

Then we have the almost obvious proposition.

Proposition 2.4. *The family $\{A_J\}, J \in \mathfrak{D}$, is an orthonormal family of $L^2(\mathbb{R}^p, \mu)$, which spans the same subspace as the family $\{1_I\}, I \in \mathfrak{I}$. Moreover we have $A_{1,1, \dots, 1} \equiv 1$.*

2.5. Notations for the coefficients of the A_J 's. We shall denote by $\alpha_J, J \in \mathfrak{D}$, the number $\alpha_J = E(A_J(X))$, and by $\widehat{\alpha_J} = \frac{1}{n} \sum_{i=1}^n A_J(X_i)$ the estimator of α_J (1.1, 2-6).

As we claim in the introduction, $\alpha_J(x_{j_1, j_2} \text{ if } p = 2)$ is the coefficient of $A_J(A_{j_1}^1 \otimes A_{j_2}^2)$ in the expansion of the density function (when there exists) in orthonormal

functions. This easy remark allows us to understand better the basic idea of the proof (see also [4]).

Since $\text{Sup}_{x \in \mathbb{R}^p} |A_J(x)| < +\infty$, we can apply the strong law of large numbers, and we obtain the following proposition.

Proposition 2.6. *Under the above hypothesis, $\widehat{\alpha}_J^n$ converges almost surely to α_J .*

2.7. Notations for the coefficients of the marginal distributions. We shall denote by \mathfrak{D}_m^s , $s \in (1, 2, \dots, p)$, the following subsets of \mathfrak{D} :

$$\mathfrak{D}_m^s = \{(i_1, i_2, \dots, i_p) \mid i_s \neq 1, i_j = 1 \text{ if } j \neq s\}.$$

We set also $\mathfrak{D}_m = \bigcup_{s=1}^p \mathfrak{D}_m^s$, and $\overline{\mathfrak{D}} = \mathfrak{D} \setminus \mathfrak{D}_m$.

Then, the coefficients of the marginal distributions — $\alpha_{j,1}$ and α_{1,j_2} as considered in the introduction — are α_{J_0} , $J_0 \in \mathfrak{D}_m$.

With these notations, we can state the following property.

Lemma 2.8. *Under the previous notations, we have*

$$\widehat{\alpha}_J^n = \langle f_e, A_J \rangle_{L^2(\mathbb{R}^p, \mu)}.$$

Proof. We can write

$$A_J = \sum_{I \in \mathfrak{J}} \lambda_I 1_I, \tag{2.3}$$

and

$$\begin{aligned} \frac{1}{n} \left(\sum_{i=1}^n A_J(X_i) \right) &= \iint \dots \int_{\mathbb{R}^p} A_J \cdot d\mu_e = \sum_{I \in \mathfrak{J}} \lambda_I \mu_e(I) \\ &= \sum_{I \in \mathfrak{J}} \lambda_I \left(\iint \dots \int_I \frac{\mu_e(I)}{\mu(I)} d\mu \right) = \langle f_e, A_J \rangle_{L^2(\mathbb{R}^p, \mu)}. \end{aligned}$$

We can now give the main proposition of this part, which tells us that hilbertian and chi-squared statistics are the same.

Theorem 2.9. *Under the previous notations, we have*

$$n \left(\sum_{J \in \mathfrak{D}^*} \widehat{\alpha}_J^n \right)^2 = \sum_{I \in \mathfrak{J}} \frac{(N_I - n\mu(I))^2}{n\mu(I)}.$$

where $N_B = n\mu_e(B)$, while B is a BOREL subset of \mathbb{R}^p .

Proof. As $\{A_J\}$, $J \in \mathfrak{D}$, and $\{1_I\}$, $I \in \mathfrak{J}$, span the same subspace, from 2-8 we obtain $\sum_{J \in \mathfrak{D}} \widehat{\alpha}_J^n A_J = f_e$ or, equivalently $\sum_{J \in \mathfrak{D}^*} \widehat{\alpha}_J^n A_J = f_e - 1$.

Finally, we have

$$\begin{aligned} n \left(\sum_{J \in \mathfrak{D}} \widehat{\alpha}_J^2 \right) &= n \left(\iint \dots \int_{\mathbb{R}^p} (f_e - 1)^2 d\mu \right) = n \left(\sum_{I \in \mathfrak{I}} \iint \dots \int_I (f_e - 1)^2 d\mu \right) \\ &= \sum_{I \in \mathfrak{I}} \frac{(N_I - n\mu(I))^2}{n\mu(I)}; \end{aligned}$$

Remark 2.10. In this paper, we use the HILBERTIAN test in $(\mathbb{R}^p, \mathfrak{B}_p, \mu)$. This test is also defined for (E, \mathfrak{E}, μ) such that $L^2(E, \mathfrak{E}, \mu)$ is separable.

It is easy to see that the theorem 2.8 holds in this generalisation. Then the chi-squared test is a HILBERTIAN test (see also BOSQ [7]).

3. Test of independence: the case of the known marginal distributions

In this chapter, we are in the case where we know the marginal laws.

For testing independence we shall test only the coefficients which measure the dependence. For this, we must eliminate the coefficients which measure the marginal distributions. The orthogonalisation gives us this result.

Hypothesis 3.1. In the sequel of this paper, we suppose that the i -th marginal law is μ_i for all $i \in (1, 2, \dots, p)$.

Proposition 3.2. *If $J_0 \in \mathfrak{D}_m$, we have $\alpha_{J_0} = 0$.* This result follows from the definition of a marginal distribution.

Since the coefficients α_{J_0} , $J_0 \in \mathfrak{D}_m$, of the marginal distributions vanish, we don't have to test them, and we shall carry out the HILBERTIAN test from the statistics $n \left(\sum \widehat{\alpha}_J^2 \right)$ which don't depend of these coefficients, contrary to the classical chi-squared test as we shall see now.

3.3. Notation for the prolongation of partitions of \mathbb{R} to \mathbb{R}^p . For $s \in (1, 2, \dots, p)$, we denote by \mathfrak{J}_s the prolongation of the partition \mathfrak{P}^s of \mathbb{R} to \mathbb{R}^p , given by

$$\mathfrak{J}_s = \{ E_{i_1}^1 \times E_{i_2}^2 \times \dots \times E_{i_p}^p \parallel E_{i_s}^s \in \mathfrak{P}^s \text{ and } E_{i_{s'}}^{s'} = \mathbb{R} \text{ if } s' \neq s \}.$$

We set also $\mathfrak{J}_0 : \mathfrak{J}_0 = \mathfrak{J}$.

3.4. Notations for the chi-squared statistics. Let $i \in (0, 1, \dots, p)$. We shall denote by χ_i^2 , the chi-squared statistics on, (\mathbb{R}^p, μ) , or (\mathbb{R}, μ_i) , $i = 1, 2, \dots, p$, respectively, prolonged to \mathbb{R}^p :

$$\chi_i^2 = \sum_{I \in \mathfrak{J}_i} \frac{(N_I - n\mu(I))^2}{n\mu(I)}, \quad i = 0, 1, \dots, p.$$

When $i \neq 0$, it is obvious that χ_i^2 is the chi-squared statistics of the i -th marginal distribution.

From 2.9., we deduce easily the following theorem.

Theorem 3.5. *With the above notation, if we denote by $\|S_n\|^2$ the HILBERTIAN statistics (1.1) defined as*

$$\|S_n\|^2 = n \left(\sum_{J \in \mathfrak{D}} \widehat{\alpha}_J^2 \right),$$

we have $\|S_n\|^2 = \chi_0^2 - \sum_{s=1}^p \chi_s^2$.

Remark 3.6. Theorem 3.5 holds without hypothesis 3.1. Then, from 1.4 and 3.5, we deduce the asymptotic distribution of $\|S_n\|^2$.

Corollary 3.7. *Under the hypothesis of independence, the statistics $\chi_0^2 - \sum_{s=1}^p \chi_s^2$ is asymptotically distributed as a chi-squared distribution with $\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1)$ degrees of freedom.*

Corollary 3.8. *We suppose $p=2$, and we denote $\chi_0^2 = \chi_{X_1, X_2}^2$, $\chi_1^2 = \chi_{X_1}^2$, $\chi_2^2 = \chi_{X_2}^2$.*

The statistics $\|S_n\|^2 = \chi_{X_1, X_2}^2 - \chi_{X_1}^2 - \chi_{X_2}^2$ is asymptotically, under the independence hypothesis, distributed as a chi-squared distribution with $(c(1) - 1)(c(2) - 1)$ degrees of freedom.

Now, as corollary of 1.3 and 2.9, we remark the asymptotic independence of above statistics.

Theorem 3.9. *Under the independence hypothesis, the statistics $\chi_1^2, \chi_2^2, \dots, \chi_p^2$ and $\|S_n\|^2$ are asymptotically independent.*

The form of $\|S_n\|^2$ looks like that of the restricted χ^2 statistics. As a matter of fact, we have the following equality.

Theorem 3.10. *With the previous notations, we have the equality.*

$$\sum_{s=1}^p \chi_s^2 = \sum_{I \in \mathfrak{S}} \frac{(N_I - n \cdot \widehat{p}_I)^2}{n \mu(I)},$$

where the \widehat{p}_I minimize

$$\sum_{I \in \mathfrak{S}} \frac{(N_I/n - q_I)^2}{\mu(I)}$$

when $q_I = \int_I dM$, M belonging to the set of measures on \mathbb{R}^p with marginal distributions μ_s .

Proof. Let $g = \sum_{I \in \mathfrak{S}} q_I 1_I / \mu(I)$. Then $\sum_{I \in \mathfrak{S}} \frac{(N_I/n - q_I)^2}{\mu(I)} = \sum_{I \in \mathfrak{S}} \mu(I) (N_I/n \mu(I) - q_I / \mu(I))^2 = \int (f_\epsilon - g)^2 d\mu$.

We can also write $g = \sum_{J \in \mathfrak{D}} g_J A_J$. Then, as $\int g \, d\mu = 1, g_{1,1,\dots,1} = 1$. Moreover, let $J_0 \in \mathfrak{D}_m$ and $A_{J_0} = \sum_{I \in \mathfrak{J}} \lambda_I 1_I$. Then,

$$\int g A_{J_0} \, d\mu = \sum_{I \in \mathfrak{J}} \lambda_I q_I = \sum_{I \in \mathfrak{J}} \lambda_I \int 1_I \, dM = \int A_{J_0} \, dM = 0 \tag{cf. 3.2}$$

Other way, by 2.8, we have $f_e - 1 = \sum_{J \in \mathfrak{D}^*} \widehat{\alpha}_J^n A_J$. Then,

$$\int (f_e - g)^2 \, d\mu = \sum_{J \in \mathfrak{D}^*} (\widehat{x}_J^n - g_J)^2 = \sum_{J \in \mathfrak{D}} (\widehat{x}_J^n - g_J)^2 + \sum_{J \in \mathfrak{D}_m} (\widehat{x}_J^n)^2.$$

We deduce that $\int (f_e - g)^2 \, d\mu$ is minimum if $g_J = \widehat{x}_J^n$ for $J \in \overline{\mathfrak{D}}$.

In this case, $\int (f_e - g)^2 \, d\mu = \sum_{J \in \mathfrak{D}_m} (\widehat{x}_J^n)^2$.

Then, by use of 2.9 for each marginal χ^2 statistics, we deduce the announced equality.

From 3.5 and 1.9 (see also [8]), we deduce the following corollary.

Corollary 3.11. *With the previous notations, we have the equality*

$$\|S_n\|^2 = \sum_{I \in \mathfrak{J}} \frac{n(\mu(I) - \widehat{p}_I)^2}{\mu(I)}$$

This corollary means that, in this case, the restricted χ^2 test of independence and the HILBERTIAN test of independence are the same (1.9).

3.12. We suppose always that the i -th marginal law is μ_i for all $i \in \{1, 2, \dots, p\}$. Then, we recall that the study by FIX, HODGES and LEHMANN (p. 98 [8]) shows that the independence test with the statistics $\|S_n\|^2$ is more powerful than the classical χ^2 independence test with χ_0^2 (3.11. 1.10). Of course, if the marginal laws are not the μ_i 's, this result is not necessarily true.

3.13. Independence test. It is obvious that if we test the independence with the help of the HILBERTIAN test, then, in the case of known marginal distributions, we shall test the hypothesis $H_0(\overline{\mathfrak{D}}) : \langle \langle \forall J \in \overline{\mathfrak{D}} : x_J = 0 \rangle \rangle$ against its negation, thanks to the statistics $\|S_n\|^2$, and theorems 1.4 and 1.5.

3.14. Remark. It is also evident that we don't test the coefficients α_{J_0} , $J_0 \in \mathfrak{D}_m$, because they are already equal to zero by proposition 3.2. However, according to 2.9 in the classical chi-squared test, we test again the vanishing of these already known coefficients. Therefore we test again that the marginals distributions are μ_1, μ_2, \dots and μ_p .

3.15. Remark. We can also express the preceding statements by remarking

that the classical chi-squared test allows us to test the hypothesis $H_0(\mathcal{D}^*) : \langle\langle \forall J \in \mathcal{D}^* : \alpha_J = 0 \rangle\rangle$, and is utilized as test of $C'_0 : \langle\langle \text{The distribution of } X \text{ is } \mu \rangle\rangle$ against its negation C'_1 . When we accept that $\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_p$, we can deduce the independence. Then, in this case, it is clear that we neglect some data: that is to say the knowledge of the marginal distributions.

On an other hand, with the HILBERTIAN test of 3.13, we can test the hypothesis $C_0 : \langle\langle \text{The variables are independent} \rangle\rangle$ against its negation. Then, the classical chi-squared test is more naturally associated to the test of C'_0 whereas the HILBERTIAN test of 3.13 is more naturally associated to the test of C_0 .

3.16. Now we compare the two tests of independence: the classical chi-squared test and the HILBERTIAN one. In fact, taking into account 2.9, we have to consider two HILBERTIAN tests:

the classical chi-squared test which is the same as the HILBERTIAN test of $H_0(\mathcal{D}^*)$ against its negation,

and the HILBERTIAN test of independence which is the test of $H_0(\overline{\mathcal{D}})$ against its negation.

Then, it is clear that the most natural test of independence is the second one. Therefore, it is not surprising that this test is the most powerful.

3.17. The above mentioned remarks point out an idea which is the key of the restricted χ^2 test, i.e. that of testing only a part of parameters. Therefore these results provide a kind example of this test.

4. Independence test: the case of unknown marginal distributions

Bosq didn't study the HILBERTIAN test with estimation of the parameters.

However, it is easily seen, in this case, $\widehat{\chi^2_{X_1}} = \widehat{\chi^2_{X_2}} = 0$. From this, it follows that the more natural independence test is always given by a statistics of the type

$$\widehat{\chi^2_{X_1, X_2}} - \widehat{\chi^2_{X_1}} - \widehat{\chi^2_{X_2}} \quad \text{or} \quad \widehat{\chi^2_{X_1, X_2}} - \widehat{\chi^2_{X_1}} - \widehat{\chi^2_{X_2}},$$

and not $\widehat{\chi^2_{X_1, X_2}}$, which, under the independence hypothesis, has asymptotically a chi-squared distribution with $(c(1) - 1)(c(2) - 1)$ degrees of freedom.

4.1. Notations. Let $s \in (1, 2, \dots, p)$. We shall denote by $\widehat{\mu}_s$ the marginal empirical probability which has a constant probability density function on each interval $I_i^s \in \mathcal{P}^s$, and is such that $\widehat{\mu}_s(I_i^s) = \mu_s(I_i^s)$.

We shall assume $\mu_s(I_i^s) > 0$.

We shall also denote by $\widehat{\mu}$ the product measure $\widehat{\mu} = \widehat{\mu}_1 \otimes \widehat{\mu}_2 \otimes \dots \otimes \widehat{\mu}_p$, by $\{\widehat{A}_J\}$,

$J \in \mathfrak{D}$, the associated orthonormal functions, and by $\hat{\alpha}_J^n$ the statistics:

$$\hat{\alpha}_J^n = \frac{1}{n} \sum_{i=1}^n \hat{A}_J(X_i).$$

4.2. Now, we recall (see [13], p. 250) that, in the case of unknown marginal distributions, the chi-squared independence test is obtained from the chi-squared statistics by estimating the parameters:

$$\hat{\chi}_0^2 = \sum_{I \in \mathfrak{I}} \frac{(N_I - n\rho(I))^2}{n\rho(I)}.$$

This statistics asymptotically, and under the hypothesis of independence, has a chi-squared distribution with $\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right)$ degrees of freedom.

Proposition 4.3. *The chi-squared statistics of the marginal distributions which are given by*

$$\hat{\chi}_s^2 = \sum_{I \in \mathfrak{I}_s} \frac{(N_I - n\hat{\mu}(I))^2}{n\rho(I)}, \quad s \neq 0,$$

vanish identically.

The proof is obvious.

Proposition 4.4. *We have the following equalities:*

$$\hat{\chi}_0^2 = n \left(\sum_{J \in \mathfrak{D}^*} \hat{\alpha}_J^2 \right) \quad \text{and} \quad \hat{\chi}_s^2 = n \left(\sum_{J_0 \in \mathfrak{D}_m^s} \hat{\alpha}_{J_0}^2 \right) = 0.$$

The proof is like in 2.9, if we replace μ by $\hat{\mu}$.

From 4.3 and 4.4, we deduce the following proposition.

Proposition 4.5. *We have $\hat{\alpha}_{J_0}^2 = 0$, for all $J_0 \in \mathfrak{D}_m$. Another formulation of the above results is given by the following proposition.*

Proposition 4.6. *If we denote by $\|\hat{S}_n\|^2$ the HILBERTIAN statistics*

$$\|\hat{S}_n\|^2 = n \left(\sum_{J \in \mathfrak{D}} \hat{\alpha}_J^2 \right),$$

then, we have

$$\|\hat{S}_n\|^2 = \hat{\chi}_0^2 - \sum_{s=1}^p \hat{\chi}_s^2 = \hat{\chi}_0^2;$$

moreover $\|\hat{S}_n\|^2$ follows asymptotically, under the independence hypothesis, a chi-

squared distribution with

$$\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right)$$

degrees of freedom.

Now, we prove that the $\hat{\alpha}_j^n$'s are almost surely consistent estimators to the α_j 's. For this purpose, we shall need to impose some obvious constraints: the A_j 's and the \hat{A}_j 's must be constructed by the same method. Therefore, we shall prove our result in a particular case, but it is easily seen that the process is applicable in the general case. And anyway, the final result does not depend on this choice. The A_j 's are useful only for the proof.

4.7. Definition of a sequence of orthonormal functions. We shall define the sequence $\{B_j^s\}$, $j \in \mathfrak{D}^s$, for all $s \in (1, 2, \dots, p)$, with $B_j^s \in L^2(\mathbb{R}, \mu_s)$, in the following way: $B_1^s \equiv 1$, and

$$B_j^s(x) = \begin{cases} \dots\dots\dots 0 & \text{if } x \in \bigcup_{t=1}^{j-1} I_t^s, \\ \dots\dots \sum_{t=j+1}^{c(s)} \mu_s(I_t^s) & \text{if } x \in I_j^s, \\ \dots\dots\dots -(\mu_s(I_j^s)) & \text{if } x \in \bigcup_{t=j+1}^{c(s)} I_t^s, \end{cases}$$

if $j \in \mathfrak{D}^{*s}$.

Then, we shall denote by A_j^s the functions B_j^s normalized in $L^2(\mathbb{R}, \mu_s)$:

$$A_j^s = \frac{B_j^s}{\sqrt{\left(\sum_{t=j+1}^{c(s)} \mu_s(I_t^s) \right) (\mu_s(I_j^s)) \left(\sum_{t=j}^{c(s)} \mu_s(I_t^s) \right)}}.$$

Then, we shall define A_J , $J \in \mathfrak{D}$, as in 2.3, and in the same way, \hat{A}_J by interchanging μ_s by $\hat{\mu}_s$ in the above definitions. It is clear that these families verify 2.3 and 2.4.

Lemma 4.8. If we denote by

$$A_J = \sum_{I \in \mathfrak{J}} \lambda_I 1_I \quad \text{and} \quad \hat{A}_J = \sum_{I \in \mathfrak{J}} \hat{\lambda}_I 1_I,$$

the decompositions of A_J and \hat{A}_J with respect to the basis $\{1_I\}$, $I \in \mathfrak{J}$, then $\hat{\lambda}_I$ converges almost surely to λ_I .

Proof. This follows directly from the almost sure convergence of $\hat{\mu}_s(I_j^s)$ to $\mu_s(I_j^s)$ (see th 8.7, [9]).

As a simple consequence of 4.8, we have the following proposition.

Proposition 4.9. For all $J \in \mathcal{D}^*$, the estimator $\widehat{\alpha}_J^n$ almost surely converges to α_J .

4.10. The independence test. We know that if we test independence by the Hilbertian test, we test the hypothesis

$$H_0(\overline{\mathcal{D}}) : \langle \langle \forall J \in \overline{\mathcal{D}} : \alpha_J = 0 \rangle \rangle .$$

It is clear that the classical chi-squared test with estimate of parameters tests these coefficients and only them. In this case, the Hilbertian statistics of independence and the chi-squared statistics are equal: $\|\widehat{S}_n\|^2 = \widehat{\chi}_0^2$. Therefore the two tests are the same.

4.11. It is also clear that this last equality comes from the equalities $\widehat{\chi}_s^2 = 0$, $s \in (1, 2, \dots, p)$,

4.12. In this case, we have still the equality

$$\|\widehat{S}_n\|^2 = \widehat{\chi}_0^2 - \sum_{s=1}^p \widehat{\chi}_s^2 ,$$

and, under the independence hypothesis, $\|\widehat{S}_n\|^2$ follows asymptotically a chi-squared distribution with $\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right)$ degrees of freedom.

4.13. It is also obvious that, in the case of estimate of parameters, we cannot test the hypothesis

$$C'_0 : \langle \langle \text{The distribution of } X \text{ is } \mu \rangle \rangle .$$

5. Conclusions

We easily deduce from the above results the following conclusions.

5.1. Whether *all* the marginals distributions μ_s , $s = 1, 2, \dots, p$, are known or unknown, the more natural independence test is always based on a statistics of the type

$$\widehat{\chi}_0^2 - \sum_{s=1}^p \widehat{\chi}_s^2 \quad \text{or} \quad \widehat{\chi}_0^2 - \sum_{s=1}^p \widehat{\chi}_s^2 \quad \text{—and not } \widehat{\chi}_0^2 .$$

Under the independence hypothesis, these statistics always follow asymptotically a chi-squared distribution with

$$\left(\left(\prod_{s=1}^p c(s) - 1 \right) - \sum_{s=1}^p (c(s) - 1) \right)$$

degrees of freedom $((c(1) - 1) (c(2) - 1) \dots (c(p) - 1))$ degrees of freedom if $p = 2$.

5.2. The classical statistics χ_0^2 is not the more natural statistics to test independence, because, in the case of known marginal distributions, it leads us to test already known coefficients: the coefficients of the marginal distributions. Therefore, in this case, we test again the marginal distribution. In fact, this statistics is more naturally associated to the test of the hypothesis C'_0 : «the distribution of X is μ » against its negation whereas the hilbertian statistics test the hypothesis C_0 : «The variables are independent». Moreover, if the marginal distributions are unknown, the HILBERTIAN test and the classical chi-squared test are the same.

