Contribution à l'étude des structures statistiques infinidimensionnelles

Jean-Louis Soler

THESE

présentée à

Université Scientifique et Médicale de Grenoble
Institut National Polytechnique de Grenoble

pour obtenir le grade de

DOCTEUR ES SCIENCES
MATHEMATIQUES

par

Jean-Louis SOLER

CONTRIBUTION A L'ETUDE DES STRUCTURES STATISTIQUES INFINIDIMENSIONNELLES

Thèse soutenue le 20 juin 1978 devant la Commission d'Examen :

Président : J. R. BARRA

Examinateurs : D. DACUNHA-CASTELLE
S. JOHANSEN
B. MAISONNEUVE
B. VAN CUTSEM

UNIVERSITE SCIENTIFIQUE
ET MEDICALE DE GRENOBLE

Monsieur Gabriel CAU : Président
Monsieur Joseph KLEIN : Vice-Président

MEMBRES DU CORPS ENSEIGNANT de l'U.S.M.G.

PROFESSEURS TITULAIRES.

MM	AMBLARD Pierre	Clinique de dermatologie
	ARNAUD Paul	Chimie
	ARVIEU Robert	I.S.N.
	AUBERT Guy	Physique
	AYANT Yves	Physique approfondie
Mme	BARBIER Marie-Jeanne	Electrochimie
MM.	BARBIER Jean-Claude	Physique Expérimentale
	BARBIER Reynold	Géologie appliquée
	BARJON Robert	Physique nucléaire
	BARNOUD Fernand	Biosynthèse de la cellulose
	BARRA Jean-René	Statistiques
	BARRIE Joseph	Clinique chirurgicale A
	BEAUDOING André	Clinique de Pédiatrie et Puériculture
	BELORIZKY Elie	Physique
	BERNARD Alain	Mathématiques Pures
Mme	BERTRANDIAS Françoise	Mathématiques Pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques Pures
	BEZES Henri	Clinique chirurgicale et Traumatologie
	BLAMBERT Maurice	Mathématiques Pures
	BOLLIET Louis	Informatique (IUT B)
	BONNET Jean-Louis	Clinique Ophtalmologique
	BONNET-EYMARD Joseph	Clinique Hépato-gastro-entérologique
Mme	BONNIER Marie-Jeanne	Chimie générale
MM.	BOUCHERLE André	Chimie et Toxicologie
	BOUCHEZ Robert	Physique nucléaire
	BOUSSARD Jean-Claude	Mathématiques appliquées
	BOUTET DE MONVEL Louis	Mathématiques Pures
	BRAVARD Yves	Géographie
	CABANEL Guy	Clinique rhumatologique et hydrologique
	CALAS François	Anatomie
	CARLIER Georges	Biologie végétale
	CARRAZ Gilbert	Biologie animale et pharmacodynamie
	CAU Gabriel	Médecine légale et toxicologie
	CAUQUIS Georges	Chimie organique
	CHABAUTY Claude	Mathématiques Pures
	CHARACHON Robert	Clinique Oto-rhino laryngologique
	CHATEAU Robert	Clinique de neurologie
	CHIBON Pierre	Biologie animale
	COEUR André	Pharmacie chimique et chimie analytique
	COUDERC Pierre	Anatomie pathologique
	DEBELMAS Jacques	Géologie générale
	DEGRANGE Charles	Zoologie
	DELORMAS Pierre	Pneumophtisiologie

../..

MM.	DEPORTES Charles	Chimie minérale
	DESRÉ Pierre	Métallurgie
	DODU Jacques	Mécanique appliquée (IUT I)
	DOLIQUE Jean-Michel	Physique des plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	FONTAINE Jean-Marc	Maths pures .
	GAGNAIRE Didier	Chimie Physique
	GALVANI Octave	Mathématiques pures
	GASTINEL Noël	Analyse numérique
	GAVEND Michel	Pharmacologie
	GEINDRE Michel	Electroradiologie
	GERBER Robert	Mathématiques pures
	GERMAIN Jean-Pierre	Mécanique
	GIRAUD Pierre	Géologie
	JANIN Bernard	Géographie
	KAHANE André	Physique générale
	KLEIN Joseph	Mathématiques pures
	KOSZUL Jean-Louis	Mathématiques pures
	KRAVTCHENKO Julien	Mécanique
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie végétale
Mme	LAJZEROWICZ Janine	Physique
MM.	LAJZEROWICZ Joseph	Physique
	LATREILLE René	Chirurgie générale
	LATURAZE Jean	Biochimie Pharmaceutique
	LAURENT Pierre	Mathématiques appliquées
	LEDRU Jean	Clinique médicale B
	LE ROY Philippe	Mécanique (IUT I)
	LLIBOUTRY Louis	Géophysique
	LOISEAUX Jean-Marie	Sciences nucléaires
	LONGEQUEUE Jean-Pierre	Physique nucléaire
	LOUP Jean	Géographie
Mle	LUTZ Elisabeth	Mathématiques pures
MM.	MALINAS Yves	Clinique obstétricale
	MARTIN-NOËL Pierre	Clinique cardiologique
	MAYNARD Roger	Physique du solide
	MAZARE Yves	Clinique Médicale A
	MICHEL Robert	Minéralogie et Pétrographie
	MICOUD Max	Clinique Maladies infectieuses
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie nucléaire
	NEGRE Robert	Mécanique
	NOZIERES Philippe	Spectrométrie Physique
	OZENDA Paul	Botanique
	PAYAN Jean-Jacques	Mathématiques pures
	PEBAY-PEYROULA Jean-Claude	Physique
	PERRET Jean	Séméiologie Médicale (Neurologie)
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
	REVOL Michel	Urologie
	RINALDI Renaud	Physique
	DE ROUGEMONT Jacques	Neuro-Chirurgie
	SARRAZIN Roger	Clinique chirurgicale B
	SEIGNEURIN Raymond	Microbiologie et Hygiène
	SENGEL Philippe	Zoologie
	SIBILLE Robert	Construction mécanique (IUT I)
	SOUTIF Michel	Physique générale
	TANCHE Maurice	Physiologie

MM.	VAILLANT François	Zoologie
	VALENTIN Jacques	Physique Nucléaire
Mme	VERAIN Alice	Pharmacie galénique
MM.	VERAIN André	Physique - Biophysique
	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale

PROFESSEURS ASSOCIES

MM.	CRABBE Pierre	CERMO
	SUNIER Jules	Physique

PROFESSEURS SANS CHAIRE

Mle	AGNIUS-DELORD Claudine	Physique pharmaceutique
	ALARY Josette	Chimie analytique
MM.	AMBROISE-THOMAS Pierre	Parasitologie
	ARMAND Gilbert	Géographie
	BENZAKEN Claude	Mathématiques appliquées
	BIAREZ Jean-Pierre	Mécanique
	BILLET Jean	Géographie
	BOUCHET Yves	Anatomie
	BRUGEL Lucien	Energétique (IUT I)
	BUISSON René	Physique (IUT I)
	BUTEL Jean	Orthopédie
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique (IUT I)
	DELOBEL Claude	M.I.A.G.
	DEPASSEL Roger	Mécanique des fluides
	GAUTRON René	Chimie
	GIDON Paul	Géologie et Minéralogie
	GLENAT René	Chimie organique
	GROULADE Joseph	Biochimie médicale
	HACQUES Gérard	Calcul numérique
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et Médecine préventive
	IDELMAN Simon	Physiologie animale
	JOLY Jean-René	Mathématiques Pures
	JULLIEN Pierre	Mathématiques Appliquées
Mme	KAHANE Josette	Physique
MM.	KRAKOWIACK Sacha	Mathématiques Appliquées
	KUHN Gérard	Physique (IUT I)
	LUU DUC Cuong	Chimie organique -Pharmacie
	MICHOULIER Jean	Physique (IUT I)
Mme	MINIER Colette	Physique (IUT I)
MM.	PELMONT Jean	Biochimie
	PERRIAUX Jean-Jacques	Géologie et Minéralogie
	PFISTER Jean-Claude	Physique du solide
Mle	PIERY Yvette	Physiologie animale

../...

MM.	RAYNAUD Hervé	M.I.A.G.
	REBECQ Jacques	Biologie (CUS)
	REYMOND Jean-Charles	Chirurgie générale
	RICHARD Lucien	Biologie végétale
Mme	RINAUDO Marguerite	Chimie macromoléculaire
MM.	SARROT-REYNAULD Jean	Géologie
	SIROT Louis	Chirurgie générale
Mme	SOUTIF Jeanne	Physique générale
MM.	STIEGLITZ Paul	Anesthésiologie
	VIALON Pierre	Géologie
	VAN CUTSEM Bernard	Mathématiques appliquées

MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

MM.	ARMAND Yves	Chimie (IUT I)
	BACHELOT Yvan	Endocrinologie
	BARGE Michel	Neuro-chirurgie
	BEGUIN Claude	Chimie organique
Mme	BERIEL Hélène	Pharmacodynamie
MM.	BOST Michel	Pédiatrie
	BOUCHARLAT Jacques	Psychiatrie adultes
Mme	BOUCHE Liane	Mathématiques (CUS)
MM.	BRODEAU François	Mathématiques (IUT B) (Personne étrangère habilitée à être directeur de thèse).
	BERNARD Pierre	Gynécologie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et organogénèse
	CHARDON Michel	Géographie
	CHERADAME Hervé	Chimie papetière
	CHIAVERINA Jean	Biologie appliquée (EFP)
	COLIN DE VERDIERE Yves	Maths pures
	CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
	CORDONNIER Daniel	Néphrologie
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie
	CYROT Michel	Physique du solide
	DENIS Bernard	Cardiologie
	DOUCE Roland	Physiologie végétale
	DUSSAUD René	Mathématiques (CUS)
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	FAURE Gilbert	Urologie
	GAUTIER Robert	Chirurgie générale
	GIDON Maurice	Géologie
	GROS Yves	Physique (IUT I)
	GUIGNIER Michel	Thérapeutique
	GUITTON Jacques	Chimie
	HICTER Pierre	Chimie
	JALBERT Pierre	Histologie
	JUNIEN-LAVILLAVROY Claude	O.R.L.
	KOLODIE Lucien	Hématologie
	LE NOC Pierre	Bactériologie-virologie
	MACHE Régis	Physiologie végétale
	MAGNIN Robert	Hygiène et médecine préventive
	MALLION Jean-Michel	Médecine du travail

../...

MM.	MARECHAL Jean	Mécanique (IUT I)
	MARTIN-BOUYER Michel	Chimie (CUS)
	MASSOT Christian	Médecine interne
	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique (IUT I)
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (IUT B) (Personnalité étrangère habilitée à être directeur de thèse).
	PEFFEN René	Métallurgie (IUT I)
	PERRIER Guy	Géophysique-Glaciologie
	PHELIP Xavier	Rhumatologie
	RACHAIL Michel	Médecine Interne
	RACINET Claude	Gynécologie et Obstétrique
	RAMBAUD Pierre	Pédiatrie
	RAPHAEL Bernard	Stomatologie
Mme	RENAUDET Jacqueline	Bactériologie (Pharmacie)
	ROBERT Jean-Bernard	Chimie-Physique
	ROMIER Guy	Mathématiques (IUT B) (Personnalité étrangère habilitée à être directeur de thèse.)
	SAKAROVITCH Michel	Maths appliquées
	SCHAERER René	Cancérologie
Mme	SEIGLE-MURANDI Françoise	Cryptogamie
	STOEBNER Pierre	Anatomie Pathologie
	STUTZ Pierre	Mécanique
	VROUSOS Constantin	Radiologie

MAITRES DE CONFERENCES ASSOCIES.

MM.	DEVINE Roderick	Spectro Physique
	KANEKO AKIRA	Maths pures
	JOHNSON Thomas	Maths appliquées
	RAY Tuhina	Physique

MAITRE DE CONFERENCES DELEGUE

M.	ROCHAT Jacques	Hygiène et Hydrologie (Pharmacie)

Fait à SAINT MARTIN D'HERES, NOVEMBRE 1977

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Année Universitaire
1977-1978

Président : M. Philippe TRAYNARD

Vice-Présidents : M. René PAUTHENET
M. Georges LESPINARD

PROFESSEURS TITULAIRES

MM BENOIT Jean	Electronique - Automatique
BESSON Jean	Chimie Minérale
BLOCH Daniel	Physique du solide - cristallographie
BONNETAIN Lucien	Génie Chimique
BONNIER Etienne	Métallurgie
*BOUDOURIS Georges	Electronique - Automatique
BRISSONNEAU Pierre	Physique du Solide - cristallographie
BUYLE-BODIN Maurice	Electronique - Automatique
COUMES André	Electronique - Automatique
DURAND Francis	Métallurgie
FELICI Noël	Electronique - Automatique
FOULARD Claude	Electronique - Automatique
LANCIA Roland	Electronique - Automatique
LONGEQUEUE Jean-Pierre	Physique Nucléaire Corpusculaire
LESPINARD Georges	Mécanique
MOREAU René	Mécanique
PARIAUD Jean-Charles	Chimie-Physique
PAUTHENET René	Electronique - Automatique
PERRET René	Electronique - Automatique
POLOUJADOFF Michel	Electronique - Automatique
TRAYNARD Philippe	Chimie - Physique
VEILLON Gérard	Informatique Fondamentale et appliquée

*en congé pour études.

PROFESSEURS SANS CHAIRE

MM BLIMAN Samuël	Electronique - Automatique
BOUVARD Maurice	Génie Mécanique
COHEN Joseph	Electronique - Automatique
GUYOT Pierre	Métallurgie Physique
LACOUME Jean-Louis	Electronique - Automatique
JOUBERT Jean-Claude	Physique du Solide - Cristallographie
ROBERT André	Chimie Appliquée et des Matériaux
ROBERT François	Analyse numérique
ZADWORNY François	Electronique - Automatique

MAITRES DE CONFERENCES

MM ANCEAU François	Informatique Fondamentale et appliquée
CHARTIER Germain	Electronique - Automatique
CHIAVERINA Jean	Biologie, biochimie, agronomie
IVANES Marcel	Electronique - Automatique
LESIEUR Marcel	Mécanique
MORET Roger	Physique Nucléaire - Corpusculaire
PIAU Jean-Michel	Mécanique
PIERRARD Jean-Marie	Mécanique
SABONNADIERE Jean-Claude	Informatique Fondamentale et appliquée
Mme SAUCIER Gabrielle	Informatique Fondamentale et appliquée
SOHM Jean-Claude	Chimie Physique

.../...

CHERCHEURS DU C.N.R.S. (Directeur et Maîtres de Recherche)

M FRUCHART Robert	Directeur de Recherche
MM ANSARA Ibrahim	Maître de Recherche
BRONOEL Guy	Maître de Recherche
CARRE René	Maître de Recherche
DAVID René	Maître de Recherche
DRIOLE Jean	Maître de Recherche
KLEITZ Michel	Maître de Recherche
LANDAU Ioan-Doré	Maître de Recherche
MATHIEU Jean-Claude	Maître de Recherche
MERMET Jean	Maître de Recherche
MUNIER Jacques	Maître de Recherche

Personnalités habilitées à diriger des travaux de recherche (Décision du Conseil Scientifique)

E.N.S.E.E.G.

MM BISCONDI Michel	Ecole des Mines ST ETIENNE (dépt.Métallurgie)
BOOS Jean-Yves	Ecole des Mines ST ETIENNE (Métallurgi
DRIVER Julian	Ecole des Mines ST ETIENNE (Métallurgi
KOBYLANSKI André	Ecole des Mines ST ETIENNE (Métallurgi
LE COZE Jean	Ecole des Mines ST ETIENNE (Métallurgi
LESBATS Pierre	Ecole des Mines ST ETIENNE (Métallurgi
LEVY Jacques	Ecole des Mines ST ETIENNE (Métallurgi
RIEU Jean	Ecole des Mines ST ETIENNE (Métallurgi
SAINFORT	C.E.N. Grenoble (Métallurgie)
SOUQUET	U.S.M.G.
CAILLET Marcel	Ecole des Mines ST ETIENNE (Chim.Min.Ph
COULON Michel	Ecole des Mines ST ETIENNE (Chim.Min.Ph
GUILHOT Bernard	Ecole des Mines ST ETIENNE (Chim.Min.Ph
LALAUZE René	Ecole des Mines ST ETIENNE (Chim.Min.Ph
LANCELOT Francis	Ecole des Mines ST ETIENNE (Chim.Min.Ph
SARRAZIN Pierre	Ecole des Mines ST ETIENNE (Chim.Min.Ph
SOUSTELLE Michel	Ecole des Mines ST ETIENNE (Chim.Min.Ph)
THEVENOT François	Ecole des Mines ST ETIENNE (Chim.Min.Ph
THOMAS Gérard	Ecole des Mines ST ETIENNE (Chim.Min.Ph
TOUZAIN Philippe	Ecole des Mines ST ETIENNE (Chim.Min.Ph
TRAN MINH Canh	Ecole des Mines ST ETIENNE (Chim.Min.Ph

E.N.S.E.R.G.

MM BOREL	Centre d'Etudes Nucléaires de GRENOBLE
KAMARINOS	Centre National Recherche Scientifique

E.N.S.E.G.P.

MM BORNARD	Centre National Recherche Scientifique
Mme CHERUY	Centre National Recherche Scientifique
MM DAVID	Centre National Recherche Scientifique
DESCHIZEAUX	Centre National Recherche Scientifique

*
* *
*

Je tiens à exprimer ma profonde reconnaissance à Monsieur J.R. BARRA, Professeur à l'Université Scientifique et Médicale de Grenoble dont les enseignements ont suscité mon goût pour les mathématiques appliquées à la Statistique, pour m'avoir accordé sa confiance en me proposant de travailler sous sa direction et pour l'honneur qu'il me fait en acceptant de présider le Jury de cette thèse.

Je remercie vivement Monsieur D. DACUNHA-CASTELLE, Professeur à l'Université de Paris-Sud de l'intérêt qu'il porte à mon travail en acceptant de faire partie de ce Jury.

Monsieur S. JOHANSEN, Professeur à l'Université de Copenhague, a bien voulu faire une lecture attentive de ce travail. Les discussions que nous avons eues, ses critiques et suggestions, qui m'ont permis de l'améliorer en plusieurs points, ainsi que ses encouragements m'ont beaucoup aidé et je l'en remercie très sincèrement.

Je suis très reconnaissant à Messieurs B. MAISONNEUVE, Maître de Conférences à l'Institut de Mathématiques pour les Sciences Sociales, et B. VAN CUTSEM, Professeur à l'Université Scientifique et Médicale de Grenoble, membres du Jury, de l'aide bienveillante qu'ils m'ont apportée.

J'aimerais que tous ceux qui m'ont encouragé par leurs témoignages d'intérêt et d'amitié partagent les remerciements que j'adresse à A. LE BRETON.

Je tiens à remercier également Madame A.M. STRANO pour le soin et la compétence qu'elle a apportés à la dactylographie du manuscrit ainsi que les membres du service de reprographie du Laboratoire de Mathématiques Appliquées qui en ont réalisé le tirage.

AVERTISSEMENT

Cette thèse qui est une contribution à un domaine particulier de la Statistique Mathématique, celui de la statistique infinidimensionnelle, nécessite la mise en place puis l'utilisation constante d'outils mathématiques divers, notamment d'analyse, d'algèbre et de théorie des probabilités sur des espaces généraux. Le texte qui suit se présente donc assez naturellement en deux parties dont la première, formée des CHAPITRES I et II est destinée à regrouper ces outils.

Cependant, même s'ils sont nouveaux où s'ils apparaissent comme des notions classiques mais adaptées au contexte, ils ne prétendent rien apporter de fondamental à ces différents domaines des mathématiques.

Toutefois, nous avons jugé indispensable et logique, dans un souci d'unité et de clarté de la présentation, de les faire figurer de pair avec la partie plus spécialement consacrée aux résultats obtenus en Statistique qui est constituée par l'essentiel des CHAPITRES III et IV, afin que l'ensemble forme un texte autonome et cohérent aussi bien dans la forme que dans le fond pour exposer un travail de mathématiques appliquées.

TABLE DES MATIERES

INTRODUCTION

Ce travail a pour objet l'étude de modèles mathématiques - appelés Structures Statistiques - pour des problèmes de statistique issus d'observations de processus stochastiques divers (fonctions aléatoires, processus ponctuels ou autres). Il s'insère donc dans le cadre de la Statistique Mathématique et il est présenté dans l'esprit et le formalisme moderne du livre de J.R. BARRA [5] . Bénéficiant de plus, de certains résultats récents du calcul des probabilités, il diffère notablement du travail de U. GRENANDER [12] paru en 1951 qui constitue la première référence sur le sujet.

Il s'agit en effet, de considérer ces observations comme celles d'éléments aléatoires à valeurs dans un espace vectoriel de dimension infinie (de fonctions, de mesures ou autres). Aussi, ces structures statistiques se présentent-elles comme des généralisations naturelles en dimension infinie, des structures classiques de l'analyse multidimensionnelle.

Les probabilistes se sont attachés depuis fort longtemps à définir et étudier des lois de probabilité sur des espaces vectoriels de dimension infinie dans le but d'étudier le comportement global de fonctions aléatoires notamment. Mais la variété des publications dans ce domaine est dûe, en partie, aux diverses hypothèses faites par leurs auteurs concernant les espaces vectoriels en question, celles-ci étant souvent de nature topologique. Ce n'est qu'assez récemment que s'est développée une théorie générale des probabilités en dimension infinie parallèlement et avec le concours de la théorie des mesures cylindriques qui, au moins dans ses fondements, est de nature algébrique.

Quant au statisticien, il s'intéresse surtout aux opérations qu'il peut faire subir à ses observations et, la possibilité d'addition, de multiplication par un scalaire, de multiplication tensorielle que leur confère la seule structure vectorielle lui fournit déjà des moyens intéressants. Il était donc naturel, compte tenu des outils apportés par le calcul des probabilités dans ses développements les plus récents et du champ d'applications possibles, de généraliser en dimension quelconque la notion de structure statistique vectorielle, afin d'envisager d'y étendre des méthodes de la statistique classique. C'est le but que nous nous sommes proposé d'atteindre dans ce travail, une telle présentation systématique ne semblant pas avoir été faite jusqu'à maintenant.

Une structure statistique vectorielle sera donc la donnée d'un triplet $(E,\mathcal{A},\mathcal{P})$, où E est un espace vectoriel réel - l'espace des observations possibles - $\mathcal{A}$, une tribu de parties de E représentant l'ensemble des événements observables et $\mathcal{P}$, une famille de probabilités sur $(E,\mathcal{A})$ c'est-à-dire une famille d'hypothèses concernant la loi de probabilité du vecteur aléatoire X que l'on observe dans E .

La PREMIERE PARTIE de ce travail, constituée par les CHAPITRES I et II, consiste à regrouper les outils mathématiques, nouveaux pour certains d'entre eux, ou classiques, que nécessite cette généralisation et auxquels il sera fait constamment appel par la suite.

En effet, considérant qu'en statistique il importe avant tout de préciser la tribu des événements observables plutôt que les ouverts d'une éventuelle topologie sur l'espace des observations (même s'il peut y avoir un lien entre les deux), nous avons choisi d'éviter le plus possible les hypothèses topologiques à priori sur cet espace. Il fallait donc en premier lieu considérer le couple $(E,\mathcal{A})$ en imposant seulement à la tribu $\mathcal{A}$ d'être compatible avec la structure d'espace vectoriel de E , ce qui nous a conduit à étudier, dans un premier chapitre, la notion d'Espace Vectoriel Mesurable (E.V.M.). Dans le deuxième, avec le même souci de généralité, nous avons introduit et étudié les notions nouvelles qui apparaissent lorsque l'on définit un Espace Vectoriel Probabilisé (E.V.P.) de dimension quelconque.

Le CHAPITRE I propose donc une étude axiomatique des E.V.M. . Toutefois, cet exposé n'en n'est qu'une esquisse, soit que l'analogie avec celle des E.V.T. soit immédiate pour certains points, soit qu'elle conduise à étudier des notions inutilisées dans la suite.

Le § 2 étudie les liens qui existent entre les concepts de mesurabilité et de dualité. L'objectif que nous nous sommes fixé plus haut nous amène en effet à définir et étudier le "Dual d'un E.V.M. $(E,\mathcal{A})$", c'est-à-dire, l'espace vectoriel $(E,\mathcal{A})^m$ de toutes les formes linéaires mesurables définies sur cet espace (pour la tribu borélienne sur $\mathbb{R}$) et qui apparaissent ici comme des "instruments linéaires d'observation". Cela nous permettra de nous référer constamment à la dualité canonique entre E et $(E,\mathcal{A})^m$ lorsqu'elle est séparante (auquel cas on dira que l'E.V.M. est séparé). On étudie aussi l'application linéaire transposée d'une application linéaire mesurable.

Le § 3 est consacré à l'étude des produits tensoriels mesurables d'E.V.M. . Dans un premier temps on profite du contexte pour rappeler des éléments d'algèbre et de la théorie abstraite de FREDHOLM [13] qui seront indispensables plus loin. Le but de cet exposé est d'étudier "l'application quadratique" $\chi_E : x \longmapsto x \otimes x$ définie sur E qui constitue une statistique sur laquelle on reviendra souvent ; il se résume à des résultats techniques.

Enfin le § 4, après la caractérisation et quelques exemples de vecteurs aléatoires les plus généraux se termine par quelques points de méthodologie concernant l'observation de ces éléments aléatoires de dimension infinie.

Le CHAPITRE II est consacré à la présentation ou à l'étude de concepts généraux associés à un E.V.P.

Le § 1 regroupe d'abord, dans un souci d'unité, les définitions et propriétés nécessaires à l'étude probabilités définies sur un E.V.M.S. de dimension quelconque, la première étant celle de probabilité cylindrique. Les rappels, tirés d'ouvrages classiques se réduisent au strict minimum utile à la clarté du texte et à l'établissement de résultats ultérieurs. Dans ce contexte, nous définissons et construisons la probabilité cylindrique "image quadratique d'une probabilité cylindrique" qui sera utilisée au chapitre suivant. Les transformées de FOURIER, de LAPLACE et autre fonction cumulante d'une probabilité sur un E.V.M.S. sont systématiquement définies sur le dual de celui-ci. On étudie également le lien qui peut exister entre le support topologique d'une probabilité borélienne régulière sur un E.V.T. et sa fonction cumulante au moyen de l'analyse convexe, sans toutefois généraliser complètement les résultats obtenus en dimension finie par O. BARNDORFF-NIELSEN dans [4] .