5.3. We have obtained the preceding results when *all* the marginal distributions are known, or when *all* the marginal distributions are unknown. But, in consideration of their identity, we can hope that they will remain valid if q marginal distributions are known and $p - q$ unknown. It seems possible that asymptotic laws for orthogonal empirical functions could give this result [5].

5.4. We must also remark that our results give us in a natural way the differences of the degrees of freedom between the several statistics that we have used.

In fact, the degree of freedom is the number of parameters that we really test; the estimate on parameters is expressed by the vanishing of the coefficients $\widehat{\alpha}_{J_0}^n$, $J_0 \in \mathcal{D}_m$. That comes not to test the corresponding α_{J_0} . In fact, the estimation of parameters look like a kind of orthogonal projection. One can ask if it is not possible to generalize this method to obtain a new way to study the effect of the estimate of parameters in the chi-squared distributions. May be we could also obtain this result with asymptotic laws for orthogonal empirical functions. Perhaps we could also solve by this method the case where we have to estimate only some of the parameters of marginal laws. Therefore, in the classical chi-squared independence tests, the difference of the degrees of freedom is fictitious. We don't test the same hypothesis in the two cases: we test C'_0 or C_0 (5.2).

References

- [1] BARUCHA-REID, A. T. (1972). *Random integral equations*. Academic Press, New York.
- [2] BICKEL, P. J. and ROSENBLATT, J. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* 1, 1071–1095.
- [3] BILLINGSLEY, P. (1968). *Convergence of probability measure*, Wiley, New York.
- [4] BLACHER, R. (1983). Indicateurs de dépendance entre deux variables aléatoires fournis par le développement en série de la densité de probabilité. Thèse 3° Cycle, Université Scientifique et Médicale de Grenoble.
- [5] BLACHER, R. (1987). Variances et coefficients de corrélation d'ordre supérieur: suite TIM3-IMAG R.R. 655 M, Université Scientifique et Médicale de Grenoble.
- [6] BOSQ, D. (1978). Tests hilbertiens et test du khi-deux. *C. R. Acad. Paris*, t 286, 946–948.
- [7] BOSQ, D. (1983). Lois limites et efficacité asymptotique des tests Hilbertiens de

- dimension finie sous des hypothèses adjacentes. *Statistiques et analyse des données*, vol 8, n° 1.
- [8] FIX, E., HODGES, J. L. and LEHMANN, E. L. (1959). The restricted chi square test. *Probability and statistics*. The HARALD CRAMÉR volume. Almqvist and Wiksell, Stockholm.
- [9] FOURGEAUD, C. and FUCHS, A. (1967). *Statistique*. Dunod, Paris.
- [10] GERLEIN, O. V. and PINCUS, R. (1978). Goodness of fit tests based on a measure of discrepancy between two scalar products. *Theory of probability and its applications*. 23, 248-258.
- [11] GREGORY, G. (1977). Large sample for U Statistics and tests of fit. *Ann. Statist.* 5, 110-123.
- [12] GREGORY, G. (1980). On efficiency and optimality of quadratic tests of fit. *Ann. Statist.* 8, 116-131.
- [13] HOEL, P. G. (1962). *Introduction to mathematical statistics*. John Wiley and Sons Inc. New-York, London.
- [14] KACAN, A. M. (1975). On a measure of divergency between two scalar products and its statistical application. *SANKHYA* 37, 211-228.
- [15] LEHMANN, E. L. (1959). *Testing statistical hypotheses*. Wiley, New York.

Received October 1985; revised July 1987.

R. BLACHER
 Université Scientifique & Médicale de Grenoble
 I. R. M. A., B. P. 68
 38402 St. Martin d'Hères
 France

DOCUMENT III

TEST HILBERTIEN

D'INDEPENDANCE

TEST HILBERTIEN D'INDEPENDANCE

Mots-clefs : Fonctions orthogonales, coefficients de corrélation, test Hilbertien, test du chi deux.

Résumé : On étudie le test Hilbertien d'indépendance, avec ou sans estimation de paramètres, et on voit que ce test généralise le test du chi-deux correspondant. Pour obtenir ces résultats, on utilise la loi asymptotique des fonctions orthogonales empiriques et on en déduit la loi des coefficients de corrélation d'ordre supérieur empiriques.

Summary : We study an Hilbertian independence test which generalizes the chi square independence test. We build this test in the cases where the marginal distributions are known or unknown. For obtain this result, we use the asymptotic distribution of orthogonal empirical functions, and we deduce the distribution of the empirical correlation coefficients of superior order.

INTRODUCTION

Lorsque l'on veut tester si la loi d'un vecteur aléatoire U est une loi M connue, on emploie souvent le test du chi deux. On peut aussi généraliser cette méthode en utilisant le Test Hilbertien dont le test du chi deux est un cas particulier.

Rappelons que le Test Hilbertien teste les coefficients de fonctions orthogonales associées à M . Par exemple, pour tester que deux variables aléatoires U et V sont indépendantes, on pourra choisir comme fonctions orthogonales les fonctions $R_i(u)T_j(v)$, où $\{R_i\}$ et $\{T_j\}$ représentent des familles orthonormales associées aux lois marginales m et m' et telles que $R_0=T_0 \equiv 1$. En effet, on sait que si (U, V) admet une densité $f \in L^2$, par rapport à la mesure produit $m \otimes m'$, alors, $f(u, v) = \sum_{i,j} \rho_{i,j} R_i(u) T_j(v)$ avec $\rho_{i,j} = \mathbb{E} [R_i(U) T_j(V)]$, l'espérance de $R_i(U) T_j(V)$. Comme $\rho_{0,0} = 1$ et $\rho_{i,0} = \rho_{0,j} = 0$ si $i \neq 0$ et $j \neq 0$, on trouvera finalement $f(u, v) = 1 + \sum_{i>0, j>0} \rho_{i,j} R_i(u) T_j(v)$. On en déduit que U et V seront indépendantes si et seulement si tous les $\rho_{i,j}$ sont nuls pour $i > 0$ et $j > 0$. Maintenant, le test Hilbertien permet de tester la valeur des $\rho_{i,j}$. Aussi, le test d'indépendance Hilbertien testera la nullité d'un nombre fini de $\rho_{i,j}$.

En fait, les $\rho_{i,j}$ forment une suite de coefficients mesurant la dépendance que l'on peut appeler coefficients de corrélation d'ordre supérieur. En effet, ils généralisent et complètent le coefficient de corrélation linéaire classique. Par exemple, si $\{R_i\}$ et $\{T_j\}$ représentent les familles de polynômes orthonormaux associés à m et m' , alors, les $\rho_{i,j}$ mesurent des dépendances polynomiales, et $\rho_{1,1}$ n'est autre que le coefficient de corrélation linéaire.

Maintenant, pour définir $\|S_n^*\|^2$, la statistique Hilbertienne, à partir d'un échantillon

$$\{(U^r, V^r)\}_{r=1, \dots, n}, \text{ on pose } \rho_{i,j}^{*n} = \frac{\sum_{r=1}^n R_i(U^r) T_j(V^r)}{n} \text{ et } \|S_n^*\|^2 = \sum_{i=1}^h \sum_{j=1}^k n (\rho_{i,j}^{*n})^2.$$

Alors, si U et V sont indépendants, $\|S_n^*\|^2$ suit asymptotiquement une loi du chi-deux à hk degrés de liberté, tandis que si un des $\rho_{i,j}$ testé est différent de 0, $\|S_n^*\|^2$ converge presque sûrement vers l'infini.

On voit donc bien que le test Hilbertien est une généralisation du test du chi-deux. D'une façon plus précise, soient $\{J_i\}_{i=1, \dots, c}$ et $\{A_j\}_{j=1, \dots, d}$, deux partitions de \mathbb{R} , et soient 1_{J_i} et 1_{A_j} , les fonctions indicatrices d'ensemble associées. On orthogonalise alors les familles $\{1_{J_i}\}$ et $\{1_{A_j}\}$ et on note toujours par $\{R_i\}_{i=0, \dots, c-1}$ et $\{T_j\}_{j=0, \dots, d-1}$ les familles orthonormales ainsi obtenues (en imposant encore $R_0=T_0 \equiv 1$).

Alors, on peut prouver que $\chi^2_{U,V}$, la statistique du chi-deux associée à la partition $\{J_i \times J_j\}$ de \mathbb{R}^2 , i.e. la statistique du test d'indépendance du chi-deux dans le cas de lois marginales

connues, est identique à la statistique Hilbertienne $n \sum_{i=0}^{c-1} \sum_{\substack{j=0 \\ (i,j) \neq (0,0)}}^{d-1} (\rho^*_{i,j})^2$.

De plus, si l'on note par χ^2_U et χ^2_V , les statistiques du chi-deux des lois marginales, alors $n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\rho^*_{i,j})^2 = \chi^2_{U,V} - \chi^2_U - \chi^2_V$.

Finalement, on constate que le test d'indépendance du chi-deux teste de nouveau la nullité des $\rho_{i,0}$ et $\rho_{0,j}$ que l'on sait déjà être nuls, i.e. il teste de nouveau que les lois marginales sont bien m et m' . Cela permet de remarquer que, plutôt qu'un test d'indépendance, le test associé à $\chi^2_{U,V}$ est, en fait, le test de l'hypothèse «la loi de (U,V) est $m \otimes m'$ ».

Il est donc normal que le test Hilbertien qui, lui, ne teste que l'indépendance, soit plus puissant que le test du chi-deux classique. On prouve ce résultat en montrant que $\chi^2_{U,V} - \chi^2_U - \chi^2_V$ n'est autre qu'une statistique du chi-deux restreint.

On vient donc de s'intéresser au test Hilbertien d'indépendance dans le cas de lois marginales connues. Maintenant, dans le cas où on ne connaît pas les lois marginales, il faudra les estimer. Pour cela, on aura recours aux fonctions orthogonales empiriques.

Par exemple, dans le cas polynômial, on obtient les polynômes orthogonaux empiriques R_i^n et T_j^n en remplaçant les moments par les moments empiriques. Ainsi, $R_1(u) = \frac{u - \mathbb{E}(U)}{\sigma_U}$ et R_1^n

$= \frac{u - \mathbb{E}_n(U)}{\sigma_n(U)}$, où $\mathbb{E}_n(U)$ représente l'espérance empirique, et σ_U et $\sigma_n(U)$, l'écart-type et l'écart-type empirique. On définit alors le coefficient de corrélation empirique d'ordre (i,j) par :

$\rho_{i,j}^n = \frac{\sum_{j'=1}^n R_i^n(U^{j'}) T_j^n(V^{j'})}{n}$. En particulier, $\rho_{1,1}^n$ n'est autre que le coefficient de corrélation linéaire

empirique classique.

Maintenant, pour obtenir un test Hilbertien d'indépendance, on utilisera la loi asymptotique des polynômes orthogonaux empiriques : on écrit R_i^n sous la forme $R_i^n = R_i + \sum_{s=0}^i \alpha_{i,s}^n R_s$, et on montre

que $\sqrt{n} \alpha_{i,s}^n$ suit asymptotiquement la même loi que $\frac{\sum_{j'=1}^n R_i(U^{j'}) R_s(U^{j'})}{\sqrt{n}}$.

En utilisant, la forme géométrique de ce résultat, on obtient alors la loi asymptotique des $\rho_{i,j}^n$ et on en déduit un test Hilbertien d'indépendance avec estimation de paramètres.

Dans cet article, on étendra ces résultats à d'autres fonctions orthogonales que les polynômes orthogonaux empiriques. En particulier, on verra que dans le cas du test du chi-deux, $\sum_{i=1}^h \sum_{j=1}^k n(\rho_{i,j}^n)^2$ est identique à la statistique du chi-deux avec estimation de paramètres.

Signalons enfin que les démonstrations de cet article sont effectuées de manière détaillée dans le rapport TIM3-IMAG 792 .

I : LOI ASYMPTOTIQUE DES COEFFICIENTS DE CORRELATION D'ORDRE SUPERIEUR EMPIRIQUES

Dans ce paragraphe, on va d'abord rappeler la définition et quelques propriétés élémentaires des coefficients de corrélation d'ordre supérieur. Ensuite, on montrera comment on peut obtenir des coefficients de corrélation d'ordre supérieur empiriques en généralisant la méthode de construction des polynômes orthogonaux empiriques, et on calculera alors leurs lois asymptotiques .

On se placera donc sous les hypothèses suivantes.

Notations 1-1 : On considère ${}^t(U, {}^tV)$ un vecteur aléatoire à valeurs dans \mathbb{R}^{p+q} , $U = (U_1, U_2, \dots, U_p)$, $V = (V_1, V_2, \dots, V_q)$. On suppose que ${}^t(U, {}^tV)$, U et V , sont de lois, respectivement, \mathcal{Q} , m et m' . Soient alors $\{R_i\}_{i=0,1,\dots}$ et $\{T_j\}_{j=0,1,\dots}$ deux familles de fonctions orthonormales de, respectivement, $L^2(\mathbb{R}^p, m)$ et $L^2(\mathbb{R}^q, m')$, vérifiant $R_0 = T_0 \equiv 1$.

On complète alors le coefficient de corrélation classique en introduisant les coefficients de corrélation d'ordre supérieur de la façon suivante.

Notations 1-2 : Pour chaque (i, j) , on pose $\rho_{i,j} = \mathbb{E} [R_i(U)T_j(V)]$. De plus, pour tout $i > 0$ et $j > 0$, on appelle coefficient de corrélation d'ordre (i, j) , le réel $\rho_{i,j}$.

Signalons que ces coefficients ont été étudiés par Lancaster dans le cadre des corrélations canoniques dès 1958 (pour leurs propriétés, voir aussi Blacher 1984 a et b).

En particulier, Lancaster avait montré que la connaissance des lois marginales et de tous les $\rho_{i,j}$ déterminent complètement \mathcal{Q} (th 1-1, p 120, Lancaster). En fait, on peut préciser cette propriété de la manière suivante.

Théorème 1-3 : On suppose que $\{R_i\}_{i=0,1,\dots}$ et $\{T_j\}_{j=0,1,\dots}$ représentent deux bases orthonormales de, respectivement, $L^2(\mathbb{R}^p, m)$ et $L^2(\mathbb{R}^q, m')$. Alors, en tout ${}^t(u, {}^tv) \in \mathbb{R}^{p+q}$, ${}^t u = (u_1, u_2, \dots, u_p)$, ${}^t v = (v_1, v_2, \dots, v_q)$, on peut écrire la fonction de répartition $F_{U,V}$ de ${}^t(U, {}^tV)$, sous la forme :

$$F_{U,V}(u, v) = F_U(u)F_V(v) + \sum_{i>0, j>0} \rho_{i,j} \left(\int_{t=-\infty}^{t=u} R_i(t).m(dt) \right) \left(\int_{w=-\infty}^{w=v} T_j(w).m'(dw) \right),$$

où F_U représente la fonction de répartition de U , et où $\int_{t=-\infty}^{t=u} R_i(t).m(dt) = \int \mathbb{1}_{]-\infty, u]}(t) R_i(t).m(dt)$,

$\mathbb{1}_{]-\infty, u]}$ représentant la fonction indicatrice de $]-\infty, u_1] \times]-\infty, u_2] \times \dots \times]-\infty, u_p]$.

En effet, on peut écrire $\mathbb{1}_{]-\infty, u]} = \sum_i \alpha_i R_i$ dans $L^2(\mathbb{R}^p, m)$ et $\mathbb{1}_{]-\infty, v]} = \sum_j \beta_j T_j$ dans $L^2(\mathbb{R}^q, m')$,

et donc, $\mathbb{1}_{]-\infty, u]}(t) = \sum_i \alpha_i R_i(t)T_0(w)$ et $\mathbb{1}_{]-\infty, v]}(w) = \sum_j \beta_j R_0(t)T_j(w)$ dans $L^2(\mathbb{R}^{p+q}, \mathcal{Q})$.

Alors, d'après la continuité du produit scalaire,
$$\iint \mathbb{1}_{]-\infty, u]}(t) \mathbb{1}_{]-\infty, v]}(w) \mathcal{Q}(dt, dw) = \sum_{(i,j)} \alpha_i \beta_j \rho_{i,j}.$$

Comme la connaissance des $\rho_{i,j}$ définit complètement la loi de (U, V) , il n'est pas étonnant qu'ils permettent de détecter divers types de dépendance fonctionnelle entre U et V . En particulier, on peut énoncer le théorème suivant.

Théorème 1-4 : On suppose $q=1$, $T_1(v) = \frac{v - \mathbb{E}(V)}{\sigma_V}$ et que $\{R_i\}$ représente une base de $L^2(\mathbb{R}^p, m)$.

Alors, $\sum_{i>0} \rho_{i,1}^2 \leq 1$, l'égalité ayant lieu si et seulement si il existe $g \in L^2(\mathbb{R}^p, m)$ tel que $V = g(U)$ p.s.

Démonstration : Notons par $\|\cdot\|$ la norme dans $L^2(\mathbb{R}^{p+q}, \mathcal{Q})$ et par $\mathcal{U}^p(T_1 | R_1, R_2, \dots)$ la projection orthogonale de $(u, v) \mapsto T_1(v)$ sur le sous-espace de $L^2(\mathbb{R}^{p+q}, \mathcal{Q})$ engendré par les $(u, v) \mapsto R_i(u)$, $i > 0$. Alors, d'après les propriétés élémentaires des projections orthogonales, $\|\mathcal{U}^p(T_1 | R_1, R_2, \dots)\|^2 \leq \|T_1\|^2 = 1$, l'égalité ayant lieu si et seulement si $\mathcal{U}^p(T_1 | R_1, R_2, \dots)(u) = T_1(v)$.

Il suffit alors de rappeler que $\mathcal{U}^p(T_1 | R_1, R_2, \dots) = \sum_{i>0} \rho_{i,1} R_i$.

D'autre part, on peut aussi facilement déduire du théorème 1-3 que, si (U, V) admet une densité f par rapport à $m \otimes m'$, alors $f(u, v) = 1 + \sum_{i>0, j>0} \rho_{i,j} R_i(u) T_j(v)$ dans $L^2(\mathbb{R}^{p+q}, m \otimes m')$

(voir aussi Lancaster th 1-3 ou Blacher 1984-b th 1-3). Enfin, il est clair que U et V seront indépendants si et seulement si tous les $\rho_{i,j}$ sont nuls (voir aussi Lancaster, cor 1.1.1). Donc, pour tester l'indépendance de U et V , il faudra tester la nullité des $\rho_{i,j}$.

On va introduire maintenant les différents estimateurs des $\rho_{i,j}$. Pour cela, on adoptera les notations suivantes.

Notations 1-5 : Soit $\{(U^j, V^j)\}_{j \in \mathbb{N}}$, $U^j = (U_1^j, U_2^j, \dots, U_p^j)$, $V^j = (V_1^j, V_2^j, \dots, V_q^j)$, une suite de vecteurs aléatoires indépendants et de même loi que (U, V) . Pour tout $n \in \mathbb{N}^*$, on note par m_n la probabilité empirique associée à l'échantillon $\{U^j\}_{j=1,2,\dots,n}$:

$$m_n(F) = \frac{\text{card}\{1 \leq j \leq n \mid U^j \in F\}}{n}.$$

On définit de la même façon \mathcal{Q}_n et m'_n , les probabilités empiriques associées aux échantillons, respectivement, $\{(U^j, V^j)\}_{j=1,2,\dots,n}$ et $\{V^j\}_{j=1,2,\dots,n}$.

Alors, dans le cas où l'on connaît les lois marginales, on retrouve le test Hilbertien de Bosq.

Théorème 1-6 : Soient h et $k \in \mathbb{N}^*$. Pour tout $(i,j) \in \{0,1,\dots,h\} \times \{0,1,\dots,k\}$, on pose $\rho^{*n}_{ij} = \int R_i(u)T_j(v)\mathcal{Q}_n(du,dv)$, et on suppose $\mathbb{E}[R_i(U)^2T_j(V)^2] < +\infty$.

Alors, $\sqrt{n}\{\rho^{*n}_{ij} - \rho_{ij}\}_{(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}}$ suit asymptotiquement la loi normale centrée de matrice de covariances $\left\{ \mathbb{E}[R_i(U)R_j(U)T_j(V)T_j(V)^2] \right\}_{((i,j),(i',j')) \in (\{1,\dots,h\} \times \{1,\dots,k\})^2}$.

Pour obtenir ce résultat, il suffit d'appliquer le Théorème de la limite Centrale. Maintenant, si l'on veut tester un nombre infini de ρ_{ij} , on peut se reporter à l'étude de Bosq.

Les ρ^{*n}_{ij} sont donc des estimateurs des ρ_{ij} que l'on peut éventuellement appeler coefficients de corrélation empiriques d'ordre (i,j) de deuxième espèce. En effet, les ρ^{*n}_{ij} correspondent au cas où l'on connaît les lois marginales, et, en général, il existe d'autres estimateurs pour le cas où il faut les estimer. Pour obtenir ceux-ci, on va devoir généraliser la méthode de construction des polynômes orthogonaux empiriques. Aussi, on va d'abord rappeler comment on construit ces derniers.

On suppose donc momentanément que U est une variable aléatoire, et que $\{P_i\}_{i=0,1,\dots,h}$ représente la famille des polynômes orthonormaux associés de degré inférieur ou égal à h . Il est bien connu que les P_i sont entièrement déterminés par les moments (cf Chihara). Aussi, pour obtenir les polynômes orthogonaux empiriques P^n_i , il suffit de remplacer les moments par les moments empiriques (cf Blacher 1984-a). En fait, on obtient les P_i et les P^n_i , en orthogonalisant $1, u, u^2, \dots, u^h$ selon le procédé de Gram-Schmidt, dans $L^2(\mathbb{R}, m)$ ou $L^2(\mathbb{R}, m_n)$.

Maintenant, supposons de nouveau que U est un vecteur aléatoire à valeurs dans \mathbb{R}^p . Le problème est donc de généraliser la construction des polynômes orthogonaux empiriques et d'obtenir des estimateurs des R_i . Alors, on peut résoudre cette question quand $\{R_i\}$ est construite selon le procédé de Gram-Schmidt.

Hypothèses 1-7 : Soient $\gamma : \mathbb{R}^p \rightarrow \mathbb{R}^{h+1}$ et $\tau : \mathbb{R}^q \rightarrow \mathbb{R}^{k+1}$, deux applications mesurables. On définit les vecteurs aléatoires X et Y par : $X = {}^t(X_0, X_1, \dots, X_h) = \gamma(U)$, et $Y = {}^t(Y_0, Y_1, \dots, Y_k) = \tau(V)$, et on suppose $\mathbb{E}[(X_s)^2] < +\infty$ et $\mathbb{E}[(Y_t)^2] < +\infty$ pour tout $(s,t) \in \{0,1,\dots,h\} \times \{0,1,\dots,k\}$.

On note par μ la probabilité de X , par \langle, \rangle et $\| \cdot \|$, le produit scalaire et la norme de $L^2(\mathbb{R}^{h+1}, \mu)$, et, pour tout ${}^t x = (x_0, x_1, \dots, x_h) \in \mathbb{R}^{h+1}$, on pose $\bar{A}_0(x) = A_0(x) \equiv 1$,

$$\text{et, pour } h \geq i > 0, \quad \bar{A}_i(x) = x_i - \sum_{s=0}^{i-1} \langle x_i, A_s(x) \rangle A_s(x) \quad \text{et} \quad A_i(x) = \frac{\bar{A}_i(x)}{\|\bar{A}_i\|}.$$

On définit de la même façon les fonctions orthonormales $\{B_j\}_{j \in \{0,1,\dots,k\}}$ associées à Y .

On suppose alors que, pour tout $(i,j) \in \{0,1,\dots,h\} \times \{0,1,\dots,k\}$, $R_i(u) = A_i[\gamma(u)]$ et $T_j(v) = B_j[\tau(v)]$.

Par exemple, si $\gamma(U) = {}^t(1, U, U^2, U^3, \dots, U^h)$, i.e $X_s = U^s$, $\{R_i\}$ est la famille des polynômes orthonormaux.

Remarquons que X est de rang $h+1$, Y de rang $k+1$, et que $X_0=Y_0 \equiv 1$ dans L^2 .

Alors, sous ces hypothèses, on peut définir des fonctions orthogonales empiriques A_i^n , estimateurs des A_i .

Notations 1-8 : On définit la suite $\{({}^tX^j, {}^tY^j)\}_{j \in \mathbb{N}}$ par $X^j = (X_0^j, X_1^j, \dots, X_h^j) = \gamma(U^j)$ et $Y^j = (Y_0^j, Y_1^j, \dots, Y_k^j) = \tau(V^j)$. Alors, pour tout $n \in \mathbb{N}^*$, on note par μ_n la mesure empirique sur \mathbb{R}^{h+1} , associée à $(X^j)_{j=1,2,\dots,n}$, et par $\langle \cdot, \cdot \rangle_n$ et $\| \cdot \|_n$, le produit scalaire et la norme de $L^2(\mathbb{R}^{h+1}, \mu_n)$.

Alors, pour tout ${}^t x = (x_0, x_1, \dots, x_h) \in \mathbb{R}^{h+1}$, on pose $\bar{A}_0^n(x) = A_0^n(x) \equiv 1$,

$$\text{et, pour } h \geq i > 0, \bar{A}_i^n(x) = x_i - \sum_{s=0}^{i-1} \langle x_i, A_s^n(x) \rangle_n A_s^n(x) \quad \text{et} \quad A_i^n(x) = \begin{cases} \frac{\bar{A}_i^n(x)}{\|\bar{A}_i^n\|_n}, & \text{si } \|\bar{A}_i^n\|_n \neq 0, \\ 0, & \text{si } \|\bar{A}_i^n\|_n = 0. \end{cases}$$

On définit de la même façon les fonctions orthogonales $\{B_j^n\}_{j \in \{0,1,\dots,k\}}$ associées à $(Y^j)_{j=1,2,\dots,n}$.

Maintenant, on obtient également des fonctions orthogonales empiriques R_i^n qui sont des estimateurs des R_i .

Fonctions orthogonales empiriques 1-9 : Pour tout $(i,j) \in \{0,1,\dots,h\} \times \{0,1,\dots,k\}$, on définit R_i^n et T_j^n par : $R_i^n(u) = A_i^n[\gamma(u)]$ et $T_j^n(v) = B_j^n[\tau(v)]$.

De cette manière, on généralise bien les polynômes orthogonaux empiriques. Par exemple, si l'on veut des polynômes orthogonaux à 2 variables associés à un vecteur aléatoire (U,W) à valeurs dans \mathbb{R}^2 , il suffit de poser $\gamma(U,W) = (1, U, U^2, \dots, U^r), (W, WU, WU^2, \dots, WU^r), \dots, (W^s, W^s U, W^s U^2, \dots, W^s U^r)$.

On peut maintenant définir des coefficients de corrélation empiriques d'ordre supérieur.

Notations 1-10 : Pour tout $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$ on définit $\rho_{i,j}^n$ le coefficient de corrélation empirique d'ordre (i,j) par $\rho_{i,j}^n = \int R_i^n(u) T_j^n(v) \mathcal{Q}_n(du, dv)$.

Remarquons que l'on peut poser aussi $\rho_{0,0}^n = 1$ et $\rho_{i,0}^n = \rho_{0,j}^n = 0$ si $i \neq 0$ et $j \neq 0$.

Maintenant, pour obtenir la loi de ces coefficients de corrélation empiriques, on aura besoin d'écrire les fonctions orthonormales empiriques sous une forme géométrique.

Notations 1-11 : Pour tout $(i,s) \in \{0,1,\dots,h\}^2$, et tout $(j,t) \in \{0,1,\dots,k\}^2$ on définit $\alpha_{i,s}^n$ et $\beta_{j,t}^n$ par : $A_i^n = A_i + \sum_{s=0}^h \alpha_{i,s}^n A_s$ et $B_j^n = B_j + \sum_{t=0}^k \beta_{j,t}^n B_t$.

De plus, pour écrire de façon simple la loi asymptotique de $\alpha^n_{i,s}$ et $\beta^n_{j,t}$, on utilisera les notations classiques suivantes

Notations 1-12 : Soient Z_n et Z^*_n deux suites de variables aléatoires définies sur le même espace probabilisé. On conserve la notation classique : on notera $Z_n = Z^*_n + o_p(n^{-1/2})$ si $\sqrt{n}(Z_n - Z^*_n)$ converge en probabilité vers 0.

Le théorème suivant nous donne alors la loi des fonctions orthogonales empiriques.

Théorème 1-13 : Pour tout $(s,t) \in \{0,1,\dots,h\}^2$, on définit les variables aléatoires $e_{s,t}$ par : $e_{s,t} = -\int A_s A_t d\mu_n$ si $s > t$, $e_{t,t} = [\int (1 - [A_t]^2) d\mu_n] / 2$, $e_{s,t} \equiv 0$ si $s < t$.

On suppose que $\mathbb{E}[(X_s)^4] < +\infty$ pour tout $s \in \{0,1,\dots,h\}$. Alors, pour tout $(s,t) \in \{0,1,\dots,h\}^2$, $\alpha_{s,t} = e_{s,t} + o_p(n^{-1/2})$.

C'est le théorème 1 de Blacher 1989-a. Pour une démonstration détaillée, on peut aussi se reporter à Blacher 1989-b.

Maintenant, on obtiendra la loi asymptotique des coefficients de corrélation empiriques sous une forme relativement simple grâce à l'introduction des projections orthogonales.

Théorème 1-14 : Soit $(i,j) \in \{0,1,\dots,h\} \times \{1,2,\dots,k\}$. On note alors par $\mathcal{P}(T_j | R_0, R_1, \dots, R_{i-1})$ la projection orthogonale de $(u,v) \mapsto T_j(v)$ sur le sous espace de $L^2(\mathbb{R}^{p+q}, \mathcal{Q})$ engendré par $R_0(u), R_1(u), \dots, R_{i-1}(u)$, et on note par $\mathcal{O}(T_j | R_0, R_1, \dots, R_{i-1})$ le vecteur de projection orthogonale : $\mathcal{O}(T_j | R_0, R_1, \dots, R_{i-1})(u,v) = T_j(v) - \mathcal{P}(T_j | R_0, R_1, \dots, R_{i-1})(u)$. De plus, on étendra de façon naturelle ces notations à d'autres projections orthogonales que celles-ci.

Soit alors $C_{i,j}(u,v) =$

$$\frac{T_j(v)\mathcal{O}[R_i | T_j](u,v) + R_i(u)\mathcal{O}[T_j | R_i](u,v)}{2}$$

$$- T_j(v)\mathcal{P}[R_i | T_0, T_1, \dots, T_{j-1}](v) - R_i(u)\mathcal{P}[T_j | R_0, R_1, \dots, R_{i-1}](u) .$$

On suppose que, pour tout $(s,t) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$, $\mathbb{E}[R_s(U)^4] < +\infty$ et $\mathbb{E}[T_t(V)^4] < +\infty$.

$$\text{Alors, } \rho^n_{i,j} - \rho_{i,j} = \int C_{i,j} d\mathcal{Q}_n + o_p(n^{-1/2}) .$$

Rappelons que $\mathcal{O}(T_j | R_0, R_1, \dots, R_{i-1})$ est orthogonal à R_0, R_1, \dots et R_{i-1} , et remarquons aussi que que $\mathcal{O}[T_j | R_i](u,v) = T_j(v) - \rho_{i,j} R_i(u)$.

Démonstration : Il est clair que $R_i^n = R_i + \sum_{s=0}^h \alpha^n_{i,s} R_s$ et $T_j^n = T_j + \sum_{t=0}^k \beta^n_{j,t} T_t$.

Comme, d'après la loi Forte de Grands Nombres, $\int R_s T_t d\mathcal{Q}_n$ converge presque sûrement vers

$\rho_{s,t}$, et comme d'après le Théorème de la Limite Centrale et 1-13, $\sqrt{n}\alpha_{i,s}^n$ suit asymptotiquement une loi normale, on peut écrire les égalités suivantes

$$\begin{aligned}
 \rho_{i,j}^n - \rho_{i,j} &= \\
 &\int R_i(u)T_j(v) \cdot \mathcal{Q}_n(du,dv) + \sum_{s=0}^i \alpha_{i,s}^n \int R_s(u)T_j(v) \cdot \mathcal{Q}_n(du,dv) + \sum_{t=0}^j \beta_{j,t}^n \int R_i(u)T_t(v) \cdot \mathcal{Q}_n(du,dv) - \rho_{i,j} \\
 &\quad + o_p(n^{-1/2}) \\
 &= \\
 &\int R_i(u)T_j(v) \cdot \mathcal{Q}_n(du,dv) - \rho_{i,j} - \rho_{i,j} \frac{\int R_i(u)R_i(u) \cdot m_n(du) - 1}{2} - \rho_{i,j} \frac{\int T_j(v)T_j(v) \cdot m'_n(dv) - 1}{2} \\
 &- \sum_{s=0}^{i-1} \mathbb{E}[R_s(U)T_j(V)] \int R_i(u)R_s(u) \cdot m_n(du) - \sum_{t=0}^{j-1} \mathbb{E}[R_i(U)T_t(V)] \int T_j(v)T_t(v) \cdot m'_n(dv) \\
 &\quad + o_p(n^{-1/2}) \\
 &= \\
 &\frac{\int R_i(u)[T_j(v) - \rho_{i,j}R_i(u)] \cdot \mathcal{Q}_n(du,dv)}{2} + \frac{\int T_j(v)[R_i(u) - \rho_{i,j}T_j(v)] \cdot \mathcal{Q}_n(du,dv)}{2} \\
 &- \int \left[\sum_{s=0}^{i-1} \mathbb{E}[R_s(U)T_j(V)]R_s(u) \right] R_i(u) \cdot m_n(du) - \int \left[\sum_{t=0}^{j-1} \mathbb{E}[R_i(U)T_t(V)]T_t(v) \right] T_j(v) \cdot m'_n(dv) \\
 &\quad + o_p(n^{-1/2}) \\
 &= \int C_{i,j} d\mathcal{Q}_n + o_p(n^{-1/2}) .
 \end{aligned}$$

On déduit de ce théorème la loi asymptotique des coefficients de corrélation empiriques

Théorème 1-15 : On garde les hypothèses de 1-14. Alors, $\sqrt{n}\{\rho_{i,j}^n - \rho_{i,j}\}_{(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}}$ suit asymptotiquement la loi normale centrée de matrice de variances-covariances $\left\{ \mathbb{E}[C_{i,j}(U,V)C_{i',j'}(U,V)] \right\}_{((i,j),(i',j')) \in \{1,\dots,h\} \times \{1,\dots,k\}^2}$.

En effet, les diverses orthogonalités existantes montrent que $C_{i,j}(U,V)$ est centrée. De plus, on déduit de l'inégalité de Holder que $\mathbb{E}[C_{i,j}(U,V)^2] < +\infty$.

Il ne reste plus qu'à appliquer le Théorème de la Limite Centrale pour obtenir 1-15.

Remarquons que si $p=q=1$, $R_1(u) = \frac{u - \mathbb{E}(U)}{\sigma(U)}$ et $T_1(v) = \frac{v - \mathbb{E}(V)}{\sigma(V)}$, c'est à dire dans le cas polynômial, on retrouve la loi du coefficient de corrélation linéaire empirique (cf Serfling p 126).

2 TEST HILBERTIEN D'INDEPENDANCE

Grâce aux théorèmes 1-6 et 1-15, on peut définir un test Hilbertien d'indépendance. On va alors remarquer que le test d'indépendance du chi-deux en est un cas particulier.

Dans le cas de lois marginales connues, on peut énoncer le théorème suivant, corollaire de 1-6.

Théorème 2-1 : On garde les hypothèses de 1-6, et on pose $\|S_n^*\|^2 = n \left\{ \sum_{i=1}^h \sum_{j=1}^k (e_{i,j}^{*n})^2 \right\}$.

Alors, si U et V sont indépendants, $\|S_n^*\|^2$ suit asymptotiquement une loi du chi-deux à hk degrés de liberté.

D'autre part, s'il existe $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$ tel que $e_{i,j} \neq 0$, alors $\|S_n^*\|^2$ converge presque sûrement vers l'infini.

Dans le cas de lois marginales inconnues, on déduit du théorème 1-15 un test d'indépendance Hilbertien avec estimation de paramètres.

Théorème 2-2 : On garde les hypothèses de 1-15, et on pose $\|S_n\|^2 = n \left\{ \sum_{i=1}^h \sum_{j=1}^k (e_{i,j}^n)^2 \right\}$.

Alors, si U et V sont indépendants, $\|S_n\|^2$ suit asymptotiquement une loi du chi-deux à hk degrés de liberté.

D'autre part, s'il existe $(i,j) \in \{1,2,\dots,h\} \times \{1,2,\dots,k\}$, tel que $e_{i,j} \neq 0$, alors $\|S_n\|^2$ converge presque-sûrement vers $+\infty$.

Remarquons que l'on peut aussi démontrer ce théorème à partir de Mardia et Jupp. De plus, en ce qui concerne le choix de h et k, rappelons que l'on peut trouver des exemples dans Blacher 1987.

Maintenant, on va montrer que le test d'indépendance du chi-deux est un test Hilbertien. Pour cela, on introduit les notations suivantes.

Hypothèses 2-3 : Soient $\{J_i\}_{i=1,2,\dots,c}$ et $\{A_j\}_{j=1,2,\dots,d}$ deux partitions de, respectivement, \mathbb{R}^p et \mathbb{R}^q , telles que $m(J_i) > 0$ et $m'(A_j) > 0$ pour tout $(i,j) \in \{1,2,\dots,c\} \times \{1,2,\dots,d\}$. On garde les notations de 1-7, et on suppose que $\gamma = \{ 1_{\mathbb{R}^p}, 1_{J_1}, 1_{J_2}, \dots, 1_{J_{c-1}} \}$ et $\tau = \{ 1_{\mathbb{R}^q}, 1_{A_1}, 1_{A_2}, \dots, 1_{A_{d-1}} \}$, 1_{J_i} et 1_{A_j} représentant les fonctions indicatrices des ensembles J_i et A_j .

On va alors approximer \mathcal{Q} par une probabilité uniforme sur tout $J_i \times A_j$.

Théorème 2-4 : Soit f définie par : $f(u,v) = \sum_{i=1}^c \sum_{j=1}^d \frac{\mathcal{Q}(J_i \times A_j)}{m(J_i)m'(A_j)} 1_{J_i}(u) 1_{A_j}(v)$.

Alors,
$$f(u,v) = 1 + \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} \rho_{i,j} R_i(u) T_j(v) .$$

Démonstration : Les R_i sont des combinaisons linéaires des $1_{J_1}, 1_{J_2}, \dots, 1_{J_c}$. On en déduit

$$\int f(u,v) R_i(u) T_j(v) . m(du) . m'(dv) = \rho_{i,j} . \text{ De plus, } \{R_i(u) T_j(v)\}_{(i,j) \in \{0,1,\dots,c-1\} \times \{0,1,\dots,d-1\}}$$

est une base orthonormale du sous-espace de $L^2(\mathbb{R}^{p+q}, m \otimes m')$ engendré par les $1_{J_i \times A_j}(u,v)$. Il ne reste plus alors qu'à écrire f dans cette base.

En fait, on approximera \mathcal{Q} par la probabilité admettant f pour densité par rapport à $m \otimes m'$. Aussi, on va maintenant définir des estimations de f et les statistiques du chi deux correspondantes.

Notations 2-5 : Pour tout $n \in \mathbb{N}^*$, on note par $n_{i,j}$ (resp, n_i, n_j) le nombre de points de l'échantillon $\{(U^l, V^l)\}_{l=1,2,\dots,n}$ appartenant à $J_i \times A_j$ (resp, $J_i \times \mathbb{R}^q, \mathbb{R}^p \times A_j$), et on suppose $n_i > 0$ et $n_j > 0$. On pose alors

$$f_n^*(u,v) = \sum_{i=1}^c \sum_{j=1}^d \frac{\frac{n_{i,j}}{n}}{m(J_i) m'(A_j)} 1_{J_i}(u) 1_{A_j}(v) , \quad f_n(u,v) = \sum_{i=1}^c \sum_{j=1}^d \frac{\frac{n_{i,j}}{n}}{\frac{n_i}{n} \frac{n_j}{n}} 1_{J_i}(u) 1_{A_j}(v) ,$$

$$\chi_{U,V}^{*2} = \sum_{i=1}^c \sum_{j=1}^d \frac{(n_{i,j} - n . m(J_i) m'(A_j))^2}{n . m(J_i) m'(A_j)} , \quad \chi_{U,V}^2 = \sum_{i=1}^c \sum_{j=1}^d \frac{\left(n_{i,j} - \frac{n_i . n_j}{n} \right)^2}{\frac{n_i . n_j}{n}} ,$$

$$\chi_U^{*2} = \sum_{i=1}^c \frac{(n_i - n . m(J_i))^2}{n . m(J_i)} \quad \text{et} \quad \chi_V^{*2} = \sum_{j=1}^d \frac{(n_j - n . m'(A_j))^2}{n . m'(A_j)} .$$

Comme, on sait que, si $i \neq 0, \rho_{i,0} = \rho_{0,i} = 0$, le test Hilbertien d'indépendance testera donc uniquement la nullité des $\rho_{i,j}$ tels que $i > 0$ et $j > 0$. Or, cela n'est pas le cas du test classique du chi-deux lorsque les lois marginales sont connues. En effet, on peut énoncer le théorème suivant. (3-8 Blacher 1988).

Théorème 2-6 : En gardant les notations précédentes, $\chi_{U,V}^{*2} - \chi_U^{*2} - \chi_V^{*2} = n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\rho_{i,j}^n)^2$,

et
$$\chi_{U,V}^{*2} = n \sum_{i=0}^{c-1} \sum_{\substack{j=0 \\ (i,j) \neq (0,0)}}^{d-1} (\rho_{i,j}^n)^2 .$$

Donc, dans ce cas le test Hilbertien d'indépendance prend une forme très simple. En

particulier, on peut énoncer le théorème suivant.

Corollaire 2-7 : Si U et V sont indépendants, $\chi_{U,V}^{*2} - \chi_U^{*2} - \chi_V^{*2}$ suit asymptotiquement la loi du chi-deux à $(c-1)(d-1)$ degrés de liberté.

En fait, on peut prouver que $\chi_{U,V}^{*2} - \chi_U^{*2} - \chi_V^{*2}$ est la statistique du chi deux restreint, et donc que le test associé est plus puissant que le test d'indépendance classique (3-10 Blacher 1988).

On en vient maintenant au cas de lois marginales inconnues, et on va voir qu'alors, la statistique du chi-deux avec estimation de paramètres et la statistique Hilbertienne sont identiques.

Théorème 2-8 : En gardant les notations précédentes, $\chi_{U,V}^2 = n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\rho_{i,j}^n)^2$.

Démonstration : En appliquant le théorème 2-4 à \mathcal{Q}_n , on voit que

$$f^n(u,v) = 1 + \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} \rho_{i,j}^n R_i^n(u) T_j^n(v) .$$

Donc, on peut écrire les égalités suivantes :

$$\begin{aligned} \sum_{i=1}^c \sum_{j=1}^d \frac{\left(n_{i,j} - \frac{n_{i.} n_{.j}}{n} \right)^2}{\frac{n_{i.} n_{.j}}{n}} &= \sum_{i=1}^c \sum_{j=1}^d \left(\frac{n_{i.} n_{.j}}{n} \right) \left(\frac{n_{i,j}}{n_{i.} n_{.j}} - 1 \right)^2 \\ &= n \sum_{i=1}^c \sum_{j=1}^d \iint_{\mathcal{D}_i \times \mathcal{D}_j} (f^n(u,v) - 1)^2 m_n(du) m'_n(dv) \\ &= n \iint (f^n(u,v) - 1)^2 m_n(du) m'_n(dv) = n \sum_{i=1}^{c-1} \sum_{j=1}^{d-1} (\rho_{i,j}^n)^2 . \end{aligned}$$

Donc le test du chi-deux d'indépendance avec estimation de paramètres est un cas particulier du test défini par les théorèmes 1-15 et 2-2 .

BIBLIOGRAPHIE

- BLACHER R. (1984-a) Coefficients de corrélation d'ordre supérieur. Statistique et Analyse des Données, Vol 9, n°2, p 48 .
- BLACHER R. (1984-b) Coefficients de corrélation d'ordre (i,j) et variances d'ordre i . Symposium international de Bar le Duc, Polynômes orthogonaux et applications, Springer-Verlag n° 1171.
- BLACHER R. (1987) Divers indices de dépendance fournis par les fonctions orthogonales. Statistique et Analyse des Données, Vol 12, n° 1-2, p 1-27 .
- BLACHER R. (1988) A New Form for the Chi-Squared Test of Independence. Statistics, 19-4, 519-536.
- BLACHER R. (1989-a) Loi asymptotique des estimateurs des moindres carrés en régression linéaire. C.R. Acad. Sci. Paris, t.309, série I, p 629-632.
- BLACHER R. (1989-b) Loi asymptotique des fonctions orthogonales empiriques et applications: II. TIM3-IMAG, R.R. n° 792-M .
- BOSQ D. (1978) Test Hilbertien et Test du chi-deux. C.R. Académie de Paris, t 286, p 946-948.
- CHIHARA T. S. (1978) An introduction to orthogonal polynomials. Gordon and Breach, Science publishers. New-york, Paris.
- JUPP.E. MARDIA K.V. (1980) A general correlation coefficient for directionnal data and related regression problems. Biometrika 163-173.
- LANCASTER H.O. (1980) Orthogonal models for contingency tables. Developments in Statistics, Academic Press, New-York.
- SERFLING R. J. (1980) Approximation theorems of mathematical statistics. Wiley. New-York.

DOCUMENT IV

**LOI DE LA SOMME DE N
VECTEURS ALEATOIRES
ET LOI DE LEURS
FORMES QUADRATIQUES**

Mots-clefs : Transformation de Fourier, polynômes d'Hermite et de Laguerre, forme quadratique de vecteurs aléatoires, vecteur aléatoire Gaussien.

Résumé : On étudie les polynômes d'Hermite et de Laguerre et on les exprime sous forme de moments. De plus, grâce à leurs transformées de Fourier, on obtient une formule de réciprocity des fonctions caractéristiques. On en déduit la loi de la somme de n vecteurs aléatoires et la loi de leurs formes quadratiques. On obtient ce résultat sous forme de séries de polynômes dont les coefficients sont des moments. On applique alors ces résultats au cas Gaussien en utilisant des lois normales de "variances négatives". On obtient ainsi la densité des formes quadratiques de vecteurs Gaussiens, non forcément indépendants, sous forme d'intégrales, ou sous forme de développements en série de Mac Laurin ou en série de polynômes de Laguerre.