Dans le § 2, on appelle "Dual de l'E.V.P. $(E,\mathcal{a},P)$" , l'espace vectoriel $L(E,\mathcal{a},P)$ des classes d'équivalence (P-p.s.) de fonctionnelles mesurables définies et linéaires sur un sous-espace vectoriel mesurable de probabilité 1 de E , que l'on appelle "Variables aléatoires réelles linéaires" (v.a.r.l.) . (De telles fonctionnelles ont d'ailleurs été déjà envisagées par divers auteurs, A.M. VERSIK, Y.V. ROZANOV [28] , A.V. SKOROHOD [33] en particulier). On étudie systématiquement les propriétés de cet espace de v.a.r. qui sera l'outil fondamental de la théorie exposée au chapitre IV. On étend notamment à cet espace le domaine de définition des fonctionnelles évoquées plus haut ce qui permet d'étendre une théorie de T. HIDA et N. IKEDA [14] basée sur le fait que ces fonctionnelles fournissent des noyaux hermitiens de type positif sur leur domaine de définition et donc des espaces de HILBERT autoreproduisants dont l'un d'eux sera utilisé avec fruit en statistique pour certains problèmes d'estimation.

Le dernier § est consacré aux notions d'intégrabilité du 1er ordre et du 2ème ordre d'un vecteur aléatoire, et parallèlement, à la définition des valeurs typiques correspondantes d'un échantillon d'observations vectorielles. On évoque des lois de grands nombres et autres généralisations utiles.

La DEUXIEME PARTIE du travail, formée des CHAPITRES III et IV, contient plus spécialement ses contributions à la Statistique Mathématique. Le premier de ces deux chapitres est essentiellement consacré à la généralisation en dimension quelconque de la loi de WISHART et à ses conséquences, le deuxième, à la généralisation de la notion de Structure Exponentielle.

Le CHAPITRE III contient en effet une liste de lois de probabilité pouvant servir d'hypothèses ou intervenant dans des structures statistiques vectorielles infinidimensionnelles. Il sert en même temps d'illustration, pour chacune d'entre-elles, de toutes les notions étudiées jusque-là.

La loi Normale la plus générale et ses propriétés qui seront directement utilisées ensuite doivent nécessairement être rappelées en tête ; c'est l'objet du § 1.

Le § 2 qui propose une extension en dimension infinie de la loi de WISHART non centrale, c'est-à-dire de la loi de l'élément aléatoire $X \otimes X$ lorsque X est un vecteur aléatoire gaussien décentré, est la véritable contribution du chapitre.

Cette généralisation nécessite le calcul de la transformée de FOURIER de l'image quadratique d'une probabilité cylindrique gaussienne canonique décentrée définie sur un espace de HILBERT réel séparable (que l'on appelle "probabilité cylindrique de WISHART canonique"). Il repose sur les résultats de la théorie de FREDHOLM et autres résultats techniques déjà évoqués. On en déduit une expression de la fonctionnelle caractéristique de $X \otimes X$ quand X est à valeurs dans un espace de FRECHET séparable muni de sa tribu boréllenne, qui généralise celle connue en dimension finie. Il est facile alors d'obtenir dans ce contexte une extension du théorème de R.A. FISHER concernant l'indépendance et la loi des statistiques : $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ et $S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}) \otimes (X_i - \overline{X})$ associées à un échantillon gaussien.

On obtient, comme corollaire, la fonction caractéristique et la moyenne de v.a.r. quadratiques telles que : $\int_T X^2(t)\, d\mu(t)$ où $(X(t))_{t \in T}$ est une f.a.r. gaussienne décentrée, p.s. à trajectoires continues sur un espace métrique σ-compact T et μ une mesure boréllenne régulière à support compact sur T.

Cette extension de résultats connus seulement dans le cas où la f.a.r. est un mouvement brownien centré standard (T.HIDA [15]) nous conduit à considérer deux sortes d'applications à la statistique : - d'abord elle permet d'améliorer l'étude asymptotique de la fonction puissance de tests basés sur un critère analogue à celui de CRAMER-Von MISES - Ensuite elle permet d'envisager des problèmes de tests d'hypothèses linéaires dans une structure statistique gaussienne à moyenne inconnue en particulier de tester la présence d'un signal déterministe dans un bruit gaussien. Cette méthode des "tests quadratiques" dont on connaît entièrement la fonction puissance mériterait à elle seule une étude plus approfondie.

Les trois autres § étudient des lois de probabilité sur des espaces de mesures : loi Multinomiale, loi de POISSON, de DIRICHLET, la deuxième intervenant dans les structures statistiques associées à l'observation de certains processus ponctuels.

Le CHAPITRE IV propose et étudie en détails une généralisation de la notion de Structure Exponentielle. Elle conduit, en particulier, à considérer des structures statistiques vectorielles $(E,\mathcal{Q},\mathcal{P})$ pour lesquelles la famille $\mathcal{P}$ est paramétrée par un vecteur qui est lui même de dimension infinie (ce qui supprime le clivage habituel entre la statistique paramétrique ou non paramétrique).

Dans un premier stade de l'élaboration de ce travail, nous avons proposé dans [34] de faire varier le paramètre dans le dual de l'E.V.M.S. $(E,\mathcal{Q})$. Il semble d'ailleurs, qu'indépendamment, d'autres auteurs et notamment S. JOHANSEN [17], aient songé à une méthode semblable qui utilise un schéma de dualité. Mais il apparaît qu'une telle généralisation ne suffit pas à recouvrir les principaux cas d'application naturels. C'est ce qui nous a conduit à concevoir de faire varier le paramètre dans un espace plus grand (de v.a.r.l. plus précisément) :

On dira que la structure statistique vectorielle $(E,\mathcal{Q},\mathcal{P})$ est une "Structure Exponentielle Canonique", si la famille $\mathcal{P}$ est constituée de probabilités équivalentes et si pour l'une quelconque d'entre-elles P^* fixée, on a :

$\left\{ \mathrm{Log} \dfrac{dP}{dP^*} \right\}_{P\in\mathcal{P}} \subset L(E,\mathcal{Q},P^*) \oplus \mathbb{R}$. Et l'on dira qu'une structure statistique est "Exponentielle" si elle admet une statistique exhaustive vectorielle, induisant une structure exponentielle canonique.

Le § 1 est donc consacré aux différentes définitions, à l'étude de la paramétrisation ainsi qu'au passage de ces structures statistiques au modèle "explicite" évoqué plus haut.

Cette extension, qui utilise le dual d'un E.V.P., est déterminante comme le montre le § 2 qui est consacré aux exemples et applications. On y remarque que les structures statistiques les plus courantes (gaussiennes à moyenne ou à variance inconnue, poissonniennes, etc.) sont des structures exponentielles comme en dimension finie ; et l'on décrit avec précision divers modèles associés à des problèmes de statistique concrets.

Le § 3 étudie les propriétés des structures exponentielles (canoniques). On généralise, en particulier, le théorème de complétion. L'étude de leur image par une statistique linéaire permet d'en donner une caractérisation intéressante. Des propriétés de conditionnement suggèrent l'application éventuelle de la méthode des tests conditionnels de E.L. LEHMANN dans des problèmes scalaires en présence d'une infinité de paramètres importuns.

Les propriétés analytiques remarquables de ce modèle permettent de développer dans le § 4 , des méthodes d'estimation par des moyens qui lui sont propres : dans le cas explicite, l'utilisation de la dualité convexe comme l'a fait O. BARNDORFF-NIELSEN [4] en dimension finie, éclaire le problème de l'estimation de maximum de vraisemblance d'un paramètre de dimension infinie. Quant à l'estimation sans biais de fonctionnelles du paramètre naturel d'une structure exponentielle implicite, elle admet une solution originale qui s'exprime au moyen d'un espace de HILBERT autoreproduisant introduit dans un chapitre précédent.

Enfin, les exemples d'illustration de chaque notion introduite, voire les applications concrètes que nous nous sommes efforcé de donner, contribueront à situer ce travail dans le cadre des mathématiques appliquées. L'aspect parfois didactique de l'exposé est motivé par la nature du sujet et le souci d'en rendre la lecture claire ; il montre en outre l'importance de l'arsenal nécessaire à la lecture du dernier chapitre dont le propos était en fait l'objectif initial de ce travail.

Toutefois, seules les propositions sont démontrées, les résultats connus figurent sous forme de théorèmes suivis de références bibliographiques. La numérotation des énoncés est propre à chaque § .

Les principaux résultats nouveaux ont fait l'objet de publications sous forme de notes [34], [35] et de conférences invitées [36], [37] respectivement au Congrès Européen des Statisticiens (GRENOBLE 1976) et à la Conférence Internationale de Statistique mathématique à WISLA, Pologne 1976 qui ont paru ou paraîtront dans leurs actes.

PREMIERE PARTIE

OUTILS MATHEMATIQUES POUR LA STATISTIQUE INFINIDIMENSIONNELLE

CHAPITRE I

ESPACES VECTORIELS MESURABLES, ESPACES D'OBSERVATIONS.

On étudie essentiellement dans ce chapitre la notion de dual d'un espace vectoriel mesurable (m-dual) introduite dans [34] , c'est-à-dire l'espace vectoriel de toutes les formes linéaires mesurables définies sur un espace vectoriel muni d'une tribu compatible avec sa structure d'espace vectoriel. Cette notion, qui s'avère utile à la construction d'un modèle statistique particulier, s'apparente évidemment à celle de dual topologique, ou à celle de dual bornologique ; il apparaît donc souhaitable de la situer dans le contexte d'une théorie axiomatique des "Espaces Vectoriels Mesurables". Toutefois, l'exposé qui suit n'en est qu'une esquisse et ne saurait être exhaustif, soit que l'analogie avec la théorie des E.V.T. s'avère immédiate et donc fastidieuse, soit qu'une telle théorie conduise à étudier pour elles-mêmes des notions qui ne seront pas directement utilisées dans la suite et qui sortent donc du cadre de ce travail. . Dans cet esprit, on étudie aussi la notion de produit tensoriel mesurable d'espaces vectoriels mesurables rendue utile par l'étude ultérieure de l'élément aléatoire constitué par le carré tensoriel d'un vecteur aléatoire ; ce contexte sert aussi de cadre à des rappels d'éléments de la théorie générale de FREDHOLM également utiles à cette étude.

Les espaces vectoriels considérés sont réels (sauf mention du contraire). Si A désigne un espace topologique quelconque, on note $\mathcal{B}(A)$ sa tribu borélienne. Enfin, deux espaces mesurables sont dits isomorphes, s'il existe une bijection bi-mesurable de l'un sur l'autre.

§ 1. - ESPACES VECTORIELS MESURABLES.

DEFINITION 1. - Soit E un espace vectoriel sur $\mathbf{R}$. Une tribu $\mathcal{a}$ de parties de E sera dite "compatible avec la structure d'espace vectoriel" si les deux applications

$$f : (\lambda, x) \longmapsto \lambda x \quad \text{de} \quad (\mathbf{R} \times E, \mathcal{B}(\mathbf{R}) \otimes \mathcal{a}) \quad \text{dans} \quad (E, \mathcal{a}),$$
$$g : (x, y) \longmapsto x+y \quad \text{de} \quad (E \times E, \mathcal{a} \otimes \mathcal{a}) \quad \text{dans} \quad (E, \mathcal{a}),$$

sont mesurables. Le couple $(E, \mathcal{a})$ est alors appelé, "Espace Vectoriel Mesurable" (E.V.M.)

Exemples. - 1. La tribu triviale $\{E, \Phi\}$ est toujours compatible avec la structure d'espace vectoriel de E, mais c'est la seule tribu finie ou dénombrable qui le soit.

2. Si E est un E.V.T. métrisable et séparable, $(E, \mathcal{B}(E))$ est un E.V.M. En effet, les applications f et g sont continues donc mesurables, puisque sous ces hypothèses $\mathcal{B}(\mathbf{R} \times E) = \mathcal{B}(\mathbf{R}) \otimes \mathcal{B}(E)$ et que $\mathcal{B}(E \times E) = \mathcal{B}(E) \otimes \mathcal{B}(E)$.

3. Soit $\mathcal{a}_0$ la tribu engendrée par les parties de E réduites à un point, alors $(E, \mathcal{a}_0)$ n'est pas un E.V.M. si $E \neq \{0\}$. En effet,

$$g^{-1}(\{0\}) = \{(x,y) \in E \times E : x+y = 0\}$$

n'appartient pas à $\mathcal{a}_0 \otimes \mathcal{a}_0$.

PROPOSITION 1. - Soit $(E, \mathcal{a})$ un E.V.M.. Quels que soient $\lambda \in \mathbf{R}$ et $x_0 \in E$, l'application $x \longmapsto \lambda x + x_0$ de $(E, \mathcal{a})$ dans lui-même est mesurable. Si $\lambda \neq 0$, cette application est un isomorphisme d'espaces mesurables. La tribu $\mathcal{a}$ est stable pour les opérations de multiplication des ensembles par un scalaire non nul et de translation des ensembles.

En effet, les applications $x \longmapsto \lambda x$ et $x \longmapsto x_0$ de $(E, \mathcal{a})$ dans lui-même sont mesurables, la première étant la section en λ de l'application mesurable f, leur somme l'est aussi. Si $\lambda \neq 0$ l'application inverse de : $x \longmapsto \lambda x + x_o$ est : $x \longmapsto \frac{1}{\lambda}(x - x_o)$ elle est donc mesurable ; la deuxième partie de la proposition en résulte directement. ■

Par analogie avec la théorie des E.V.T. il est facile de définir les notions de sous-espace vectoriel mesurable et d'espace quotient d'un E.V.M. ; il suffit de s'y reporter en remplaçant topologie par tribu et continuité par mesurabilité, du moins pour les définitions.

On constate également, que si $(E_i, \mathcal{A}_i)_{i \in I}$ est une famille d'E.V.M., la tribu produit : $\bigotimes_{i \in I} \mathcal{A}_i$ est compatible avec la structure d'espace vectoriel de $\prod_{i \in I} E_i$; par suite, il est également facile de définir la somme directe mesurable d'E.V.M.

On peut cependant rappeler dans ce contexte une propriété utile (cf. [19],p.44) :

<u>Si</u> E <u>est un E.V.T.</u> <u>et</u> E_1 <u>un sous-espace (vectoriel) de</u> E , <u>la tribu trace de</u> $\mathcal{B}(E)$ <u>sur</u> E_1 <u>(tribu induite par</u> $\mathcal{B}(E)$ <u>sur</u> E_1<u>) coïncide avec la tribu borélienne</u> $\mathcal{B}(E_1)$ <u>pour la topologie induite sur</u> E_1 <u>par celle de</u> E .

§ 2. - DUALITE ET MESURABILITE.

1. Tribus faibles.

Soit (E,F) un couple d'espaces vectoriels réels mis en dualité séparante par une forme bilinéaire sur $E \times F$: $(x,y) \longmapsto \langle x,y \rangle$, $x \in E$, $y \in F$. En identifiant tout élément $y \in F$ à la forme linéaire définie sur E par : $x \longmapsto \langle x,y \rangle$, $x \in E$, on considèrera F comme un sous-espace vectoriel du dual algébrique E^* de E.

DEFINITION 1. - On appelle "tribu faible" sur E relativement à la dualité entre E et F, et l'on note $\mathcal{C}(E,F)$, la tribu engendrée par les éléments de F considérés comme applications réelles (linéaires) sur E.

On appelle "faiblement mesurable", un ensemble, une application $\mathcal{C}(E,F)$- mesurable. On utilisera le fait que $\mathcal{C}(E,F)$ est aussi la tribu engendrée par la famille de sous-ensembles de E : $\{x \in E : \langle x,y \rangle < \alpha\}_{\alpha \in \mathbb{R}, y \in F}$.

REMARQUE 1. - On constate que cette tribu est identique à la tribu engendrée seulement par les éléments d'une base algébrique de F ou aussi par les éléments d'une partie de F engendrant un sous-espace vectoriel séquentiellement dense pour la topologie faible $\sigma(F,E)$ puisque la limite simple de toute suite de fonctions mesurables est encore une fonction mesurable.

Si E est un espace de FRECHET séparable ou de BANACH réflexif et F est son dual topologique E', les notions de fermeture et de fermeture séquentielle d'une partie convexe dans E' coïncident pour la topologie faible $\sigma(E',E)$; il s'ensuit que dans ces cas, la tribu faible $\mathcal{C}(E,E')$ est identique à la tribu engendrée seulement par les éléments d'un sous-ensemble total de E' pour la topologie faible. ■

On peut donner une construction utile de la tribu faible :

Soit N un sous-espace vectoriel de dimension finie de F et soit N^o son polaire dans E. N^o est un sous-espace vectoriel de E, fermé pour la topologie faible $\sigma(E,F)$, de codimension finie et l'espace quotient E/N^{o} est isomorphe au dual algébrique N^* de N . Soit $\mathfrak{M}(E,F)$ la base de filtre constituée des sous-espaces vectoriels M de E, $\sigma(E,F)$ - fermés, de codimension finie (pour la relation d'inclusion); on note π_M l'application canonique de E sur E/M . Pour tout sous-ensemble borélien de l'espace vectoriel de dimension finie E/M , soit $B \in \mathcal{B}(E/M)$, on appelle "cylindre de base B" , le sous-ensemble : $\pi_M^{-1}(B)$ de E ; on a alors :

THEOREME 1. - <u>La réunion des tribus</u> $\pi_M^{-1}(\mathcal{B}(E/M))$ <u>lorsque</u> M <u>parcourt</u> $\mathfrak{M}(E,F)$ <u>est une algèbre</u> $\mathcal{A}(E,F)$ <u>de parties de</u> E <u>engendrant la tribu faible</u> $\mathcal{C}(E,F)$.

Ce théorème classique (cf. [3]) justifie que l'on appelle le plus souvent $\mathcal{C}(E,F)$ la "tribu cylindrique sur E relativement à la dualité entre E et F" .

Il est clair que cette tribu est compatible avec la structure d'espace vectoriel de E et donc que E est un E.V.M.

REMARQUE 2. - Avant de poursuivre l'étude des tribus faibles, il convient de préciser que si N et M sont deux éléments de $\mathfrak{M}(E,F)$ tels que $N \supset M$, en notant π_{NM} l'application canonique de E/M sur E/N , la famille $\pi(E,F) = (E/M , \pi_{NM})$ est un système projectif d'espaces localement convexes indexé par $\mathfrak{M}(E,F)$, les E/M étant munis de la topologie $\sigma(E/M , M^{o})$, les π_{NM} étant continues et satisfaisant la relation de cohérence : $\pi_{LM} = \pi_{LN} \circ \pi_{NM}$ pour $L \supset N \supset M$ dans $\mathfrak{M}(E,F)$.

DEFINITION 2. - <u>On appelle</u> $\pi(E,F)$, <u>le "système projectif des quotients de dimension finie de</u> E (<u>relativement à la dualité entre</u> E <u>et</u> F)".
(cf. [8],p.70).

On utilisera d'ailleurs la caractérisation de sa limite projective :

THEOREME 2. - (cf. [3],p.37) <u>La limite projective du système projectif</u>

$\pi(E,F)$ est isomorphe au dual algébrique F^* de F muni de la topologie $\sigma(F^*,F)$, qui est aussi identique au complété $\hat{E}_\sigma$ de E pour la topologie faible $\sigma(E,F)$; ($F^* = \hat{E}_\sigma = \varprojlim \pi(E,F)$ algébriquement et topologiquement).

2. Relations avec d'autres tribus.

Si E est un E.V.T., en prenant pour F le dual topologique E' en dualité avec E pour la forme bilinéaire canonique, on a évidemment :

$$\mathcal{C}(E,E') \subset \widetilde{\mathcal{B}}(E) \subset \mathcal{B}(E)$$

en notant $\widetilde{\mathcal{B}}(E)$, la tribu de BAIRE de E (c'est-à-dire la tribu engendrée par toutes les fonctions numériques continues sur E). Il existe des relations plus précises entre ces différentes tribus qu'il est utile de rappeler (cf. [3], pp.43-45) :

THEOREME 3. - Soit E un E.V.T. . Si $(K_n)_{n\in\mathbb{N}}$ est une suite croissante de sous-ensembles compacts de E, alors en posant $K = \bigcup_{n\in\mathbb{N}} K_n$ on a : $\mathcal{C}(E,E') \cap K = \widetilde{\mathcal{B}}(E) \cap K$ (tribus traces) (Lemme de SAZONOV).

Si E est métrisable et séparable on a : $\mathcal{C}(E,E') = \mathcal{B}(E)$.

Si E est le dual d'un espace de FRECHET séparable F muni de la topologie faible, (ou de la topologie de la convergence compacte) on a encore : $\mathcal{C}(E,F) = \mathcal{B}(E)$.

Dans les applications, on aura à considérer les situations suivantes :

EXEMPLE 1. - Soit T un ensemble quelconque. Prenons $E = \mathbb{R}^T$ muni de la topologie de la convergence simple ; son dual E' est l'espace $\sum_{t\in T} \mathbb{R}_t$ somme directe de droites. Alors, $\mathcal{C}(E,E')$ n'est autre que la tribu produit : $\bigotimes_{t\in T} \mathcal{B}(\mathbb{R}_t)$, $(\mathbb{R}_t = \mathbb{R}, \forall t\in T)$. Si T est un intervalle de la droite par exemple, on peut remarquer que le sous-espace $C(T)$ des fonctions réelles continues sur T n'est pas faiblement mesurable. (cf. [3], p.47).

EXEMPLE 2. - Soit T un espace métrique σ - compact. Prenons pour E l'espace vectoriel des fonctions numériques continues sur T qui, muni de la topologie de la convergence uniforme sur les parties compactes de T, est un espace de FRECHET séparable. Son dual E' peut être identifié à l'espace vectoriel $M_C(T)$ des mesures boréliennes régulières à support compact sur T, la forme bilinéaire canonique sur $C(T) \times M_C(T)$ étant définie par : $x \in C(T)$, $y \in M_C(T)$, $\langle x, y \rangle = \int_T x(t)\, dy(t)$.

D'après le THEOREME 2, $\mathcal{C}(C(T), M_C(T)) = \mathcal{B}(C(T))$. Cette tribu est d'ailleurs engendrée par la famille $\{\delta_t\}_{t \in T}$ d'éléments de $M_C(T)$ où δ_t est la mesure de DIRAC au point $t \in T$, c'est-à-dire par les formes linéaires $x \longmapsto \langle x, \delta_t \rangle = x(t)$ sur E, lorsque t parcourt T. De fait, l'espace vectoriel engendré par cette famille est séquentiellement dense dans $M_C(T)$ pour la topologie faible $\sigma(M_C(T), C(T))$.

EXEMPLE 3. - Soit $(T, \mathcal{T})$ un espace mesurable quelconque. Prenons pour E, l'espace vectoriel $\mathcal{M}(T, \mathcal{T})$ des mesures bornées x sur $(T, \mathcal{T})$ qui, muni de la norme $\|x\|_1 = |x|(T)$ est un espace de BANACH. Soit $F = \mathcal{L}_\infty(T, \mathcal{T})$ l'espace vectoriel des fonctions numériques mesurables bornées sur $(T, \mathcal{T})$. E et F sont mis en dualité séparante par la forme bilinéaire définie par :

$$x \in E \ , \ y \in F \ , \ \langle x, y \rangle = \int_T y \, dx \ .$$

Soit $\mathcal{A}$, la tribu de parties de E engendrée par les applications : $x \longmapsto \langle x, 1_A \rangle = x(A)$ quand A parcourt $\mathcal{T}$. Alors $\mathcal{C}(E, F) = \mathcal{A}$, d'après la REMARQUE 1.

3. Dual d'un espace vectoriel mesurable.

DEFINITION 3. - On appelle "Dual (ou m-dual) d'un espace vectoriel mesurable $(E,\mathcal{A})$", l'espace vectoriel de toutes les formes linéaires $\mathcal{A}$-mesurables définies sur E on le note $(E,\mathcal{A})^m$ (ou E^m s'il n'y a pas d'ambiguité).

$(E,\mathcal{A})^m$ est donc un sous-espace vectoriel du dual algébrique E^* de E, égal à E^* si $\mathcal{A} = \mathfrak{P}(E)$ ou a $\{0\}$ si $\mathcal{A} = \{\phi, E\}$. Si $\mathcal{A}_1 \subset \mathcal{A}_2$, $(E,\mathcal{A}_1)^m \subset (E,\mathcal{A}_2)^m$.

On peut maintenant considérer la dualité entre E et $(E,\mathcal{A})^m$ induite par la dualité entre E et E^* pour la restriction à $E \times (E,\mathcal{A})^m$ de la forme bilinéaire canonique : $(x, x^*) \longmapsto \langle x,x^*\rangle = x^*(x)$ sur $E \times E^*$. Cette dualité est toujours séparante en $(E,\mathcal{A})^m$; elle peut ne pas l'être en E, comme lorsque $\mathcal{A}$ est la tribu triviale, mais elle peut l'être même si $\mathcal{A} \neq \mathfrak{P}(E)$, d'où :

DEFINITION 4. - On dira que l'espace vectoriel mesurable $(E,\mathcal{A})$ est "séparé" (E.V.M.S.) si la dualité entre E et $(E,\mathcal{A})^m$ est séparante (en E).

PROPOSITION 1. - Si $(E,\mathcal{A})$ est un E.V.M.S. tout sous-espace vectoriel mesurable de E de dimension 1 est isomorphe à $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

En effet, pour tout $x \neq 0$ dans E, il existe $y \in (E,\mathcal{A})^m$ tel que $\langle x,y\rangle \neq 0$, la restriction de y au sous-espace vectoriel E_o de E engendré par x est $E_o \cap \mathcal{A}$ - mesurable ; comme on peut choisir y tel que $\langle x,y\rangle = 1$, cette restriction est de la forme $u_y : \lambda x \longmapsto \lambda$ qui est un isomorphisme d'espaces topologiques entre E_o et $\mathbb{R}$. Donc $u_y^{-1}(\mathcal{B}(\mathbb{R})) = \mathcal{B}(E_o) \subset E_o \cap \mathcal{A}$; mais $u_y(E_o \cap \mathcal{A}) \subset \mathcal{B}(\mathbb{R})$, car u_y^{-1} qui est l'application $\lambda \longmapsto \lambda x$ de $\mathbb{R}$ dans E_o est la section en x de l'application f de la DEFINITION 1, §1 ; donc $u_y^{-1}(u_y(E_o \cap \mathcal{A})) = E_o \cap \mathcal{A} \subset u_y^{-1}(\mathcal{B}(\mathbb{R}))$ par suite, $E_o \cap \mathcal{A} = \mathcal{B}(E_o)$, ce qui signifie que la tribu trace de $\mathcal{A}$ sur tout sous-espace vectoriel de dimension 1 de E coïncide avec sa tribu borélienne. ■

Si $(E,\mathcal{A})$ est un E.V.M.S., $\mathcal{A}$ contient les parties de E réduites à un point.