Summary : We study the Hermite polynomials and the Laguerre polynomials, and we express them in the forms of moments. Moreover we use their Fourier Transforms in order to obtain a formula of reciprocity to the characteristic functions. We deduce the distributions of a sum of n random vectors and of their quadratic forms: their densities are expanded in terms of orthogonal polynomials whose coefficients are moments. We apply these results to the normal distributions. For this, we need normal distributions with negative variances. By this way, we obtain the expansion of the density of quadratic forms in Mac Laurin series or in series of Laguerre polynomials. We obtain also this density in the form of an integral.

LOI DE LA SOMME DE N VECTEURS ALEATOIRES ET LOI DE LEURS FORMES QUADRATIQUES

SOMMAIRE

Introduction	p 3
Notations	p 5
I : Quelques propriétés des polynômes d'Hermite et de Laguerre.	p 6
II : Quelques propriétés des Transformées de Fourier.	p 10
III : Application au calcul d'une densité.	p 13
IV : Loi de la somme de n vecteurs aléatoires et lois de leurs formes quadratiques	p 15
IV-I Développement de la densité en polynômes orthogonaux	p 15
IV-II Développement de la densité en série de Mac Laurin	p 21
IV-III Densité sous forme d'intégrale.	p 23
V : Application au cas Gaussien	p 24
V-I Quelques propriétés des lois normales.	p 24
V-II Variances négatives dans les lois normales	p 28
V-III Formes quadratiques de vecteurs Gaussiens	p 30
Bibliographie	p 36

INTRODUCTION

Dans ce rapport on considère un vecteur aléatoire $X = (X_1, X_2, \dots, X_n)$ à valeurs dans \mathbb{R}^n , admettant une densité appartenant à L^2 par rapport à une loi normale centrée sur \mathbb{R}^n , et on va étudier la loi des vecteurs aléatoires $\Sigma = (X_1 + X_2 + \dots + X_n)$ et $\Sigma^2 = ((X_1)^2 + (X_2)^2 + \dots + (X_n)^2)$: on obtiendra la densité f de Σ par rapport à une loi normale $N(m, \sigma^2)$, et la densité g de Σ^2 par rapport à la loi $\gamma(n/2, 2)$. Remarquons que l'on n'impose pas de conditions d'indépendance entre les X_s .

Pour obtenir f et g , on va étudier les polynômes d'Hermite H_j , et les polynômes de Laguerre

L_j^a . On montrera en particulier que $H_j(x) = \frac{\mathbb{E} [(iU+x)^j]}{\sqrt{j!}}$ ($i^2 = -1$) et $L_j^a(x) = C_j \mathbb{E} [H_{2j}(\sqrt{Bx})]$,

où \mathbb{E} représente l'espérance mathématique, C_j , une constante, et U et B , deux variables aléatoires indépendantes, U de loi $N(0,1)$, et B de loi Béta.

On déduit alors immédiatement de ces égalités que l'on peut écrire f sous forme d'un développement en série de polynômes d'Hermite, les coefficients des H_j étant exprimés sous

forme de moments :
$$f(y) = \sum_{j=0}^{\infty} \frac{\mathbb{E} \left\{ \left(iU + \frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}} \right)^j \right\}}{\sqrt{j!}} H_j \left(\frac{y}{\sqrt{n}} \right).$$

On obtiendrait de la même façon une écriture de g en polynômes de Laguerre. Mais, dans ce cas, il est plus intéressant de faire intervenir les variables complexes $Z_s = iU_s + X_s$, où les U_s sont des variables aléatoires de loi $N(0,1)$, indépendantes entre elles, et indépendantes de X . En effet, on montrera que

$$f(y) = \sum_{j=0}^{\infty} \mathbb{E} \left\{ \frac{[Z_1 + Z_2 + \dots + Z_n]^j}{\sqrt{n^j j!}} \right\} H_j \left(\frac{y}{\sqrt{n}} \right) \text{ et } g(y) = \sum_{j=0}^{\infty} \mathbb{E} \left\{ \frac{[Z_1^2 + Z_2^2 + \dots + Z_n^2]^j}{\sqrt{2^j j!}} \right\} \mathbb{E} \{ H_{2j}(\sqrt{By}) \}.$$

Pour démontrer ces égalités, il nous faudra une formule explicite de réciprocity des fonctions caractéristiques. On obtiendra celle-ci grâce aux Transformées de Fourier des polynômes

d'Hermite :
$$\int e^{itx} H_j(x) \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = \frac{(it)^j}{\sqrt{j!}} e^{-\frac{t^2}{2}}.$$

Maintenant, dans les égalités ci dessus, écrivons H_j sous la forme : $H_j(x) = \frac{\mathbb{E} [(iV+x)^j]}{\sqrt{j!}}$.

On voit alors que
$$g(y) = \sum_{j=0}^{\infty} \mathbb{E} \left\{ \frac{[(Z_1^2 + Z_2^2 + \dots + Z_n^2)(iV + \sqrt{By})^2]^j}{2^j j!} \right\}.$$

Donc, si on peut intervertir les signes Σ et \int , on obtiendra g sous une forme intégrale, et on aura le même type de résultats en série de Mac Laurin. Comme exemple d'applications, on étudiera le cas Gaussien, et on arrivera donc ainsi à obtenir g sous forme d'une intégrale de U et de X .

Cependant, si l'on procède directement, il faudra imposer un ordre d'intégration: d'abord par rapport à X , puis par rapport à U , les intégrales obtenues n'étant pas définies autrement.

Cependant, on pourra tourner cette difficulté en constatant que les Z_s se comportent comme n variables Gaussiennes de variances négatives. En effet, on verra que, si X_1 suit la loi $N(0, \sigma^2)$, alors $Z_1 = iU_1 + X_1$ a les mêmes moments qu'une variable aléatoire Y_1 de loi $N(0, S^2)$, où $S^2 = \sigma^2 - 1$, et on aura le même type de résultats dans le cas vectoriel. Finalement, on pourra remplacer $Z = iU + X$ par un vecteur aléatoire $Y = iT + m$, T étant Gaussien centré réel. Alors, on pourra intervertir les signes Σ et \int , et on trouvera que

$$g(y) = \mathbb{E} \left\{ e^{-\frac{Y_1^2 + Y_2^2 + \dots + Y_n^2}{2}} \operatorname{ch}(\sqrt{B(Y_1^2 + Y_2^2 + \dots + Y_n^2)} y) \right\}$$

$$\text{et } g(y) = \sum_{j=0}^{\infty} \frac{\Gamma(n/2) \Gamma(1/2+j)}{\Gamma(1/2) \Gamma(n/2+j) 2^j j!} \mathbb{E} \left\{ e^{-\frac{Y_1^2 + Y_2^2 + \dots + Y_n^2}{2}} (Y_1^2 + Y_2^2 + \dots + Y_n^2)^j \right\} y^j .$$

Signalons pour terminer que tous ces résultats restent vrais dans le cas vectoriel : dans ce rapport, nous allons obtenir la loi des vecteurs aléatoires

$$\{(X_{1,1} + X_{1,2} + \dots + X_{1,n}), (X_{2,1} + X_{2,2} + \dots + X_{2,n}), \dots, (X_{p,1} + X_{p,2} + \dots + X_{p,n})\}$$

$$\text{et } \{[(X_{1,1})^2 + (X_{1,2})^2 + \dots + (X_{1,n})^2], \dots, [(X_{p,1})^2 + (X_{p,2})^2 + \dots + (X_{p,n})^2]\},$$

et ce, toujours sans hypothèse d'indépendance entre les $X_{s,r}$.

NOTATIONS

Dans ce rapport, on considérera un vecteur aléatoire X à valeurs dans \mathbb{R}^{np} défini sur un espace probabilisé (Ω, \mathcal{F}, P) . On supposera qu'il existe un vecteur aléatoire Gaussien centré $U \in \mathbb{R}^{np}$, indépendant de X , de matrice de variances-covariances, la matrice identité de \mathbb{R}^{np} , et un vecteur aléatoire B indépendant de X et U , $B = (B_1, B_2, \dots, B_p)$, les B_p représentant des variables aléatoires indépendantes de loi béta :

$$X = \{(X_{1,1}, X_{1,2}, \dots, X_{1,n}), (X_{2,1}, X_{2,2}, \dots, X_{2,n}), \dots, (X_{p,1}, X_{p,2}, \dots, X_{p,n})\} ,$$

$$U = \{(U_{1,1}, U_{1,2}, \dots, U_{1,n}), (U_{2,1}, U_{2,2}, \dots, U_{2,n}), \dots, (U_{p,1}, U_{p,2}, \dots, U_{p,n})\} .$$

Remarquons qu'en fait, U et B ne servent qu'à introduire des intégrales par rapport aux lois de Gauss ou Béta, et l'on aurait pu énoncer les résultats sans supposer cette existence. Mais, en procédant ainsi, cela permet de mettre ces intégrales sous forme probabiliste. De toutes façons, quitte à agrandir (Ω, \mathcal{F}, P) , on peut toujours supposer l'existence de U et B .

De plus, parfois on sera amené à considérer les cas particuliers : $p=1$, ou $n=1$, voire $n=p=1$. Dans ces cas, on simplifiera alors, par exemple, $(U_{1,1}, U_{1,2}, \dots, U_{1,n})$ en (U_1, U_2, \dots, U_n) .

D'autre part, $N(m, \sigma^2)$ représentera la loi normale sur \mathbb{R} , d'espérance m et de variance σ^2 , et $\gamma(a, \lambda)$ représentera la loi Gamma de paramètres a et λ . De plus, les lois Bétas que l'on considérera seront des lois Bétas de première espèce (pour leur définition, Voir Johnson, Kotz, 1970, ch 24).

De plus, on gardera les notations classiques : on notera par $E(\cdot)$ et $\sigma(\cdot)$, respectivement, l'espérance et l'écart-type. Pour toute matrice A , on notera par A' sa transposée.

Pour $a > 0$, on pose $\Gamma(a) = \int_0^{\infty} e^{-x} x^{a-1} dx$ et, pour $0 < a$ et $0 < b$, $\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Pour tout réel x , $[x]$ représentera la partie entière de x . Enfin, on notera par i la racine de -1 : $i^2 = -1$.

Signalons aussi que l'on adoptera la convention $0^0 = 1$.

Enfin, pour définir les limites des séries indicées par \mathbb{N}^n , on utilisera des bases de filtre $\{\mathfrak{B}\}$ telles que les éléments \mathfrak{B} soient des complémentaires de parties finies de \mathbb{N}^n , et telles que, pour tout $q \in \mathbb{N}$, il existe toujours un \mathfrak{B} de la base de filtre vérifiant $\{0, 1, \dots, q\}^n \subset \mathbb{N}^n \setminus \mathfrak{B}$.

Alors, les résultats que nous donnerons dans ce rapport seront vrais quelle que soit la base de filtre choisie vérifiant ces conditions, ou alors, on le précisera.

I : QUELQUES PROPRIETES DES POLYNOMES D'HERMITE ET DE LAGUERRE

Dans ce chapitre on va étudier quelques propriétés des polynômes d'Hermite et de Laguerre. On montrera en particulier qu'ils peuvent se mettre sous forme de moments et on en déduira les conséquences.

On introduit d'abord les notations.

Notations 1-1 : Soit $a > 0$. On note par H_j et L_j^a , les polynômes orthonormaux d'Hermite et de Laguerre associés aux lois $N(0,1)$ et $\gamma(a,2)$, i.e. si on note par $\delta_{j,k}$, le symbole de Kronecker,

$$\int_{x=-\infty}^{x=+\infty} H_j\left(\frac{x-m}{\sigma}\right) H_k\left(\frac{x-m}{\sigma}\right) \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = \int_{x=0}^{x=+\infty} L_j^a\left(\frac{2x}{\lambda}\right) L_k^a\left(\frac{2x}{\lambda}\right) \frac{x^{a-1} e^{-\frac{x}{\lambda}}}{\lambda^a \Gamma(a)} dx = \delta_{j,k} .$$

De plus, lorsque on considérera les polynômes de Laguerre, on supposera toujours $x \geq 0$.

En utilisant les relations 2-11 page 145, 2-15 page 146, et 2-18 page 148, de Chihara, on démontre facilement la proposition suivante.

Proposition 1-2 : Pour tout $j \in \mathbb{N}$, on a les égalités :

$$H_j\left(\frac{x-m}{\sigma}\right) = \frac{(-1)^j \sigma^j}{\sqrt{j!}} \left[e^{\frac{(x-m)^2}{2\sigma^2}} \frac{d^j \left(e^{-\frac{(x-m)^2}{2\sigma^2}} \right)}{dx^j} \right] = \sqrt{\frac{j!}{2^j}} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{(-1)^k}{[(j-2k)!][k!]} \left(\frac{\sqrt{2}(x-m)}{\sigma} \right)^{j-2k} ,$$

$$L_j^a\left(\frac{2x}{\lambda}\right) = (-1)^j \sqrt{\frac{\Gamma(a)}{j! \Gamma(j+a)}} \left[x^{-a+1} e^{\frac{x}{\lambda}} \frac{d^j \left(x^{j+a-1} e^{-\frac{x}{\lambda}} \right)}{dx^j} \right] = (-1)^j \sum_{k=0}^j \frac{\sqrt{\Gamma(a) \Gamma(j+a)!} (-x)^k}{(j-k)! \Gamma(k+a) \lambda^k k!} .$$

Remarquons que, si l'on note par ρ_j^α , les polynômes de Laguerre classiques (cf 2-1 Blacher 1989), alors on a les relations suivantes : $\sqrt{\Gamma(j+a)} L_j^a\left(\frac{2x}{\lambda}\right) = (-1)^j \sqrt{\Gamma(a) j!} \rho_j^{a-1}\left(\frac{x}{\lambda}\right)$.

On déduit immédiatement de 1-2 la proposition suivante (voir aussi 2-43, page 156 de Chihara).

Corollaire 1-3 : Pour tout $j \in \mathbb{N}$, H_{2j} est pair, et H_{2j+1} impair. De plus, $H_{2j}(x) = L_j^{1/2}(x^2)$.

Maintenant, on va voir que les polynômes d'Hermite et de Laguerre peuvent être considérés comme des moments. Cette propriété est un des résultats fondamentaux de ce rapport.

Théorème 1-4 : Soit $a > 1/2$, et soient U et B deux variables aléatoires indépendantes, U suivant la loi $N(0,1)$, et B suivant la loi Béta $(1/2, a-1/2)$.

Alors, pour tout $j \in \mathbb{N}$, $H_j(x) = \frac{\mathbb{E} [(iU+x)^j]}{\sqrt{j!}}$ et $L_j^a(x) = \frac{2^j}{(2j)!} \sqrt{\frac{j! \Gamma(j+a)}{\Gamma(a)}} \mathbb{E} [(iU - \sqrt{Bx})^{2j}]$.

Démonstration : D'abord, $\mathbb{E} \{ (iU+x)^j \} = \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{j!}{[(j-2k)!][2k!]} \frac{2k!}{2^k [k!]} (-1)^k x^{j-2k}$.

De plus, $\mathbb{E} [(U + i\sqrt{Bx})^{2j}] = \sum_{k=0}^j \frac{2j!}{(2k)!(2j-2k)!} \frac{(2j-2k)!}{2^{j-k} (j-k)!} \frac{\Gamma(k+1/2)}{\Gamma(k+a)} \frac{\Gamma(a)}{\Gamma(1/2)} (-x)^k$
 $= \frac{2j! \Gamma(a)}{2^j} \sum_{k=0}^j \frac{1}{(j-k)! \Gamma(k+a)} \frac{(-x)^k}{2^k (k!)}$.

Il ne reste plus alors qu'à appliquer les égalités de 1-2.

Rappelons que l'on connaît d'autres écritures des H_j et des L_j^a sous formes d'intégrale (cf Szego th 5-4, p102, 5-5-12, p105).

Grâce au théorème précédent, on peut maintenant établir une relation entre les polynômes d'Hermite et de Laguerre.

Corollaire 1-5 : Avec les notations de 1-4, $L_j^a(x) = 2^j \sqrt{\frac{j! \Gamma(j+a)}{2j! \Gamma(a)}} \mathbb{E} \{ H_{2j}(\sqrt{Bx}) \}$.

On obtient aussi une majoration des H_j et des L_j^a .

Théorème 1-6 : Soit $a > 1/2$. Pour tout $K > 1$, et tout compact \mathcal{K} , il existe $J \in \mathbb{N}$, tel que, pour

tout $j \geq J$, et tout $x \in \mathcal{K}$, $|H_{2j}(x)| \leq \frac{K^{2j} \sqrt{2j!}}{2^{2j}}$ et $|L_j^a(x)| \leq K^{2j} \sqrt{\frac{\Gamma(j+a)}{j! \Gamma(a)}}$.

Démonstration : Soit $K > 1$. Définissons S par $K = 1 + 2S^2$. Soient U et W , deux variables aléatoires indépendantes, de lois, respectivement, $N(0,1)$ et $N(0,S^2)$. Soit $\mu = \sup \{ |x| \mid x \in \mathcal{K} \}$, et soit T une variable aléatoire indépendante de U et W et telle que $|T| \leq 1$.

On sait qu'il existe s_0 tel que, pour tout $s \geq s_0$, $\mu^{2s} \leq \frac{S^{2s} 2s!}{2^s s!} = \mathbb{E} [W^{2s}]$.

Alors, il existe j_0 tel que $\mu^{2s} \frac{(\sqrt{1+S^2})^{2s}}{\mathbb{E} [W^{2s}]} \leq (\sqrt{1+S^2})^{2j_0}$ pour tout $s=0,1,\dots,s_0-1$.

Soit $j \geq \text{Max}(j_0, s_0)$. Alors, $|\mathbb{E} ((iU+Tx)^{2j})| \leq \sum_{s=0}^j \frac{2j!}{2s!(2j-2s)!} \mathbb{E} (U^{2j-2s}) \mu^{2s}$

$$\leq \sum_{s=0}^{s_0-1} \frac{2j!}{2s!(2j-2s)!} \mathbb{E}(U^{2j-2s}) (\sqrt{1+S^2})^{2j-2s} \mathbb{E}(W^{2s}) + \sum_{s=s_0}^j \frac{2j!}{2s!(2j-2s)!} \mathbb{E}(U^{2j-2s}) \mathbb{E}(W^{2s})$$

$$\leq \mathbb{E} [((\sqrt{1+S^2})U + W)^{2j}] = \frac{K^{2j}(2j)!}{2^j j!} .$$

On aura également besoin de la fonction génératrice des polynômes d'Hermite.

Théorème 1-7 : Pour tout t et $x \in \mathbb{C}$, $e^{xt - \frac{t^2}{2}} = \sum_{j=0}^{\infty} H_j(x) \frac{t^j}{\sqrt{j!}}$.

En fait, cette propriété est surtout connue si x et t sont réels (Chihara 2-14, p 145). Mais, il est clair qu'elle reste vérifiée si t et x sont complexes.

Maintenant, comme conséquence de 2-21, page 149 de Chihara, on a le lemme suivant.

Lemme 1-8 : On a l'égalité : $\frac{dH_j(x)}{dx} = \sqrt{j} H_{j-1}(x)$.

Ce résultat permet alors de démontrer la propriété suivante.

Lemme 1-9 : Pour tous entiers k et j , $k \geq j$, on peut écrire les égalités suivantes :

$$\int H_k\left(\frac{x-\mu}{S}\right) H_j\left(\frac{x-m}{\sigma}\right) \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = \sqrt{\frac{k!}{(k-j)!j!}} \left(\frac{\sigma}{S}\right)^j \left(\frac{\sqrt{\sigma^2 - S^2}}{iS}\right)^{k-j} H_{k-j}\left(\frac{i(m-\mu)}{\sqrt{\sigma^2 - S^2}}\right) .$$

Démonstration : On sait que ce résultat est vrai si $j=0$, $\mu=0$ et $S=1$ (3-9, Blacher 1989). Pour les cas où $j \neq 0$, mais $\mu \neq 0$ et/ou $S \neq 1$, il suffit d'effectuer un changement de variable. Donc, si $j \neq 0$, en intégrant par partie, et en utilisant 1-2 et 1-8,

$$\int H_k\left(\frac{x-\mu}{S}\right) H_j\left(\frac{x-m}{\sigma}\right) \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = -\frac{\sqrt{k}}{S} \frac{(-1)^j \sigma^j}{\sqrt{j!} \sqrt{2\pi\sigma^2}} \int H_{k-1}\left(\frac{x-\mu}{S}\right) \frac{d^{j-1} \left(e^{-\frac{(x-m)^2}{2\sigma^2}} \right)}{dx^{j-1}} dx$$

$$= \frac{\sqrt{k(k-1)\dots(k-j+1)}}{S^j} \frac{\sigma^j}{\sqrt{j!} \sqrt{2\pi\sigma^2}} \int H_{k-j}\left(\frac{x-\mu}{S}\right) \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx .$$

Enfin, on aura aussi besoin des propriétés suivantes

Lemme 1-10 : On a l'égalité : $\int x^j \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = \frac{\sqrt{j!} \sigma^j}{i^j} H_j\left(\frac{im}{\sigma}\right)$.

En effet, $\int x^j \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx = \frac{\sigma^j}{i^j} \int (ix + im/\sigma)^j \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$. Il suffit alors d'appliquer 1-4.

Lemme 1-11 : On a l'égalité : $x^j = \sum_{s=0}^{\lfloor \frac{j}{2} \rfloor} \frac{j!}{2^s s! \sqrt{(j-2s)!}} H_{j-2s}(x)$.

En effet, $\int x^j \frac{d^k \left(e^{-\frac{x^2}{2}} \right)}{dx^k} dx = (-j) \int x^{j-1} \frac{d^{k-1} \left(e^{-\frac{x^2}{2}} \right)}{dx^{k-1}} dx = (-1)^k j(j-1)\dots(j-k+1) \int x^{j-k} e^{-\frac{x^2}{2}} dx$.

Il suffit alors d'employer 1-2, et la formule $x^j = \sum_{k=0}^j \left(\int y^j H_k(y) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \right) H_k(x)$.

Lemme 1-12 : On a l'égalité : $H_j\left(\frac{x-m}{\sigma}\right) = \sum_{s=0}^j \sqrt{\frac{j!}{(j-s)!}} \frac{H_{j-s}\left(-\frac{m}{\sigma}\right)}{s! \sigma^s} x^s$.

En effet, $\mathbb{E} \left\{ \left(iU + \frac{x-m}{\sigma} \right)^j \right\} = \sum_{s=0}^j \frac{j!}{s! (j-s)!} \mathbb{E} \left\{ \left(iU - \frac{m}{\sigma} \right)^{j-s} \right\} \left(\frac{x}{\sigma} \right)^s$.

Il ne reste plus alors qu'à appliquer 1-4.

II : QUELQUES PROPRIETES DES TRANSFORMEES DE FOURIER

Dans ce chapitre, on étudie les Transformées de Fourier des polynômes d'Hermite et des polynômes de Laguerre. Leurs formes très particulières permettent alors d'obtenir une formule de réciprocity explicite des fonctions caractéristiques de certaines lois de probabilité.

On introduit d'abord la notation des Transformées de Fourier.