Si E est un E.V.T.L.C.S., $(E,\mathcal{B}(E))$ est un E.V.M.S.

Si $(E,\mathcal{A})$ est un E.V.M.S., la tribu faible $\mathcal{C}(E,(E,\mathcal{A})^m)$ sur E relativement à cette dualité est contenue dans $\mathcal{A}$. On l'appelle "<u>Tribu affaiblie</u>" de $\mathcal{A}$, et l'on a : $(E,\mathcal{C}(E,(E,\mathcal{A})^m)^m = (E,\mathcal{A})^m$; l'espace vectoriel mesurable et l'espace vectoriel mesurable affaibli ont donc même dual. De plus, la tribu affaiblie d'une tribu faible est identique à celle-ci ; en effet si E et F sont deux espaces vectoriels en dualité séparante $(E,\mathcal{C}(E,F))^m \supset F$ par suite la dualité entre E et $(E,\mathcal{C}(E,F))^m$ est aussi séparante et $\mathcal{C}(E,F) \subset \mathcal{C}(E,(E,\mathcal{C}(E,F))^m)$ comme on a l'inclusion inverse, ces deux tribus sont identiques.

Cependant, alors que F coïncide avec le dual topologique de E muni de la topologie faible $\sigma(E,F)$, il n'en est pas de même en général du dual de l'espace faiblement mesurable séparé $(E,\mathcal{C}(E,F))$, on peut seulement affirmer que F est dense dans $(E,\mathcal{C}(E,F))^m$ pour la topologie faible de ce dernier espace ; ils peuvent donc être très différents. On se contentera d'énoncer un théorème caractérisant ce dual dans un cas suffisamment général pour les applications ultérieures, à savoir lorsque E est un espace de FRECHET séparable et F son dual topologique E' ; on sait en effet d'après le THEOREME 3 que $\mathcal{B}(E) = \mathcal{C}(E,E')$ et DOUADY a démontré (d'ailleurs dans un cas plus général) que

THEOREME 4. - <u>Si</u> E <u>est un espace de FRECHET séparable toute forme linéaire borélienne sur</u> E <u>est continue</u>.

On a donc $(E,\mathcal{B}(E))^m = (E,\mathcal{C}(E,E'))^m = E'$ ce qui s'applique en particulier à l'EXEMPLE 2.

(*) cf. L.SCHWARTZ - Sur le théorème du graphe fermé. C.R. Acad. Sc. Paris (1966) t.263, pp.602-605.

Compte-tenu de ces observations,

DEFINITION 5. - Si E et F sont deux espaces vectoriels réels en dualité séparante, on dira qu'une tribu $\mathcal{a}$ sur E est "compatible avec cette dualité", si F coïncide avec le dual de l'E.V.M. $(E,\mathcal{a})$. Si de telles tribus existent, $\mathcal{C}(E,F)$ est la plus petite et l'on dira alors que "E et F sont en dualité séparante et mesurable (en E)".

Il en est ainsi d'un E.V.M.S. et de son dual. Le THEOREME 4 exprime que E et E' sont en dualité séparante et mesurable quand E est un espace de FRECHET séparable.

4. Transposée d'une application linéaire mesurable.

PROPOSITION 4. - Soient (E_1, F_1) et (E_2, F_2) deux couples d'espaces vectoriels en dualité séparante. Pour qu'une application linéaire u soit mesurable de $(E_1, \mathcal{C}(E_1,F_1))$ dans $(E_2, \mathcal{C}(E_2,F_2))$, il faut et il suffit qu'il existe une application linéaire v de $(E_2, \mathcal{C}(E_2,F_2))^m$ dans $(E_1, \mathcal{C}(E_1,F_1))^m$ telle que $\langle u(x_1), y_2\rangle = \langle x_1, v(y_2)\rangle$ quels que soient $x_1 \in E_1$ et $y_2 \in (E_2, \mathcal{C}(E_2,F_2))^m$.

On utilise en effet la propriété classique suivante : si $(\Omega_1, \mathcal{a}_1)$ et $(\Omega_2, \mathcal{a}_2)$ sont deux espaces mesurables et si $\mathcal{a}_2$ est engendrée par une famille $\{f_i\}_{i\in I}$ de fonctions numériques sur Ω_2, une condition nécessaire et suffisante pour qu'une application f de Ω_1 dans Ω_2 soit $\mathcal{a}_1/\mathcal{a}_2$ - mesurable est que pour tout $i \in I$, l'application $f_i \circ f$ soit $\mathcal{a}_1$ - mesurable. Supposons d'abord les dualités mesurables : si la condition de la proposition est vérifiée, la forme linéaire $x_1 \mapsto \langle u(x_1), y_2\rangle$ est $\mathcal{C}(E_1,F_1)$ - mesurable pour tout $y_2 \in F_2$ donc u est $\mathcal{C}(E_1,F_1)/\mathcal{C}(E_2,F_2)$ mesurable. Inversement si u est mesurable, $y_2 \circ u$ l'est pour tout $y_2 \in F_2$ donc appartient à $(E_1, \mathcal{C}(E_1,F_1))^m = F_1$. Il existe donc un élément (unique car la dualité est séparante) $v(y_2) \in F_1$ tel que :

$$(y_2 \circ u)(x_1) = \langle u(x_1), y_2\rangle = \langle x_1, v(y_2)\rangle, \text{ quel que soit } x_1 \in E_1 .$$

L'application v ainsi définie de F_2 dans F_1 est la transposée de u . Si les dualités ne sont pas mesurables l'application transposée de u est une application de $(E_2, \mathcal{C}(E_2,F_2))^m$ dans $(E_1, \mathcal{C}(E_1,F_1))^m$; il suffit de remplacer F_i par $(E_i, \mathcal{C}(E_i,F_i))^m$, dans ce qui précède, la mesurabilité de u ne changeant pas puisque : $\mathcal{C}(E_i,F_i) = \mathcal{C}(E_i,(E_i,\mathcal{C}(E_i,F_i))^m$. ∎

PROPOSITION 5. - Soient $(E_1,\mathcal{a}_1)$ un E.V.M.S. et $(E_2,\mathcal{a}_2)$ un E.V.M., pour toute application linéaire mesurable u de $(E_1,\mathcal{a}_1)$ dans $(E_2,\mathcal{a}_2)$, il existe une application linéaire unique : ${}^t u$ de $(E_2,\mathcal{a}_2)^m$ dans $(E_1,\mathcal{a}_1)^m$ appelée "transposée de u " , telle que : $\langle u(x_1),y_2\rangle = \langle x_1, {}^t u(y_2)\rangle$ quels que soient $x_1 \in E_1$ et $y_2 \in (E_2,\mathcal{a}_2)^m$.

En effet, si u est $\mathcal{a}_1/\mathcal{a}_2$-mesurable, u est $\mathcal{a}_1/\mathcal{C}(E_2,\mathcal{a}_2)^m)$-mesurable et $y_2 \circ u$ est $\mathcal{a}_1$-mesurable pour tout $y_2 \in (E_2,\mathcal{a}_2)^m$, donc appartient à $(E_1,\mathcal{a}_1)^m$; il existe donc un élément ${}^t u(y_2) \in (E_1,\mathcal{a}_1)^m$ tel que $(y_2 \circ u)(x_1) = \langle u(x_1),y_2\rangle = \langle x_1, {}^t u(y_2)\rangle$ quels que soient $x_1 \in E_1$ et $y_2 \in (E_2,\mathcal{a}_2)^m$; l'unicité de ${}^t u$ résulte de l'hypothèse faite sur la dualité entre E_1 et $(E_1,\mathcal{a}_1)^m$ et cette unicité entraîne que ${}^t u$ est linéaire. ∎

§ 3. PRODUITS TENSORIELS D'ESPACES VECTORIELS MESURABLES.

1. Rappels d'algèbre et d'éléments de la théorie de FREDHOLM.

Nous rappelons dans ce n° des définitions et quelques théorèmes d'algèbre tensorielle utiles pour la suite, en les particularisant souvent au cas d'espaces de HILBERT (cf. [22]) ; de même, nous rappelons les éléments nécessaires de la théorie de FREDHOLM tirés de [13]. Ces rappels se limitent toutefois au minimum utile dans ce § pour les premiers, dans le chapitre III pour les seconds.

a) - Puissances tensorielles symétriques et extérieures d'un espace vectoriel ou d'un espace Hilbertien.

Rappelons pour fixer les notations, que si E_1 et E_2 sont deux espaces vectoriels sur un corps K et $B(E_1, E_2)$ est l'espace vectoriel de toutes les formes bilinéaires sur $E_1 \times E_2$, pour tout couple $(x_1, x_2) \in E_1 \times E_2$ l'application : $f \mapsto f(x_1, x_2)$ de $B(E_1, E_2)$ dans $\mathbf{R}$ est un élément $u_{(x_1, x_2)}$ du dual algébrique : $B(E_1, E_2)^*$ et l'application $\tau : (x_1, x_2) \mapsto u_{(x_1, x_2)}$ est bilinéaire.

On appelle "produit tensoriel (algébrique) de E_1 par E_2" et l'on note : $E_1 \otimes E_2$, l'espace vectoriel engendré par l'ensemble $\tau(E_1 \times E_2)$ dans $B(E_1, E_2)^*$

On note plutôt $x_1 \otimes x_2$ l'élément $u_{(x_1, x_2)}$ de $E_1 \otimes E_2$ et tout élément de cet espace vectoriel est une somme finie : $u = \sum_i \lambda_i \, x_1^i \otimes x_2^i$; cette représentation n'est pas unique, et le plus petit nombre possible de termes de cette somme s'appelle "rang de u".

Si F_1 et F_2 sont deux espaces vectoriels respectivement en dualité séparante avec E_1 et E_2 ; $E_1 \otimes E_2$ et $F_1 \otimes F_2$ le sont aussi pour la forme bilinéaire : $\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle = \langle x_1, y_1 \rangle . \langle x_2, y_2 \rangle$ (notations évidentes).

Soit $\mathcal{L}(E_1, E_2)$ l'espace vectoriel de toutes les applications linéaires de E_1 dans E_2 ; le sous-espace $\mathcal{L}_f(E_1, E_2)$ des applications linéaires de <u>rang fini</u> peut s'identifier au produit tensoriel $E_1^* \otimes E_2$ en faisant correspondre au tenseur $y_1 \otimes x_2$ l'application linéaire :

$$x_1 \longmapsto \langle x_1, y_1 \rangle \, x_2 \; ; \; x_1 \in E_1 \; , \; x_2 \in E_2 \; , \; y_1 \in E_1^* \; .$$

Si $E_1 = E_2 = E$, on notera $B(E)$, $E^{2\otimes}$, $\mathcal{L}(E)$, $\mathcal{L}_f(E)$ les espaces vectoriels correspondant respectivement à ceux définis plus haut et le produit tensoriel $\underbrace{E \otimes \ldots \otimes E}_{n \text{ fois}} = E^{n\otimes}$ qui se définit de façon analogue au cas où $n = 2$, s'appelle "<u>n^e puissance tensorielle de</u> E " .

Si u est une application linéaire de E_1 dans E_2 , il existe pour tout entier n , une application linéaire de $E_1^{n\otimes}$ dans $E_2^{n\otimes}$ que l'on appelle : "<u>n^e puissance tensorielle de</u> u " notée $u^{n\otimes}$ et caractérisée par :

$$u^{n\otimes}(x_1 \otimes \ldots \otimes x_n) = u(x_1) \otimes \ldots \otimes u(x_n) \; , \; x_i \in E_1 \; , \; i = 1 \; , \; \ldots , \; n \; .$$

Si F est un espace vectoriel en dualité séparante avec E , $E^{n\otimes}$ et $F^{n\otimes}$ le sont aussi pour la forme bilinéaire définie pour des couples de tenseurs simples par :

$$\langle x_1 \otimes \ldots \otimes x_n \; , \; y_1 \otimes \ldots \otimes y_n \rangle = \prod_{i=1}^{n} \langle x_i, y_i \rangle \; , \; x_i \in E \; , \; y_i \in F \; , \; i = 1, \ldots, n \; .$$

En particulier, si E est un espace hilbertien séparé, soit H , en prenant $F = E = H$ et $\langle \cdot , \cdot \rangle$ le produit scalaire de H , la formule précédente définit un produit scalaire sur $H^{n\otimes}$ qui en fait <u>un espace préhilbertien séparé</u> (cf.[22]). Le cas $n = 2$ a une importance particulière, notons ici $L(H)$ et $L_f(H)$ l'espace vectoriel des opérateurs linéaires continus et celui des opérateurs continus de rang fini dans H .

THEOREME 1. - <u>Soit</u> H <u>un espace préhilbertien séparé</u>. <u>Il existe un isomorphisme</u> ψ <u>d'espaces préhilbertiens séparés entre</u> $H^{2\otimes}$ <u>et</u> $L_f(H)$ <u>défini par la relation</u> $\psi(x \otimes y)(x_o) = \langle x_o, y \rangle x$; $x_o, x, y \in H$. <u>Le produit scalaire dans</u> $L_f(H)$ <u>étant donné par</u> : $\langle u, v \rangle = \operatorname{trace}(u \circ {}^t v) = \operatorname{trace}(v \circ {}^t u)$.

La définition de l'isomorphisme est classique ; remarquons que ${}^t\psi(x\otimes y) = \psi(y\otimes x)$ et que ψ étant une bijection, elle induit sur $L_f(H)$ un produit scalaire tel que $\langle\psi(x\otimes y)\,,\,\psi(x'\otimes y')\rangle = \langle x\otimes y, x'\otimes y'\rangle = \langle x,x'\rangle\langle y,y'\rangle$ on obtient alors la formule du théorème en appliquant la définition de la trace d'un endomorphisme de rang fini :

DEFINITION 1. - Soit E un espace vectoriel, E^* son dual algébrique et $u \in \mathcal{L}_f(E)$; on appelle trace de u le nombre : $\mathrm{trace}(u) = \sum_{i\in I}\langle y_i, x_i^*\rangle$ pour tout couple de familles $(y_i)_{i\in I}$ et $(x_i^*)_{i\in I}$ resp. dans E et E^* telles que $u(x) = \sum_{i\in I}\langle x, x_i^*\rangle y_i$.

Remarquons que u étant de rang fini, I est fini puisque $u(E)$ est un sous-espace de dimension finie de E .

COROLLAIRE. - Pour tout endomorphisme continu $u \in L(H)$, il existe une application linéaire u^f de $L_f(H)$ dans $L_f(H)$ définie par

$$u^f(v) = u\circ v\circ {}^t u \;;\; v \in L_f(H) .$$

Nous pouvons maintenant définir les puissances tensorielles symétriques et extérieures d'un espace vectoriel ou d'un espace de HILBERT, d'une application linéaire :

LEMME 1. - Soit E un espace vectoriel, et soit un entier $n \geq 2$. Pour toute permutation σ appartenant au n^e groupe symétrique $\mathfrak{S}_n$, il existe un opérateur linéaire inversible U_σ sur $E^{n\otimes}$, unique, tel que :

$U_\sigma(x_1\otimes\ldots\otimes x_n) = x_{\sigma(1)}\otimes\ldots\otimes x_{\sigma(n)}$; $x_1,\ldots,x_n \in E$ et $(U_\sigma)^{-1} = U_{\sigma^{-1}}$.

DEFINITION 2. - On appelle n^e puissance tensorielle symétrique [resp. extérieure] de l'espace vectoriel E et l'on note $E^{n\odot}$ [resp. $E^{n\wedge}$] le sous-espace vectoriel de $E^{n\otimes}$ constitué des tenseurs symétriques [resp. anti-symétriques] soit :

$E^{n\odot} = \{x : x \in E^{n\otimes},\ U_\sigma(x) = x,\ \forall\sigma\in\mathfrak{S}_n\}$

$E^{n\wedge} = \{x : x \in E^{n\otimes},\ U_\sigma(x) = \epsilon_\sigma . x,\ \forall\sigma\in\mathfrak{S}_n$ (avec ϵ_σ = signature de σ)$\}$.

Si $x_1,\dots,x_n$ sont des éléments de E, on pose dans $E^{n\odot}$ [resp. dans $E^{n\wedge}$] :

$$x_1\odot\dots\odot x_n = \frac{1}{n!}\sum_\sigma U_\sigma(x_1\otimes\dots\otimes x_n)\ \text{[resp.}\ x_1\wedge\dots\wedge x_n = \frac{1}{n!}\sum_\sigma \epsilon_\sigma U_\sigma(x_1\otimes\dots\otimes x_n)] .$$

DEFINITION 3. - <u>Soit</u> u <u>une application linéaire d'un espace vectoriel</u> E_1 <u>dans un espace vectoriel</u> E_2 ; <u>on appelle</u> n^e <u>puissance tensorielle symétrique</u> [<u>resp. extérieure</u>] <u>de</u> u <u>l'application linéaire, notée</u> $u^{n\odot}$, <u>de</u> $E_1^{n\odot}$ <u>dans</u> $E_2^{n\odot}$ [<u>resp.</u> $u^{n\wedge}$ <u>de</u> $E_1^{n\wedge}$ <u>dans</u> $E_2^{n\wedge}$] <u>restriction de</u> $u^{n\otimes}$ <u>au sous-espace vectoriel</u> $E_1^{n\odot}$ [resp. $E_1^{n\wedge}$] .

On a donc : $u^{n\odot}(x_1\odot\dots\odot x_n) = u(x_1)\odot\dots\odot u(x_n)$, [resp. $u^{n\wedge}(x_1\wedge\dots\wedge x_n) = u(x_1)\wedge\dots\wedge u(x_n)$] , pour $x_1,\dots,x_n \in E$.

Les définitions ci-dessus sont valables si $n \geq 1$; il convient de poser pour $n = 0$, $E^{0\odot} = E^{0\wedge} = K$ et $u^{0\odot} = u^{0\wedge} = \mathbb{I}$ (application identique de E) .

Si E est de dimension finie r , $E^{n\wedge} = \{0\}$ si $n > r$, et $\dim(E^{n\wedge}) = C_r^n$ si $n \leq r$. (En particulier $E^{r\wedge} \simeq K$) .

Si $u \in \mathcal{L}_f(E_1,E_2)$ alors $u^{n\wedge} = 0$ pour $n > \text{rang}(u)$; $u^{n\wedge}$ est donc de rang fini et $\text{rang}(u^{n\wedge}) = C_r^n$ si $n \leq r$.

THEOREME 2. - <u>Si</u> E <u>et</u> F <u>sont en dualité séparante</u>, $E^{n\odot}$ [<u>resp.</u> $E^{n\wedge}$] <u>et</u> $F^{n\odot}$ [<u>resp.</u> $F^{n\wedge}$] <u>le sont aussi et l'on a</u> :

$$\langle x_1\odot\dots\odot x_n, y_1\odot\dots\odot y_n\rangle = \sum_\sigma \langle x_1,y_{\sigma(1)}\rangle\dots\langle x_n,y_{\sigma(n)}\rangle$$

$$\text{[resp.}\ \langle x_1\wedge\dots\wedge x_n, y_1\wedge\dots\wedge y_n\rangle = \sum_\sigma \epsilon_\sigma \langle x_1,y_{\sigma(1)}\rangle\dots\langle x_n,y_{\sigma(n)}\rangle$$

$$= \text{dét}(\langle x_i,y_j\rangle)\ ,\ \begin{matrix}1\leq i\leq n\ ,\\ 1\leq j\leq n\ .\end{matrix}$$

En particulier, si E est un espace hilbertien séparé, soit H , en prenant $F = E = H$ et $\langle .,.\rangle$ le produit scalaire ces formules sont des spécifications des produits scalaires des espaces préhilbertiens $H^{n\odot}$ et $H^{n\wedge}$ respectivement (cf. [22]) . De plus le projecteur orthogonal de l'espace préhilbertien $H^{n\otimes}$ sur le sous-espace préhilbertien $H^{n\odot}$ [resp. $H^{n\wedge}$] est donné par : $\frac{1}{n!}\sum_{\sigma} U_{\sigma}$ [resp. $\frac{1}{n!}\sum_{\sigma} \epsilon_{\sigma} U_{\sigma}$] . Dans le cas où $n = 2$ on constate que $\psi(H^{2\odot}) = L_{f_*}(H)$ espace vectoriel des opérateurs continus, de rang fini, autoadjoints de H :

THEOREME 3. - Il existe un isomorphisme $\tilde{\psi}$ d'espaces préhilbertiens séparés entre $H^{2\odot}$ et $L_{f_*}(H)$ caractérisé par la formule :

$$\tilde{\psi}(x\odot y) = \frac{1}{2}(\langle x_o,x\rangle y+\langle x_o,y\rangle x) \; ; \; x_o,x,y\in H \; ,$$

lorsqu'on munit $L_{f_*}(H)$ du produit scalaire : $\langle u,v\rangle = \mathrm{trace}(u\circ v) = \mathrm{trace}(v\circ u)$; $u,v \in L_{f_*}(H)$.

b) - Formes fondamentales et déterminant de FREDHOLM d'un endomorphisme de rang fini.

DEFINITION 4. - Soit E un espace vectoriel sur K ; pour tout entier $n > 0$, on appelle forme fondamentale, la forme α_n définie sur $\mathcal{L}_f(E)$ par $\alpha_n(u) = \mathrm{trace}\,(u^{n\wedge})$.

Comme $u^{n\wedge}$ est de rang fini, sa trace est bien définie ; en particulier, $\alpha_1(u) = \mathrm{trace}(u)$; il convient de poser $\alpha_o(u) = 1$ (élément unité de K) . On montre facilement que $\alpha_n(Au) = \alpha_n(uA)$; $u \in \mathcal{L}_f(E)$, $A \in \mathcal{L}(E)$ et dans le cas où A est un automorphisme de E , $\alpha_n(AuA^{-1}) = \alpha_n(u)$.

DEFINITION 5. - Pour tout $u \in \mathcal{L}_f(E)$ on pose $\mathrm{dét}(\mathbb{1}+u) = \sum_{n\geq 0} \alpha_n(u)$. On appelle "Déterminant de FREDHOLM" de u", la fonction de la variable $z \in K$: $z\longmapsto \mathrm{dét}(\mathbb{1} - zu)$.

La somme $\sum_{n\geq 0} \alpha_n(u)$ est en fait finie puisque $u^{n\wedge} = 0$ pour $n > \mathrm{rang}(u)$. Supposons E de dimension finie r ; alors $u^{r\wedge}$ est un endomorphisme de $E^{r\wedge} \simeq K$ donc de la forme : $z\longmapsto\lambda(u)z$, le scalaire $\lambda(u)$ est par définition, le déterminant de u (cf. [7]) . Nous avons donc, $\alpha_r(u) = \mathrm{dét}(u)$ et $\alpha_n(u) = 0$ pour $n > r$; on a aussi :

$$\alpha_n(\mathbb{I}) = \mathrm{trace}(\mathbb{I}^{n\wedge}) = \dim(E^{n\wedge}) = C_r^n \quad \text{si} \quad n \leq r \;,\quad 0 \quad \text{sinon} \;;$$

on peut montrer alors que : $\mathrm{dét}(\mathbb{I}+u) = \alpha_r(\mathbb{I}+u) = \sum_{n=0}^{r} \alpha_n(u)$, et puisque $\alpha_n(u) = 0$ pour tout $n > r$, $\mathrm{dét}(\mathbb{I}+u) = \sum_{n\geq 0} \alpha_n(u)$. Supposons maintenant E de dimension quelconque et $u \in \mathcal{L}_f(E)$; comme la somme $\sum_{n\geq 0} \alpha_n(u)$ est encore finie, on l'appelle encore $\mathrm{dét}(\mathbb{I}+u)$, en remarquant qu'elle coïncide avec le déterminant de la restriction de $\mathbb{I}+u$ à tout sous-espace vectoriel de E contenant $u(E)$. On en conclut que les propriétés usuelles du déterminant sont conservées, en particulier,

$\mathrm{dét}[(\mathbb{I}+u)\circ(\mathbb{I}+v)] = \mathrm{dét}(\mathbb{I}+u).\mathrm{dét}(\mathbb{I}+v)$; $u, v \in \mathcal{L}_f(E)$; et, désignant par $\lambda_1 , \ldots , \lambda_r$ les valeurs propres de $u \in \mathcal{L}_f(E)$ on a :

$\mathrm{dét}(\mathbb{I}+zu) = (1+\lambda_1 z)\ldots(1+\lambda_r z)$ d'où : $\alpha_n(u) = \sum_{i_1<\ldots<i_n} \lambda_{i_1}\ldots\lambda_{i_n}$. Transcrivons de [13] le résultat fondamental de ce n° :

THEOREME 4. - Pour tout $u \in \mathcal{L}_f(E)$, il existe un élément unique $R(u) \in \mathcal{L}(E)$ tel que $\mathrm{trace}[v\circ R(u)] = \mathrm{dét}(\mathbb{I}+u).v$; $v \in \mathcal{L}_f(E)$. On a :

$(\mathbb{I}+u)\circ R(u) = R(u)\circ(\mathbb{I}+u) = \mathrm{dét}(\mathbb{I}+u).\mathbb{I}$; on peut écrire : $R(u) = \sum_{n\geq o} R_n(u)$ où $R_n(u)$ est donné par la relation de récurrence :

$$R_n(u) = \alpha_n(u)\mathbb{I} - u\; R_{n-1}(u) \;;\; R_o(u) = \mathbb{I} \;.$$

Pour que $\mathbb{I} + u$ soit inversible, il faut et il suffit que $\mathrm{dét}(\mathbb{I}+u) \neq 0$, dans ce cas, $(\mathbb{I}+u)^{-1} = \dfrac{R(u)}{\mathrm{dét}(\mathbb{I}+u)} = \mathbb{I} - \dfrac{u}{\mathrm{dét}(\mathbb{I}+u)} \circ R(u)$; soit, en posant :

$r(u) = R(u)\circ u = u\circ R(u)$, $(\mathbb{I}+u)^{-1} = \mathbb{I} - \dfrac{r(u)}{\mathrm{dét}(\mathbb{I}+u)}$; $u \in \mathcal{L}_f(E)$.

c) - La théorie de FREDHOLM dans les espaces hilbertiens.

Dans cet alinéa, on adapte au cas d'un espace hilbertien, qui est un contexte suffisant pour la suite, la théorie de FREDHOLM dans les espaces de BANACH exposée dans [13]. Il s'agit essentiellement de l'extension du déterminant de FREDHOLM et du théorème fondamental (THEOREME 4, ci-dessus) aux opérateurs nucléaires autoadjoints d'un tel espace.

Soit H un espace de HILBERT séparable (réel ou complexe) dont on identifie le dual H' à H lui-même dans le cas réel, ou a son conjugué dans le cas complexe. On note $L_*(H)$ l'espace vectoriel des opérateurs linéaires continus autoadjoints dans H . Les théorèmes suivants, sont des particularisations de théorèmes démontrés dans [22] par exemple.