Notations 2-1 : Pour tout $n \in \mathbb{N}^*$, et tout vecteur aléatoire X , à valeurs dans \mathbb{R}^n , on note par ϕ_X , sa fonction caractéristique. De plus, pour toute probabilité M sur \mathbb{R}^n , et tout $g \in L^2(\mathbb{R}^n, M)$, on note par $\phi_{g,dM}$, la Transformée de Fourier :

$$\phi_{g,dM}(t_1, t_2, \dots, t_n) = \int e^{i(t_1 x_1 + t_2 x_2 + \dots + t_n x_n)} g(x_1, x_2, \dots, x_n) \cdot M(dx_1, dx_2, \dots, dx_n) .$$

Enfin, si $g \equiv 1$, on simplifie $\phi_{g,dM}$ en $\phi_M = \phi_{g,dM}$.

On peut maintenant rappeler la forme particulièrement simple des Transformées de Fourier des polynômes d'Hermite et de Laguerre.

Théorème 2-2 : Pour tout $j \in \mathbb{N}$, on définit \mathcal{L}_j^a par : $\mathcal{L}_j^a(x) = \sqrt{\frac{j! \Gamma(a)}{\Gamma(a+j)}} L_j^a(x)$.

Alors, pour tous entiers j et j' , $\phi_{H_j\left(\frac{x-m}{\sigma}\right)dN(m,\sigma^2)}(t) = \frac{\sigma^j}{\sqrt{j!}} \left[(it)^j \phi_{N(m,\sigma^2)}(t) \right]$,

$$\phi_{\mathcal{L}_j^a\left(\frac{2x}{\lambda}\right)\gamma^{a,\lambda}(dx)}(t) = \left(\frac{\lambda it}{1-\lambda it}\right)^j \phi_{\gamma^{a,\lambda}}(t) , \quad \phi_{L_j^a\left(\frac{2x}{\lambda}\right)\gamma^{a,\lambda}(dx)}(t) = \sqrt{\frac{\Gamma(a+j)}{j! \Gamma(a)}} \left(\frac{\lambda it}{1-\lambda it}\right)^j \phi_{\gamma^{a,\lambda}}(t) ,$$

$$\left[\phi_{\mathcal{L}_j^a\left(\frac{2x}{\lambda}\right)\gamma^{a,\lambda}(dx)}(t) \right] \left[\phi_{\mathcal{L}_{j'}^{a'}\left(\frac{2x}{\lambda}\right)\gamma^{a',\lambda}(dx)}(t) \right] = \phi_{\mathcal{L}_{j+j'}^{a+a'}\left(\frac{2x}{\lambda}\right)\gamma^{a+a',\lambda}(dx)}(t) , \quad \text{où } \gamma^{a,\lambda}(x) = \frac{x^{a-1} e^{-\frac{x}{\lambda}}}{\lambda^a \Gamma(a)} .$$

La première égalité, a été démontrée en 2-11 de Blacher 1985. Pour les autres, on emploie la même méthode que celle utilisée pour les démonstrations des lemmes 2-15 à 2-18 de Blacher 1985.

L'avantage d'écrire les densités sous forme de développements en polynômes orthogonaux dans les Transformées de Fourier, c'est qu'on peut alors intervertir les signes \sum et \int comme nous l'assure le théorème suivant.

Théorème 2-3 : Soit M une probabilité sur \mathbb{R}^n . et soit $\{B_{k_1, k_2, \dots, k_n}\}_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n}$, une base orthonormale de $L^2(\mathbb{R}^n, M)$. Soit $\{\alpha_{k_1, k_2, \dots, k_n}\}_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n}$, une suite réelle telle que

$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} (\alpha_{k_1, k_2, \dots, k_n})^2 < +\infty .$$

Alors, en posant $f = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \cdot B_{k_1, k_2, \dots, k_n}$ dans $L^2(\mathbb{R}^n, M)$,

$$\phi_{f, dM} = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \phi_{B_{k_1, k_2, \dots, k_n}, dM} \text{ pour la topologie de la convergence uniforme,}$$

et ce, pour n'importe quelle base de filtre définie sur \mathbb{N}^n , vérifiant les conditions imposées dans l'introduction.

La démonstration de ce théorème est identique à celle de 4-2 de Blacher 1989 (en remplaçant le carré de la norme par le carré de la somme.

En appliquant ces résultats, on obtient immédiatement une formule de réciprocity de la Transformation de Fourier.

Théorème 2-4 : Soient $N_n = N(m_1, \sigma_1^2) \otimes N(m_2, \sigma_2^2) \otimes \dots \otimes N(m_n, \sigma_n^2)$ et $\gamma_n = \gamma(a_1, \lambda_1) \otimes \gamma(a_2, \lambda_2) \otimes \dots \otimes \gamma(a_n, \lambda_n)$. De plus, soit $\{\alpha_{k_1, k_2, \dots, k_n}\}_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n}$, une suite réelle telle que

$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} (\alpha_{k_1, k_2, \dots, k_n})^2 < +\infty .$$

Soient f, g, φ et Φ les fonctions définies par :

$$f(x_1, x_2, \dots, x_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} H_{k_1} \left(\frac{x_1 - m_1}{\sigma_1} \right) H_{k_2} \left(\frac{x_2 - m_2}{\sigma_2} \right) \dots H_{k_n} \left(\frac{x_n - m_n}{\sigma_n} \right) \in L^2(\mathbb{R}^n, N_n)$$

$$g(x_1, x_2, \dots, x_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} L_{k_1}^{a_1} \left(\frac{2x_1}{\lambda_1} \right) L_{k_2}^{a_2} \left(\frac{2x_2}{\lambda_2} \right) \dots L_{k_n}^{a_n} \left(\frac{2x_n}{\lambda_n} \right) \in L^2(\mathbb{R}^n, \gamma_n),$$

$$\varphi(t_1, t_2, \dots, t_n) = \left(\prod_{s=1}^n e^{i m_s t_s - \frac{\sigma_s^2 t_s^2}{2}} \right) \left[\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \left(\prod_{s=1}^n \frac{(\sigma_s i t_s)^{k_s}}{\sqrt{k_s!}} \right) \right],$$

$$\Phi(t_1, t_2, \dots, t_n) = \left\{ \prod_{s=1}^n \frac{1}{(1 - \lambda_s i t_s)^{a_s}} \right\} \left\{ \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \left(\prod_{s=1}^n \sqrt{\frac{\Gamma(a_s + k_s)}{\Gamma(a_s) \cdot k_s!}} \left(\frac{\lambda_s i t_s}{1 - \lambda_s i t_s} \right)^{k_s} \right) \right\}$$

Alors $\varphi = \phi_{f, dN_n}$ et $\Phi = \phi_{g, d\gamma_n}$.

On utilisera particulièrement ce théorème pour obtenir explicitement les lois de probabilité. En particulier, on aura besoin des théorèmes suivants, corollaires immédiats de 2-4.

Théorème 2-5 : Soit M une probabilité sur \mathbb{R}^n , et soit N_n la loi $N_n = N(m_1, \sigma_1^2) \otimes N(m_2, \sigma_2^2) \otimes \dots \otimes N(m_n, \sigma_n^2)$. Alors, pour que M admette une densité f par rapport à N_n , $f \in L^2(\mathbb{R}^n, N_n)$:

$$f(x_1, x_2, \dots, x_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} H_{k_1} \left(\frac{x_1 - m_1}{\sigma_1} \right) H_{k_2} \left(\frac{x_2 - m_2}{\sigma_2} \right) \dots H_{k_n} \left(\frac{x_n - m_n}{\sigma_n} \right),$$

il faut et il suffit que

$$\left(\prod_{s=1}^n e^{-\frac{\sigma_s^2 t_s^2}{2} - i m_s t_s} \right) \phi_M(t_1, t_2, \dots, t_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \left(\prod_{s=1}^n \frac{(\sigma_s i t_s)^{k_s}}{\sqrt{k_s!}} \right),$$

avec
$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} (\alpha_{k_1, k_2, \dots, k_n})^2 < +\infty.$$

Théorème 2-6 : Soit M une probabilité sur \mathbb{R}^n , et soit γ_n la loi $\gamma_n = \gamma(a_1, \lambda_1) \otimes \gamma(a_2, \lambda_2) \otimes \dots \otimes \gamma(a_n, \lambda_n)$. Alors, pour que M admette une densité g par rapport à γ_n , $g \in L^2(\mathbb{R}^n, \gamma_n)$:

$$g(x_1, x_2, \dots, x_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} L_{k_1}^{a_1} \left(\frac{2x_1}{\lambda_1} \right) L_{k_2}^{a_2} \left(\frac{2x_2}{\lambda_2} \right) \dots L_{k_n}^{a_n} \left(\frac{2x_n}{\lambda_n} \right),$$

il faut et il suffit que

$$\left\{ \prod_{s=1}^n (1 - \lambda_s i t_s)^{a_s} \right\} \phi_M(t_1, t_2, \dots, t_n) = \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} \left(\prod_{s=1}^n \sqrt{\frac{\Gamma(a_s + k_s)}{\Gamma(a_s) k_s!}} \left(\frac{\lambda_s i t_s}{1 - \lambda_s i t_s} \right)^{k_s} \right),$$

avec
$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} (\alpha_{k_1, k_2, \dots, k_n})^2 < +\infty.$$

III : APPLICATION AU CALCUL D'UNE DENSITE

Supposons que l'on aie une variable aléatoire X admettant une densité f , appartenant à L^2 par rapport à $N(0,1)$. On sait que $f(x) = \sum_{j=0}^{\infty} \alpha_j H_j(x)$ dans $L^2(\mathbb{R}, N(0,1))$, avec $\alpha_j = \mathbb{E}[H_j(X)]$. En

écrivant alors H_j et α_j sous la forme : $H_j(x) = \frac{\mathbb{E}[(iU+x)^j]}{\sqrt{j!}}$ et $\alpha_j = \frac{\mathbb{E}[(iV+X)^j]}{\sqrt{j!}}$, on voit que

$$f(x) = \sum_{j=0}^{\infty} \mathbb{E} \left\{ \frac{[(iV+X)(iU+x)]^j}{j!} \right\}.$$

Ce résultat suggère alors d'étudier f lorsque l'on peut intervertir les signes \sum et \int . C'est ce que l'on va faire dans ce chapitre, après avoir remarqué que cette étude sera plus simple si l'on utilise la fonction génératrice des polynômes d'Hermite, plutôt que le développement en série de $e^{(iV+X)(iU+x)}$, c'est à dire si l'on écrit f sous la forme : $f(x) = \sum_{j=0}^{\infty} \mathbb{E} \left\{ \frac{(iV+X)^j}{\sqrt{j!}} H_j(x) \right\}$ (cf 1-7).

Maintenant, les remarques que nous venons de faire pour une variable aléatoire s'étendent sans difficulté au cas vectoriel. En effet, on peut énoncer le théorème suivant.

Théorème 3-1 : Soit \mathcal{N}_p la loi : $\mathcal{N}_p = N(\mu_1, S_1^2) \otimes N(\mu_2, S_2^2) \otimes \dots \otimes N(\mu_p, S_p^2)$, et soit $Y = (Y_1, Y_2, \dots, Y_p)$, un vecteur aléatoire à valeurs dans \mathbb{R}^p . On suppose qu'il existe un vecteur aléatoire Gaussien centré $U = (U_1, U_2, \dots, U_p) \in \mathbb{R}^p$, indépendant de Y , de matrice de variances-covariances, la matrice identité de \mathbb{R}^p .

On suppose que Y admet une densité de probabilité $f \in L^2(\mathbb{R}^p, \mathcal{N}_p)$ par rapport à \mathcal{N}_p et on pose

$$\begin{aligned} & \sum_{k_1 \leq j, k_2 \leq j, \dots, k_p \leq j} \mathbb{E} \left\{ \prod_{s=1}^p \frac{\left(iU_s + \frac{Y_s - \mu_s}{S_s} \right)^{k_s}}{\sqrt{k_s!}} H_{k_1} \left(\frac{y_1 - \mu_1}{S_1} \right) H_{k_2} \left(\frac{y_2 - \mu_2}{S_2} \right) \dots H_{k_p} \left(\frac{y_p - \mu_p}{S_p} \right) \right\} \\ &= \sum_{r_1 \leq j, r_2 \leq j, \dots, r_p \leq j} d_{r_1, r_2, \dots, r_p}^j y_1^{r_1} y_2^{r_2} \dots y_p^{r_p} = f_j(y_1, y_2, \dots, y_p). \end{aligned}$$

Alors, lorsque j tends vers $+\infty$, f_j converge vers f dans $L^2(\mathbb{R}^p, \mathcal{N}_p)$.

C'est une simple application de 1-4 et des propriétés élémentaires des espaces L^2 .

Dans l'égalité ci-dessus, on reconnaît la fonction génératrice de Taylor. Donc, si l'on peut inverser les signes \sum et \int , on va obtenir l'intégrale d'une fonction exponentielle.

Théorème 3-2: On garde les notations et hypothèses de 3-1 et on suppose que $f_{\lambda}(y_1, y_2, \dots, y_p)$ converge vers $f(y_1, y_2, \dots, y_p)$ en tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}^p$. On suppose aussi que, dans le premier terme des égalités du théorème 3-1, lorsque $\lambda = +\infty$, pour tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}^p$, on peut intervertir les signes Σ et \mathbb{E} , l'intégration étant effectuée d'abord par rapport à Y , puis par rapport à U .

Alors, l'intégrale

$$\int_{u_1=-\infty}^{u_1=+\infty} \int_{u_2=-\infty}^{u_2=+\infty} \dots \int_{u_p=-\infty}^{u_p=+\infty} \mathbb{E} \left[\prod_{s=1}^p \frac{e^{-iu_s \frac{Y_s - y_s}{S_s} - \frac{(Y_s - y_s)^2}{2S_s^2}}}{2\pi S_s} \right] du_1 \cdot du_2 \dots du_p$$

existe et est égale à la densité de Y par rapport à la mesure de Lebesgue sur \mathbb{R}^p .

En effet, d'après 1-7, il est clair que, si l'on peut intervertir les signes Σ et \mathbb{E} ,

$$f(y_1, y_2, \dots, y_p) = \mathbb{E}_U \left\{ \mathbb{E}_Y \left\{ \prod_{k=1}^p \exp \left[-\frac{\left(iU_k + \frac{Y_k - \mu_k}{S_k} \right)^2}{2} + \frac{y_k - \mu_k}{S_k} \left(iU_k + \frac{Y_k - \mu_k}{S_k} \right) \right] \right\} \right\}.$$

D'autre part, on aura le même type de propriétés pour les développements en séries de Mac Laurin.

Théorème 3-3 : On garde les notations de 3-1. Alors, pour tout $(r_1, r_2, \dots, r_p) \in \mathbb{N}^p$,

$$d'_{r_1, r_2, \dots, r_p} = \sum_{r_1 \leq k_1 \leq \lambda, \dots, r_p \leq k_p \leq \lambda} \mathbb{E} \left\{ \prod_{t=1}^p \frac{\left(iU_t + \frac{Y_t - \mu_t}{S_t} \right)^{r_t}}{r_t! S_t^{r_t}} \frac{\left(iU_t + \frac{Y_t - \mu_t}{S_t} \right)^{k_t - r_t}}{\sqrt{(k_t - r_t)!}} H_{k_t - r_t} \left(-\frac{\mu_t}{S_t} \right) \right\}.$$

De plus, si $d'_{r_1, r_2, \dots, r_p}$ converge vers un réel d_{r_1, r_2, \dots, r_p} , et si dans le terme de droite de l'égalité précédente, on peut intervertir les signes Σ et \mathbb{E} lorsque $\lambda = +\infty$, l'intégration étant effectuée d'abord par rapport à Y , puis par rapport à U , alors

$$d_{r_1, r_2, \dots, r_p} = \int_{u_1=-\infty}^{u_1=+\infty} \int_{u_2=-\infty}^{u_2=+\infty} \dots \int_{u_p=-\infty}^{u_p=+\infty} \mathbb{E} \left\{ \prod_{k=1}^p \frac{e^{\frac{\mu_k^2}{2S_k^2}}}{r_k! S_k^{r_k} \sqrt{2\pi}} \left(iu_k + \frac{Y_k - \mu_k}{S_k} \right)^{r_k} e^{-\frac{iu_k Y_k}{S_k} - \frac{Y_k^2}{2S_k^2}} \right\} du_1 du_2 \dots du_p$$

Pour démontrer la première égalité, on utilise 1-12. Donc, si l'on peut inverser Σ et \mathbb{E} , il ne reste plus qu'à employer 1-7 pour obtenir le résultat.

Les résultats que nous venons d'énoncer maintenant sont purement théoriques. Mais, nous verrons des exemples d'applications dans le chapitre V.

IV : LOI DE LA SOMME DE N VECTEURS ALEATOIRES ET LOIS DE LEURS FORMES QUADRATIQUES

Les théorèmes des chapitres précédents vont nous permettre maintenant d'obtenir la loi de la somme de n vecteurs aléatoires et la loi de leurs formes quadratiques définies positives. En fait, on sait déjà que, si le vecteur aléatoire somme admet une densité appartenant à L^2 , on peut développer celle-ci en série de polynômes d'Hermite, les coefficients des H_j étant donnés sous forme de moments. En effet, c'est une conséquence immédiate du théorème 1-4, et l'on obtient le même type de résultats pour les formes quadratiques avec les polynômes de Laguerre.

Donc, en fait, dans ce chapitre, nous allons montrer que si le vecteur aléatoire X admet une densité dans \mathbb{R}^{np} par rapport à une loi normale N_{np} , alors on se retrouvera sous ces hypothèses. De plus, on verra que ces résultats se simplifient lorsque N_{np} est centrée, et que, dans ce cas, on les obtient en fonction des moments des vecteurs complexes $iU+X$. Il sera alors intéressant de comparer la loi de la somme et celles des formes quadratiques.

Enfin, comme dans le chapitre précédent, on obtiendra aussi ces résultats sous formes d'intégrales et de séries de Mac Laurin.

On introduit d'abord les notations que l'on utilisera tout au long de ce chapitre.

Notations 4-1 : Dans ce chapitre, on note par N_{np} et γ_{np} les lois produits sur \mathbb{R}^{np} :

$$N_{np} = N(m_{1,1}, \sigma_{1,1}^2) \otimes \dots \otimes N(m_{1,n}, \sigma_{1,n}^2) \otimes \dots \otimes N(m_{p,1}, \sigma_{p,1}^2) \otimes \dots \otimes N(m_{p,n}, \sigma_{p,n}^2) ,$$

$$\gamma_{np} = \gamma(a_{1,1}, \lambda_1) \otimes \dots \otimes \gamma(a_{1,n}, \lambda_1) \otimes \dots \otimes \gamma(a_{p,1}, \lambda_p) \otimes \dots \otimes \gamma(a_{p,n}, \lambda_p) .$$

On considère alors un vecteur aléatoire X à valeurs dans \mathbb{R}^{np} , et on suppose qu'il existe un vecteur aléatoire Gaussien centré $U \in \mathbb{R}^{np}$, indépendant de X , de matrice de covariances la matrice identité de \mathbb{R}^{np} :

$$X = \{ (X_{1,1}, X_{1,2}, \dots, X_{1,n}), (X_{2,1}, X_{2,2}, \dots, X_{2,n}), \dots, (X_{p,1}, X_{p,2}, \dots, X_{p,n}) \} ,$$

$$U = \{ (U_{1,1}, U_{1,2}, \dots, U_{1,n}), (U_{2,1}, U_{2,2}, \dots, U_{2,n}), \dots, (U_{p,1}, U_{p,2}, \dots, U_{p,n}) \} .$$

IV-1 : DEVELOPPEMENT DE LA DENSITE EN POLYNOMES ORTHOGONAUX

Dans ce paragraphe, on va donc appliquer le théorème 1-4, en montrant que l'on peut l'utiliser dès qu'il existe une densité dans \mathbb{R}^{np} . On commence par étudier les résultats les plus généraux, et on passera ensuite au cas centré.

On commence par étudier la loi de la somme de n vecteurs aléatoires.

Théorème 4-2 : On garde les hypothèses précédentes et, pour tout $s \in \{1, 2, \dots, p\}$, on pose $\mu_s = m_{s,1} + m_{s,2} + \dots + m_{s,n}$, et $(S_s)^2 = (\sigma_{s,1})^2 + (\sigma_{s,2})^2 + \dots + (\sigma_{s,n})^2$, et on définit la loi Gaussienne \mathcal{N}_p par : $\mathcal{N}_p = N(\mu_1, S_1^2) \otimes N(\mu_2, S_2^2) \otimes \dots \otimes N(\mu_p, S_p^2)$. On suppose aussi qu'il existe un vecteur aléatoire Gaussien centré $V = (V_1, V_2, \dots, V_p) \in \mathbb{R}^p$, indépendant de U et X , de matrice de covariances, la matrice identité.

On suppose de plus que X admet une densité f par rapport à \mathcal{N}_{np} appartenant à $L^2(\mathbb{R}^{np}, \mathcal{N}_{np})$.

Alors, $\Sigma = \{(X_{1,1} + X_{1,2} + \dots + X_{1,n}), (X_{2,1} + X_{2,2} + \dots + X_{2,n}), \dots, (X_{p,1} + X_{p,2} + \dots + X_{p,n})\}$ admet une densité par rapport à \mathcal{N}_p , $f \in L^2(\mathbb{R}^p, \mathcal{N}_p)$:

$$f(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} T_{q_1, q_2, \dots, q_p} \cdot H_{q_1} \left(\frac{y_1 - \mu_1}{S_1} \right) H_{q_2} \left(\frac{y_2 - \mu_2}{S_2} \right) \dots H_{q_p} \left(\frac{y_p - \mu_p}{S_p} \right),$$

$$\text{où } T_{q_1, q_2, \dots, q_p} = \mathbb{E} \left[\frac{\prod_{j=1}^p \left(iV_j + \frac{(X_{j,1} - m_{j,1}) + (X_{j,2} - m_{j,2}) + \dots + (X_{j,n} - m_{j,n})}{\sqrt{\sigma_{j,1}^2 + \sigma_{j,2}^2 + \dots + \sigma_{j,n}^2}} \right)^{q_j}}{\sqrt{q_j!}} \right].$$

Démonstration : Rappelons d'abord comment, on peut écrire la densité.

Lemme 4-3 : Sous les hypothèses de 4-2, on peut écrire f de la façon suivante

$$\sum_{\{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})\} \in \mathbb{N}^{np}} \alpha_{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})} \left[\prod_{i=1}^p \prod_{s=1}^n H_{k_{i,s}} \left(\frac{x_{i,s} - m_{i,s}}{\sigma_{i,s}} \right) \right],$$

$$\text{où } \alpha_{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})} = \mathbb{E} \left[\frac{\prod_{i=1}^p \prod_{s=1}^n \left(iU_{i,s} + \frac{X_{i,s} - m_{i,s}}{\sigma_{i,s}} \right)^{k_{i,s}}}{\sqrt{k_{i,s}!}} \right].$$

Ce résultat est une simple application des propriétés des espaces L^2 et de 1-4.

Maintenant, on va démontrer que Σ admet une densité par rapport à \mathcal{N}_p . Il n'y aura plus alors qu'à appliquer 1-4 pour en déduire le résultat.

Lemme 4-4 : Sous les hypothèses de 4-2, Σ admet une densité f^* par rapport à \mathcal{N}_p :

$$f^*(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathcal{C}_{q_1, q_2, \dots, q_p} H_{q_1} \left(\frac{y_1 - \mu_1}{S_1} \right) H_{q_2} \left(\frac{y_2 - \mu_2}{S_2} \right) \dots H_{q_p} \left(\frac{y_p - \mu_p}{S_p} \right) \in L^2(\mathbb{R}^p, \mathcal{N}_p)$$

$$\text{où } \mathcal{C}_{q_1, q_2, \dots, q_p} = \sum_{\{k_{t,1}, k_{t,2}, \dots, k_{t,n}\}} \alpha_{(k_{1,1}, k_{1,2}, \dots, k_{1,n}), \dots, (k_{p,1}, k_{p,2}, \dots, k_{p,n})} \left\{ \prod_{t=1}^p \frac{\sqrt{q_t!} \prod_{s=1}^n (\sigma_{t,s})^{k_{t,s}}}{S_t^{q_t} \sqrt{k_{t,s}!}} \right\},$$

la somme portant sur tous les $\{(k_{1,1}, k_{1,2}, \dots, k_{1,n}), \dots, (k_{p,1}, k_{p,2}, \dots, k_{p,n})\} \in \mathbb{N}^{np}$ tels que $k_{t,1} + k_{t,2} + \dots + k_{t,n} = q_t$ pour tout $t=1, 2, \dots, p$.

Démonstration : Cette démonstration est la généralisation évidente de 5-1 de Blacher 1989.

On remarque d'abord que

$$\sum_{\substack{k_{1,1}+k_{1,2}+\dots+k_{1,n}=q_1, \dots, k_{p,1}+k_{p,2}+\dots+k_{p,n}=q_p}} \left\{ \prod_{t=1}^p \frac{q_t!}{S_t^{2q_t}} \prod_{s=1}^n \frac{(\sigma_{t,s}^2)^{k_{t,s}}}{k_{t,s}!} \right\}$$

$$= \prod_{t=1}^p \frac{q_t!}{S_t^{2q_t}} \left(\sum_{k_{t,1}+k_{t,2}+\dots+k_{t,n}=q_t} \frac{(\sigma_{t,1}^2)^{k_{t,1}} (\sigma_{t,2}^2)^{k_{t,2}} \dots (\sigma_{t,n}^2)^{k_{t,n}}}{k_{t,1}! k_{t,2}! \dots k_{t,n}!} \right) = 1 .$$

On en déduit que

$$\sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} (\mathcal{C}_{q_1, q_2, \dots, q_p})^2 < +\infty .$$

En utilisant 2-2 et 2-3, on calcule alors $\phi_{X_{1,1}+X_{1,2}+\dots+X_{1,n}, \dots, X_{p,1}+X_{p,2}+\dots+X_{p,n}}$.
 Il ne reste plus qu'à utiliser 2-5 pour conclure.

Maintenant, on obtient le même type de résultats si on utilise des développements en polynômes de Laguerre.

Théorème 4-5 : On garde les hypothèses de 4-1, et pour tout $s \in \{1, 2, \dots, p\}$, on pose $A_s = a_{s,1} + a_{s,2} + \dots + a_{s,n}$, et on définit la loi \mathcal{G}_p sur \mathbb{R}^p , par : $\mathcal{G}_p = \gamma(A_1, \lambda_1) \otimes \gamma(A_2, \lambda_2) \otimes \dots \otimes \gamma(A_p, \lambda_p)$.

On suppose que X admet une densité appartenant à $L^2(\mathbb{R}_+^{np}, \gamma_{np})$ par rapport à γ_{np} :

$$\sum_{\{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})\} \in \mathbb{N}^{np}} \beta_{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})} \left[\prod_{t=1}^p \prod_{s=1}^n L_{k_{t,s}}^{a_{t,s}} \left(\frac{2x_{t,s}}{\lambda_t} \right) \right] .$$

Alors, le vecteur aléatoire à valeurs dans \mathbb{R}_+^p :

$$\Sigma = \{(X_{1,1}+X_{1,2}+\dots+X_{1,n}), (X_{2,1}+X_{2,2}+\dots+X_{2,n}), \dots, (X_{p,1}+X_{p,2}+\dots+X_{p,n})\}$$

admet une densité par rapport à la loi Gamma \mathcal{G}_p :

$$\sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} B_{q_1, q_2, \dots, q_p} L_{q_1}^{A_1} \left(\frac{2x_1}{\lambda_1} \right) L_{q_2}^{A_2} \left(\frac{2x_2}{\lambda_2} \right) \dots L_{q_p}^{A_p} \left(\frac{2x_p}{\lambda_p} \right) \in L^2(\mathbb{R}^p, \mathcal{G}_p) ,$$

où $B_{q_1, q_2, \dots, q_p} = \sum_{\substack{k_{1,1}, \dots, k_{1,n} \\ k_{1,1}+\dots+k_{1,n}=q_1}} \beta_{(k_{1,1}, \dots, k_{1,n}), \dots, (k_{p,1}, \dots, k_{p,n})} \left\{ \prod_{j=1}^p \left(\prod_{s=1}^n \sqrt{\frac{\Gamma(a_{j,s} + k_{j,s})}{k_{j,s}! \Gamma(a_{j,s})}} \right) \sqrt{\frac{q_j! \Gamma(A_j)}{\Gamma(A_j + q_j)}} \right\} ,$

la somme portant sur tous les $\{(k_{1,1}, k_{1,2}, \dots, k_{1,n}), \dots, (k_{p,1}, k_{p,2}, \dots, k_{p,n})\} \in \mathbb{N}^{np}$ tels que $k_{t,1} + k_{t,2} + \dots + k_{t,n} = q_t$ pour tout $t=1, 2, \dots, p$.

Démonstration : Cette démonstration est la généralisation évidente de 5-5 de Blacher 1989.

On sait que $(1-x)^{-A_j} = \sum_{j=0}^{\infty} \frac{\Gamma(A_j+j)}{j! \Gamma(A_j)} x^j$ et $(1-x)^{-a_{j,s}} = \sum_{j=0}^{\infty} \frac{\Gamma(a_{j,s}+j)}{j! \Gamma(a_{j,s})} x^j$ si $|x| < 1$.

Comme $(1-x)^{-A_j} = \prod_{s=1}^n (1-x)^{-a_{j,s}}$, il suffit d'égaliser le coefficient de x^{q_j} pour en conclure

$$\sum_{k_{1,1}+k_{1,2}+\dots+k_{1,n}=q_1, \dots, k_{p,1}+k_{p,2}+\dots+k_{p,n}=q_p} \left\{ \prod_{j=1}^p \left(\prod_{s=1}^n \frac{\Gamma(a_{j,s}+k_{j,s})}{k_{j,s}! \Gamma(a_{j,s})} \right) \frac{q_j! \Gamma(A_j)}{\Gamma(A_j+q_j)} \right\} = 1.$$

On emploie alors la même méthode qu'en 4-4.

Finalement, en utilisant 1-4, on obtiendra la loi de Σ sous forme de développement en série de polynômes de Laguerre, les coefficients des L^a ; étant donnés sous forme de moments, à une constante multiplicative près. Maintenant, on pourra bien sûr appliquer ces résultats pour connaître la loi des formes quadratiques de vecteurs aléatoires. Pour cela, il nous faudra d'abord calculer la loi de $((X_1)^2, (X_2)^2, \dots, (X_n)^2)$. C'est ce résultat que nous donne le théorème suivant.

Théorème 4-6 : Soit $X = (X_1, X_2, \dots, X_n)$, un vecteur aléatoire à valeurs dans \mathbb{R}^n admettant une densité de probabilité par rapport à $N_n = N(0, \sigma_1^2) \otimes N(0, \sigma_2^2) \otimes \dots \otimes N(0, \sigma_n^2)$:

$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{k_1, k_2, \dots, k_n} H_{k_1} \left(\frac{x_1}{\sigma_1} \right) H_{k_2} \left(\frac{x_2}{\sigma_2} \right) \dots H_{k_n} \left(\frac{x_n}{\sigma_n} \right) \in L^2(\mathbb{R}^n, N_n).$$

Alors, $((X_1)^2, (X_2)^2, \dots, (X_n)^2)$ admet une densité de probabilité par rapport à $\gamma_n = \gamma(1/2, 2\sigma_1^2) \otimes \gamma(1/2, 2\sigma_2^2) \otimes \dots \otimes \gamma(1/2, 2\sigma_n^2)$:

$$\sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{2k_1, 2k_2, \dots, 2k_n} L_{k_1}^{1/2} \left(\frac{x_1}{\sigma_1} \right) L_{k_2}^{1/2} \left(\frac{x_2}{\sigma_2} \right) \dots L_{k_n}^{1/2} \left(\frac{x_n}{\sigma_n} \right) \in L^2(\mathbb{R}^n, \gamma_n).$$

Démonstration : En utilisant la même méthode qu'en 2-3, on peut écrire les égalités suivantes :

$$\phi_{X_1^2, X_2^2, \dots, X_n^2}(t_1, t_2, \dots, t_n)$$

$$\begin{aligned}
 &= \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{2k_1, 2k_2, \dots, 2k_n} \left\{ \prod_{s=1}^n \left(\int_{x_s=-\infty}^{x_s=+\infty} e^{it_s x_s^2} L_{k_s}^{1/2} \left(\frac{x_s^2}{\sigma_s^2} \right) \frac{e^{-\frac{x_s^2}{2\sigma_s^2}}}{\sqrt{2\pi\sigma_s^2}} dx_s \right) \right\} \quad (\text{cf 1-3}) \\
 &= \sum_{(k_1, k_2, \dots, k_n) \in \mathbb{N}^n} \alpha_{2k_1, 2k_2, \dots, 2k_n} \left\{ \prod_{s=1}^n \phi_{L_{k_s}^{1/2} \left(\frac{2z_s}{2\sigma_s^2} \right) \gamma^{1/2, 2\sigma_s^2} (t_s)} \right\} .
 \end{aligned}$$

Il ne reste plus qu'à appliquer 2-6 pour achever la démonstration.

Grâce à tous ces résultats, on obtient maintenant la loi des formes quadratiques de vecteurs aléatoires. On obtient ce résultat sous une forme relativement simple grâce à l'introduction des fonctions complexes $iU+X$, et il est alors intéressant de comparer ces résultats à ceux du théorème 4-2. C'est pourquoi nous énonçons maintenant le théorème général suivant.

Théorème 4-7 : On garde les hypothèses de 4-2 et on suppose de plus que, pour tout $(i,s) \in \{1,2,\dots,p\} \times \{1,2,\dots,n\}$, $m_{i,s}=0$ et $\sigma_{i,s}=\sigma_i$. On définit alors la loi gamma produit \mathcal{G}_p , par : $\mathcal{G}_p = \gamma(n/2, 2\sigma_1^2) \otimes \gamma(n/2, 2\sigma_2^2) \otimes \dots \otimes \gamma(n/2, 2\sigma_p^2)$. On suppose de plus qu'il existe p variables aléatoires $B_s, s=1,2,\dots,p$, de loi béta $(1/2, (n-1)/2)$, indépendantes entre elles et indépendantes de U et X .

Pour tout $(i,s) \in \{1,2,\dots,p\} \times \{1,2,\dots,n\}$, on définit alors la variable aléatoire complexe $Z_{i,s}$ par $Z_{i,s} = iU_{i,s} + \frac{X_{i,s}}{\sigma_i}$ avec $i^2=-1$.

Alors, $\Sigma = \{(X_{1,1}+X_{1,2}+\dots+X_{1,n}), (X_{2,1}+X_{2,2}+\dots+X_{2,n}), \dots, (X_{p,1}+X_{p,2}+\dots+X_{p,n})\}$ admet une densité $f \in L^2(\mathbb{R}^p, \mathcal{N}_p)$ par rapport à \mathcal{N}_p :

$$\begin{aligned}
 &f(y_1, y_2, \dots, y_p) = \\
 &\sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathbb{E} \left\{ \prod_{j=1}^p \frac{[Z_{j,1}+Z_{j,2}+\dots+Z_{j,n}]^{q_j}}{\sqrt{n^{q_j} q_j!}} \right\} H_{q_1} \left(\frac{y_1}{\sqrt{n\sigma_1}} \right) H_{q_2} \left(\frac{y_2}{\sqrt{n\sigma_2}} \right) \dots H_{q_p} \left(\frac{y_p}{\sqrt{n\sigma_p}} \right) .
 \end{aligned}$$

D'autre part, $\Sigma^2 = \{[(X_{1,1})^2+(X_{1,2})^2+\dots+(X_{1,n})^2], \dots, [(X_{p,1})^2+(X_{p,2})^2+\dots+(X_{p,n})^2]\}$
 admet une densité $g \in L^2(\mathbb{R}^p, \mathcal{G}_p)$ par rapport à \mathcal{G}_p :

$$g(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathbb{E} \left\{ \prod_{j=1}^p [Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2]^{q_j} \right\} \frac{\mathbb{E} \left\{ H_{2q_1} \left(\frac{\sqrt{B_1 y_1}}{\sigma_1} \right) H_{2q_2} \left(\frac{\sqrt{B_2 y_2}}{\sigma_2} \right) \dots H_{2q_p} \left(\frac{\sqrt{B_p y_p}}{\sigma_p} \right) \right\}}{\sqrt{2q_1! 2q_2! \dots 2q_p!}}$$

Démonstration : La première assertion est un cas particulier de 4-2.

Pour démontrer la seconde, on remarque que $\frac{\Gamma(1/2+k)}{k! \Gamma(1/2)(2k)!} = \frac{1}{(2^k k!)^2}$, et on applique alors

ce résultat aux théorèmes 4-6 et 4-5, avec $a_{i,s} = 1/2$.

$$\begin{aligned} & \text{On trouve que } \left(\prod_{j=1}^p 2^{q_j} \sqrt{\frac{\Gamma(n/2+q_j)}{\Gamma(n/2)(q_j!)}} \right) B_{q_1, q_2, \dots, q_p} \\ &= \mathbb{E} \left\{ \sum_{\substack{k_{1,1}+k_{1,2}+\dots+k_{1,n}=q_1 \\ \dots \\ k_{p,1}+k_{p,2}+\dots+k_{p,n}=q_p}} \prod_{j=1}^p \prod_{s=1}^n \frac{\left(iU_{j,s} + \frac{X_{j,s}}{\sigma_j} \right)^{2k_{j,s}}}{k_{j,s}!} \right\} \\ &= \mathbb{E} \left\{ \prod_{j=1}^p \left[\sum_{k_{j,1}+k_{j,2}+\dots+k_{j,n}=q_j} \prod_{s=1}^n \frac{\left(iU_{j,s} + \frac{X_{j,s}}{\sigma_j} \right)^{2k_{j,s}}}{k_{j,s}!} \right] \right\} = \mathbb{E} \left\{ \prod_{j=1}^p \frac{\left[\sum_{s=1}^n \left(iU_{j,s} + \frac{X_{j,s}}{\sigma_j} \right)^2 \right]^{q_j}}{q_j!} \right\} \end{aligned}$$

Il suffit alors d'appliquer 1-5.

IV-II : DEVELOPPEMENT DE LA DENSITE EN SERIE DE MAC LAURIN

On va maintenant appliquer la même méthode que dans le chapitre III et obtenir ainsi, sous forme de séries de Mac Laurin, la densité de la somme de n vecteurs aléatoires et la densité de leurs formes quadratiques.

On va d'abord expliciter l'écriture des densités lorsque on limite l'ordre de leurs développements.

Théorème 4-8 : On garde les notations et hypothèses du théorème 4-7, et on pose

$$\sum_{q_1 \leq r, q_2 \leq r, \dots, q_p \leq r} \mathbb{E} \left\{ \prod_{j=1}^p \frac{[Z_{j,1} + Z_{j,2} + \dots + Z_{j,n}]^{q_j}}{\sqrt{n^{q_j} q_j!}} \right\} H_{q_1} \left(\frac{y_1}{\sqrt{n\sigma_1}} \right) H_{q_2} \left(\frac{y_2}{\sqrt{n\sigma_2}} \right) \dots H_{q_p} \left(\frac{y_p}{\sqrt{n\sigma_p}} \right)$$

$$= \sum_{r_1 \leq r, r_2 \leq r, \dots, r_p \leq r} d'_{r_1, r_2, \dots, r_p} y_1^{r_1} y_2^{r_2} \dots y_p^{r_p} = f_r(y_1, y_2, \dots, y_p) \quad (4-8-1),$$

$$\sum_{q_1 \leq r, q_2 \leq r, \dots, q_p \leq r} \mathbb{E} \left\{ \prod_{j=1}^p [Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2]^{q_j} \frac{H_{2q_1} \left(\frac{\sqrt{B_1} y_1}{\sigma_1} \right) H_{2q_2} \left(\frac{\sqrt{B_2} y_2}{\sigma_2} \right) \dots H_{2q_p} \left(\frac{\sqrt{B_p} y_p}{\sigma_p} \right)}{\sqrt{2^{q_1} 2^{q_2} \dots 2^{q_p} q_1! q_2! \dots q_p!}} \right\}$$

$$= \sum_{r_1 \leq r, r_2 \leq r, \dots, r_p \leq r} c'_{r_1, r_2, \dots, r_p} y_1^{r_1} y_2^{r_2} \dots y_p^{r_p} = g_r(y_1, y_2, \dots, y_p) \quad (4-8-2),$$

Alors, pour tout $(r_1, r_2, \dots, r_p) \in \mathbb{N}^p$,

$$d'_{r_1, r_2, \dots, r_p} = \sum_{r_1 \leq q_1 \leq r, \dots, r_p \leq q_p \leq r} \mathbb{E} \left\{ \prod_{j=1}^p \frac{(Z_{j,1} + Z_{j,2} + \dots + Z_{j,n})^{r_j}}{r_j! n^{r_j} \sigma_j^{r_j}} \frac{(Z_{j,1} + Z_{j,2} + \dots + Z_{j,n})^{q_j - r_j}}{\sqrt{n^{q_j - r_j} (q_j - r_j)!}} H_{q_j - r_j}(0) \right\} \quad (4-8-3),$$

et $c'_{r_1, r_2, \dots, r_p} =$

$$\sum_{r_1 \leq q_1 \leq r, \dots, r_p \leq q_p \leq r} \left(\prod_{j=1}^p \frac{\mathbb{E}[B_j^{r_j}]}{(2r_j)! \sigma_j^{2r_j}} \right) \mathbb{E} \left\{ \prod_{j=1}^p (Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2)^{r_j} \frac{(Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2)^{q_j - r_j}}{(q_j - r_j)!} \right\} \quad (4-8-4).$$

Pour démontrer 4-8-4, on utilise 1-2. Quant à 4-8-3, c' est un cas particulier de 3-3.

Maintenant, si on peut intervertir les signes Σ et \int , et si l'on peut développer f et g en séries de Mac Laurin, alors les coefficients de ces séries pourront s'exprimer en termes de moments.

Théorème 4-9 : On garde les notations et hypothèses de 4-8.

Si on suppose que $d_{r_1, r_2, \dots, r_p}^{\rho}$ converge vers un réel d_{r_1, r_2, \dots, r_p} , et que, dans le deuxième terme de l'égalité 4-8-3, lorsque $\rho = +\infty$, on peut intervertir les signes Σ et \int , l'intégration étant effectuée d'abord par rapport à X, puis par rapport à U, alors

$$d_{r_1, r_2, \dots, r_p} = \mathbb{E}_U \left\{ \mathbb{E}_X \left\{ \prod_{j=1}^p \frac{(Z_{j,1} + Z_{j,2} + \dots + Z_{j,n})^{r_j}}{r_j! n^{r_j} \sigma_j^{r_j}} \exp \left\{ - \frac{(Z_{j,1} + Z_{j,2} + \dots + Z_{j,n})^2}{2n} \right\} \right\} \right\} .$$

Si on suppose que $c_{r_1, r_2, \dots, r_p}^{\rho}$ converge vers un réel c_{r_1, r_2, \dots, r_p} , et que, dans le deuxième terme de l'égalité 4-8-4, lorsque $\rho = +\infty$, on peut intervertir les signes Σ et \int , l'intégration étant effectuée d'abord par rapport à X, puis par rapport à U, alors

$$c_{r_1, r_2, \dots, r_p} = \left(\prod_{j=1}^p \frac{\beta \left(\frac{1}{2} + r_j, \frac{n-1}{2} \right)}{(2r_j)! \sigma_j^{2r_j} \beta \left(\frac{1}{2}, \frac{n-1}{2} \right)} \right) \mathbb{E}_U \left\{ \mathbb{E}_X \left\{ \prod_{j=1}^p (Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2)^{r_j} e^{-\frac{Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2}{2}} \right\} \right\} .$$

La première égalité se déduit de 1-7, la deuxième est évidente.

IV-III : DENSITE SOUS FORME D'INTEGRALE

On poursuit l'étude du paragraphe précédent, et comme dans le chapitre III, on obtient maintenant les densités sous forme intégrale.

Théorème 4-10 : On garde les notations et hypothèses de 4-8.

Si on suppose que $f_p(y_1, y_2, \dots, y_p)$ converge vers $f(y_1, y_2, \dots, y_p)$ en tout point $(y_1, y_2, \dots, y_p) \in \mathbb{R}^p$, et que, dans le premier terme de l'égalité 4-8-1, lorsque $p \rightarrow +\infty$, pour tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}^p$, on peut intervertir les signes Σ et \mathbb{E} , l'intégration étant effectuée d'abord par rapport à X, puis par rapport à U, alors, pour tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}^p$,

$$f(y_1, y_2, \dots, y_p) = \mathbb{E}_U \left\{ \mathbb{E}_X \left\{ \prod_{j=1}^p \exp \left(- \frac{(Z_{j,1} + \dots + Z_{j,n})^2}{2n} + \frac{y_j (Z_{j,1} + \dots + Z_{j,n})}{n\sigma_j} \right) \right\} \right\} .$$

Si on suppose que $g_p(y_1, y_2, \dots, y_p)$ converge vers $g(y_1, y_2, \dots, y_p)$ en tout point $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$, et que, dans le premier terme de l'égalité 4-8-2, lorsque $p \rightarrow +\infty$, pour tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$, on peut intervertir les signes Σ et \mathbb{E} , l'intégration étant effectuée d'abord par rapport à B, puis par rapport à X, enfin par rapport à U, alors, pour tout $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$,

$$g(y_1, y_2, \dots, y_p) = \mathbb{E}_U \left\{ \mathbb{E}_X \left\{ \prod_{j=1}^p \left(\int_{t_j=0}^{t_j=1} e^{-\frac{Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2}{2}} \frac{2 \operatorname{ch} \left(\frac{t_j \sqrt{(Z_{j,1}^2 + Z_{j,2}^2 + \dots + Z_{j,n}^2)} y_j}{\sigma_j} \right)}{\beta \left(\frac{1}{2}, \frac{n-1}{2} \right)} (1-t_j^2)^{\frac{n-3}{2}} dt_j \right) \right\} \right\} .$$

La première assertion est un cas particulier de 3-2, (en fait, il suffit d'appliquer 1-7). Pour démontrer la seconde, il faut utiliser la définition des lois Bétas.

Comme exemple d'application de ces résultats, nous allons maintenant étudier le cas Gaussien. On pourra remarquer que, grâce à l'introduction des lois normales de "variances négatives", on n'aura plus à imposer un ordre d'intégration comme ci dessus.