THEOREME 5. - Soit $H^{n\tilde{\odot}}$ l'espace de HILBERT complété de l'espace préhilbertien $H^{n\otimes}$. Le sous-espace $H^{n\tilde{\odot}}$ [resp. $H^{n\tilde{\wedge}}$] défini de même coïncide avec la fermeture dans $H^{n\tilde{\otimes}}$ du sous-espace $H^{n\odot}$ [resp. $H^{n\wedge}$] de $H^{n\otimes}$; on l'appelle "espace de HILBERT n^e puissance tensorielle symétrique [resp. extérieure]" de l'espace de HILBERT H . En particulier, pour $n = 2$ si $\{e_i\}_{i\in I}$ est une base orthonormale de H, la famille $\{e_i^{2\odot}\}_{i\in I}$ est une base orthonormale de $H^{2\tilde{\odot}}$.

THEOREME 6. - L'isomorphisme $\tilde{\psi}$ du THEOREME 3 se prolonge en un isomorphisme d'espaces hilbertiens entre $H^{2\tilde{\odot}}$ et l'espace de HILBERT $HS_*(H)$ des opérateurs de HILBERT-SCHMIDT autoadjoints dans H muni du produit scalaire : $\langle u,v\rangle = Tr(u\circ v)$ où $Tr(.)$ désigne l'unique forme linéaire sur le sous-espace $N_*(H)$ de $HS_*(H)$ formé des opérateurs nucléaires autoadjoints dans H qui à $\tilde{\psi}(x\odot y)$ fasse correspondre le scalaire $\langle x,y\rangle_H$.

Rappelons que tout opérateur nucléaire est le produit de deux opérateurs de HILBERT-SCHMIDT et inversement, d'où le produit scalaire de $HS_*(H)$.

Alors que $H^{2\tilde{\odot}}$ [resp. $HS_*(H)$] est le complété de $H^{2\odot}$ [resp. $L_{f*}(H)$] pour la norme hilbertienne induite par les produits scalaires de ces deux espaces, notons maintenant $H^{2\hat{\odot}}$ l'espace de BANACH complété de $H^{2\odot}$ pour la norme (dite projective) définie par :

$x \in H^{2\odot}$, $\|x\|_1 = \inf \sum_i \|x_1^i\|.\|x_2^i\|$, l'inf. étant pris sur toutes les suites finies (x_1^i, x_2^i) telles que $x = \sum_i x_1^i \odot x_2^i$.

On a $H^{2\hat{\odot}} \subset H^{2\tilde{\odot}}$ et la restriction de l'isomorphisme $\tilde{\psi}$ à $H^{2\hat{\odot}}$ est une bijection de cet espace sur le sous-espace $N_*(H)$ qui permet encore d'identifier ces deux espaces de BANACH, auxquels s'étend la théorie de FREDHOLM (cf. [13]) :

THEOREME 7. - Soit H un espace de HILBERT réel ou complexe ; pour tout $n \geq 0$ les applications $\alpha_n(u)$ de $H^{2\odot}$ dans le corps des scalaires sont continues pour la norme $\|.\|_1$ et se prolongent en des applications continues sur $H^{2\hat{\odot}}$ (on a les majorations : $|\alpha_n(u)| < \frac{1}{n!}\|u\|_1^n$) ; de plus, la série $\sum_{n\geq 0} \alpha_n(u)$ est absolument convergente. Pour tout $u \in H^{2\hat{\odot}} (= N_*(H))$ on pose : $\text{dét}(\mathbb{I}+u) = \sum_{n\geq 0} \alpha_n(u)$ et on appelle "déterminant de FREDHOLM de l'opérateur nucléaire autoadjoint u", la fonction $z \mapsto \text{dét}(\mathbb{I}-zu)$ sur le corps des scalaires de H.

THEOREME 8. - Soient H un espace de HILBERT réel ou complexe et $u \in H^{2\hat{\odot}}$. Pour que $\mathbb{I}+u$ soit inversible, il faut et il suffit que : $\text{dét}(\mathbb{I}+u) \neq 0$; alors, on a la formule de FREDHOLM :

$$(\mathbb{I}+u)^{-1} = \frac{R(u)}{\text{dét}(\mathbb{I}+u)} = \mathbb{I} - \frac{u}{\text{dét}(\mathbb{I}+u)} \circ R(u) \ ; \ R(u) \in L_*(H) \ .$$

2. Produits tensoriels mesurables d'espaces vectoriels mesurables.

DEFINITION 6. - On appelle "produit tensoriel mesurable des deux E.V.M. $(E_1, \mathcal{a}_1)$ et $(E_2, \mathcal{a}_2)$" l'espace vectoriel $E_1 \otimes E_2$ muni de la plus grande tribu rendant mesurable l'application bilinéaire canonique $\tau : (x_1, x_2) \mapsto x_1 \otimes x_2$, de $(E_1 \times E_2, \mathcal{a}_1 \otimes \mathcal{a}_2)$ dans $E_1 \otimes E_2$. On note $\mathcal{a}_1 \otimes_\tau \mathcal{a}_2$ (*) cette tribu et $(E_1, \mathcal{a}_1) \otimes (E_2, \mathcal{a}_2) = (E_1 \otimes E_2, \mathcal{a}_1 \otimes_\tau \mathcal{a}_2)$ l'E.V.M. ainsi obtenu.

PROPOSITION 1. - Le dual de l'E.V.M. $(E_1 \otimes E_2, \mathcal{a}_1 \otimes_\tau \mathcal{a}_2)$ est isomorphe à l'espace vectoriel $BM(E_1, E_2)$ des formes bilinéaires mesurables définies sur l'E.V.M. produit : $(E_1 \times E_2, \mathcal{a}_1 \otimes \mathcal{a}_2)$.

(*) Cette notation permet de distinguer cette tribu de la tribu produit sur $E_1 \times E_2$ dont la notation : $\mathcal{a}_1 \otimes \mathcal{a}_2$ a été consacrée par l'usage mais qui eut été plus adaptée pour désigner la tribu projective sur $E_1 \otimes E_2$ définie ici.

On sait en effet que l'application : $u \mapsto u \circ \tau$ est un isomorphisme du dual algébrique $(E_1 \otimes E_2)^*$ sur l'espace vectoriel $B(E_1, E_2)$ des formes bilinéaires sur $E_1 \times E_2$ montrons que sa restriction au m-dual $(E_1 \otimes E_2, \mathcal{a}_1 \otimes_\tau \mathcal{a}_2)$ est un isomorphisme de ce sous-espace sur le sous-espace vectoriel $BM(E_1, E_2)$ de $B(E_1, E_2)$. Si u est une forme linéaire mesurable sur le produit tensoriel mesurable, comme τ est mesurable, par définition de la tribu $\mathcal{a}_1 \otimes_\tau \mathcal{a}_2$, la forme bilinéaire $u \circ \tau = f$ est mesurable sur $(E_1 \times E_2, \mathcal{a}_1 \otimes \mathcal{a}_2)$; il suffit alors de montrer que la restriction de $u \mapsto u \circ \tau$ est surjective. Soit f une forme bilinéaire mesurable, définissons u par : $u(\sum_i x_1^i \otimes x_2^i) = \sum_i f(x_1^i, x_2^i)$; alors u est une forme linéaire sur $E_1 \otimes E_2$ telle que : $u \circ \tau = f$; de plus, elle est mesurable car $\mathcal{a}_1 \otimes_\tau \mathcal{a}_2$ est la plus grande tribu rendant τ mesurable. ■

COROLLAIRE 1. - <u>Le produit tensoriel</u> $(E_1, \mathcal{a}_1)^m \otimes (E_2, \mathcal{a}_2)^m$ <u>peut être identifié à un sous-espace vectoriel de</u> $(E_1 \otimes E_2, \mathcal{a}_1 \otimes_\tau \mathcal{a}_2)^m$.

En effet, si l'on pose pour $y_i \in (E_i, \mathcal{a}_i)^m$, $i = 1,2$, $y_1 \otimes y_2(x_1, x_2) = y_1(x_1).y_2(x_2)$, $x_i \in E_i$, $i = 1,2$, on définit une application bilinéaire mesurable sur $(E_1 \times E_2, \mathcal{a}_1 \otimes \mathcal{a}_2)$; l'espace vectoriel engendré par ces fonctions est donc contenu dans $BM(E_1, E_2)$, il suffit alors d'appliquer la proposition. ■

COROLLAIRE 2. - <u>Si</u> $(E_1, \mathcal{a}_1)$ <u>et</u> $(E_2, \mathcal{a}_2)$ <u>sont des E.V.M.S., en notant</u> $\mathcal{C}_i$ <u>la tribu affaiblie de</u> $\mathcal{a}_i$ <u>sur</u> E_i, $i = 1,2$, <u>on peut identifier le produit tensoriel</u> $(E_1, \mathcal{a}_1)^m \otimes (E_2, \mathcal{a}_2)^m$ <u>à un sous-espace vectoriel de</u> $(E_1 \otimes E_2, \mathcal{C}_1 \otimes_\tau \mathcal{C}_2)^m$.

En effet, comme $(E_i, \mathcal{a}_i)^m = (E_i, \mathcal{C}_i)^m$, $i = 1,2$, (cf.§2,n°3), on a : $(E_1, \mathcal{a}_1)^m \otimes (E_2, \mathcal{a}_2)^m = (E_1, \mathcal{C}_1)^m \otimes (E_2, \mathcal{C}_2)^m \subsetneq (E_1 \otimes E_2, \mathcal{C}_1 \otimes_\tau \mathcal{C}_2)^m$ en utilisant le COROLLAIRE 1. Comme $\mathcal{C}_1 \otimes_\tau \mathcal{C}_2 \subset \mathcal{a}_1 \otimes_\tau \mathcal{a}_2$, cette inclusion est plus précise que celle du COROLLAIRE 1. ■

REMARQUE 1. - Soient (E_1,F_1) et (E_2,F_2) deux couples d'espaces vectoriels réels respectivement en dualité séparante, alors $E_1 \otimes E_2$ et $F_1 \otimes F_2$ le sont (cf. n° 1, a)) . On peut donc considérer sur $E_1 \otimes E_2$ la tribu faible $\mathcal{C}(E_1 \otimes E_2, F_1 \otimes F_2)$ relativement à cette dualité et l'on a :
$\mathcal{C}(E_1 \otimes E_2, F_1 \otimes F_2) \subset \mathcal{C}(E_1,F_1) \otimes_\tau \mathcal{C}(E_2,F_2)$, car τ est mesurable de $(E_1 \times E_2, \mathcal{C}(E_1,F_1) \otimes \mathcal{C}(E_1,F_2))$ dans $(E_1 \otimes E_2, \mathcal{C}(E_1 \otimes E_2, F_1 \otimes F_2))$ comme on le vérifie facilement. ∎

Il est aisé de définir la n^e puissance tensorielle mesurable [resp. mesurable symétrique, resp. mesurable extérieure] d'un E.V.M. $(E,\mathcal{Q})$. On posera : $(E,\mathcal{Q})^{n\otimes} = (E^{n\otimes}, \mathcal{Q}^{n\otimes_\tau})$ et on prendra pour tribu sur $E^{n\odot}$ [resp. sur $E^{n\wedge}$] la tribu trace de $\mathcal{Q}^{n\otimes_\tau}$ sur $E^{n\odot}$ [resp. sur $E^{n\wedge}$] que l'on notera : $\mathcal{Q}^{n\odot_\tau}$ [resp. $\mathcal{Q}^{n\wedge_\tau}$] ; on aura donc : $(E,\mathcal{Q})^{n\odot} = (E^{n\odot}, \mathcal{Q}^{n\odot_\tau})$ et $(E,\mathcal{Q})^{n\wedge} = (E^{n\wedge}, \mathcal{Q}^{n\wedge_\tau})$.

Si u est une application linéaire mesurable de $(E_1,\mathcal{Q}_1)$ dans $(E_2,\mathcal{Q}_2)$, il est facile de montrer que la n^e puissance tensorielle $u^{n\otimes}$ de $(E_1,\mathcal{Q}_1)^{n\otimes}$ dans $(E_2,\mathcal{Q}_2)^{n\otimes}$ [resp. la n^e puissance tensorielle symétrique $u^{n\odot}$ de $(E_1,\mathcal{Q}_1)^{n\odot}$ dans $(E_2,\mathcal{Q}_2)^{n\odot}$, resp. la n^e puissance tensorielle extérieure $u^{n\wedge}$ de $(E_1,\mathcal{Q}_1)^{n\wedge}$ dans $(E_2,\mathcal{Q}_2)^{n\wedge}$] est mesurable.

Dans la suite on envisagera uniquement le cas $n = 2$. On peut montrer que le dual de l'E.V.M. $(E,\mathcal{Q})^{2\odot}$ [resp. $(E,\mathcal{Q})^{2\wedge}$] est isomorphe à l'espace vectoriel des formes bilinéaires symétriques [resp. anti-symétriques] mesurables sur $(E \times E, \mathcal{Q} \otimes \mathcal{Q})$.

REMARQUE 2. - Comme on l'a déjà fait, la DEFINITION 6 est calquée sur la définition du produit tensoriel topologique de deux espaces vectoriels topologiques E_1 et E_2 : on appelle "topologie projective sur $E_1 \otimes E_2$" , la topologie la plus fine rendant continue l'application bilinéaire canonique τ , et l'on note $E_1 \hat{\otimes} E_2$ l'espace vectoriel complété de $E_1 \otimes E_2$ pour cette topologie appelé "produit tensoriel projectif complété" de E_1 par E_2 . La PROPOSITION 1 et le COROLLAIRE 1 sont alors les équivalents ("mesurables") des résultats topologiques classiques à savoir que $(E_1 \otimes E_2)'$ est isomorphe à l'espace vectoriel des formes bilinéaires continues sur $E_1 \times E_2$ et que $E_1' \otimes E_2'$ peut être identifié à un sous-espace vectoriel de $(E_1 \otimes E_2)'$. (Cependant alors que $E_1' \otimes E_2'$ peut être identifié au dual $(E_{1_\sigma} \otimes E_{2_\sigma})'$ du produit tensoriel projectif de E_1 par E_2 munis de leur topologie affaiblie, (cf. [31], p.190, exercice 2), on ne sait pas si, de façon analogue, dans le COROLLAIRE 2, on peut identifier en général $(E_1, \mathcal{C}_1)^m \otimes (E_2, \mathcal{C}_2)^m$ au dual $(E_1 \otimes E_2, \mathcal{C}_1 \otimes_\tau \mathcal{C}_2)^m$ tout entier.) ∎

Nous aurons besoin plus loin du résultat suivant :

PROPOSITION 2. - <u>Soient</u> E <u>un E.V.T.L.C. métrisable et séparable et</u> E' <u>son dual topologique. La tribu faible</u> $\mathcal{C}(E^{2\otimes}, E'^{2\otimes})$ [<u>resp</u>. $\mathcal{C}(E^{2\odot}, E'^{2\odot})$] <u>relativement à la dualité séparante entre</u> $E^{2\otimes}$ et $E'^{2\otimes}$ [<u>resp</u>. $E^{2\odot}$ <u>et</u> $E'^{2\odot}$] (<u>cf. n°1, a</u>)) <u>coïncide avec la tribu des boréliens</u> $\mathcal{B}(E^{2\otimes})$ [<u>resp</u>. $\mathcal{B}(E^{2\odot})$] <u>du produit tensoriel projectif</u> $E^{2\otimes}$ [<u>resp. du produit tensoriel symétrique projectif</u> $E^{2\odot}$].

Démontrons d'abord que $\mathcal{C}(E^{2\otimes}, E'^{2\otimes}) = \mathcal{B}(E^{2\otimes})$.

a) $E'^{2\otimes}$ est dense dans $(E^{2\otimes})'$ pour la topologie faible $\sigma((E^{2\otimes})', E^{2\otimes})$, puisque la dualité entre $E^{2\otimes}$ et $E'^{2\otimes}$ est séparante (cf. [31], p.123).

b) En utilisant le THEOREME 3, §2, $\mathcal{B}(E^{2\otimes}) = \mathcal{C}(E^{2\otimes}, (E^{2\otimes})')$ puisque $E^{2\otimes}$ est métrisable et séparable comme E (cf. [31]). Il suffit alors de montrer que $\mathcal{C}(E^{2\otimes}, E'^{2\otimes}) = \mathcal{C}(E^{2\otimes}, (E^{2\otimes})')$. L'inclusion de la première tribu dans la seconde est évidente puisque $E'^{2\otimes} \subset (E^{2\otimes})'$.

Pour démontrer l'inclusion contraire, il suffit que $E'^{2\otimes}$ soit séquentiellement dense dans $(E^{2\otimes})'$; or $(E^{2\otimes})'$ est une réunion dénombrable de compacts $K_n, n = 1, 2, \ldots$ métrisables pour la topologie faible $\sigma((E^{2\otimes})', E^{2\otimes})$ comme dual d'un E.V.T.L.C. métrisable séparable (cf. [6]) ; par suite, pour tout élément $t_o \in (E^{2\otimes})'$, il existe $n(t_o)$ tel que $t_o \in K_{n(t_o)}$ et comme $E'^{2\otimes}$ est dense dans $(E^{2\otimes})'$, $E'^{2\otimes} \cap K_{n(t_o)}$

est dense dans $K_{n(t_o)}$ et donc il existe une suite $\{t_k\}_{k\geq 1}$ d'éléments de $E'^{2\otimes} \cap K_{n(t_o)}$ donc de $E'^{2\otimes}$ qui converge faiblement vers t_o quand $k \to \infty$.

Dans le cas du produit tensoriel symétrique la démonstration est identique ; on peut aussi obtenir le résultat à partir du premier cas, par passage aux tribus-trace. ■

COROLLAIRE. - $\mathcal{B}(E^{2\otimes}) = \mathcal{C}(E^{2\otimes}, E'^{2\hat{\otimes}})$ [resp. $\mathcal{B}(E^{2\odot}) = \mathcal{C}(E^{2\odot}, E'^{2\hat{\odot}})$] . (<u>en notant</u> $E'^{2\hat{\otimes}}$ [resp. $E'^{2\hat{\odot}}$] <u>le produit tensoriel</u> [resp. symétrique] <u>projectif complété de</u> E <u>par lui-même</u>).

Il suffit de remarquer que $\mathcal{C}(E^{2\otimes}, E'^{2\otimes}) = \mathcal{C}(E^{2\otimes}, E'^{2\hat{\otimes}})$ et d'appliquer la proposition. ■

3. <u>L'application quadratique</u>.

DEFINITION 7. - <u>Soit</u> E <u>un espace vectoriel</u>. <u>On appellera</u> "<u>Application quadratique sur</u> E " , <u>l'application</u> χ_E <u>de</u> E <u>dans</u> $E^{2\odot}$ <u>définie par</u> : $\chi_E(x) = \tau(x,x) = x \otimes x$, $x \in E$.

Cette application jouera un rôle particulier en statistique. Si $(E,\mathcal{A})$ est un E.V.M. elle est mesurable dans $(E^{2\odot}, \mathcal{A}^{2\odot\tau})$ puisqu'elle peut s'écrire $\chi_E = \tau \circ \delta$ où δ est l'application : $x \mapsto (x,x)$ de $(E,\mathcal{A})$ dans $(E \times E, \mathcal{A} \otimes \mathcal{A})$.

Si E et F sont deux espaces vectoriels en dualité séparante, $E^{2\odot}$ et $F^{2\odot}$ le sont aussi et χ_E est mesurable de $(E, \mathcal{C}(E,F))$ dans $(E^{2\odot}, \mathcal{C}(E^{2\odot}, F^{2\odot}))$ ce qui résulte de la REMARQUE 1 ou plus directement du fait que pour tout $u \in F^{2\odot}$ de la forme $u_1 \odot u_2$, $u_1, u_2 \in F$, on a d'après le THEOREME 2, $u \circ \chi_E = x \mapsto u_1 \odot u_2(x \otimes x) = 2\langle x, u_1\rangle \langle x, u_2\rangle$ qui est $\mathcal{C}(E,F)$-mesurable. (L'application χ_E est d'ailleurs continue pour les topologies faibles $\sigma(E,F)$ et $\sigma(E^{2\odot}, F^{2\odot})$.

Au chapitre suivant nous utiliserons le système projectif : $\Pi(E^{2\odot}, F^{2\odot})$

des quotients de dimension finie de $E^{2\odot}$ relativement à la dualité entre $E^{2\odot}$ et $F^{2\odot}$ (cf. DEFINITION 2, §2), la proposition suivante qui utilise le rapport entre ce système projectif et l'application quadratique, sera déterminante pour cette utilisation.

PROPOSITION 3. - <u>Avec les notations du §2 n°1 et celle de la DEFINITION §3, la famille</u> $((E/M)^{2\odot}, \Pi^{2\odot}_{N,M})_{M\in\mathfrak{M}(E,F)}$ <u>est un système projectif d'espaces vectoriels de dimension finie tel que pour tout</u> $K \in \mathfrak{M}(E^{2\odot}, F^{2\odot})$, <u>il existe</u> $M_K \in \mathfrak{M}(E,F)$ <u>tel que</u> $(E/M_K)^{2\odot}$ <u>soit isomorphe à un élément du système projectif</u> $\Pi(E^{2\odot}, F^{2\odot})$ <u>contenant l'élément</u> $E^{2\odot}/K$ <u>de ce système</u>.

Pour démontrer la première assertion, il suffit de remarquer que si N et M sont deux éléments de $\mathfrak{M}(E,F)$ tels que $N \supset M$, et si $\Pi_{N,M}$ est l'application linéaire canonique de E/M sur E/N, le diagramme suivant est commutatif :

$$\begin{array}{ccc} E/M & \xrightarrow{\chi_{E/M}} & (E/M)^{2\odot} \\ {\scriptstyle \Pi_{N,M}}\downarrow & & \downarrow{\scriptstyle \Pi^{2\odot}_{N,M}} \\ E/N & \xrightarrow{\chi_{E/N}} & (E/N)^{2\odot} \end{array}$$

par suite, si $L \supset M \supset N$ dans $\mathfrak{M}(E,F)$ la relation de compatibilité : $\Pi^{2\odot}_{NL} = \Pi^{2\odot}_{NM} \circ \Pi^{2\odot}_{ML}$ pour les puissances tensorielles symétriques des applications linéaires canoniques se déduit de celle que vérifient ces dernières. De plus, $\Pi^{2\odot}_{N,M}$ est continue. (En quelque sorte $((E/M)^{2\odot}, \Pi^{2\odot}_{N,M})_{M\in\mathfrak{M}(E,F)}$ est l'image par χ_E de $\Pi(E,F)$).

La deuxième partie de la proposition résulte du lemme d'algèbre suivant, (cf. [7]) : Si M est un sous-espace vectoriel de E, le produit tensoriel $(E/M)^{2\otimes}$ est isomorphe à l'espace quotient $E^{2\otimes}/\Gamma(M)$ où $\Gamma(M)$ est l'espace vectoriel engendré dans $E^{2\otimes}$ par les tenseurs $x \otimes y$ tels que $x \in M$ ou $y \in M$.

Montrons d'abord que si $M \in \mathfrak{M}(E,F)$, $\Gamma(M) \in \mathfrak{M}(E^{2\otimes}, F^{2\otimes})$ c'est-à-dire que $\Gamma(M)$ est de codimension finie dans $E^{2\otimes}$ (relativement à la dualité entre $E^{2\otimes}$ et $F^{2\otimes}$) et que $\Gamma(M)$ est $\sigma(E^{2\otimes},F^{2\otimes})$-fermé. Le premier point résulte du fait que $\Gamma(M)$ contient l'espace $M^{2\otimes}$ qui est de codimension finie ; pour démontrer le deuxième il suffit de montrer que $\Gamma(M)$ est fermé pour la topologie projective de $E^{2\otimes}$ quand E est muni de la topologie faible $\sigma(E,F)$, en effet, elle est compatible avec la dualité entre $E^{2\otimes}$ et $F^{2\otimes}$ puisqu'en vertu de la propriété rappelée à la REMARQUE 2 $(E_\sigma \otimes E_\sigma)'$ peut être identifié à $F^{2\otimes}$. Comme cette topologie est une topologie finale c'est-à-dire la plus fine rendant continue l'application τ , il suffit que $\tau^{-1}(\Gamma(M))$ soit fermé dans $E \times E$; or, $\tau^{-1}(\Gamma(M)) = \{(x,y) \in E \times E : \tau(x,y) = x \otimes y \in \Gamma(M)\} = \{(x,y) \in E\times E : x\in M \text{ ou } y\in M\}$ $= (M\times E) \cup (E\times M)$ est fermé.

En passant aux produits tensoriels symétriques on peut alors montrer que pour tout $M \in \mathcal{M}(E,F)$, $(E/M)^{2\odot}$ est isomorphe à l'espace quotient : $E^{2\odot}/\Gamma_o(M)$ où $\Gamma_o(M)$ est l'espace vectoriel engendré dans $E^{2\odot}$ par l'ensemble des tenseurs symétriques $x \odot y$ tels que $x \in M$ ou $y \in M$ et que $\Gamma_o(M) \in \mathcal{M}(E^{2\odot},F^{2\odot})$.

Soit maintenant $K \in \mathcal{M}(E^{2\odot},F^{2\odot})$ et notons M_K l'espace vectoriel engendré dans E par la famille : $\mathcal{F}_K = \{M \in \mathcal{M}(E,F) : \Gamma_o(M) \subset K\}$(*). Alors $\Gamma_o(M_K) \subset K$; en effet, $\Gamma_o(M_K) = \mathrm{sp}\{x\odot x' \in E^{2\odot} : x \in M_K \text{ ou } x' \in M_K\}$; soit $x \odot x'$ tel que $x \in M_K$ par exemple, alors $x = \sum_{i=1}^{n} \lambda_i x_i$ avec $x_i \in M_i$ et $\Gamma_o(M_i) \subset K$ donc $x_i \odot x' \in \Gamma_o(M_i) \subset K$ pour tout $i = 1,\dots n$ et $x \odot x' = \sum_{i=1}^{n} \lambda_i x_i \odot x' \in K$ ce qui entraîne que $\Gamma_o(M_K) \subset K$.