V : APPLICATION AU CAS GAUSSIEN

Dans ce chapitre, on va appliquer les résultats précédents au cas Gaussien. La somme de la loi de n vecteurs aléatoires Gaussiens ne présentant pas beaucoup d'intérêt, on va donc s'intéresser uniquement aux formes quadratiques.

Bien entendu, les résultats du chapitre précédent restent vrais sous l'hypothèse Gaussienne. On peut donc écrire immédiatement la densité sous forme de série de polynômes de Laguerre. De plus, on constatera que cette série converge aussi pour la topologie de la convergence uniforme sur tout compact. Enfin, on obtiendra également cette densité sous forme de série de Mac Laurin, ou sous forme d'intégrale.

Pour cela, on commencera d'abord par rappeler quelques propriétés des vecteurs aléatoires Gaussiens. Ensuite on introduira les "variances négatives" dans les lois normales, et on appliquera enfin ces résultats aux formes quadratiques de vecteurs Gaussiens.

On commence par introduire les hypothèses que nous utiliserons tout au long de ce chapitre.

Hypothèses 5-1 : Dans ce chapitre, on supposera que X est un vecteur aléatoire Gaussien non dégénéré. Pour tout $(j,t) \in \{1,2,\dots,p\} \times \{1,2,\dots,n\}$, on note par $m_{j,t}$, l'espérance de $X_{j,t}$, et par $\sigma_{j,t}$, son écart-type. De plus, si $p=1$, on simplifie $m_{1,t}$ en m_t , et $\sigma_{1,t}$ en σ_t , et pour tout $(t,t') \in \{1,2,\dots,n\}^2$, on note $C_{t,t'}$, la covariance de X_t et $X_{t'}$, et $\rho_{t,t'}$, leur coefficient de corrélation linéaire. On note toujours par U , le vecteur aléatoire Gaussien indépendant de X défini en 4-1.

V-1 : QUELQUES PROPRIETES DES LOIS NORMALES

Dans ce paragraphe, on va rappeler quelques propriétés des lois Gaussiennes : d'abord, le développement de la densité en polynômes d'Hermite, ensuite la valeur des moments, enfin à quelle condition il existe une densité appartenant à L^2 .

D'abord, rappelons quel est le développement de la densité en polynômes d'Hermite.

Théorème 5-2 : On suppose $p=1$. Soit $N_{\otimes n} = N(0,1) \otimes N(0,1) \otimes \dots \otimes N(0,1)$. On suppose que, pour tout $k=1,2,\dots,n$, $m_k=0$ et $\sigma_k=1$. On suppose de plus que X admet une densité f par rapport à $N_{\otimes n}$, $f \in L^2(\mathbb{R}^n, N_{\otimes n})$. Pour tout $j \in \mathbb{N}$, on pose $\bar{H}_j(x) = \sqrt{j!} H_j(x)$, et pour tout $k \in \mathbb{N}^{n(n-1)/2}$ $k = ((k_{1,2}, k_{1,3}, \dots, k_{1,n}), (k_{2,3}, k_{2,4}, \dots, k_{2,n}), \dots, (k_{n-2,n-1}, k_{n-2,n}), (k_{n-1,n}))$, on définit le réel \mathfrak{H}_k par

$$\mathfrak{H}_k = \frac{\begin{pmatrix} k_{1,2} & k_{1,3} & & k_{1,n} \\ \rho_{1,2} & \rho_{1,3} & \dots & \rho_{1,n} \end{pmatrix} \begin{pmatrix} k_{2,3} & k_{2,4} & & k_{2,n} \\ \rho_{2,3} & \rho_{2,4} & \dots & \rho_{2,n} \end{pmatrix} \dots \begin{pmatrix} k_{n-2,n-1} & k_{n-2,n} \\ \rho_{n-2,n-1} & \rho_{n-2,n} \end{pmatrix} \begin{pmatrix} k_{n-1,n} \\ \rho_{n-1,n} \end{pmatrix}}{(k_{1,2}! k_{1,3}! \dots k_{1,n}!) (k_{2,3}! k_{2,4}! \dots k_{2,n}!) \dots (k_{n-2,n-1}! k_{n-2,n}!) (k_{n-1,n}!)}$$

Alors, $f(x_1, x_2, \dots, x_n) =$

$$\sum_{k \in \mathbb{N}^{n(n-1)/2}} \mathfrak{H}_k \bar{H}_{k_{1,2} + k_{1,3} + \dots + k_{1,n}}(x_1) \bar{H}_{k_{1,2} + k_{2,3} + \dots + k_{2,n}}(x_2) \dots \bar{H}_{k_{1,n} + k_{2,n} + \dots + k_{n-1,n}}(x_p) \cdot$$

Pour ce théorème, on peut se reporter à Johnson et Kotz, 1972 (21, p 45).

Comme exemple d'application, on peut voir que si $n=2$ $E[H_j(X_1)H_k(X_2)] = \delta_{j,k}e^j$, ou bien

$$\text{si } n=3, \quad f(x_1, x_2, x_3) = \sum_{i,j,k=0}^{\infty} \frac{e^{j_1} e^{j_2} e^{j_3}}{j_1! j_2! j_3!} \bar{H}_{j_1+k}(x_1) \bar{H}_{j_2+j_3}(x_2) \bar{H}_{k+j_3}(x_3).$$

Rappelons maintenant quelle est la valeur des moments des lois normales.

Théorème 5-3 : Soit X une variable aléatoire normale de loi $N(0, \sigma^2)$. Alors, pour tout $k \in \mathbb{N}$,

$$E[X^{2k+1}] = 0 \quad \text{et} \quad E[X^{2k}] = \frac{\sigma^{2k} 2k!}{2^k k!}.$$

Nous aurons également besoin de connaître les moments dans le cas multidimensionnel.

Théorème 5-4 : On suppose $p=1$, et que, pour tout $k=1,2,3,\dots,n$, $m_k=0$. Alors,

$$E[X_1^{j_1} X_2^{j_2} \dots X_n^{j_n}] = \sum_{r_1=0}^{\lfloor \frac{j_1}{2} \rfloor} \sum_{r_2=0}^{\lfloor \frac{j_2}{2} \rfloor} \dots \sum_{r_n=0}^{\lfloor \frac{j_n}{2} \rfloor} \mathcal{D}_{r_1, r_2, \dots, r_n} \frac{j_1! \sigma_1^{2r_1}}{2^{r_1} r_1!} \frac{j_2! \sigma_2^{2r_2}}{2^{r_2} r_2!} \dots \frac{j_n! \sigma_n^{2r_n}}{2^{r_n} r_n!},$$

où $\mathcal{D}_r = \sum_{|k|=j-2r} \frac{(C_{1,2}^{k_{1,2}} C_{1,3}^{k_{1,3}} \dots C_{1,n}^{k_{1,n}}) \cdot (C_{2,3}^{k_{2,3}} C_{2,4}^{k_{2,4}} \dots C_{2,n}^{k_{2,n}}) \dots (C_{n-2,n-1}^{k_{n-2,n-1}} C_{n-2,n}^{k_{n-2,n}}) (C_{n-1,n}^{k_{n-1,n}})}{(k_{1,2}! k_{1,3}! \dots k_{1,n}!) (k_{2,3}! k_{2,4}! \dots k_{2,n}!) \dots (k_{n-2,n-1}! k_{n-2,n}!) (k_{n-1,n}!)}$

la somme portant sur tous les $((k_{1,2}, k_{1,3}, \dots, k_{1,n}), (k_{2,3}, k_{2,4}, \dots, k_{2,n}), \dots, (k_{n-2,n-1}, k_{n-2,n}), (k_{n-1,n}))$ tels que $k_{1,2} + k_{1,3} + \dots + k_{1,n} = j_1 - 2r_1$, $k_{1,2} + k_{2,3} + \dots + k_{2,n} = j_2 - 2r_2, \dots$, $k_{1,n} + k_{2,n} + \dots + k_{n-1,n} = j_n - 2r_n$.

On démontre ce résultat en remarquant que, lorsque les conditions de 5-2 sont vérifiées,

$$\frac{\sigma_1^{j_1} \sigma_2^{j_2} \dots \sigma_n^{j_n}}{\sqrt{j_1! j_2! \dots j_n!}} E\{H_{j_1}(X_1) H_{j_2}(X_2) \dots H_{j_n}(X_n)\}$$

$$= \sum_{|k|=j} \frac{(C_{1,2}^{k_{1,2}} C_{1,3}^{k_{1,3}} \dots C_{1,n}^{k_{1,n}}) \cdot (C_{2,3}^{k_{2,3}} C_{2,4}^{k_{2,4}} \dots C_{2,n}^{k_{2,n}}) \dots (C_{n-2,n-1}^{k_{n-2,n-1}} C_{n-2,n}^{k_{n-2,n}}) (C_{n-1,n}^{k_{n-1,n}})}{(k_{1,2}! k_{1,3}! \dots k_{1,n}!) (k_{2,3}! k_{2,4}! \dots k_{2,n}!) \dots (k_{n-2,n-1}! k_{n-2,n}!) (k_{n-1,n}!)},$$

la somme portant sur tous les $((k_{1,2}, k_{1,3}, \dots, k_{1,n}), (k_{2,3}, k_{2,4}, \dots, k_{2,n}), \dots, (k_{n-2,n-1}, k_{n-2,n}), (k_{n-1,n}))$ tels que $k_{1,2} + k_{1,3} + \dots + k_{1,n} = j_1$, $k_{1,2} + k_{2,3} + \dots + k_{2,n} = j_2, \dots$ et $k_{1,n} + k_{2,n} + \dots + k_{n-1,n} = j_n$.

Il ne reste plus alors qu'à employer 1-11 pour en déduire le résultat.

On remarque alors, que, même s'il n'existe pas une densité, les moments de X sont une fonction polynômiale des σ_i et des $C_{i,j}$. On en déduit que ces résultats restent vrais même si les hypothèses de 5-2 ne sont pas vérifiées.

En particulier, si $n=2$, $\mathbb{E}[X^J X^K] = 0$ si $J+K$ impair,

$$\mathbb{E}[X_1^{2j} X_2^{2k}] = \frac{2^j 2^k}{2^{j+k}} \sum_{t=0}^j \frac{\sigma_1^{2j-2t} \sigma_2^{2k-2t}}{(j-t)!(k-t)!2^t} \left(\frac{2\mathbb{E}(X_1 X_2)}{\sigma_1 \sigma_2} \right)^{2t} \quad \text{si } j \leq k,$$

$$\mathbb{E}[X_1^{2j+1} X_2^{2k+1}] = \frac{(2j+1)!(2k+1)!}{2^{j+k}} \sum_{t=0}^j \frac{4^t \sigma_1^{2j+1-2t} \sigma_2^{2k+1-2t}}{(j-t)!(k-t)!(2t+1)!} \left(\frac{\mathbb{E}(X_1 X_2)}{\sigma_1 \sigma_2} \right)^{2t+1} \quad \text{si } j \leq k,$$

(Voir aussi Johnson et Kotz, 1972, p 91).

Rappelons maintenant à quelle condition il existe une densité appartenant à L^2 .

Théorème 5-5 : Soit X une variable aléatoire $N(m, \sigma^2)$. Alors si $\sigma^2 \neq 0$, X admet une densité f par rapport à la loi $N(0,1)$. De plus, pour que $f \in L^2(\mathbb{R}, N(0,1))$, il faut et il suffit que $2 \geq \sigma^2$ si $m < 0$, et que $2 > \sigma^2$ si $m \geq 0$.

En effet, $f(x) = \frac{e^{-\frac{(1-\sigma^2)x^2 - 2mx + m^2}{2\sigma^2}}}{\sigma}$, donc, $f(x)^2 e^{-\frac{x^2}{2}} = \frac{e^{-\frac{2x^2 - \sigma^2 x^2 - 4mx + 2m^2}{2\sigma^2}}}{\sigma^2}$.

Remarquons que si $1 \geq \sigma^2$, le cas est encore plus simple vu que f est bornée.

Dans le cas multidimensionnel, ce résultat est plus compliqué à obtenir. Aussi, on va normaliser X de façon à être toujours sous les bonnes hypothèses.

Théorème 5-6 : Pour tout $j=1,2,\dots,p$, on pose $\mathcal{J}_j^2 = \sup_{s=1,2,\dots,n} ((np+1)\sigma_{j,s}^2)$, et on définit \tilde{X} par :

$$\tilde{X} = \left\{ \frac{X_{1,1}}{\mathcal{J}_1}, \frac{X_{1,2}}{\mathcal{J}_1}, \dots, \frac{X_{1,n}}{\mathcal{J}_1}, \frac{X_{2,1}}{\mathcal{J}_2}, \frac{X_{2,2}}{\mathcal{J}_2}, \dots, \frac{X_{2,n}}{\mathcal{J}_2}, \dots, \frac{X_{p,1}}{\mathcal{J}_p}, \frac{X_{p,2}}{\mathcal{J}_p}, \dots, \frac{X_{p,n}}{\mathcal{J}_p} \right\}.$$

On note par $\mathcal{C} = \{\mathcal{C}_{i,t,j,s}\}$ la matrice de covariances de \tilde{X} , et par I la matrice identité sur \mathbb{R}^{np} . Alors, $I - \mathcal{C}$ est la matrice de covariances d'un vecteur aléatoire Gaussien centré non dégénéré, $T = \{(T_{1,1}, T_{1,2}, \dots, T_{1,n}), (T_{2,1}, T_{2,2}, \dots, T_{2,n}), \dots, (T_{p,1}, T_{p,2}, \dots, T_{p,n})\}$.

Démonstration : Il est clair que l'on a l'inégalité : $\left| \mathbb{E} \left\{ \frac{X_{i,t} - \mathbb{E}(X_{i,t})}{\mathcal{J}_i} \frac{X_{j,s} - \mathbb{E}(X_{j,s})}{\mathcal{J}_j} \right\} \right| \leq \frac{1}{np+1}$.

On en déduit que $\left| \frac{\mathcal{C}_{i,t,j,s}}{\sqrt{(1-\mathcal{C}_{i,t,i,t})(1-\mathcal{C}_{j,s,j,s})}} \right| \leq \frac{1}{np}$ si $(i,t) \neq (j,s)$.

Aussi, on utilise alors le lemme suivant.

Lemme 5-7 : Soit $(.,.)$ une forme bilinéaire symétrique sur un espace vectoriel réel E de dimension n . On suppose qu'il existe une base de E , e_1, e_2, \dots, e_n , telle que $(e_s, e_s) = 1$ et $|(e_s, e_t)| \leq 1/n$ si $s \neq t$. Alors, il existe une base b_1, b_2, \dots, b_n de E telle que $(b_s, b_t) = \delta_{s,t}$.

Démonstration : On démontre ce résultat par récurrence sur n . Pour $n = 1$, il est évident.

On suppose ce résultat vrai pour tout $q < n$. On pose $g_j = \frac{e_j - (e_1, e_j)e_1}{\sqrt{1 - (e_1, e_j)^2}}$ pour $j \geq 2$.

$$\text{Alors, si } i \neq j, \quad |(g_i, g_j)| = \left| \frac{(e_i, e_j) - (e_1, e_i)(e_1, e_j)}{\sqrt{1 - (e_1, e_i)^2} \sqrt{1 - (e_1, e_j)^2}} \right| \leq \frac{|(e_i, e_j)| + |(e_1, e_i)(e_1, e_j)|}{\frac{n^2 - 1}{n^2}} \leq \frac{1}{n-1},$$

et, $(g_i, g_i) = 1$.

Soit G le sous espace de E engendré par g_2, g_3, \dots, g_n . D'après l'hypothèse de récurrence, il existe une base b_s de G , $s \geq 2$, telle que $(b_s, b_t) = \delta_{s,t}$. On en déduit le résultat.

On en déduit que $I-\mathcal{C}$ est définie positive. Donc, il existe une matrice inversible D telle que $D'(I-\mathcal{C})D = I$. Alors, le vecteur aléatoire UD^{-1} répond à la question.

Maintenant, on peut démontrer que \tilde{X} admet une densité appartenant à L^2 .

Théorème 5-8 : On garde les notations de 5-6. Alors, \tilde{X} admet une densité f par rapport à $N_{\otimes np}$, la loi normale centrée sur \mathbb{R}^{np} , de matrice de variances-covariances, la matrice identité : $f \in L^2(\mathbb{R}^{np}, N_{\otimes np})$.

Démonstration : Posons $m = \mathbb{E}(X)$. Comme X est non dégénéré, on sait que \tilde{X} admet une densité par rapport à la mesure de Lebesgue. Donc, \tilde{X} admet une densité f par rapport à $N_{\otimes np}$:

$$f(x) = |\mathcal{C}|^{-1/2} e^{-\frac{xx'}{2}} e^{-\frac{(x-m)\mathcal{C}^{-1}(x-m)'}{2}} = |\mathcal{C}|^{-1/2} e^{-\frac{x(\mathcal{C}^{-1}-I)x'}{2}} e^{-\frac{m\mathcal{C}^{-1}x' + x\mathcal{C}^{-1}m'}{2}} e^{-\frac{m\mathcal{C}^{-1}m'}{2}}.$$

De plus, il existe une matrice B inversible telle que $\mathcal{C} = BB'$. Donc, $B'(\mathcal{C}^{-1}-I)B = I - B'B$. Remarquons alors que $\det(I - BB' - xI) = \det(BB^{-1}(1-x) - BB')$ $= \det(B) \cdot \det(B^{-1}(1-x) - B')$ $= \det([B^{-1}(1-x) - B']B) = \det(I - B'B - xI)$. Donc, les polynômes caractéristiques $\det(I - BB' - xI)$ et $\det(I - B'B - xI)$ ont les mêmes racines. Donc, comme $I - \mathcal{C}$ est définie positive, $I - B'B$ est aussi définie positive. Donc, il existe une matrice D telle que $D'(I - B'B)D = I$. Donc, $(BD)'(\mathcal{C}^{-1}-I)BD = I$. Donc, $\mathcal{C}^{-1}-I$ est définie positive. Donc, $|e^{-x'(\mathcal{C}^{-1}-I)x}| \leq 1$. On en déduit le résultat.

V-II: VARIANCES NEGATIVES DANS LES LOIS NORMALES

On va étudier maintenant les variances négatives dans les lois normales. Bien sûr, cette appellation est abusive. Cependant, c'est elle qui résume le mieux les résultats que nous allons énoncer maintenant.

En fait, ce que l'on va constater, c'est que la variable complexe $iU+X$ a les mêmes moments qu'une variable réelle Y , ou purement imaginaire, iY .

Théorème 5-9 : Soit X une variable aléatoire normale de loi $N(0, \sigma^2)$. Soit U une variable aléatoire normale de loi $N(0,1)$ indépendante de X . On pose $S^2 = \sigma^2 - 1$. Alors, pour tout $k \in \mathbb{N}$,

$$\mathbb{E} [(iU+X)^{2k+1}] = 0 \quad \text{et} \quad \mathbb{E} [(iU+X)^{2k}] = \frac{S^{2k} 2k!}{2^k k!} .$$

Pour démontrer ce résultat, il suffit d'utiliser 5-3.

Ce résultat suggère immédiatement que l'on doit pouvoir considérer $iU+X$ comme une variable normale de variance S , S étant éventuellement négative.

Remarquons que l'on généralise facilement l'énoncé de ce théorème au cas non centré: en effet, si $\mathbb{E}(X) = m \neq 0$, on obtient les moments de $iU+X$ exactement comme ceux d'une variable aléatoire réelle en remplaçant les variances σ^2 par S^2 . On peut aussi remarquer que cela reste vrai lorsque $S=0$, c'est à dire pour une variable normale dégénérée concentrée en m : dans ce cas, $\mathbb{E} ((iU+X)^k) = m^k$.

Donc, en fait, ce que l'on entend par "variances éventuellement négatives", c'est tout simplement que les moments d'ordre deux puissent être négatifs. Si on a adopté ici cet abus de langage, c'est à cause du théorème 5-9 qui suggère immédiatement cette appellation. De plus, les lois normales centrées sont entièrement définies par un seul paramètre: la variance. Aussi, la façon la plus simple de résumer 5-9, c'est bien de dire que $iU+X$ se comporte comme une loi normale de "variance éventuellement négative".

Cette remarque s'imposait particulièrement dans le cadre de cette étude où l'on n'a justement besoin que de la valeur des moments. Or, c'est cette constatation qui a donné l'idée de remplacer ces variables $iU+X$ par des variables iY , et c'est à partir de là que l'on a pu obtenir les résultats du paragraphe suivant, et ce, sous une forme que l'on ne pouvait espérer autrement.

D'autre part, il faut aussi signaler que les variances négatives ont également permis de démontrer le théorème suivant.

Théorème 5-10 : Sous les hypothèses de 5-9, et si $S \neq 0$,

$$\mathbb{E} \left[H_j \left(\frac{iU+X}{S} \right) H_k \left(\frac{iU+X}{S} \right) \right] = \delta_{j,k} .$$

Pour démontrer ce résultat, on remarque que $\mathbb{E} [(iU+X)^k] = \frac{S_k}{k} \mathbb{E} [X^k]$ et on écrit $H_j(x) = \sum_{s=0}^j a_{j,s} x^s$.

Ce théorème signifie donc que les H_j sont aussi orthonormaux pour les variables aléatoires $(iU+X)/S$. Ce résultat est à rapprocher de la définition des polynômes d'Hermite complexes (qui sont des polynômes à deux variables : cf Neveu, p 149).

Maintenant, pour mener à bien l'étude des formes quadratiques de vecteurs Gaussiens, il nous faudra les moments de $iU+X$ dans le cas vectoriel. Pour calculer ceux-ci, on aura besoin de la propriété suivante

Lemme 5-11 : On suppose $p=1$, et que, pour tout $k=1,2,\dots,n$, $m_k=0$ et $S_k^2 = \sigma_k^2 - 1 \neq 0$.

$$\text{Alors, pour tout } (j_1, j_2, \dots, j_n) \in \mathbb{N}^n, \quad \mathbb{E} \left\{ \prod_{k=1}^n H_{j_k} \left(\frac{iU_k + X_k}{S_k} \right) \right\} = \left(\prod_{k=1}^n \left(\frac{\sigma_k}{S_k} \right)^{j_k} \right) \mathbb{E} \left\{ \prod_{k=1}^n H_{j_k} \left(\frac{X_k}{\sigma_k} \right) \right\}.$$

$$\text{Pour démontrer ce théorème, il suffit d'écrire } H_{j_k} \left(\frac{iU_k + X_k}{S_k} \right) = \frac{1}{\sqrt{j_k!}} \mathbb{E}_{v_k} \left\{ \left(iV_k + \frac{iU_k + X_k}{S_k} \right)^{j_k} \right\}.$$

On va voir maintenant que, dans le cas multidimensionnel aussi, on obtient les moments de $iU+X$ en remplaçant les variances σ_k^2 par les S_k^2 .

Théorème 5-12 : On garde les notations et hypothèses de 5-4 et 5-11. Alors, pour tout $(j_1, j_2, \dots, j_n) \in \mathbb{N}^n$, les moments de $iU+X$ vérifient les égalités suivantes:

$$\mathbb{E} [(iU_1 + X_1)^{j_1} \dots (iU_n + X_n)^{j_n}] = \sum_{r_1=0}^{\lfloor \frac{j_1}{2} \rfloor} \sum_{r_2=0}^{\lfloor \frac{j_2}{2} \rfloor} \dots \sum_{r_n=0}^{\lfloor \frac{j_n}{2} \rfloor} \mathfrak{D}_{r_1, r_2, \dots, r_n} \frac{j_1! S_1^{2r_1}}{2^{r_1} r_1!} \frac{j_2! S_2^{2r_2}}{2^{r_2} r_2!} \dots \frac{j_n! S_n^{2r_n}}{2^{r_n} r_n!}.$$

On démontre ce résultat en utilisant 5-11 et en raisonnant de la même manière qu'en 5-4.