Maintenant, compte tenu de ce qui précède, si $K \in \mathcal{M}(E^{2\odot},F^{2\odot})$, il existe bien $M_K \in \mathcal{M}(E,F)$ tel que $(E/M_K)^{2\odot}$ soit isomorphe à $E^{2\odot}/\Gamma_o(M_K)$ qui est bien un élément du système projectif $\Pi(E^{2\odot},F^{2\odot})$ contenant l'élément $E^{2\odot}/K$. ■

REMARQUE 3. - Dans le cas où E est un espace de HILBERT réel séparable, soit H , et F est son dual H' identifié à H lui-même, la donnée du système projectif $\Pi(H,H)$ est équivalente à la donnée du système projectif $(V,P_{V,W})_{V\in\mathcal{F}(H)}$ où $\mathcal{F}(H)$ désigne la famille des sous-

(*) - On peut montrer que cette famille est non vide, comme nous l'a fait remarquer S. JOHANSEN :

espaces vectoriels de dimension finie de H et où pour V et W dans $\mathcal{F}(H)$ tels que $V \supset W$, $P_{V,W}$ désigne le projecteur orthogonal de V sur W. Les faits précédents s'expriment alors plus simplement :

La famille $(V^{2\odot}, P^{2\odot}_{V,W})_{V \in \mathcal{F}(H)}$ est un système projectif d'espaces vectoriels de dimension finie et pour tout sous-espace K de dimension finie de $H^{2\odot}$, il existe un élément $V^{2\odot}_K$ de ce système contenant K.

En effet, si $V \supset W$, $V^{2\odot} \supset W^{2\odot}$ et le diagramme suivant est commutatif :

$$\begin{array}{ccc} V & \xrightarrow{\chi_V} & V^{2\odot} \\ {\scriptstyle P_{V,W}}\downarrow & & \downarrow{\scriptstyle P^{2\odot}_{V,W}} \\ W & \xrightarrow{\chi_W} & W^{2\odot} \end{array}$$

et $P^{2\odot}_{V,W}$ coïncide avec le projecteur orthogonal de $V^{2\odot}$ sur $W^{2\odot}$.
Si $K \in \mathcal{F}(H^{2\odot})$ il suffit de prendre pour V_K l'espace vectoriel engendré par $\chi^{-1}_H(K)$ dans H, pour constater que $V^{2\odot}_K$ est le plus petit élément du système $(V^{2\odot}, P^{2\odot}_{V,W})_{V \in \mathcal{F}(H)}$ contenant K ; ce système peut être considéré comme l'image par l'application quadratique χ_H du système projectif $(V, P_{V,W})_{V \in \mathcal{F}(H)}$. ■

§4. - VECTEURS ALEATOIRES ET LEURS OBSERVATIONS.

1. Définitions et exemples.

Un vecteur aléatoire est un élément aléatoire à valeurs dans un E.V.M.

Une application X définie sur un espace probabilisé $(\Omega,\mathcal{F},P)$, à valeurs dans un espace vectoriel E, mis en dualité séparante avec un espace vectoriel F est "un vecteur aléatoire dans E relativement à cette dualité" si X est $\mathcal{F}/\mathcal{C}(E,F)$-mesurable. Si X est un vecteur aléatoire à valeurs dans l'E.V.M.S.$(E,\mathcal{a})$, il l'est relativement à la dualité (séparante) entre E et $(E,\mathcal{a})^m$.

Avec les mêmes notations qu'au §2, on a évidemment :

THEOREME 1. - Une condition nécessaire et suffisante pour que X soit un vecteur aléatoire dans E relativement à la dualité entre E et F est que l'application $\langle X,y\rangle$ définie sur le même espace probabilisé soit une variable aléatoire réelle quel que soit $y \in F$.

La REMARQUE 1, §2 s'avère très utile dans l'application pratique de ce théorème : il suffit, en effet, de vérifier cette condition pour un sous-ensemble de fonctionnelles y qui peut être bien plus restreint que F.

L'exemple le plus important de vecteur aléatoire de dimension quelconque est constitué par les fonctions aléatoires ou les processus stochastiques dont presque toutes les trajectoires appartiennent à un espace vectoriel de fonctions :

Soit $(X_t)_{t\in T}$ un processus stochastique défini sur $(\Omega,\mathcal{F},P)$ et supposé réel pour simplifier. Si E est un espace vectoriel de fonctions réelles sur T tel que pour tout $\omega \in \Omega$, la trajectoire $X.(\omega)$ appartienne à E, alors on pourra considérer l'application $X : \omega \longmapsto X.(\omega)$ de Ω dans E comme un vecteur aléatoire dans E, s'il existe par exemple un espace vectoriel F en dualité séparante avec E et tel que X soit $\mathcal{F}/\mathcal{C}(E,F)$-mesurable. (Le plus souvent E est muni d'une topologie localement convexe séparée et F est le dual de E) .

Il en est ainsi sans autre hypothèse sur le processus dès que F contient l'ensemble des fonctionnelles d'évaluation sur E c'est-à-dire les formes linéaires : $\delta_t : x \longmapsto x(t)$ quand t parcourt T et que cet ensemble engendre un espace vectoriel séquentiellement dense dans F pour la topologie faible $\sigma(F,E)$; il suffit en effet d'appliquer le THEOREME 1 et la REMARQUE 1, §2 .

En reprenant les exemples du §2 , on obtient :

EXEMPLE 1. - Il est clair que l'application X associée à un processus réel dont on ne connaît rien à priori des trajectoires est un vecteur aléatoire dans $\mathbb{R}^T$ muni de la topologie de la convergence ponctuelle relativement à la dualité entre cet espace et son dual topologique dont le sous-ensemble $\{\delta_t\}_{t\in T}$ est une base algébrique. (cf.EXEMPLE 1, §2) . ∎

EXEMPLE 2. - Si T est un espace métrique σ-compact et si le processus $(X_t)_{t\in T}$ est P - presque sûrement à trajectoires continues sur T , on prend alors $E = C(T)$ et $F = E' = M_c(T)$ (cf.EXEMPLE 2, §2) . L'application $X : \omega \longmapsto X.(\omega)$ est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{C}(E,E') = \mathcal{B}(E)$ puisque cette dernière tribu est engendrée par les formes linéaires continues : $x \longrightarrow \delta_t(x) = x(t) = \int_T x\, d\delta_t$ (où δ_t est la mesure de DIRAC au point t) quand t parcourt T . Il suffit donc que les $X_t = \langle X,\delta_t\rangle$ soient des v.a.r. (ce qui est le cas) pour que $\langle X,y\rangle$ soit une v.a.r. quel que soit $y \in M_c(T)$.

On obtient d'autres exemples de ce type en prenant pour T un intervalle de la droite et E l'espace de BANACH séparable $C^n(T)$ des fonctions réelles n fois continûment dérivables, ou bien encore l'espace de BANACH séparable $AC(T)$ des fonctions réelles absolument continues sur T etc... ; il suffit que le processus ait P-presque toutes ses trajectoires dans ces espaces pour que l'application X associée, soit un vecteur aléatoire relativement à la dualité entre ceux-ci et leur dual topologique. ∎

EXEMPLE 3. - Un processus ponctuel sur un ensemble T est le plus souvent défini directement comme un élément aléatoire X à valeurs dans l'ensemble $M_p(T)$ des mesures dites ponctuelles sur T. On suppose le plus souvent que T est un espace topologique localement compact à base dénombrable, muni de sa tribu borélienne $\mathcal{B}(T)$ - ce qui recouvre la plupart des applications - (cf. [24] pour un exposé récent et complet). On peut toujours considérer X comme un exemple de vecteur aléatoire à valeurs dans (un sous-ensemble d') un espace vectoriel de mesures sur $(T,\mathcal{B}(T))$. ∎

2. Observations d'un vecteur aléatoire.

Ayant reconnu qu'un problème de statistique donné concerne un élément aléatoire X à valeurs dans un E.V.M. $(E,\mathcal{A})$ la question se pose de savoir comment s'effectue l'observation de X, de façon à préciser la tribu des événements observables sur E.

Lorsque E est de dimension finie, muni de sa tribu borélienne, on dispose généralement d'une (ou plusieurs) observation(s) de X, élément(s) de E ce qui revient à admettre $\mathcal{B}(E)$ comme tribu des ensembles observables.

En dimension infinie, cela peut être le cas également, par exemple lorsque E est un sous-espace vectoriel de $\mathbb{R}^T$, T étant une partie bornée de $\mathbb{R}$, voire de $\mathbb{R}^k$, $k > 1$, l'observation est alors constituée par une trajectoire x, élément de E, de la f.a.r. à laquelle correspond l'élément aléatoire X.

De façon générale on peut toujours admettre que l'observation de X se ramène à l'observation d'une famille : $\{\langle X,y\rangle\}_{y \in \Phi \subset (E,\mathcal{A})^m}$ de v.a.r., fonctionnelles linéaires mesurables sur $(E,\mathcal{A})$, qui apparaissent comme "instruments linéaires d'observation"(*). La tribu des événements observables est donc celle engendrée sur E par les éléments de Φ.

(*) Ce terme est de GUELFAND.

Cette tribu peut coïncider avec la tribu initiale ; dans ce cas on dira que X est "complètement observable" ce qui revient à dire que ce système d'observations est équivalent à la donnée d'un élément (observation) x dans E . C'est ce qui se passe en général dans l'exemple que nous venons de citer, où l'observation de la trajectoire x de la f.a.r. à laquelle X correspond, est équivalente à l'observation de la famille infinie : $\{\langle X,\delta_t\rangle\}_{t\in T}$ de fonctionnelles d'évaluation de X .

Mais il ne saurait en être de même, pour des raisons pratiques évidentes lorsque E est par exemple un espace de distributions, de suites, de mesures. Comme la structure statistique doit rendre compte de l'observation effective, il faut dans ce cas considérer que l'espace des observations possibles est isomorphe à $\mathbf{R}^{\Phi}$ et peut donc différer de E .

Il y a là un point de méthodologie qu'il fallait préciser. Dans toute la suite on admettra que X est complètement observable, quitte à remplacer l'espace vectoriel initial par celui des observations réelles en supposant toutefois que cette correction nous conduise encore à un problème de statistique concernant un vecteur aléatoire de dimension infinie afin qu'il demeure dans le cadre de ce travail. ∎

CHAPITRE II

ESPACES VECTORIELS PROBABILISES, STRUCTURES STATISTIQUES VECTORIELLES.

Une structure statistique vectorielle est la donnée d'un triplet $(E,\mathcal{a},\mathcal{P})$ où $\mathcal{P}$ est une famille de lois de probabilité définies sur l'E.V.M. $(E,\mathcal{a})$ des observations possibles, c'est-à-dire la donnée d'une famille d'"Espaces Vectoriels Probabilisés".

L'étude des lois de probabilité sur un espace vectoriel de dimension infinie utilise celle des probabilités cylindriques ; il est donc nécessaire de rappeler la définition et les principales propriétés de cet outil fondamental. De plus, les différentes fonctionnelles associées à une probabilité sur un E.V.M.S. sont définies ici sans hypothèses topologiques grâce à l'utilisation systématique de la dualité séparante entre cet espace vectoriel et son m-dual. Outre le rappel de résultats classiques, on définit la notion d'image quadratique d'une probabilité cylindrique dont on verra une utilisation plus loin. On essaie également, d'étendre en dimension quelconque des caractérisations du support d'une probabilité au moyen de l'analyse convexe (cf. [4]).

Le principal apport de ce chapitre est constitué par la définition et l'étude de la notion de dual d'un espace vectoriel probabilisé, qui jouera un rôle essentiel dans la construction des structures statistiques exponentielles généralisées. Il permet aussi de définir des espaces de HILBERT auto-reproduisants associés à un espace vectoriel probabilisé.

§ 1. - MESURES CYLINDRIQUES ET PROBABILITES.

Soient E et F deux espaces vectoriels réels mis en dualité séparante par une forme bilinéaire : $(x,y) \longmapsto \langle x,y \rangle$, $x \in E$, $y \in F$ sur $E \times F$. Dans ce contexte, on utilisera les notations du CHAPITRE I, § 2 .

1. Définition d'une mesure cylindrique.

DEFINITION 1. - On appelle "mesure cylindrique sur E , relativement à la dualité entre E et F" , un système projectif d'espaces vectoriels mesurés $(E/M , \mathcal{B}(E/M) , \mu_M)_{M \in \mathcal{M}(E,F)}$ où $\mu = (\mu_M)_{M \in \mathcal{M}(E,F)}$ est une famille de mesures bornées définies respectivement sur les éléments du système projectif $\pi(E,F)$ des quotients de dimension finie de E relativement à la dualité entre E et F , munis de leur tribu borélienne (cf. CHAPITRE I, § 2, DEFINITION 2) et vérifiant la relation de cohérence : $\mu_N = \pi_{N,M}(\mu_M)$ dès que $N \supset M$ dans $\mathcal{M}(E,F)$.

On ne considère, dans la suite, que des mesures μ_M qui sont des probabilités, alors le système projectif, ou plus simplement, la famille $\mu = (\mu_M)_{M \in \mathcal{M}(E,F)}$ est appelée "probabilité cylindrique sur E relativement à la dualité entre E et F" . Si E est un E.V.T.L.C.S., on appelle "probabilité cylindrique sur E" , toute probabilité cylindrique relativement à la dualité séparante entre E et son dual topologique E' .

THEOREME 1. - (cf. [3] ou [8]) Il existe une bijection entre :

- L'ensemble des probabilités cylindriques sur E relativement à la dualité entre E et F .
- L'ensemble des fonctions d'ensembles μ , simplement additives sur l'algèbre $\mathcal{A}(E,F)$ des ensembles cylindriques telles que : $0 \leq \mu(A) \leq 1$ pour tout $A \in \mathcal{A}(E,F)$, $\mu(E) = 1$ et dont la restriction à chaque espace mesurable $(E,\pi_M^{-1}(\mathcal{B}(E/M)))$ soit une probabilité.

Cette fonction d'ensembles μ sur $\mathcal{A}(E,F)$, associée à chaque probabilité cylindrique sur E relativement à la dualité entre E et F n'est pas nécessairement σ-additive et n'est donc pas toujours prolongeable en une probabilité sur l'espace vectoriel faiblement mesurable $(E,\mathcal{C}(E,F))$. Mais les théorèmes de BOCHNER et de KOLMOGOROV permettent de démontrer le résultat suivant :

THEOREME 2. - <u>Si</u> $\mu = (\mu_M)_{M\in\mathcal{M}(E,F)}$ <u>est une probabilité cylindrique relativement à la dualité entre</u> E <u>et</u> F , <u>il existe une probabilité unique</u> P^* <u>sur l'espace vectoriel faiblement mesurable</u> $(F^*,\mathcal{C}(F^*,F))$, <u>où</u> F^* <u>désigne le dual algébrique de</u> F <u>et telle que si</u> π^*_M <u>désigne l'application linéaire canonique de</u> F^* <u>sur</u> E/M <u>on ait</u> $\mu_M = \pi^*_M(P^*)$ <u>quel que soit</u> $M \in \mathcal{M}(E,F)$.

On peut, alors, donner une condition pour que la probabilité cylindrique $\mu = (\mu_M)_{M\in\mathcal{M}(E,F)}$, définisse une probabilité sur l'E.V.M.S. $(E,\mathcal{C}(E,F))$ en considérant E comme un sous-ensemble de F^* ; d'après un résultat connu, pour qu'il en soit ainsi, il faut et il suffit que l'ensemble E soit de mesure extérieure 1 relativement à P^* , c'est-à-dire : $P^*(C) = 1$ pour tout $C \in \mathcal{C}(F^*,F)$, $C \supset E$; (E n'appartient pas nécessairement à $\mathcal{C}(F^*,F)$) . Alors cette probabilité notée encore μ est induite par P^* , c'est la restriction de P^* à la tribu trace : $E \cap \mathcal{C}(F^*,F)$ identique d'ailleurs à $\mathcal{C}(E,F)$ sur E . Il existe un grand nombre de résultats concernant ce problème ; on ne les rappelle pas, puisqu'ils ne seront pas utilisés dans la suite (cf. [3], [8], [1]).

EXEMPLE 1. - Dans l'EXEMPLE 1, CHAPITRE I, §2 où $E = R^T$ muni de la topologie de la convergence simple, toute probabilité cylindrique sur E induit une probabilité sur $(E,\mathcal{C}(E,E'))$ puisque $E' = \sum_{t\in T} R$ et que $(E')^* = E$. ■

EXEMPLE 2. - Soit H un espace de HILBERT réel séparable (identifié à son dual topologique) et $\mu = (\mu_M)_{M\in\mathcal{M}(H,H)}$ la probabilité cylindrique sur H telle que μ_M soit la probabilité gaussienne centrée de matrice de covariance unité sur l'E.V.M.S. de dimension finie $(H/M, \mathcal{B}(H/M))$, $M \in \mathcal{M}(H,H)$.

On peut montrer que si H n'est pas de dimension finie, μ n'induit pas une probabilité sur $(H, \mathcal{C}(H,H)) = (H,\mathcal{B}(H))$. ∎

REMARQUE 1. - Il convient de noter à la suite de ce dernier exemple, compte tenu de la REMARQUE 3 , CHAPITRE I , §3 , que si H est un espace de HILBERT réel séparable, la donnée d'une probabilité cylindrique sur H est équivalente à la donnée du système projectif d'espaces de HILBERT probabilisés : $(V, \mathcal{B}(V), \mu_V)_{V\in\mathcal{F}(H)}$ où $\mathcal{F}(H)$ désigne la famille des sous-espaces vectoriels de dimension finie de H , les μ_V vérifiant la relation de cohérence : $\mu_W = P_{V,W}(\mu_V)$ dès que $V \supset W$ dans $\mathcal{F}(H)$, $P_{V,W}$ désignant le projecteur orthogonal de V sur W . ∎

REMARQUE 2. - Toute probabilité P sur l'E.V.M.S. $(E,\mathcal{C}(E,F))$ est en particulier une fonction d'ensembles ayant les propriétés requises au THEOREME 1 il lui correspond donc une probabilité cylindrique unique : $(P_M)_{M\in\mathcal{M}(E,F)}$ relativement à la dualité entre E et F . En particulier, si X est un vecteur aléatoire à valeurs dans $(E,\mathcal{C}(E,F))$, la probabilité cylindrique : $[(P_X)_M]_{M\in\mathcal{M}(E,F)}$ associée à sa loi de probabilité P_X est appelée : <u>loi de probabilité cylindrique de</u> X" . D'après ce qui précède, elle caractérise entièrement la loi de probabilité de X ; mais inversement, la donnée d'une probabilité cylindrique sur E relativement à la dualité entre E et F n'est pas équivalente à la donnée d'un vecteur aléatoire à valeurs dans $(E,\mathcal{C}(E,F))$, mais à celle d'un élément aléatoire X^* dans $(F^*,\mathcal{C}(F^*,F))$ (il suffit en effet de prendre comme espace fondamental $(F^*,\mathcal{C}(F^*,F),P^*)$ avec les notations du THEOREME 2 et pour X^* l'application identique sur F^*) .

2. <u>Image linéaire d'une probabilité cylindrique</u>.

Soient (E_1,F_1) un autre couple d'espaces vectoriels réels en dualité séparante, $\mu = (\mu_M)_{M\in\mathcal{M}(E,F)}$, une probabilité cylindrique sur E relativement à la dualité entre E et F , et u une application linéaire de E dans E_1 continue pour les topologies faibles $\sigma(E,F)$ et $\sigma(E_1,F_1)$.

Si $M_1 \in \mathcal{M}(E_1,F_1)$, le sous-espace $M = u^{-1}(M_1)$ appartient à $\mathcal{M}(E,F)$ et u définit par passage aux quotients une application linéaire u_{M_1} de E/M sur E/M_1 ; si $N_1 \supset M_1$ dans $\mathcal{M}(E_1,F_1)$, $N = u^{-1}(N_1) \supset M = u^{-1}(M_1)$ dans $\mathcal{M}(E,F)$. Alors, si l'on pose $\nu_{M_1} = u_{M_1}(\mu_{u^{-1}(M_1)})$, la famille : $\nu = (\nu_{M_1})_{M_1 \in \mathcal{M}(E_1,F_1)}$ est une probabilité cylindrique sur E_1 relativement à la dualité entre E_1 et F_1 que l'on appelle : "<u>image de la probabilité cylindrique</u> μ <u>par l'application linéaire</u> u" <u>et que l'on note</u> $u(\mu)$, (cf. [8],p.72).

REMARQUE 3. - Il peut se faire que l'application linéaire u transforme la probabilité cylindrique μ en une probabilité cylindrique ν qui induit une probabilité sur $(E_1, \mathcal{C}(E_1,F_1))$; c'est le cas, évidemment, chaque fois que E_1 est de dimension finie. Un autre exemple important est le suivant :

Soient H l'espace de HILBERT réel séparable $L_2([0,1], \mathcal{B}([0,1]), dt)$ des fonctions réelles boréliennes et de carré intégrable par rapport à la mesure de LEBESGUE sur l'intervalle $[0,1]$ et E l'espace de BANACH séparable $C_0[0,1]$ des fonctions numériques continues et nulles en 0 sur cet intervalle. L'application linéaire u de H dans E définie par $u(f)(t) = \int_0^t f(s)ds = \langle f, \mathbb{1}_{[0,t[}\rangle_H$, $t \in [0,1]$, $f \in H$ est continue, elle est donc continue pour les topologies faibles $\sigma(H,H')$ et $\sigma(E,E')$ (elle est d'ailleurs mesurable pour les tribus boréliennes de H et de E qui sont aussi les tribus faibles). On peut montrer (cf. [8],§7) qu'elle transforme la probabilité cylindrique sur H définie à l'EXEMPLE 2 en une probabilité sur $(E, \mathcal{C}(E,E')) = (E, \mathcal{B}(E))$, que l'on appelle "mesure de WIENER" et sur laquelle nous reviendrons d'ailleurs.

3. <u>Image quadratique d'une probabilité cylindrique</u>.

Avec les définitions et les notations du CHAPITRE I, §3 nous pouvons énoncer :

PROPOSITION 1. - <u>Soit</u> $\mu = (\mu_M)_{M \in \mathcal{M}(E,F)}$ <u>une probabilité cylindrique sur</u> E <u>relativement à la dualité entre</u> E <u>et</u> F . <u>L'application quadratique</u>

$\chi_E : x \mapsto x \otimes x$, $x \in E$ induit une probabilité cylindrique unique $\chi_E(\mu)$ sur $E^{2\odot}$ relativement à la dualité entre $E^{2\odot}$ et $F^{2\odot}$ que l'on appellera "image quadratique de μ" .

La démonstration repose sur la PROPOSITION 3, CHAPITRE I, §3 : soit $K \in \mathcal{M}(E^{2\odot}, F^{2\odot})$; il existe $M_K \in \mathcal{M}(E,F)$ tel que $(E/M_K)^{2\odot}$ puisse s'identifier à l'espace quotient $E^{2\odot}/\Gamma_o(M_K)$ où $\Gamma_o(M_K) \in \mathcal{M}(E^{2\odot}, F^{2\odot})$ et $\Gamma_o(M_K) \subset K$.

On constate d'abord que la famille :

$$((E/M)^{2\odot} , \mathcal{B}[(E/M)^{2\odot}] , \nu_M = \chi_{E/M}(\mu_M))_{M \in \mathcal{M}(E,F)}$$

est un système projectif d'espaces probabilisés ; en effet on sait déjà, par la proposition citée que la famille : $((E/M)^{2\odot} , \pi^{2\odot}_{N,M})_{M \in \mathcal{M}(E,F)}$ est un système projectif d'espaces vectoriels de dimension finie ; il suffit de vérifier la relation de cohérence pour les ν_M , $M \in \mathcal{M}(E,F)$; soit $N \supset M$ dans $\mathcal{M}(E,F)$ on a en effet :

$$\nu_N = \chi_{E/N}(\mu_N) = \chi_{E/N}(\pi_{N,M}(\mu_M)) = \pi^{2\odot}_{N,M}(\chi_{E/M}(\mu_M)) = \pi^{2\odot}_{N,M}(\nu_M) .$$

Soit maintenant $\underset{\sim}{\pi}_{K,\Gamma_o(M_K)}$ l'application canonique de $E^{2\odot}/\Gamma_o(M_K)$ sur $E^{2\odot}/K$; posons : $\tilde{\nu}_K = \pi_{K,\Gamma_o(M_K)}(\nu_{M_K})$ et montrons que la famille :

$$(E^{2\odot}/K , \mathcal{B}(E^{2\odot}/K) , \tilde{\nu}_K)_{K \in \mathcal{M}(E^{2\odot}, F^{2\odot})}$$

est un système projectif d'espaces probabilisés qui définit donc une probabilité cylindrique sur $E^{2\odot}$ relativement à la dualité entre $E^{2\odot}$ et $F^{2\odot}$. Soient $K_1 \supset K_2$ dans $\mathcal{M}(E^{2\odot}, F^{2\odot})$ et π_{K_1,K_2} l'application canonique de $E^{2\odot}/K_2$ sur $E^{2\odot}/K_1$; on a :

$$\tilde{\nu}_{K_1} = \pi_{K_1,\Gamma_o(M_{K_1})}\left(\nu_{M_{K_1}}\right) = \pi_{K_1,\Gamma_o(M_{K_1})} \circ \left(\pi_{M_{K_1},M_{K_2}}\right)^{2\odot} \left(\nu_{M_{K_2}}\right)$$

$$= \pi_{K_1,\Gamma_o(M_{K_1})} \circ \pi_{\Gamma_o(M_{K_1}),\Gamma_o(M_{K_2})} \left(\nu_{M_{K_2}}\right)$$

$$= \pi_{K_1,\Gamma_o\left(M_{K_2}\right)}\left(\nu_{M_{K_2}}\right) \quad , \text{ car } K_1 \supset \Gamma_o\left(M_{K_1}\right) \supset \Gamma_o\left(M_{K_2}\right)$$

$$= \pi_{K_1,K_2} \circ \pi_{K_2,\Gamma_o\left(M_{K_2}\right)}\left(\nu_{M_{K_2}}\right) \quad , \text{ car } K_1 \supset K_2 \supset \Gamma_o\left(M_{K_2}\right) .$$

soit,

$\tilde{\nu}_{K_1} = \pi_{K_1,K_2}\left(\tilde{\nu}_{K_2}\right)$. Enfin, la mesure $\tilde{\nu}_K$ est entièrement déterminée par la probabilité ν_{M_K} ; l'unicité résulte de la construction de l'espace M_K. ∎

REMARQUE 4. - (Cas d'un espace de HILBERT séparable). Compte tenu de la REMARQUE 1 de ce § et de la REMARQUE 3, CHAPITRE I, §3, la construction de l'image quadratique d'une probabilité cylindrique définie sur un espace de HILBERT réel séparable se simplifie : Avec les notations habituelles, pour $V \in \mathcal{F}(H)$, soit χ_V la restriction de l'application quadratique χ_H à V. On constate d'abord que la famille :

$$\left(V^{2\odot}, \mathcal{B}(V^{2\odot}), \nu_{V^{2\odot}} = \chi_V(\mu_V)\right)_{V \in \mathcal{F}(H)}$$

est un système projectif d'espaces probabilisés (constitué de sous-espaces vectoriels probabilisés de dimension finie de $H^{2\odot}$) car pour $V \supset W$ dans $\mathcal{F}(H)$, on a : $V^{2\odot} \supset W^{2\odot}$ dans $F(H^{2\odot})$ et
$\nu_{W^{2\odot}} = P^{2\odot}_{V,W}(\nu_{V^{2\odot}}) = P_{V^{2\odot},W^{2\odot}}(\nu_{V^{2\odot}})$; en effet, on a :
$\nu_{W^{2\odot}} = \chi_W(\mu_W) = \chi_W(P_{V,W}(\mu_V)) = P^{2\odot}_{V,W}(\chi_V(\mu_V)) = P^{2\odot}_{V,W}(\nu_{V^{2\odot}})$; de plus l'application $P^{2\odot}_{VW}$ coïncide avec le projecteur orthogonal $P_{V^{2\odot},W^{2\odot}}$ de $V^{2\odot}$ sur $W^{2\odot}$ dans l'espace préhilbertien $H^{2\odot}$. Soit maintenant K un sous-espace de dimension finie de $H^{2\odot}$; si K est de la forme $K = V^{2\odot}$ pour un $V \in \mathcal{F}(H)$, on pose $\tilde{\nu}_K = \nu_{V^{2\odot}}$; sinon, en appelant V_K l'espace vectoriel engendré par $\chi_H^{-1}(K)$, et en remarquant que $V_K^{2\odot}$ est le plus petit espace vectoriel de la forme précédente contenant K, on pose :
$\tilde{\nu}_K = P_{V_K^{2\odot},K}\left(\nu_{V_K^{2\odot}}\right)$ et l'on constate que la famille :

$$\left(K, \mathcal{B}(K), \tilde{\nu}_K\right)_{K \in \mathcal{F}(H^{2\odot})}$$

est <u>un système projectif de sous-espaces probabilisés de dimension finie de $H^{2\odot}$ uniquement déterminé à partir de la probabilité cylindrique</u>

$\mu = (\mu_V)_{V \in \mathcal{F}(H)}$ sur H et l'application quadratique χ_H. C'est la probabilité cylindrique $\chi_H(\mu)$ sur $H^{2\odot}$ relativement à la dualité entre $H^{2\odot}$ et lui-même que lui confère sa structure d'espace préhilbertien séparé, que nous appellerons "image quadratique de la probabilité cylindrique $\mu = (\mu_V)_{V \in \mathcal{F}(H)}$ sur H" . ∎

Nous étudierons au chapitre suivant l'image quadratique de la probabilité cylindrique sur un espace de HILBERT définie à l'EXEMPLE 2 , et nous utiliserons le résultat suivant qui combine l'image quadratique et l'image linéaire d'une probabilité cylindrique.

PROPOSITION 2. - Soient (E_1, F_1) un autre couple d'espaces vectoriels réels en dualité séparante et u une application linéaire de E dans E_1, continue pour les topologies faibles $\sigma(E,F)$ et $\sigma(E_1,F_1)$. Pour toute probabilité cylindrique μ sur E relativement à la dualité entre E et F , on a : $\chi_{E_1}(u(\mu)) = u^{2\odot}(\chi_E(\mu))$ sur $E_1^{2\odot}$. De plus, si u transforme la probabilité cylindrique en une probabilité cylindrique qui induit une probabilité sur $(E_1, \mathcal{C}(E_1,F_1))$, cette égalité définit une probabilité cylindrique qui induit une probabilité sur $(E_1^{2\odot}, \mathcal{C}(E_1^{2\odot}, F_1^{2\odot}))$.

En effet, cette proposition résulte directement de la construction de l'image quadratique d'une probabilité cylindrique, de la commutativité du diagramme :

$$\begin{array}{ccc} E & \xrightarrow{\chi_E} & E^{2\odot} \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle u^{2\odot}} \\ E_1 & \xrightarrow{\chi_{E_1}} & E_1^{2\odot} \end{array}$$

et de la continuité de l'application linéaire $u^{2\odot}$ pour les topologies faibles $\sigma(E^{2\odot}, F^{2\odot})$ et $\sigma(E_1^{2\odot}, F_1^{2\odot})$. Enfin si $u(\mu)$ induit une probabilité sur $(E_1, \mathcal{C}(E_1, F_1))$, la probabilité cylindrique $\chi_{E_1}(u(\mu))$ induit une probabilité sur $(E_1^{2\odot}, \mathcal{C}(E_1^{2\odot}, F_1^{2\odot}))$ qui coïncide nécessairement avec la probabilité image de la probabilité précédente par l'application χ_{E_1} , mesurable pour les tribus $\mathcal{C}(E_1, F_1)$ et $\mathcal{C}(E_1^{2\odot}, F_1^{2\odot})$ d'après la REMARQUE 2 . Ainsi, si X est un vecteur aléatoire à valeurs dans $(E_1, \mathcal{C}(E_1, F_1))$ de loi de probabilité cylindrique $u(\mu)$, alors $\chi_{E_1}(u(\mu))$ est la loi de probabilité cylindrique de l'élément aléatoire $X \otimes X$ à valeurs dans $(E_1^{2\odot}, \mathcal{C}(E_1^{2\odot}, F_1^{2\odot}))$. ∎

4. Fonctionnelles associées à une probabilité cylindrique ou à une probabilité.

a) Transformée de FOURIER.

DEFINITION 2. - Soit $\mu = (\mu_M)_{M \in \mathcal{M}(E,F)}$ une probabilité cylindrique sur E relativement à la dualité entre E et F et pour tout $y \in F$ soit μ_y la probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, image de μ par la forme linéaire $x \longmapsto \langle x,y \rangle$, $x \in E$. La transformée de FOURIER de μ est la fonctionnelle complexe φ_μ, définie sur F par : $\varphi_\mu(y) = \int_{\mathbb{R}} e^{it} d\mu_y(t)$, $y \in F$.

Soient $(E, \mathcal{a})$ un E.V.M.S., et P une probabilité sur $(E, \mathcal{a})$; alors P est une probabilité sur l'E.V.M.S. affaibli $(E, \mathcal{C}(E, (E,\mathcal{a})^m))$ et comme $(E,\mathcal{a})^m$ coïncide avec le dual topologique de E pour la topologie faible $\sigma(E, (E,\mathcal{a})^m)$, la transformée de FOURIER de la probabilité cylindrique associée à P sur E relativement à la dualité entre E et $(E,\mathcal{a})^m$ peut se calculer pour tout $y \in (E,\mathcal{a})^m$ par : $\varphi_P(y) = \int_E e^{i\langle x,y \rangle} dP(x)$. (En notant $\langle .,. \rangle$ la forme bilinéaire mettant E et $(E,\mathcal{a})^m$ en dualité).

On appellera cette fonctionnelle associée à P, la transformée de FOURIER de la probabilité P définie sur l'E.V.M.S.$(E,\mathcal{a})$" . ∎

En se reportant à la définition de l'image linéaire d'une probabilité cylindrique on vérifie facilement, avec les notations du n° 2, que $\varphi_{u(\mu)} = \varphi_\mu \circ {}^t u$ où ${}^t u$ désigne l'application linéaire transposée de l'application linéaire faiblement continue u.

Le théorème suivant permet de caractériser une probabilité cylindrique par sa transformée de FOURIER :

THEOREME 3. - (cf. [3], p.19) Il existe une bijection entre l'ensemble des probabilités cylindriques sur E relativement à la dualité entre E et F

<u>et l'ensemble des fonctionnelles</u> φ <u>de</u> F <u>dans</u> $\mathbb{C}$, <u>de type positif,</u> <u>telles que</u> $\varphi(0) = 1$ <u>et dont les restrictions aux sous-espaces vectoriels de dimension finie de</u> F <u>sont continues</u>.

De ce théorème, on déduit en particulier que la transformée de FOURIER d'une probabilité définie sur un E.V.M.S., décrite plus haut, caractérise entièrement cette probabilité.

On en déduit aussi, compte tenu du THEOREME 2 , que l'ensemble des fonctionnelles φ de F dans $\mathbb{C}$, de type positif, telles que $\varphi(0) = 1$ et continues sur les sous-espaces vectoriels de dimension finie de F est en bijection avec l'ensemble des probabilités P^* définies sur $(F^*, \mathcal{C}(F^*, F))$; par suite il existe un prolongement unique d'une telle fonctionnelle φ à l'espace vectoriel $(F^*, \mathcal{C}(F^*, F))^m \supset F$ à savoir la transformée de FOURIER de la probabilité P^* correspondante. ∎

b) <u>Transformée de LAPLACE réelle</u>.

Soient $\mu = (\mu_M)_{M \in \mathcal{M}(E,F)}$ une probabilité cylindrique sur E relativement à la dualité entre E et F et pour tout $y \in F$, μ_y la probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ image de μ par la forme linéaire $x \longmapsto \langle x, y\rangle$, $x \in E$. Alors le sous-ensemble de F : $D_\mu = \{y \in F : \int_{\mathbb{R}} e^t d\mu_y(t) < \infty\}$ est un convexe non vide. En effet, soient P^* la probabilité unique sur $(F^*, \mathcal{C}(F^*, F))$ associée à μ et P^*_y la probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ image de P^* par la forme linéaire mesurable $y^* \longmapsto \langle y^*, y\rangle$ définie sur F^* ; on a nécessairement : $P^*_y = \mu_y$ et donc : $\int_{\mathbb{R}} e^t d\mu_y(t) < \infty \Longleftrightarrow \int_{\mathbb{R}} e^t dP^*_y(t) = \int_{F^*} e^{\langle y^*, y\rangle} dP^*(y^*) < \infty$. Mais cette dernière intégrale définit une fonctionnelle convexe l_μ sur F (à cause de la convexité de la fonction exponentielle) par conséquent, $D_\mu = l_\mu^{-1}(\mathbb{R})$ est un convexe contenant l'origine dans F .

DEFINITION 3. - <u>On appelle</u> "<u>transformée de LAPLACE réelle de la probabilité cylindrique</u> μ <u>sur</u> E <u>relativement à la dualité entre</u> E <u>et</u> F" , <u>la fonctionnelle</u> (<u>convexe</u>) l_μ <u>définie sur le sous-ensemble convexe non vide</u> D_μ <u>de</u> F <u>par</u> :

$$y \in D_\mu \ , \ l_\mu(y) = \int_{\mathbf{R}} e^t d\mu_y(t) \ .$$

Comme on l'a justifié pour la transformée de FOURIER, il convient d'appeler "transformée de LAPLACE réelle d'une probabilité P définie sur l'E.V.M.S. $(E,\mathcal{a})$", la fonctionnelle convexe l_P définie sur le sous-ensemble convexe D_P de $(E,\mathcal{a})^m$ comme étant la transformée de LAPLACE réelle de la probabilité cylindrique sur E relativement à la dualité entre E et $(E,\mathcal{a})^m$ associée à P. ■

Soient E_1 et F_1 un autre couple d'espaces vectoriels réels en dualité séparante et u une application linéaire de E dans E_1 continue pour les topologies faibles $\sigma(E,F)$ et $\sigma(E_1,F_1)$. On constate que $l_{u(\mu)} = l_\mu \circ {}^t u$ sur le convexe $D_{u(\mu)} = {}^t u^{-1}(D_\mu)$ de F_1. En particulier, si $M \in \mathcal{M}(E,F)$, en prenant pour u l'application canonique π_M de E sur E/M, ${}^t\pi_M$ est un isomorphisme du dual de E/M sur le sous-espace de dimension finie M^o, polaire de M, de F et pour la probabilité $\mu_M = \pi_M(\mu)$ sur $(E/M, C(E/M,M^o)) = (E/M, \mathcal{B}(E/M))$ on a :
$D_{\mu_M} = {}^t\pi_M^{-1}(D_\mu) = D_\mu \cap M^o$ et donc pour tout $y \in D_\mu \cap M^o$ on a :
$l_\mu(y) = \int_{E/M} e^{\langle x,y\rangle} d\mu_M(x)$ et comme $F = \bigcup_{M\in\mathcal{M}(E,F)} M^o$, $D_\mu = \bigcup_{M\in\mathcal{M}(E,F)} D_\mu \cap M^o$;
de sorte que cette formule caractérise entièrement la fonctionnelle l_μ sur D_μ.

La transformée de LAPLACE complexe s'obtient en compléxifiant l'espace F : soit $\tilde{F}$ l'espace $F \times F$ qui, muni de l'opération d'addition habituelle de ses éléments et de la multiplication par un nombre complexe définie par : $z = \alpha + i\beta$, $\tilde{y} = (y_1,y_2) \in F\times F$, $z\tilde{y} = (\alpha y_1 - \beta y_2, \beta y_1 + \alpha y_2)$ est un espace vectoriel sur $\mathbb{C}$ dont les éléments peuvent s'écrire : $(y_1,0) + i(y_2,0)$ et dont F peut être considéré comme un sous-espace à condition d'identifier $y \in F$ à $(y,0) \in \tilde{F}$. On appelle "transformée de LAPLACE complexe de μ", la fonctionnelle $\tilde{l}_\mu$ définie sur la bande $\{\tilde{y} \in \tilde{F} : Re(\tilde{y}) \in D_\mu \subset F\}$ par :

$$\tilde{y} = (y_1,y_2) \ , \ \tilde{l}_\mu(\tilde{y}) = \int_{\mathbf{R}^2} e^{t_1 + it_2} \, d\mu_{y_1,y_2}(t_1,t_2)$$

en notant μ_{y_1,y_2} la probabilité sur $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ image de μ par l'application linéaire (faiblement continue) $x \longmapsto (\langle x,y_1\rangle, \langle x,y_2\rangle)$, $x \in E$.

Pour $y \in F$ on a donc : $\varphi_\mu(y) = \tilde{l}_\mu(0,iy)$ et pour $y \in D_\mu$, $y \in D_\mu$, $l_\mu(y) = \varphi_\mu(-iy)$.

c) <u>Fonction cumulante</u>.

DEFINITION 4. - <u>On appelle "fonction cumulante d'une probabilité cylindrique μ sur E relativement à la dualité entre E et F [resp. d'une probabilité P définie sur un E.V.M.S. $(E,\mathcal{A})$]" la fonctionnelle c_μ [resp. c_P] définie sur le convexe D_μ de F [resp. D_P de $(E,\mathcal{A})^m$] par</u> :

$$y \in D_\mu \text{ , } c_\mu(y) = \text{Log}\, l_\mu(y) \text{ [resp. } y \in D_P, c_P(y) = \text{Log}\, l_P(y)] \text{ .}$$

Cette fonctionnelle, logarithme de la transformée de LAPLACE réelle, se trouve être également convexe sur son domaine de définition, (cf. [4]).

Les trois fonctionnelles ainsi définies associées à une mesure cylindrique ou a une probabilité seront des outils utilisés à différents stades de ce travail.

REMARQUE 5. - Lorsque P est une probabilité (ou une probabilité cylindrique) sur un E.V.M.S. $(E,\mathcal{A})$, et u est une application linéaire mesurable de $(E,\mathcal{A})$ dans $(E_1,\mathcal{A}_1)$, les formules pour la loi de probabilité (cylindrique) $u(P) = P \circ u^{-1}$ de u : $\varphi_{u(P)} = \varphi_P \circ {}^t u$; $l_{u(p)} = l_P \circ {}^t u$ et $c_{u(P)} = c_P \circ {}^t u$ sont valables à condition de prendre pour ${}^t u$ l'application linéaire de $(E_1,\mathcal{A}_1)^m$ dans $(E,\mathcal{A})^m$ transposée de l'application mesurable u , compte-tenu de la PROPOSITION 5, CHAPITRE I, §2 . ∎

5. <u>Analyse convexe d'une probabilité</u>.

Il s'agit essentiellement de caractériser le support d'une probabilité définie sur un E.V.T. muni de sa tribu borélienne au moyen de l'analyse convexe de la fonction cumulante de cette probabilité ; c'est un problème d'importance pratique car lorsque cette probabilité est la loi d'un élément aléatoire, il est nécessaire de savoir dans quel domaine il prend presque-sûrement ses valeurs.

Cette étude a été faite en dimension finie par O. BARNDORFF-NIELSEN dans [4]. Bien que cette théorie n'ait pas pu être étendue de façon complète en dimension quelconque, il semble que le cadre dans lequel on se place ici convienne tout à fait à cette extension et que cette étude puisse se poursuivre de façon féconde ; c'est pourquoi nous en donnons ici quelques éléments.

Rappelons d'abord les définitions concernant le support d'une probabilité :

DEFINITION 5. - Soit Ω un espace topologique. Une probabilité P sur $(\Omega, \mathcal{B}(\Omega))$ est dite intérieurement régulière si pour tout $A \in \mathcal{B}(\Omega)$, $P(A) = \sup P(K)$ lorsque K parcourt l'ensemble des parties compactes de A.

On peut montrer, (cf. [26]), que toute probabilité borélienne P définie sur un espace métrique séparable et complet Ω est intérieurement régulière.

Il existe alors un sous-ensemble fermé unique S de Ω tel que (i) : $P(S) = 1$, (ii) : $S \subset D$ pour tout fermé D vérifiant $P(D) = 1$. De plus, S est l'ensemble de tous les points $\omega \in \Omega$ tels que $P(U) > 0$ pour tout ouvert U contenant ω. On appelle S le "support topologique de P", noté : $\mathrm{supp}(P)$.

Soient E un E.V.T.L.C.S. et P une probabilité sur $(E, \mathcal{B}(E))$; compte tenu de ce qui précède nous nous restreindrons à un cas où nous sommes assurés de l'existence du support de P, c'est-à-dire au cas où E est un espace de FRECHET séparable. Nous savons alors que $\mathcal{B}(E) = \mathcal{C}(E,E')$ et que $(E, \mathcal{B}(E))^m = E'$ (où E' est le dual topologique de E) d'après le THEOREME 4, CHAPITRE I, §2.

On déduit du n°3 que la transformée de LAPLACE réelle l_P et la fonction cumulante c_P associées à P sont des fonctions convexes définies sur un domaine convexe non vide D_P de E'. On considère plus particulièrement ici la fonction cumulante.

Suivant [20] on étend le domaine de la fonction c_P à l'espace E' tout entier et on la considère comme une fonctionnelle convexe sur E' à valeurs dans $\mathbb{R} \cup \{+\infty\}$ en posant $c_P(y) = +\infty$ pour $y \notin D_P$; par suite, D_P est le "domaine effectif" de cette fonctionnelle, noté encore : $\mathrm{dom}(c_P)$. On a la propriété de continuité suivante :

PROPOSITION 3. - La fonction cumulante c_P sur E' est une fonctionnelle strictement convexe sur son domaine effectif D_P et elle est semi-continue inférieurement (s.c.i.) pour toute topologie sur E' compatible avec la dualité entre E et E' .

En effet, la stricte convexité provient du fait que la fonction Log est strictement croissante et la fonction exponentielle strictement convexe. c_P est s.c.i. si et seulement si $\{y \in E' : c_P(y) \leq r\}$ est fermé pour tout $r \in \mathbb{R}$; comme les convexes fermés sont les mêmes pour toutes les topologies localement convexes sur E' compatibles avec la dualité, il suffit de le vérifier pour la topologie faible $\sigma(E',E)$. Cette définition est équivalente à : pour tout $y \in E'$, $c_P(y) = \liminf_{y' \to y} c_P(y')$ et comme $c_P(y) \geq \liminf_{y' \to y} c_P(y')$ dès que $D_P = \mathrm{dom}(c_P)$ contient plus d'un point, il suffit de vérifier l'inégalité contraire, elle résulte du lemme de FATOU :

$$\text{si } y \in D_P \text{ , } c_P(y) = \mathrm{Log} \int_E e^{\langle x,y\rangle} dP(x) = \mathrm{Log} \int_E \liminf_{y' \to y} e^{\langle x,y'\rangle} dP(x)$$

$$\leq \mathrm{Log} \liminf_{y' \to y} \int_E e^{\langle x,y'\rangle} dP(x) = \liminf_{y' \to y} c_P(y') .$$

si $y \notin D_P$, l'inégalité est triviale. ■

Cette propriété est équivalente à $c_P \in \Gamma(E')$ = ensemble des fonctions qui sont enveloppes supérieures de familles de fonctions affines continues (cf. [20], pp.27,28).

On définit alors la "polaire" c_P^* de c_P par :

$$x \in E \ , \quad c_P^*(x) = \sup_{y \in E'}\{\langle x,y\rangle - c_P(y)\} = \sup_{y \in D_P}\{\langle x,y\rangle - c_P(y)\} \ .$$

La fonctionnelle convexe sur E, c_P^* appartient à $\Gamma(E)$ et $(c_P^*)^* = c_P$. On dit que c_P <u>et</u> c_P^* <u>sont deux fonctions convexes duales</u>" et que les points $x \in E$ et $y \in E'$ sont "<u>conjugués</u> par rapport au couple (c_P, c_P^*)" si l'on a : $c_P^*(x) + c_P(y) = \langle x,y\rangle$. Enfin, si A est une partie convexe de E' (par exemple), on note ψ_A son indicatrice convexe, c'est-à-dire la fonction dans $\mathbb{R} \cup \{+\infty\}$ définie par $\psi_A(y) = 0$ si $y \in A$, $+\infty$ sinon ; sa polaire $\psi_A^*(x) = \sup_{y \in A}\{\langle x,y\rangle\}$ est appelée la "<u>fonction d'appui de</u> A".

Lorsque E est de dimension finie il existe des relations entre le support de P et le domaine effectif de c_P^* que l'on résume ainsi :

THEOREME 4. - (cf. [4],p.6.3). <u>Si</u> P <u>est une probabilité sur</u> $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ <u>telle que l'intérieur</u> $\overset{\circ}{D}_P$ <u>de</u> D_P <u>soit non vide</u>, <u>en notant</u> Ω <u>l'enveloppe convexe fermée de supp</u>(P) <u>on a</u> :

(i)- $\overset{\circ}{\Omega} \neq \emptyset$,

(ii)- $\mathrm{dom}(c_P^*) \subset \Omega$,

(iii)- $\overset{\circ}{\mathrm{dom}}(c_P^*) = \overset{\circ}{\Omega}$ et donc $\overline{\mathrm{dom}(c_P^*)} = \Omega$.

Lorsque E est de dimension infinie, on peut espérer étendre certains de ces résultats on a déjà :

PROPOSITION 4. - <u>Soient</u> E <u>un espace de FRECHET séparable et</u> P <u>une probabilité sur</u> $(E, \mathcal{B}(E))$. <u>Le domaine effectif de la polaire de la fonction cumulante de</u> P <u>est contenu dans l'enveloppe convexe fermée du support de</u> P , <u>ce que l'on écrit</u> : $\mathrm{dom}(c_P^*) \subset \Omega$.

Montrons, en effet, que $\Omega^c \subset [\mathrm{dom}(c_P^*)]^c$; c'est-à-dire que si $x_o \notin \Omega$, $c_P^*(x_o) = +\infty$. Comme Ω est un convexe fermé dans E, il existe d'après le théorème de HAHN-BANACH un hyperplan fermé séparant strictement x_o et Ω, ce que l'on peut traduire par : il existe un réel $\alpha > 0$ et un élément $y_o \in E'$ tels que : $\langle x_o, y_o\rangle > \alpha$ et $\langle x, y_o\rangle < \alpha$ pour tout $x \in \Omega$. Alors, pour tout $r \in \mathbb{R}_+$, $c_P(ry_o) = \mathrm{Log} \int_E e^{r\langle x,y_o\rangle}\, dP(x) \le r\alpha$;

or $c_P^*(x_o) \geq [\langle x_o, r y_o\rangle - c_P(r y_o)]$ quel que soit $r \in \mathbb{R}_+$ donc :
$c_P^*(x_o) \geq \lim\limits_{r\to+\infty} [\langle x_o, r y_o\rangle - c_P(r y_o)] = +\infty$. ∎

Pour généraliser (iii) du THEOREME 4 , il suffit de montrer que $\overset{o}{\Omega} \subset \mathrm{dom}(c_P^*)$ ce qui peut découler du lemme suivant :

LEMME 1. - <u>Une condition suffisante pour que</u> $\overset{o}{\Omega} \subset \mathrm{dom}(c_P^*)$ <u>est que pour tout</u> $x_o \in \overset{o}{\Omega}$ <u>on ait</u> : $\zeta(x_o) = \inf\limits_{y \in E'} P\{x \in E : \langle x - x_o, y\rangle \geq 0\} > 0$.

En effet, pour tout $x_o \in E$ et tout $y \in E'$ on a, en posant $E(x_o, y) = \{x \in E : \langle x - x_o, y\rangle \geq 0\}$, $\int_E e^{\langle x,y\rangle} dP(x) = \int_{E(x_o,y)} e^{\langle x,y\rangle} dP(x)$

$$+ \int_{E(x_o,y)^c} e^{\langle x,y\rangle} dP(x) \geq e^{\langle x_o, y\rangle} - P[E(x_o, y)]$$

donc : $c_P(y) \geq \langle x_o, y\rangle + \mathrm{Log}\, P[E(x_o, y)]$,
soit : $\langle x_o, y\rangle - c_P(y) \leq - \mathrm{Log}\, P[E(x_o, y)]$,
et $c_P^*(x_o) \leq - \mathrm{Log}\, \zeta(x_o)$. Il est donc suffisant, lorsque $x_o \in \overset{o}{\Omega}$ que l'on ait $\zeta(x_o) > 0$ pour que $x_o \in \mathrm{dom}(c_P^*)$.

Il n'est pas surprenant que le THEOREME 4 ne se généralise pas systématiquement en dimension infinie puisque la fonction cumulante se définit (aussi) pour une probabilité cylindrique, elle est donc spécifique de cette notion. Nous verrons toutefois que dans le cas gaussien ce théorème se généralise. ∎

§ 2. - DUAL D'UN ESPACE VECTORIEL PROBABILISE.

1. Définitions et caractérisations.

Soit P une probabilité sur un E.V.M.S. $(E,\mathfrak{a})$.

Tout élément du dual $(E,\mathfrak{a})^m$ est une variable aléatoire réelle définie sur cet espace probabilisé ; si une suite $\{f_n\}_{n\geq 1}$ de telles fonctionnelles linéaires mesurables converge simplement vers f , cette limite est encore une fonctionnelle linéaire mesurable ; mais si l'on suppose que la suite $\{f_n\}_{n\geq 1}$ converge seulement P-presque partout simplement, on constate que sa limite f est encore une fonctionnelle mesurable, mais qu'elle n'est définie que sur un sous-espace vectoriel mesurable de probabilité 1 de E , sur lequel elle est d'ailleurs linéaire, puisque pour tous α et β réels, x_1 et x_2 dans E on a :

$$\alpha \lim_{n\to\infty} f_n(x_1) + \beta \lim_{n\to\infty} f_n(x_2) = \lim_{n\to\infty} f_n(\alpha x_1 + \beta x_2) .$$

On est donc amené à considérer un nouveau type de fonctionnelles linéaires, déjà envisagé d'ailleurs par divers auteurs dans certains cas particuliers. (cf. [9], [15], par ex.)