Malheureusement, dans le cas multidimensionnel, il n'est pas forcé qu'il existe un vecteur aléatoire Gaussien purement réel ou purement imaginaire ayant les même moments que $iU+X$. En effet, la matrice de covariances de $iU+X$ est C-I qui n'est pas forcément définie positive ou définie négative. Par contre, ce sera le cas de $iU+\tilde{X}$ qui, dans le cas centré, a les même moments que iT . Plus, précisément, en utilisant 5-4 et 5-12, on démontre le théorème suivant.

Théorème 5-13 : On garde les notations de 5-6. Alors, pour tout $(j_{1,1}, j_{1,2}, \dots, j_{p,n}) \in \mathbb{N}^{np}$,

$$\mathbb{E} \left\{ \prod_{j=1}^p \prod_{s=1}^n \left(iU_{j,t} + \frac{X_{j,t}}{\mathcal{F}_j} \right)^{j_{j,t}} \right\} = \mathbb{E} \left\{ \prod_{j=1}^p \prod_{s=1}^n \left(iT_{j,t} + \frac{m_{j,t}}{\mathcal{F}_j} \right)^{j_{j,t}} \right\}.$$

V-III : FORMES QUADRATIQUES DE VECTEURS GAUSSIENS

On applique maintenant les résultats précédents et on va obtenir ainsi, la loi des formes quadratiques de vecteurs Gaussiens, et ce, sous forme de séries de polynômes de laguerre, de séries de Mac Laurin, ou sous forme d'intégrales.

Rappelons d'abord que divers résultats de ce type sont connus dans la littérature dans le cas unidimensionnel. Ainsi, Tiku a donné sous forme de séries de polynômes de laguerre, la loi du chi deux décentré. Plus généralement, Gurland et Kotz, Johnson, Boyd ont montré que, si l'on a n variables aléatoires de lois Gaussiennes centrées réduites indépendantes, U_i , alors la densité de

$$\sum_{i=1}^n \lambda_i (U_i - \omega_i)^2 \text{ par rapport à la mesure de Lebesgue est égale à } \left(\sum_{j=0}^{\infty} d_j L_j^{n/2} \left(\frac{y}{\beta} \right) \right) \frac{y^{n/2-1} e^{-y/2}}{(2\beta)^{n/2} \Gamma(n/2)},$$

avec $d_j = (-1)^j \sqrt{\frac{j! \Gamma(n/2)}{\Gamma(n/2+j)}} c'_j$, où c'_j est défini par récurrence par : $c'_0=1$, $c'_r = \frac{1}{2r} \sum_{j=0}^{r-1} S'_{r-j} c'_j$,

avec $S'_r = -\frac{r}{\beta} \sum_{j=1}^n \lambda_j \omega_j^2 \left(1 - \frac{\lambda_j}{\beta}\right)^{r-1} + \sum_{j=1}^n \left(1 - \frac{\lambda_j}{\beta}\right)^r$ (Johnson, Kotz, 1970, p 172).

Pour faciliter la lecture des résultats de ce paragraphe, on récapitule d'abord les différentes notations et hypothèses utilisées tout au long de ce chapitre.

Notations 5-14 : Soit $X = \{(X_{1,1}, X_{1,2}, \dots, X_{1,n}), \dots, (X_{p,1}, X_{p,2}, \dots, X_{p,n})\}$, un vecteur aléatoire Gaussien non dégénéré à valeurs dans \mathbb{R}^{np} . On note toujours par $m_{j,t}$ et $\sigma_{j,t}$, l'espérance et l'écart-type de $X_{j,t}$. Pour tout $j \in \{1, 2, \dots, p\}$, on pose alors $\mathcal{J}_j^2 = \sup_{s=1, 2, \dots, n} ((np+1)\sigma_{j,s}^2)$. De

plus, on définit la matrice \mathcal{C} par : $\mathcal{C}_{i,s;j,t} = \mathbb{E} \left\{ \frac{X_{i,s} - \mathbb{E}(X_{i,s})}{\mathcal{J}_i} \frac{X_{j,t} - \mathbb{E}(X_{j,t})}{\mathcal{J}_j} \right\}$.

Soit I la matrice identité d'ordre np , et soit T un vecteur aléatoire Gaussien centré réel non dégénéré, $T = \{(T_{1,1}, T_{1,2}, \dots, T_{1,n}), (T_{2,1}, T_{2,2}, \dots, T_{2,n}), \dots, (T_{p,1}, T_{p,2}, \dots, T_{p,n})\}$, T admettant $I - \mathcal{C}$ pour matrice de covariances. Soient alors p variables aléatoires B_1, B_2, \dots, B_p , indépendantes entre elles et indépendantes de T , et suivant chacune la loi Béta $(1/2, (n-1)/2)$. Alors, pour tout $(i,s) \in \{1, 2, \dots, p\} \times \{1, 2, \dots, n\}$, on définit la variable complexe $Y_{i,s}$ par : $Y_{i,s} = iT_{i,s} + \frac{m_{i,s}}{\mathcal{J}_s}$.

On peut alors énoncer les résultats en commençant par affiner le théorème 4-7 .

Théorème 5-15: On garde les notations et hypothèses de 5-14. Alors le vecteur aléatoire

$$\left\{ \frac{X_{1,1}^2 + X_{1,2}^2 + \dots + X_{1,n}^2}{\varphi_1^2}, \frac{X_{2,1}^2 + X_{2,2}^2 + \dots + X_{2,n}^2}{\varphi_2^2}, \dots, \frac{X_{p,1}^2 + X_{p,2}^2 + \dots + X_{p,n}^2}{\varphi_p^2} \right\}$$

admet une densité g par rapport à la loi du chi-deux produit $\chi^2_{\otimes p} = \gamma(n/2, 2) \otimes \gamma(n/2, 2) \otimes \dots \otimes \gamma(n/2, 2)$, $g \in L^2(\mathbb{R}^p, \chi^2_{\otimes p})$.

Alors, pour la topologie de la convergence uniforme sur tout compact de \mathbb{R}_+^p ,

$$g(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathbb{E} \left\{ \prod_{j=1}^p [Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2]^{q_j} \right\} \frac{\mathbb{E} \{ H_{2q_1}(\sqrt{B_1 y_1}) H_{2q_2}(\sqrt{B_2 y_2}) \dots H_{2q_p}(\sqrt{B_p y_p}) \}}{\sqrt{2^{q_1}! 2^{q_2}! \dots 2^{q_p}!}}$$

Maintenant, on peut écrire la densité sous forme d'intégrale.

Théorème 5-16 : On garde les hypothèses de 5-15. Alors, en tout point $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$,

$$g(y_1, y_2, \dots, y_p) = \mathbb{E} \left\{ \prod_{j=1}^p e^{-\frac{Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2}{2}} \operatorname{ch} \left(\sqrt{B_j (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)} y_j \right) \right\}$$

Enfin, on obtient aussi g sous forme de série de Mac Laurin.

Théorème 5-17 : On garde les hypothèses de 5-15. Alors, en tout point $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$,

$$g(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} d_{q_1, q_2, \dots, q_p} y_1^{q_1} y_2^{q_2} \dots y_p^{q_p}$$

$$\text{où } d_{q_1, q_2, \dots, q_p} = \mathbb{E} \left\{ \prod_{j=1}^p e^{-\frac{Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2}{2}} (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \right\} \left(\prod_{j=1}^p \frac{\Gamma(n/2) \Gamma(1/2 + q_j)}{\Gamma(1/2) \Gamma(n/2 + q_j) 2^{q_j}!} \right)$$

Remarquons que la forme de ce théorème rappelle celle de certains résultats du cas unidimensionnel (cf Robbins. Voir aussi Johnston, kotz, 1970, p 162).

On déduit alors de 5-17 la densité par rapport à la mesure de Lebesgue.

Théorème 5-18 : On garde les notations et hypothèses de 5-17. Alors le vecteur aléatoire

$$\left\{ \frac{X_{1,1}^2 + X_{1,2}^2 + \dots + X_{1,n}^2}{\varphi_1^2}, \frac{X_{2,1}^2 + X_{2,2}^2 + \dots + X_{2,n}^2}{\varphi_2^2}, \dots, \frac{X_{p,1}^2 + X_{p,2}^2 + \dots + X_{p,n}^2}{\varphi_p^2} \right\}$$

admet une densité par rapport à la mesure de Lebesgue égale à, en tout point $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$,

$$\left(\sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} d_{q_1, q_2, \dots, q_p} y_1^{q_1} y_2^{q_2} \dots y_p^{q_p} \right) \frac{(y_1 y_2 \dots y_p)^{n/2-1} e^{-\frac{y_1 + y_2 + \dots + y_p}{2}}}{2^{\frac{np}{2}} \left[\Gamma\left(\frac{n}{2}\right) \right]^p}$$

Remarquons qu'il n'aurait pas été possible d'obtenir ces résultats sans utiliser les lois normales de variances négatives: si l'on remplaçait les Y par les Z (i.e. si l'on n'avait pas fait l'opération inverse) pour que les intégrales des théorèmes 5-16 et 5-17 soient définies, il faudrait imposer un ordre d'intégration comme en 4-9 et 4-10.

Démonstration de 5-15 et 5-16 : On démontre d'abord le lemme suivant.

Lemme 5-19 : Soit $\delta_K = \mathbb{E} \left\{ e^{K \frac{T_{1,1}^2 + T_{1,2}^2 + \dots + T_{1,n}^2 + T_{2,1}^2 + T_{2,2}^2 + \dots + T_{2,n}^2 + \dots + T_{p,1}^2 + T_{p,2}^2 + \dots + T_{p,n}^2}{2}} \right\}$.

Alors, il existe $K_0 > 1$ tel que, pour tout $K \leq K_0$, $\delta_K \leq +\infty$.

Démonstration : Soit $K > 0$. On sait qu'il existe une matrice G inversible telle que $I - \mathcal{C} = GG'$. On effectue le changement de variables $t = Gu$.

$$\text{Alors, } \delta_K = \frac{|I - \mathcal{C}|^{-1/2}}{\sqrt{(2\pi)^{np}}} \int e^{K \frac{t't}{2}} e^{-\frac{t'(I - \mathcal{C})^{-1}t}{2}} dt = \frac{|I - \mathcal{C}|^{-1/2} |G|}{\sqrt{(2\pi)^{np}}} \int e^{K \frac{u'G'Gu}{2}} e^{-\frac{u'u}{2}} du.$$

On démontre, comme en 5-8 que $KG'G - I$ et $KGG' - I$ ont les mêmes valeurs propres.

Donc, $KG'G - I$ sera définie négative si et seulement si $K\mathcal{C} - (K-1)I$ est définie positive, ce qui est le cas dès que $K \leq 1$. Supposons $K > 1$, et faisons tendre K vers 1. Considérons un intervalle ouvert $]a, b[$, $a > 0$, contenant toutes les racines du polynôme caractéristique Q de \mathcal{C} . Sur $]a, b[$, Q_K , le polynôme caractéristique de $K\mathcal{C} - (K-1)I$ tend uniformément vers Q. On en déduit qu'il existe un K_0 tel que pour tout $K \leq K_0$, les racines de Q_K appartiennent toutes à $]a, b[$, i.e. $K\mathcal{C} - (K-1)I$ est définie positive. Donc, il existe une matrice D telle que $D'(KG'G - I)D = -I$.

$$\text{Alors, } \int e^{K \frac{u'G'Gu}{2}} e^{-\frac{u'u}{2}} du = |D| \int e^{-\frac{v'v}{2}} dv < +\infty.$$

Maintenant, d'après 4-7, 5-8 et 5-13, il est clair que, dans $L^2(\mathbb{R}^p, \chi^2_{\otimes p})$,

$$g(y_1, y_2, \dots, y_p) = \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathbb{E} \left\{ \prod_{j=1}^p (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \frac{H_{2q_j}(\sqrt{B_j y_j})}{\sqrt{2q_j!}} \right\}.$$

On se place alors en un point $(y_1, y_2, \dots, y_p) \in \mathbb{R}_+^p$, appartenant à un compact \mathcal{K}^p , $\mathcal{K} \subset \mathbb{R}_+$.

On sait qu'il existe J tel que, pour tout $q \geq J$, et tout $s=1, 2, \dots, p$, $|H_{2q}(\sqrt{B_s y'_s})| \leq \frac{K_0^q \sqrt{2q!}}{2^q q!}$,

pour tout $(y'_1, y'_2, \dots, y'_p) \in \mathcal{K}^p$ (cf 1-6).

$$\text{Donc, } \mathbb{E} \left\{ \sum_{(q_1 \geq J, q_2 \geq J, \dots, q_p \geq J)} \left| \prod_{j=1}^p (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \frac{H_{2q_j}(\sqrt{B_j y'_j})}{\sqrt{2q_j!}} \right| \right\}$$

$$\leq \mathbb{E} \left\{ \sum_{(q_1 \geq J, q_2 \geq J, \dots, q_p \geq J)} \prod_{j=1}^p \left(T_{j,1}^2 + \frac{m_{j,1}^2}{\varphi_j^2} + T_{j,2}^2 + \frac{m_{j,2}^2}{\varphi_j^2} + \dots + T_{j,n}^2 + \frac{m_{j,n}^2}{\varphi_j^2} \right)^{q_j} \frac{K_0^{q_j}}{2^{q_j} q_j!} \right\}$$

$$\leq e^{K_0 \frac{\sum_{j=1}^p \sum_{s=1}^n \frac{m_{j,s}^2}{\varphi_j^2}}{2}} \mathbb{E} \left\{ e^{K_0 \frac{\sum_{j=1}^p \sum_{s=1}^n T_{j,s}^2}{2}} \right\} < +\infty.$$

$$\text{Donc, } \mathbb{E} \left\{ \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \left| \prod_{j=1}^p (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \frac{H_{2q_j}(\sqrt{B_j y'_j})}{\sqrt{2q_j!}} \right| \right\} \text{ est majorée par une}$$

constante finie ne dépendant que de \mathcal{K} pour tout $(y'_1, y'_2, \dots, y'_p) \in \mathcal{K}^p$.

$$\text{Donc, en employant 1-7, } \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \mathbb{E} \left\{ \prod_{j=1}^p (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \frac{H_{2q_j}(\sqrt{B_j y_j})}{\sqrt{2q_j!}} \right\}$$

$$= \mathbb{E} \left\{ \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \left(\prod_{j=1}^p (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)^{q_j} \frac{H_{2q_j}(\sqrt{B_j y_j})}{\sqrt{2q_j!}} \right) \right\}$$

$$= \mathbb{E} \left\{ \prod_{j=1}^p e^{-\frac{Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2}{2}} \text{ch} \left(\sqrt{B_j (Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2)} y_j \right) \right\} \stackrel{\text{déf}}{=} g^*(y_1, y_2, \dots, y_p).$$

De plus, comme la majoration ci-avant est vraie sur tout compact \mathcal{K} , la convergence est uniforme sur tout \mathcal{K}^p . Donc, $g=g^*$, $\chi^2_{\otimes p}$ -presque partout.

Démonstration de 5-17 : On peut écrire les inégalités suivantes :

$$\begin{aligned} & \mathbb{E} \left\{ \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \left| \prod_{j=1}^p e^{-\frac{Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2}{2}} \frac{\sqrt{B_j(Y_{j,1}^2 + Y_{j,2}^2 + \dots + Y_{j,n}^2) y_j}}{2q_j!} \right| \right\} \\ & \leq e^{-\frac{\sum_{j=1}^p \sum_{s=1}^n \frac{m_{j,s}^2}{\varphi_j^2}}{2}} \mathbb{E} \left\{ e^{\frac{\sum_{j=1}^p \sum_{s=1}^n T_{j,s}^2}{2}} \sum_{(q_1, q_2, \dots, q_p) \in \mathbb{N}^p} \prod_{j=1}^p \frac{y_j^{q_j} \left(\sum_{s=1}^n \left(T_{j,s}^2 + \frac{m_{j,s}^2}{\varphi_j^2} \right) \right)^{q_j}}{2q_j!} \right\} \\ & = e^{-\frac{\sum_{j,s} \frac{m_{j,s}^2}{\varphi_j^2}}{2}} \mathbb{E} \left\{ e^{\frac{\sum_{j,s} T_{j,s}^2}{2}} \prod_{j=1}^p \operatorname{ch} \left\{ \sqrt{\sum_{s=1}^n \left(T_{j,s}^2 + \frac{m_{j,s}^2}{\varphi_j^2} \right)} y_j \right\} \right\} \\ & \leq e^{-\frac{\sum_{j,s} \frac{m_{j,s}^2}{\varphi_j^2}}{2}} \mathbb{E} \left\{ \prod_{j=1}^p e^{\sqrt{y_j} \sqrt{\sum_{s=1}^n T_{j,s}^2 + \frac{\sum_{s=1}^n m_{j,s}^2}{\varphi_j^2} + \frac{\sum_{s=1}^n T_{j,s}^2}{2}}} \right\} \\ & \leq e^{\sum_j \sqrt{y_j} \sqrt{\sum_s \frac{m_{j,s}^2}{\varphi_j^2}}} e^{-\frac{\sum_{s,j} \frac{m_{j,s}^2}{\varphi_j^2}}{2}} \mathbb{E} \left\{ \prod_{j=1}^p e^{\sqrt{y_j} \sqrt{\sum_s T_{j,s}^2 + \frac{\sum_s T_{j,s}^2}{2}}} \right\} . \end{aligned}$$

$$\text{Alors, } \mathbb{E} \left\{ \prod_{j=1}^p e^{\sqrt{y_j} \sqrt{\sum_s T_{j,s}^2 + \frac{\sum_s T_{j,s}^2}{2}}} \right\} = \sum_{\mathcal{D} \subset \{1, 2, \dots, p\}} \int_{\mathcal{J}_{\mathcal{D}}} \left(\prod_{j=1}^p e^{\sqrt{y_j} \sqrt{\sum_s T_{j,s}^2 + \frac{\sum_s T_{j,s}^2}{2}}} \right) d\mathbf{P} ,$$

$$\text{où } \mathcal{J}_{\mathcal{D}} = \left\{ \sqrt{\sum_s T_{j,s}^2} < \frac{2\sqrt{y_j}}{K_0 - 1} \text{ si } j \in \mathcal{D} \text{ et } \frac{2\sqrt{y_j}}{K_0 - 1} \leq \sqrt{\sum_s T_{j,s}^2} \text{ si } j \in \{1, \dots, p\} \setminus \mathcal{D} \right\} .$$

$$\begin{aligned}
\text{Or, } \int_{J_{\mathfrak{D}}} \left(\prod_{j=1}^p e^{\sqrt{\frac{\sum y_j T_{j,s}^2}{s} + \frac{\sum T_{j,s}^2}{2}}} \right) d\mathbf{P} &= \int_{J_{\mathfrak{D}}} \left(\prod_{j \in \mathfrak{D}} e^{\sqrt{\frac{\sum y_j T_{j,s}^2}{s} + \frac{\sum T_{j,s}^2}{2}}} \right) \left(\prod_{j \notin \mathfrak{D}} e^{\sqrt{\frac{\sum y_j T_{j,s}^2}{s} + \frac{\sum T_{j,s}^2}{2}}} \right) d\mathbf{P} \\
&\leq \int_{J_{\mathfrak{D}}} C_0^{te} \left(\prod_{j \in \mathfrak{D}} e^{K_0 \frac{\sum T_{j,s}^2}{2}} \right) d\mathbf{P} \leq C_0^{te} \mathbb{E} \left(\prod_{j \in \mathfrak{D}} e^{K_0 \frac{\sum T_{j,s}^2}{2}} \right) < +\infty \quad (C_0^{te} \text{ représentant une constante}).
\end{aligned}$$

On en déduit le résultat.

BIBLIOGRAPHIE

- BLACHER R. (1985) Loi de la somme des carrés de n variables aléatoires. TIM3-IMAG , R.R. n° 522 .
- BLACHER R. (1989) Loi de la somme de n variables aléatoires et théorème de la limite centrale par les coefficients de dépendance polynômiale et les moments. TIM3-IMAG , R.R. n° 768-M .
- CHIHARA T. S. (1978) An introduction to orthogonal polynomials. Gordon and Breach, Science publishers. New-york, Paris.
- CUICHARDET (1969) Calcul Intégral (2° édition). Armand Colin, Paris.
- GURLAND J (1953) . Distribution of Quadratic Forms and Ratios of Quadratic Forms. Annals of Mathematical Statistics, 24, 416-427.
- GURLAND J (1955, 1962) . Distribution of Definite and of Indefinite Quadratic Forms . Annals of Mathematical Statistics, 26, 122-127 et 33, 813.
- GURLAND J (1956) . Quadratic Forms in Normally Distributed Random Variables. Sankhya, Ser A, 17, 37-50.
- GURLAND J et GIDEON R.A. (1976) . Series Expansions for Quadratic Forms in Normal Variables. Journal of the American Statistical Association, vol 71, n° 353, 227-232.
- JOHNSON N.L. et KOTZ S. (1970) Distributions in Statistics, Continuous Univariate Distributions - 2 . Wiley, New York.
- JOHNSON N.L. et KOTZ S. (1972) Distributions in Statistics, Continuous Multivariate Distributions - 2 . Wiley, New York.
- KOTZ S., JOHNSON N.L et BOYD D.W. (1967) Series representations of distributions of quadratic forms in normal variables II, non central case. Annals of Mathematical Statistics, 38, 838-848.
- NEVEU J. (1968) Processus Aléatoires Gaussiens. Cours d'été de l'OTAN, Presses de l'Université de Montréal.
- ROBBINS H.E. (1948) The distribution of a definite quadratic form. Annals of Mathematical Statistics 19, 266-270.
- SZEGO G (1959) Orthogonal polynomials. American Mathematical Society.
- TIKU M.L. (1965) Laguerre Series Forms of non central χ^2 and F Distributions. Biometrika, 52 3 et 4 , p 415-427.

QUELQUES APPLICATIONS DES FONCTIONS ORTHOGONALES EN PROBABILITE ET STATISTIQUE

Mots-clefs : Fonctions orthogonales, coefficient de corrélation, test Hilbertien, test du chi-deux, polynômes d'Hermite, forme quadratique de vecteurs aléatoires.

Résumé : On poursuit l'étude des coefficients de corrélation d'ordre supérieur. On obtient la loi asymptotique des coefficients empiriques. On en déduit un test Hilbertien d'indépendance. De plus, on exprime les polynômes d'Hermite et de Laguerre sous forme de moments. On en déduit la loi de la somme de n vecteurs aléatoires et la loi de leurs formes quadratiques.

SOME APPLICATIONS OF ORTHOGONAL FUNCTIONS TO PROBABILITY AND STATISTICS

Summary : We continue the study of the correlation coefficients of superior order. We obtain the asymptotic distribution of the empirical correlation coefficients. We deduce an Hilbertian independence test. Moreover, we show that the Hermite polynomials and the Laguerre polynomials are equal to any moments. We deduce the distributions of a sum of n random vectors and of their quadratic forms.