DEFINITION 1. - <u>On appellera "variable aléatoire réelle linéaire" (v.a.r.l.) sur l'espace vectoriel probabilisé</u> $(E,\mathfrak{a},P)$, <u>toute fonctionnelle</u> $\mathfrak{a}$-<u>mesurable définie et linéaire sur un sous-espace vectoriel (mesurable) de probabilité</u> 1 <u>de</u> E .

Si f_1 et f_2 sont deux telles v.a.r.l. de domaines de définition respectifs E_{f_1} et E_{f_2} , quels que soient les réels α et β , la fonctionnelle $\alpha f_1 + \beta f_2$ est linéaire, mesurable et définie sur le sous-espace vectoriel mesurable $E_{f_1} \cap E_{f_2}$ de probabilité 1 de E . Cependant, l'ensemble des v.a.r.l. que l'on notera $\mathcal{L}(E,\mathfrak{a},P)$, qui contient d'ailleurs $(E,\mathfrak{a})^m$, n'est pas en général un espace vectoriel car si la fonction 0 définie sur E est l'élément neutre pour l'addition, étant donnée une v.a.r.l. f il n'existe pas nécessairement de v.a.r.l. g telle que $f + g = 0$.

Par contre, on ne distinguera pas entre elles deux telles v.a.r.l. P-presque sûrement égales, d'où :

DEFINITION 2. - On appellera "dual de l'espace vectoriel probabilisé $(E,\mathcal{a},P)$" l'espace vectoriel, que l'on notera : $L(E,\mathcal{a},P)$, des classes d'équivalences des variables aléatoires réelles linéaires sur cet espace.

C'est donc l'espace quotient de $\mathcal{L}(E,\mathcal{a},P)$ par la relation de P-équivalence et l'on a aussi : $(E,\mathcal{a})^m/P \subset \mathcal{L}(E,\mathcal{a},P)/P = L(E,\mathcal{a},P)$.

REMARQUE 1. - Si la tribu $\mathcal{a}$ est supposée contenir tous les ensembles P-négligeables, on peut facilement prolonger toute v.a.r.l. f en une v.a.r.l. définie partout sur E (ceci est incompatible avec l'hypothèse faite dans [28] , chap. IV, § 2) . Par suite, en notant $\overline{\mathcal{a}}$ la tribu complétée de $\mathcal{a}$ par rapport à P et $\overline{P}$ l'extension de P à $\overline{\mathcal{a}}$ on obtient que $L(E,\overline{\mathcal{a}},\overline{P}) = (E,\overline{\mathcal{a}})^m/\overline{P}$ (en remarquant que la complétion des tribus sur E n'altère pas leur caractère de compatibilité avec la structure d'espace vectoriel de E . ∎

DEFINITION 3. - On appelle "tribu engendrée par la classe d'équivalence d'une v.a.r.l. définie sur $(E,\mathcal{a},P)$" , la plus petite sous-tribu de $\mathcal{a}$ qui rende mesurable chacune de ses versions.

La tribu engendrée par une famille de classes d'équivalence de v.a.r. se définit aisément ; en particulier,

DEFINITION 4. - La sous-tribu $\mathcal{a}_L$ de $\mathcal{a}$ engendrée par les éléments de $L(E,\mathcal{a},P)$ sera appelée "sous-tribu linéaire de $\mathcal{a}$ relativement à P" et l'on appellera "espace vectoriel probabilisé affaibli de $(E,\mathcal{a},P)$" l'espace vectoriel probabilisé $(E,\mathcal{a}_L,P\mathcal{a}_L)$ où $P\mathcal{a}_L$ désigne la restriction de P à $\mathcal{a}_L$. On a évidemment : $L(E,\mathcal{a}_L,P\mathcal{a}_L) = L(E,\mathcal{a},P)$.

REMARQUE 2. - Si P' est une probabilité équivalente à P sur $(E,\mathcal{A})$, $L(E,\mathcal{A},P') = L(E,\mathcal{A},P)$, ce qui nous permettra de parler de dual d'une structure statistique vectorielle $(E,\mathcal{A},\mathcal{P})$ dont la famille $\mathcal{P}$ est constituée de lois de probabilité équivalentes.

REMARQUE 3. - Si $\xi \in L(E,\mathcal{A},P)$, les éléments de la classe d'équivalence de ξ dans $\mathcal{L}(E,\mathcal{A},P)$ n'ont pas nécessairement même domaine de définition et cette classe n'étant pas nécessairement finie ou dénombrable, on ne peut pas parler de domaine de définition unique associé à ξ . ■

REMARQUE 4. - Malgré la terminologie employée, les espaces vectoriels E et $L(E,\mathcal{A},P)$ ne peuvent pas en général être mis en dualité par une forme bilinéaire canonique puisque $L(E,\mathcal{A},P)$ n'est pas un sous-espace vectoriel du dual algébrique E^* de E , sauf si E' est un espace vectoriel de dimension finie muni de sa tribu borélienne, et si l'on suppose que tout sous-espace vectoriel strict de E est de probabilité nulle alors, $\mathcal{L}(E,\mathcal{A},P) = L(E,\mathcal{A},P) = E^*$.

L'espace vectoriel $L(E,\mathcal{A},P)$ est un sous-espace vectoriel de l'espace $L_o(E,\mathcal{A},P)$ de toutes les (classes d'équivalence de) v.a.r. définies sur $(E,\mathcal{A},P)$. En munissant ce dernier de la topologie métrisable de la convergence en probabilité on a :

PROPOSITION 1. - $L(E,\mathcal{A},P)$ <u>est un sous-espace vectoriel fermé de</u> $L_o(E,\mathcal{A},P)$ <u>et</u> $L(E,\mathcal{A},P) \cap L_p(E,\mathcal{A},P)$ <u>est un sous-espace vectoriel fermé de</u> $L_p(E,\mathcal{A},P)$, $1 \leq p < \infty$.

En effet, soit $\{\xi_n\}_{n\geq 1}$ une suite d'éléments de $L(E,\mathcal{A},P)$ convergeant vers $\xi \in L_o(E,\mathcal{A},P)$ soit en probabilité, soit, le cas échéant, en moyenne d'ordre p . D'après des théorèmes classiques, on peut en extraire une sous-suite $\{\xi_{n_k}\}_{k\geq 1}$ qui converge presque-sûrement vers la même limite ξ ; mais alors, ξ est limite presque-sûre d'une suite de (classes d'équivalence de) v.a.r.l. , elle est donc elle-même une (classe d'équivalence de) v.a.r.l. donc un élément de $L(E,\mathcal{A},P)$. ■

COROLLAIRE. - $L(E,\mathcal{A},P)$ contient la fermeture en probabilité de l'espace vectoriel $(E,\mathcal{A})^m/P$; il contient également sa fermeture pour la topologie de $L_p(E,\mathcal{A},P)$ pour un $p : 1 \le p < \infty$ dès que toute fonctionnelle linéaire mesurable sur $(E,\mathcal{A})$ est de p^{eme} puissance intégrable par rapport à P. Si de plus, $(E,\mathcal{A})^m/P$ est dense dans $L(E,\mathcal{A},P)$ pour l'une quelconque de ces topologies, $L(E,\mathcal{A},P)$ est exactement constitué des (classes d'équivalence de) v.a.r. qui sont limites simples presque-sûrement d'éléments du dual $(E,\mathcal{A})^m$ de l'E.V.M.S. $(E,\mathcal{A})$.

Nous savons en effet, que $(E,\mathcal{A})^m \subset \mathcal{L}(E,\mathcal{A},P)$ [resp. $(E,\mathcal{A})^m/P \subset L(E,\mathcal{A},P)$] la première partie du corollaire résulte directement de la proposition. Si la fermeture $\overline{(E,\mathcal{A})^m/P}$ pour la topologie de $L_o(E,\mathcal{A},P)$ ou de $L_p(E,\mathcal{A},P)$, le cas échéant, coïncide avec $L(E,\mathcal{A},P)$, tout élément de ce dernier espace est limite presque-sûre d'une sous-suite extraite d'une suite de (classes d'équivalence de) v.a.r.l. appartenant à $(E,\mathcal{A})^m/P$ qui converge vers cet élément pour la topologie considérée, en vertu du même argument que celui employé dans la démonstration de la proposition ; mais nous savons qu'une telle limite est une (classe d'équivalence de) v.a.r.l. ce qui implique la caractérisation annoncée de $L(E,\mathcal{A},P)$. ∎

Nous aurons l'occasion au chapitre suivant de considérer plusieurs exemples de duals d'E.V.P.

2. Fonctionnelles associées à un espace vectoriel probabilisé.

La fonction complexe définie sur $L(E,\mathcal{A},P)$ par :

$$\xi \in L(E,\mathcal{A},P)\ ,\ \Phi_P(\xi) = \int_E e^{i\xi}\,dP \qquad (1)$$

peut être interprétée comme un prolongement de la transformée de FOURIER φ_P de la probabilité P sur l'E.V.M.S. $(E,\mathcal{A})$ puisque, comme on l'a vu à la suite de la DEFINITION 2, §1, celle-ci est définie sur $(E,\mathcal{A})^m$ par : $\varphi_P(y) = \int_E e^{i\langle x,y\rangle} dP(x)$ (où $\langle .,. \rangle$ est la forme bilinéaire canonique mettant E et $(E,\mathcal{A})^m$ en dualité) et que cette intégrale ne dépend que de la classe d'équivalence de y dans $(E,\mathcal{A})^m/P$.

La fonctionnelle Φ définie par (1) est donc uniquement déterminée par l'E.V.P. $(E,\mathcal{A},P)$:

DEFINITION 5. - On appellera "transformée de FOURIER de l'E.V.P. $(E,\mathcal{A},P)$" la fonctionnelle complexe Φ_P définie sur le dual de cet E.V.P. par :
$\xi \in L(E,\mathcal{A},P)$, $\Phi_P(\xi) = E_P(e^{i\xi})$.

La v.a.r.l. ξ a donc pour fonction caractéristique : $\varphi_\xi(t) = \Phi_P(t\xi)$, $t \in \mathbb{R}$. On vérifie que Φ_P est une fonctionnelle de type positif sur $L(E,\mathcal{A},P)$ telle que $\Phi_P(0) = 1$.

DEFINITION 6. - On appellera "transformée de LAPLACE réelle de l'E.V.P. $(E,\mathcal{A},P)$" [resp. "fonction cumulante de l'E.V.P. $(E,\mathcal{A},P)$"] la fonctionnelle convexe L_P [resp. C_P] définie par :

$$\xi \in L(E,\mathcal{A},P) \ , \ L_P(\xi) = \int_E e^{\xi} dP \ \text{[resp. } C_P(\xi) = \operatorname{Log} L_P(\xi)\text{]}$$

sur le sous-ensemble convexe non vide $D(E,\mathcal{A},P)$ du dual de cet E.V.P. où cette intégrale est finie.

Ces fonctionnelles peuvent être respectivement interprétées comme le prolongement au sous-ensemble convexe $D(E,\mathcal{A},P)$ de $L(E,\mathcal{A},P)$ de la transformée de LAPLACE réelle l_P et de la fonction cumulante C_P définies sur le sous-ensemble convexe D_P/P de $(E,\mathcal{A})^m/P$. On peut définir la transformée de LAPLACE complexe de l'E.V.P. $(E,\mathcal{A},P)$ de façon analogue à celle du §1 n°3.

REMARQUE 5. - Puisque la valeur en ξ des fonctions Φ_P , L_P et C_P ne dépend que de la loi de la v.a.r. ξ , il est clair que ces fonctions sont continues sur leur domaine de définition pour la topologie de $L_o(E,\mathcal{A},P)$ de la convergence en probabilité ou, le cas échéant (si $L(E,\mathcal{A},P) \subset L_p(E,\mathcal{A},P)$ pour un $p : 1 \leq p < \infty$) pour la topologie de la convergence en moyenne d'ordre p . ∎

L'étude des propriétés du convexe $D(E,\mathcal{A},P)$ peut apporter des précisions intéressantes ; nous utiliserons par exemple :

PROPOSITION 2. - Si l'origine de $L(E,\mathcal{A},P)$ est un point interne du convexe $D(E,\mathcal{A},P)$, alors, $L(E,\mathcal{A},P) \subset L_1(E,\mathcal{A},P)$.

En effet, 0 est un point interne de $D(E,\mathcal{A},P)$ si et seulement si tout sous-espace vectoriel de dimension 1 de $L(E,\mathcal{A},P)$ coupe ce convexe en un segment dont 0 est un point intérieur. Soit donc $\xi \in L(E,\mathcal{A},P)$ et montrons que $E(|\xi| < \infty$. Pour tout réel non nul δ fixé,

$$E(|\xi|) = \frac{1}{|\delta|} E(|\delta\xi|) = \frac{1}{|\delta|}\left(\int_{\delta\xi\geq 0} \delta\xi \, dP + \int_{\delta\xi<0} -\delta\xi \, dP\right) \quad ; \text{ comme } \quad u < e^u$$

pour tout u réel, on a :

$$E(|\xi|) \leq \frac{1}{|\delta|}\left(\int_{\delta\xi\geq 0} e^{\delta\xi} dP + \int_{\delta\xi<0} e^{-\delta\xi} dP\right) \leq \frac{L_P(\delta\xi) + L_P(-\delta\xi)}{|\delta|} \quad .$$

Mais, d'après l'hypothèse faite, il existe un réel $\varepsilon > 0$ tel que pour tout réel δ, vérifiant $|\delta| < \varepsilon$, $\delta\xi$ et $-\delta\xi$ appartiennent à $D(E,\mathcal{A},P)$, donc $E(|\xi|) < \infty$. ∎

REMARQUE 6. - Si l'on note $f'(x,h)$, la dérivée selon la direction h, en un point x d'une fonction réelle f définie sur un espace vectoriel, c'est-à-dire la quantité, si elle existe : $\lim\limits_{\varepsilon>0,\varepsilon\downarrow 0} \frac{f(x+\varepsilon h)-f(x)}{\varepsilon}$, on montre facilement que sous l'hypothèse faite à la PROPOSITION 2, pour tout $\xi \in L(E,\mathcal{A},P)$, $E(\xi) = L'_P(0,\xi)$. ∎

3. Application transposée d'un vecteur aléatoire linéaire.

Par extension de la DEFINITION 1, on appellera "vecteur aléatoire linéaire" sur l'E.V.P. $(E,\mathcal{A},P)$ toute (classe d'équivalence d'une) application U à valeurs dans un E.V.M.S. $(E_1,\mathcal{A}_1)$ qui soit mesurable, définie et linéaire sur un sous-espace vectoriel mesurable E_U, de probabilité 1 de E. Soit $P_1 = P \circ U^{-1}$ la loi de probabilité de U sur $(E_1,\mathcal{A}_1)$; alors $P_1[U(E_U)] = P[U^{-1}(U(E_U))] = 1$.

Comme pour les v.a.r. on définira la "tribu engendrée par la classe d'équivalence de U", comme étant la plus petite sous-tribu de $\mathcal{A}$ rendant mesurable chaque version de U .

PROPOSITION 3. - <u>Il existe une application linéaire unique,</u> ${}^tU$ <u>de</u> $L(E_1,\mathcal{A}_1,P_1)$ <u>dans</u> $L(E,\mathcal{A},P)$ <u>continue pour la topologie de la convergence en probabilité sur ces espaces telle que pour tout</u> $\xi_1 \in L(E_1,\mathcal{A}_1,P_1)$, ${}^tU(\xi_1) = \xi_1 \circ U$ <u>et que l'on appellera "application transposée du vecteur aléatoire linéaire</u> U" .

En effet, soit $f_1 \in \mathcal{L}(E_1,\mathcal{A}_1,P_1)$ une v.a.r.l. sur $(E_1,\mathcal{A}_1,P_1)$; $f_1 \circ U$ est une fonctionnelle mesurable et linéaire lorsqu'elle est définie, c'est-à-dire, en notant E_{f_1} le domaine de définition de f_1 dans E_1 , sur l'ensemble : $\{x \in E : U(x) \in E_{f_1}\} = U^{-1}(E_{f_1})$; mais $P[U^{-1}(E_{f_1})] = P_1(E_{f_1}) = 1$; donc $f_1 \circ U \in \mathcal{L}(E,\mathcal{A},P)$. Soit maintenant f_1' appartenant à la classe d'équivalence de f_1 ; alors $P_1\{f_1' \neq f_1\} = 0 = P[U^{-1}\{f_1' \neq f_1\}] = P\{f_1' \circ U \neq f_1 \circ U\}$, donc $f_1' \circ U$ appartient à la classe d'équivalence de $f_1 \circ U$; par conséquent, pour tout $\xi_1 \in L(E_1,\mathcal{A}_1,P_1)$, $\xi_1 \circ U$ a un sens et définit un élément unique de $L(E,\mathcal{A},P)$ noté ${}^tU(\xi_1)$ et l'application linéaire ${}^tU$ de $L(E_1,\mathcal{A}_1,P_1)$ dans $L(E,\mathcal{A},P)$ ainsi définie est unique. Elle est continue pour la topologie de la convergence en probabilité sur ces deux espaces puisque si $\xi_n^1 \xrightarrow{P_1} \xi^1$, alors, $\forall \delta > 0$ $P\{|\xi_n^1 \circ U - \xi^1 \circ U| > \delta\} = P_1\{|\xi_n^1 - \xi^1| > \delta\} \underset{n\to\infty}{\longmapsto} 0$ donc ${}^tU(\xi_n^1) \overset{P}{\longmapsto} {}^tU(\xi^1)$. ∎

4. <u>Propriétés projectives</u>.

Soit $\mathcal{F}(L)$ la famille des sous-espaces vectoriels de dimension finie de $L(E,\mathcal{A},P)$. Si $N \in \mathcal{F}(L)$ et si $\{\xi_1,\dots,\xi_n\}$ est une base de N , le n-uplet $\xi_N = (\xi_1,\dots,\xi_n)$ définit un vecteur aléatoire linéaire sur $(E,\mathcal{A},P)$. On peut identifier l'espace $L(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n), P \circ \xi_N^{-1})$ à $\mathbb{R}^n$ d'après la REMARQUE 4, et si $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{R}^n$ on a :

$$ {}^t\xi_N(\alpha) = \sum_{j=1}^{n} \alpha_j \, \xi_j = \alpha \circ \xi_N \ , $$

ce qui montre que ${}^t\xi_N$ est un isomorphisme de $L(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n), P \circ \xi_N^{-1})$ sur N .

Soit maintenant $\mathcal{A}_N$ la sous-tribu de $\mathcal{A}$ engendrée par ξ_N , il est clair que $\mathcal{A}_L = \bigvee_{N \in \mathcal{F}(L)} \mathcal{A}_N$ puisque $L(E,\mathcal{A},P) = \bigcup_{N \in \mathcal{F}(L)} N$. Alors, en notant plus généralement E_N l'espace euclidien des valeurs de ξ_N

PROPOSITION 4. - <u>Il existe un système projectif d'E.V.P. de dimension finie</u>

$$(E_N, \mathcal{B}(E_N), P \circ \xi_N^{-1})_{N \in \mathcal{F}(L)}$$

<u>qui détermine entièrement l'E.V.P. affaibli</u> $(E, \mathcal{A}_L, P_{\mathcal{A}_L})$.

REMARQUE 7. - Cela signifie que la transformée de FOURIER Φ_P [resp. la transformée de LAPLACE réelle L_P , la fonction cumulante C_P] de l'E.V.P. $(E, \mathcal{A}, P)$ est entièrement déterminée par ses valeurs sur les sous-espaces vectoriels de dimension finie N de $L(E, \mathcal{A}, P)$ [resp. sur les convexes $D(E, \mathcal{A}, P) \cap N$] et que, pour la transformée de FOURIER par exemple,

si $$\xi \in N \ , \ \Phi_P(\xi) = \int_E e^{i\xi} \, dP = \int_{E_N} e^{i\,{}^t\xi_N^{-1}(\xi)} \, dP \circ \xi_N^{-1} \ . \quad \blacksquare$$

5. <u>Variables aléatoires réelles quadratiques</u>.

Sur l'E.V.M.S. $(E, \mathcal{A})$ une forme quadratique mesurable est une application mesurable q de $(E, \mathcal{A})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que :

$$\text{i} - q(\alpha x) = \alpha^2 q(x) \quad \text{pour tout réel } \alpha \ ,$$

$$\text{ii} - \text{l'application} \quad (x,y) \longmapsto \frac{1}{2} [q(x+y) - q(x) - q(y)] \ ; \ x, y \in E \ ,$$

soit une forme bilinéaire (symétrique) sur $E \times E$.

Sur l'E.V.P. $(E, \mathcal{A}, P)$, si une suite $\{q_n\}_{n \geq 1}$ de formes quadratiques mesurables converge presque-sûrement simplement vers une fonctionnelle q , celle-ci est mesurable et définie sur un sous-ensemble de probabilité 1 de E sur lequel elle vérifie (i) .

Soit alors $\mathcal{B}(E \times E, \mathcal{A} \otimes \mathcal{A}, P \otimes P)$ l'ensemble des fonctionnelles f définies sur un sous-espace de la forme $E_f \times E_f$ de $E \times E$ sur lequel f est une forme bilinéaire symétrique mesurable, ou E_f est un sous-espace vectoriel mesurable de probabilité 1 de E .

DEFINITION 7. - <u>On appellera "variable aléatoire réelle quadratique"</u> (<u>v.a.r.q.</u>) <u>sur l'E.V.P.</u> $(E, \mathcal{A}, P)$ <u>toute fonctionnelle mesurable</u> q <u>telle que</u>

$q(x) = f(x,x)$, $x \in E_f$ <u>pour un élément</u> f <u>de</u> $\mathcal{B}(E\times E, \mathcal{A}\otimes\mathcal{A}, P\otimes P)$.

L'ensemble $\mathcal{Q}(E,\mathcal{A},P)$ des v.a.r.q. définies seulement presque-partout n'est pas en général un espace vectoriel, mais l'ensemble des classes d'équivalence de v.a.r.q. est un espace vectoriel que l'on notera $Q(E,\mathcal{A},P)$.

Comme pour le dual d'un E.V.P. on a :

PROPOSITION 5. - $Q(E,\mathcal{A},P)$ <u>est un sous-espace vectoriel fermé de</u> $L_o(E,\mathcal{A},P)$ <u>et</u> $Q(E,\mathcal{A},P)\cap L_p(E,\mathcal{A},P)$ <u>est un sous-espace vectoriel fermé de</u> $L_p(E,\mathcal{A},P)$, $1 \leq p < \infty$.

La démonstration est la même que celle de la PROPOSITION 1. ■

Le but de ce qui suit est d'expliciter le lien entre v.a.r.q. et application quadratique.

Soient χ_E l'application quadratique de E dans $E^{2\odot}$ définie par $\chi_E(x) = x\otimes x$, $x \in E$ (cf.CHAPITRE I, §3, n°3), $\mathcal{A}^{2\chi}$ la plus grande tribu sur $E^{2\odot}$ rendant cette application mesurable et $P_{\chi_E} = P\circ \chi_E^{-1}$ la loi de probabilité du vecteur aléatoire ainsi défini sur $(E,\mathcal{A},P)$.

Notons $\tilde{\mathcal{L}}(E^{2\odot},\mathcal{A}^{2\chi},P_{\chi_E})$ le sous-ensemble de $\mathcal{L}(E^{2\odot},\mathcal{A}^{2\chi},P_{\chi_E})$ constitué des v.a.r.l. u définies sur un sous-espace vectoriel de $E^{2\odot}$ de la forme $E_u^{2\odot}$ où E_u désigne un sous-espace vectoriel de probabilité 1 de E et $\tilde{L}(E^{2\odot},\mathcal{A}^{2\chi},P_{\chi_E})$ l'espace vectoriel de leurs classes de P_{χ_E}-équivalence.

PROPOSITION 6. - <u>Il existe un isomorphisme de</u> $\tilde{L}(E^{2\odot},\mathcal{A}^{2\chi},P_{\chi_E})$ <u>sur</u> $Q(E,\mathcal{A},P)$.

Considérons en effet l'application j définie sur $\tilde{\mathcal{L}}(E^{2\odot},\mathcal{A}^{2\chi},P_{\chi_E})$ par $j(u) = u\circ\chi_E$. Elle est à valeurs dans l'ensemble $\mathcal{Q}(E,\mathcal{A},P)$ puisque la fonction $u\circ\chi_E$ est mesurable et définie sur l'ensemble $\{x\in E : \chi_E(x) \in E_u^{2\odot}\}$ qui est un sous-espace de probabilité 1 de E sur lequel elle vérifie la condition (i) ci-dessus ; de plus l'application

$$(x,y)\longmapsto \frac{1}{2}\,[u\circ\chi_E(x+y) - u\circ\chi_E(x) - u\circ\chi_E(y)] = u(x\odot y)\,;\ x,y \in E_u$$

est une forme bilinéaire symétrique définie sur le sous-espace $E_u \times E_u$ de $E \times E$.

On vérifie aisément que si u_1 et u_2 sont deux éléments de $\tilde{\mathfrak{L}}(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$, $j(u_1 + u_2) = j(u_1) + j(u_2)$ en prenant $E_{u_1 + u_2} = E_{u_1} \cap E_{u_2}$.

Si $q \in \mathfrak{Q}(E,\mathcal{a},P)$, il existe $u \in \tilde{\mathfrak{L}}(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$ tel que $q = u \circ \chi_E$; il suffit de poser $u(\sum_{i=1}^{n} \lambda_i \, x_i \otimes x_i) = \sum_{i=1}^{n} \lambda_i \, q(x_i)$; $\lambda_i \in \mathbb{R}$, $x_i \in E_q$, $i = 1,\dots,n$, pour constater que u est linéaire mesurable définie sur le sous-espace $E_q^{2\odot}$ de $E^{2\odot}$ de P_{χ_E}-probabilité 1 . Ceci montre que j est surjective.

Enfin il est clair que si $u \circ \chi_E$ est nulle sur son domaine de définition u l'est aussi sur le sien.

Par conséquent l'application j définit par passage aux quotients une application linéaire injective et surjective de $\tilde{L}(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$ sur $Q(E,\mathcal{a},P)$ qui est l'isomorphisme annoncé. ∎

COROLLAIRE 1. - <u>Le produit tensoriel symétrique</u> $L(E,\mathcal{a},P)^{2\odot}$ <u>est isomorphe à un sous-espace vectoriel de</u> $L(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$.

En effet, si ξ_1 et ξ_2 sont deux éléments de $L(E,\mathcal{a},P)$, en posant $i(\xi_1 \odot \xi_2) = \xi_1 \cdot \xi_2$ on définit une injection linéaire i de $L(E,\mathcal{a},P)^{2\odot}$ dans $Q(E,\mathcal{a},P)$ et compte tenu de la proposition, on peut construire un isomorphisme de $L(E,\mathcal{a},P)^{2\odot}$ sur un sous-espace vectoriel de $L(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$.

De plus on constate que :

COROLLAIRE 2. - <u>Tout élément</u> $\zeta \in L(E,\mathcal{a},P)^{2\odot}$ <u>admet une version de la forme</u> $u \circ \chi_E$ <u>où</u> u <u>est une v.a.r.l. unique à une équivalence près sur l'E.V.P.</u> $(E^{2\odot}, \mathcal{a}^{2\chi}, P_{\chi_E})$ <u>où</u> χ_E <u>est l'application quadratique sur</u> E .

Ce dernier corollaire, dont nous verrons une illustration au chapitre suivant, jouera un rôle décisif dans une application de la théorie statistique exposée au CHAPITRE IV.

§ 3. - ESPACES DE HILBERT ASSOCIES A UN ESPACE VECTORIEL PROBABILISE.

Ayant étendu le domaine de définition de la transformée de FOURIER d'une probabilité définie sur un espace vectoriel mesurable à l'espace vectoriel des v.a.r.l. nous pouvons élargir également le cadre d'une théorie classique (cf. [14]) basée sur les propriétés analytiques de cette fonctionnelle. Mais nous pouvons, alors, élaborer une théorie semblable à partir de la transformée de LAPLACE réelle de l'E.V.P. qui sera utile dans certains problèmes d'estimation de paramètres en statistique.

Soient $(E,\mathcal{A},P)$ un E.V.P. faible c'est-à-dire tel que $\mathcal{A}$ coïncide avec la tribu linéaire $\mathcal{A}_L$ (cf. DEFINITION 4, §2) - ce que l'on peut toujours supposer sans restreindre la généralité - et Φ_P et L_P les transformées de FOURIER et de LAPLACE de cet E.V.P. définies respectivement sur son dual $L(E,\mathcal{A},P)$ et sur le convexe $D(E,\mathcal{A},P)$ de ce dual.

1. <u>L'espace de HILBERT autoreproduisant associé à la transformée de FOURIER</u>.

LEMME 1. - <u>La fonction complexe</u> K_{Φ_P} <u>définie sur</u> $L(E,\mathcal{A},P) \times L(E,\mathcal{A},P)$ <u>par</u> : $K_{\Phi_P}(\xi,\xi') = \Phi_P(\xi-\xi')$, $\xi,\xi' \in L(E,\mathcal{A},P)$ <u>est un noyau hermitien de type positif</u>. ∎

Ceci résulte directement de la définition d'un noyau hermitien de type positif (à savoir que : $K_{\Phi_P}(\xi,\xi') = \overline{K_{\Phi_P}(\xi',\xi)}$ et que $\sum_{i,j=1}^{n} \lambda_i \overline{\lambda_j}\, K_{\Phi_P}(\xi_i,\xi_j) \geq 0$ quels que soient l'entier n , les nombres complexes $\lambda_1,\ldots,\lambda_n$ et les éléments $\xi_1,\ldots,\xi_n$ de $L(E,\mathcal{A},P)$) et de la propriété pour la transformée de FOURIER d'être de type positif sur $L(E,\mathcal{A},P)$. ∎

LEMME 2. - <u>L'algèbre engendrée par les variables aléatoires complexes</u> (v.a.c.) $\{e^{i\xi}\}_{\xi \in L(E,\mathcal{A},P)}$ <u>sur l'E.V.P.</u> $(E,\mathcal{A},P)$ <u>est dense dans l'espace de HILBERT complexe</u> $L_2^{\mathbb{C}}(E,\mathcal{A},P)$ <u>constitué des v.a.c.</u> Z <u>telles que</u> $E_P(|Z|^2) < \infty$ <u>muni du produit scalaire</u> : $\langle Z,Z'\rangle = E_P(Z\bar{Z}')$.

Il suffit de prouver que toute v.a.c. $Z \in L_2^{\mathbb{C}}(E,\mathcal{A},P)$, orthogonale aux v.a.c. de la forme : $\prod_{k=1}^{n} e^{i\alpha_k \xi_k}$, $\alpha_1,\dots,\alpha_n \in \mathbb{R}$, $\xi_1,\dots,\xi_k \in L(E,\mathcal{A},P)$ est nécessairement nulle. Supposons donc que :

$$E_P(e^{i\sum_{j=1}^{n} \alpha_j \xi_j} \cdot \bar{Z}) = 0$$

et soit $\mathcal{A}_n$ la sous-tribu de $\mathcal{A}$ engendrée par les v.a.r.l. $\xi_1,\dots,\xi_n$; alors,

$$E_P(e^{i\sum_{j=1}^{n} \alpha_j \xi_j} E_P^{\mathcal{A}_n}(\bar{Z})) = 0 \ , \ \forall \alpha_1,\dots,\alpha_n \in \mathbb{R}$$

ce qui implique que : $E_P^{\mathcal{A}_n}(\bar{Z}) = 0$, P - p.s . Comme $\mathcal{A}$ est la tribu linéaire, il existe une suite croissante $\{\mathcal{A}_n\}_{n\geq 1}$ de tribus engendrant $\mathcal{A}$, alors $E_P^{\mathcal{A}_n}(\bar{Z}) = 0$, $\forall n \geq 1$ entraîne que $Z = 0$. ∎

PROPOSITION 1. - <u>Il existe un espace de HILBERT</u> $\mathcal{H}(K_{\Phi_P})$ <u>de fonctionnelles complexes définies sur</u> $L(E,\mathcal{A},P)$ <u>tel que</u> :

1°) - <u>Les fonctionnelles</u> $K_{\Phi_P}(\cdot,\xi) = \Phi_P(\cdot - \xi)$ <u>appartiennent à</u> $\mathcal{H}(K_{\Phi_P})$ <u>et l'engendrent</u>.

2°) - <u>Si</u> $\varphi \in \mathcal{H}(K_{\Phi_P})$, $\varphi(\xi) = \langle \varphi, K_{\Phi_P}(\cdot,\xi)\rangle_{\mathcal{H}(K_{\Phi_P})}$, $\xi \in L(E,\mathcal{A},P)$.

3°) - <u>La transformation</u> τ_P <u>définie par</u> :

$Z \in L_2^{\mathbb{C}}(E,\mathcal{A},P)$, $\xi \in L(E,\mathcal{A},P)$, $[\tau_P(Z)](\xi) = E_P(Z e^{i\xi})$

<u>est une isométrie de l'espace de HILBERT</u> $L_2^{\mathbb{C}}(E,\mathcal{A},P)$ <u>sur l'espace de HILBERT</u> $\mathcal{H}(K_{\Phi_P})$.

Les 1°) et 2°) constituent un théorème fondamental de ARONSZAJN (cf. [22], par ex.) concernant l'espace de HILBERT autoreproduisant associé à un noyau hermitien de type positif, ils résultent du LEMME 1. Le 3°) résulte des 1°) et 2°) et du LEMME 2 et de la remarque que : $\tau_P(e^{-i\xi})(\cdot) = K_{\Phi_P}(\cdot,\xi)$. ∎

On peut donc étendre à ce cadre général, la théorie exposée dans [14] qui est basée sur ces résultats. En ce qui nous concerne nous utiliserons cette proposition pour établir une caractérisation intéressante de l'équivalence de deux probabilités sur un E.V.M. :

PROPOSITION 2. - <u>Soit</u> P' <u>une probabilité sur</u> $(E,\mathcal{A})$, <u>telle que</u> $L(E,\mathcal{A},P) = L(E,\mathcal{A},P')$. <u>Alors</u>, P' <u>est absolument continue par rapport à</u> P <u>et</u> $\frac{dP'}{dP} \in L_2^{\mathbb{C}}(E,\mathcal{A},P)$ <u>si et seulement si</u> $\Phi_{P'} \in \mathcal{H}(K_{\Phi_P})$, <u>auquel cas</u> , $\frac{dP'}{dP} = \tau_P^{-1}(\Phi_{P'})$.

D'après l'hypothèse, Φ_P et $\Phi_{P'}$ sont définies sur le même espace ; si $P' \ll P$ et si, $\frac{dP'}{dP} \in L_2^{\mathbb{C}}(E,\mathcal{A},P)$, $\tau_P(\frac{dP'}{dP})(\xi) = \int_E e^{i\xi} \frac{dP'}{dP} dP = \Phi_{P'}(\xi)$ pour tout $\xi \in L(E,\mathcal{A},P)$, ce qui prouve que la condition est nécessaire. Inversement, si $\Phi_{P'} \in \mathcal{H}(K_{\Phi_P})$, posons : $U = \tau_P^{-1}(\Phi_{P'})$ dans $L_2^{\mathbb{C}}(E,\mathcal{A},P)$; nous avons donc :

(1) $$\int_E U e^{i\xi} dP = \int_E e^{i\xi} dP' \text{ , pour tout } \xi \in L(E,\mathcal{A},P)$$

et en prenant les conjugués des deux membres,

$$\int_E \overline{U} e^{-i\xi} dP = \int_E e^{-i\xi} dP' \text{ , pour tout } \xi \in L(E,\mathcal{A},P)$$

ce qui est équivalent à :

(2) $$\int_E \overline{U} e^{i\xi} dP = \int_E e^{i\xi} dP' \text{ , pour tout } \xi \in L(E,\mathcal{A},P) \text{ ;}$$

En retranchant (2) de (1) on obtient : $\int_E (U-\overline{U}) e^{i\xi} dP = 0$, $\forall \xi \in L(E,\mathcal{A},P$ en utilisant le même argument que pour la démonstration du LEMME 2 , ceci implique que $U = \overline{U}$, ce qui montre que U est une v.a.r. Posons alors : $U^+ = \mathrm{Sup}(U,0)$ et $U^- = -\mathrm{inf}(U,0)$. Pour tout $\xi \in L(E,\mathcal{A},P)$, on a donc :

(3) $$\int_E U e^{i\xi} dP = \int_E e^{i\xi} dP'$$

en particulier pour $\xi = 0$, $\int_E U dP = \int_E U^+ dP - \int_E U^- dP = \int_E dP' = 1$

on peut poser : $c = \int_E U^+ dP > 0$ et pour tout $A \in \mathcal{A}$, $P^+(A) = \frac{1}{c}\int_A U^+ dP$ et

$P^-(A) = \frac{1}{c}\,[\int_A U^- dP + P'(A)]$. Alors pour tout $\xi \in L(E,\mathcal{A},P)$, on a, d'après (3) :

$$\frac{1}{c}\int_E e^{i\xi}dP' = \frac{1}{c}\int_E U^+ e^{i\xi}dP - \frac{1}{c}\int_E U^- e^{i\xi}dP$$

$$= \int_E e^{i\xi}dP^+ - \int_E e^{i\xi}dP^- + \frac{1}{c}\int_E e^{i\xi}dP'$$

d'où : $\Phi_{P+} = \Phi_{P-}$ et d'après l'unicité de la transformée de FOURIER, $P^+ = P^-$ c'est-à-dire que pour tout $A \in \mathcal{A}$, $P'(A) = \int_A U^+ dP - \int_A U^- dP = \int_A U\,dP$, donc $P' << P$ et $\frac{dP'}{dP} = U \in L_2^{\mathbb{C}}(E,\mathcal{A},P)$ par définition de U . ∎

2. <u>L'espace de HILBERT autoreproduisant associé à la transformée de LAPLACE réelle</u>.

On peut obtenir des résultats analogues à partir de la transformée de LAPLACE réelle à condition toutefois d'apporter quelques restrictions puisque celle-ci n'est pas définie partout sur $L(E,\mathcal{A},P)$.

LEMME 3. - <u>La fonction complexe</u> $K^1_{L_P}$ <u>définie sur</u> $D(E,\mathcal{A},P) \times D(E,\mathcal{A},P)$ <u>par</u> : $K^1_{L_P}(\xi,\xi') = L_P(\frac{\xi+\xi'}{2})$; $\xi, \xi \in D(E,\mathcal{A},P)$ <u>est un noyau symétrique de type positif</u>.

En effet, soient $\lambda_1,\ldots,\lambda_n$ des réels et $\xi_1,\ldots,\xi_n$ des éléments de $D(E,\mathcal{A},P)$.

$$\sum_{i,j=1}^{n} \lambda_i\lambda_j\, K^1_{L_P}(\xi_i,\xi_j) = \sum_{i,j=1}^{n} \lambda_i\lambda_j\, L_P(\frac{\xi_i+\xi_j}{2}) = \int_E (\sum_{i=1}^{n} \lambda_i e^{\frac{\xi_i}{2}})^2\, dP \geq 0$$

remarquons que $L_P(\frac{\xi+\xi'}{2})$ est bien défini puisque $D(E,\mathcal{A},P)$ est convexe. ∎

PROPOSITION 3. - <u>Si l'origine de</u> $L(E,\mathcal{A},P)$ <u>est un point interne du convexe</u> $D(E,\mathcal{A},P)$ <u>l'espace vectoriel engendré par la famille de v.a.r.</u>

$$\{\prod_{i=1}^{n} e^{\frac{\lambda_i \xi_i}{2}} \ ;\ n \geq 1 \ ,\ \xi_1,\ldots,\xi_n \in D(E,\mathcal{A},P) \ ,\ \lambda_1,\ldots,\lambda_n \in \mathbb{R} \ ,\ \sum_{i=1}^{n} \lambda_i \xi_i \in D(E,\mathcal{A},P)\}$$

est dense dans $L_2(E,\mathcal{A},P)$. (on ne peut pas parler d'algèbre engendrée comme au LEMME 2).

En effet, il suffit de montrer que toute v.a.r. orthogonale à toutes les v.a.r. de la famille est identiquement nulle. Soit donc $X \in L_2(E,\mathcal{A},P)$ telle que :

$$\int_E X\, e^{\sum_{i=1}^{n} \lambda_i \frac{\xi_i}{2}}\, dP = 0 \ , \ \forall n \geq 1 \ , \ \xi_1,\dots,\xi_n \in D(E,\mathcal{A},P) \ , \ \lambda_1,\dots,\lambda_n \in \mathbf{R} :$$
$$\sum \lambda_i \xi_i \in D(E,\mathcal{A},P) \ .$$

Soient N le sous-espace vectoriel de $L(E,\mathcal{A},P)$ engendré par les v.a.r.l. $\xi_1,\dots,\xi_n$ $\mathcal{A}_N$ la sous-tribu de $\mathcal{A}$ engendrée par ces (classes d'équivalence de) v.a.r.l. et ξ_N le vecteur aléatoire linéaire sur $(E,\mathcal{A},P)$ à valeurs dans $(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n))$ défini par : $\xi_N = (\xi_1,\dots,\xi_n)$ dont la loi de probabilité sera notée P_{ξ_N} . On peut toujours supposer que $L(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n),P_{\xi_N}) = \mathbf{R}^n$(cf. REMARQUE 4 ,§2). Alors, l'application transposée de ξ_N(cf. §2,n°3) de $\mathbf{R}^n$ dans $L(E,\mathcal{A},P)$ est comme au §2, n°4 définie par : $\lambda = (\lambda_1,\dots,\lambda_n) \in \mathbf{R}^n$, ${}^t\xi_N(\lambda) = \sum_{i=1}^{n} \lambda_i \xi_i$ et est un isomorphisme de $\mathbf{R}^n$ sur N et donc l'hypothèse est que :

$$\int_E X\, e^{{}^t\xi_N(\frac{\lambda}{2})}\, dP = 0 \ , \ \forall \lambda \in {}^t\xi_N^{-1}(D(E,\mathcal{A},P) \cap N) = D(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n),P_{\xi_N}) \subset \mathbf{R}^n \ .$$

ceci implique que :

$$\int_E E^{\mathcal{A}_N}(X)\, e^{{}^t\xi_N(\frac{\lambda}{2})}\, dP = 0 \ , \ \forall \lambda \in D(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n),P_{\xi_N})$$

ou encore :

$$\int_{\mathbf{R}^n} f(t)\, e^{\langle t,\lambda\rangle}\, dP_{\xi_N}(t) = 0 \ , \ \forall \lambda \in \frac{1}{2} D(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n),P_{\xi_N}) \ .$$

mais cette hypothèse entraîne que $f = 0$, P_{ξ_N} - p.s. dès que $D(\mathbf{R}^n,\mathcal{B}(\mathbf{R}^n),P_{\xi_N})$ est d'intérieur non vide dans $\mathbf{R}^n$ d'après un théorème connu (cf. [5], p.165).

Or 0 est un point interne de $D(E,\mathcal{a},P)$ si et seulement si l'intersection de $D(E,\mathcal{a},P)$ avec tout sous-espace vectoriel de dimension finie de $L(E,\mathcal{a},P)$ est d'intérieur non vide contenant l'origine dans ce sous-espace. Par conséquent l'intérieur de $D(R^n,\mathcal{B}(R^n),P_{\xi_N})$ image réciproque de l'intérieur de $D(E,\mathcal{a},P)\cap N$ est non vide. On en déduit que : $E^{\mathcal{a}_N}(X) = 0$ P.p.S pour tout élément N de la famille $\mathcal{F}(L)$ des sous-espaces vectoriels de dimension finie de $L(E,\mathcal{a},P)$ comme $\mathcal{a} = \mathcal{a}_L = \vee_{N\in\mathcal{F}(L)} \mathcal{a}_N$, $X = 0$ dans $L_2(E,\mathcal{a},P)$. ■

PROPOSITION 4. - <u>Il existe un espace de HILBERT réel</u> $\mathcal{H}(K^1_{L_P})$ <u>de fonctionnelles définies sur</u> $D(E,\mathcal{a},P)$ <u>tel que</u> :

1°) - <u>Les fonctionnelles</u> $K^1_{L_P}(.,\xi) = L_P(\frac{.+\xi}{2})$ <u>appartiennent à</u> $\mathcal{H}(K^1_{L_P})$ <u>et l'engendrent</u>.

2°) - <u>Si</u> $f \in \mathcal{H}(K^1_{L_P})$, $f(\xi) = \langle f, K^1_{L_P}(.,\xi)\rangle_{\mathcal{H}(K^1_{L_P})}$, $\xi \in D(E,\mathcal{a},P)$.

3°) - <u>Si l'origine de</u> $L(E,\mathcal{a},P)$ <u>est un point interne du convexe</u> $D(E,\mathcal{a},P)$, <u>la transformation</u> λ^1_P <u>définie par</u> :

$$X \in L_2(E,\mathcal{a},P)\ ,\ \xi \in D(E,\mathcal{a},P)\ ,\ [\lambda^1_P(X)](\xi) = E_P(Xe^{\frac{\xi}{2}})$$

<u>est un isomorphisme de l'espace de HILBERT</u> $L_2(E,\mathcal{a},P)$ <u>sur</u> $\mathcal{H}(K^1_{L_P})$.

Cette proposition résulte de la théorie de ARONSZAJN et de la PROPOSITION 3, ayant remarqué que

$$\lambda^1_P(e^{\frac{\xi}{2}})(.) = K^1_{L_P}(.,\xi) .$$

REMARQUE 1. - Si $D(E,\mathcal{a},P) + D(E,\mathcal{a},P) \subset D(E,\mathcal{a},P)$ il est plus commode d'utiliser le noyau $K^2_{L_P}$ défini sur $D(E,\mathcal{a},P) \times D(E,\mathcal{a},P)$ par : $K^2_{L_P}(\xi,\xi') = L_P(\xi+\xi')$. On pourra alors lui associer l'espace de HILBERT autoreproduisant $\mathcal{H}(K^2_{L_P})$ engendré par les fonctionnelles $L_P(.+\xi) = K^2_{L_P}(.,\xi)$ lorsque ξ parcourt $D(E,\mathcal{a},P)$ isométrique à un sous-espace fermé de

$L_2(E,\mathcal{a},P)$ ou à $L_2(E,\mathcal{a},P)$ tout entier si par exemple l'origine de $L(E,\mathcal{a},P)$ est un point interne de $D(E,\mathcal{a},P)$; les démonstrations sont les mêmes que pour $K^1_{L_P}$. ∎

REMARQUE 2. - On pourrait énoncer un théorème d'absolue continuité faisant intervenir la transformée de LAPLACE réelle ; mais en pratique il ferait double emploi avec celui énoncé en PROPOSITION 2 . ∎

§ 4. - ELEMENTS CARACTERISTIQUES D'UN VECTEUR ALEATOIRE. LOI DES GRANDS NOMBRES. VALEURS TYPIQUES D'UN ECHANTILLON.

Soient E et F deux espaces vectoriels réels en dualité séparante pour la forme bilinéaire notée $\langle .,. \rangle$ sur $E \times F$, X un vecteur aléatoire dans E relativement à cette dualité (cf. CHAPITRE I, §4), défini sur un espace probabilisé quelconque et P_X sa loi de probabilité sur $(E, \mathcal{C}(E,F))$. Comme tout élément aléatoire X à valeurs dans un E.V.M.S. $(E, \mathcal{a})$ est un vecteur aléatoire dans E relativement à la dualité séparante entre E et $(E,\mathcal{a})^m$, tout ce qui suit reste valable dans ce cas en prenant $F = (E,\mathcal{a})^m$

DEFINITION 1. - On appelle "fonction caractéristique de X" la transformée de FOURIER φ_X de sa loi de probabilité (ou ce qui revient au même, de sa loi de probabilité cylindrique (cf. REMARQUE 2, §1)) qui est donc définie par :

$$y \in F \ , \ \varphi_X(y) = E(e^{i\langle X,y\rangle}) = \int_E e^{i\langle x,y\rangle} \, dP_X(x) = \int_{\mathbb{R}} e^{it} \, dP_{\langle X,y\rangle}(t) \ .$$

on dira qu'une suite $\{X_n\}_{n\geq 1}$ de vecteurs aléatoires dans E relativement à la dualité entre E et F converge en loi cylindrique vers X si pour tout $y \in F$, $\varphi_{X_n}(y) \underset{n\to\infty}{\longmapsto} \varphi_X(y)$.

1. Caractéristiques du 1er ordre.

DEFINITION 2. - X est dit "scalairement intégrable" si la v.a.r. $\langle X,y\rangle$ est intégrable pour tout $y \in F$. Dans ce cas l'application : $y \longmapsto E(\langle X,y\rangle)$ est une forme linéaire sur F ; si elle appartient à E, considéré comme sous-espace vectoriel du dual algébrique F^* de F (c'est-à-dire, si elle est continue pour la topologie faible $\sigma(F,E)$), elle définit un élément unique $E(X)$ de E que l'on appelle "espérance mathématique de X". On dit, alors, que "X est intégrable" et l'on a : $\langle E(X),y\rangle = E(\langle X,y\rangle)$, pour tout $y \in F$; si $E(X) = 0$ on dit que X est centré.

Il existe un grand nombre de théorèmes assurant l'intégrabilité de X sous des conditions très générales qui font l'objet du CHAPITRE I, §2 de

[1] . Nous retiendrons le suivant qui recouvre déjà bon nombre d'applications :

THEOREME 1. - *Soit* E *un espace de FRECHET séparable dont la topologie est définie par une suite croissante* $\{p_n\}_{n\geq 1}$ *de semi-normes. Si* X *est un vecteur aléatoire dans* $(E,\mathcal{B}(E))$ *(ou ce qui revient au même, relativement à la dualité entre* E *et* E' *d'après le THEOREME 4 , CHAPITRE I , §2) , il est intégrable dès que les v.a.r.* $p_n(X)$ *le sont pour tout* $n \geq 1$. *En particulier, si* E *est un espace de BANACH séparable* $E(X)$ *existe dès que* $\|X\|$ *est une v.a.r. intégrable.*

2. Une loi des grands nombres.

Nous profiterons de ce contexte pour rappeler une loi forte des grands nombres. (cf. [1], CHAPITRE I, §3 ou [25] qui fait le point sur la question et donne aussi des lois faibles).

THEOREME 2. - *Dans les mêmes conditions qu'au THEOREME 1 , soit* $\{X_n\}_{n\geq 1}$ *une suite de vecteurs aléatoires à valeurs dans* $(E,\mathcal{B}(E))$ *indépendants et de même loi de probabilité. Si* $E(p_n(X_1)) < \infty$ *pour tout* $n \geq 1$, *il existe un vecteur aléatoire* X *dans* $(E,\mathcal{B}(E))$ *tel que* $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ *converge vers* X *presque sûrement pour la topologie de* E *quand* $n\to\infty$. *En particulier, si* E *est un espace de BANACH séparable et si* $E(\|X_1\| < \infty$, *alors* $\|S_n - E(X_1)\| \to 0$ p.s. *quand* $n\to\infty$.

3. Caractéristiques du 2e ordre.

DEFINITION 3. - *Dans les même conditions qu'aux DEFINITIONS* 1 *et* 2 , X *est dit "scalairement de carré intégrable" si la v.a.r.* $\langle X,y\rangle$ *est de carré intégrable pour tout* $y \in F$. *Dans ce cas, l'application* :

$$(y,y')\longmapsto E(\langle X-E(X),y\rangle\ \langle X-E(X),y'\rangle) = E[\langle (X-E(X) \otimes (X-E(X)),y\otimes y'\rangle]$$

est une forme bilinéaire symétrique sur $F\times F$, *donc un élément du dual algébrique* $(F^{2\odot})^{*}$ *de* $F^{2\odot}$ *noté* Λ_X *et que l'on appellera "covariance faible de* X" . *On appellera souvent "variance de* X" *la forme quadratique* positive sur F associée à la forme bilinéaire précédente,

$$Q_X : y \longmapsto E(\langle X,y\rangle^2) \ , \ y \in F \ .$$

Il est rare en pratique que Λ_X appartienne à l'espace vectoriel $E^{2\odot}$ sauf évidemment, si E est de dimension finie, auquel cas Λ_X s'identifie à la matrice de covariance de X : $E[(X-E(X)).{}^t(X-E(X))]$. Cela revient à dire, compte tenu du CHAPITRE I, §3, n° 3 que le vecteur aléatoire $\chi_E(X-E(X))$ dans $E^{2\odot}$ relativement à la dualité entre $E^{2\odot}$ et $F^{2\odot}$ est rarement intégrable au sens de la DEFINITION 2. On peut cependant apporter quelques précisions sur Λ_X dans un cas assez général :

PROPOSITION 1. - Soient E un espace de FRECHET séparable dont la topologie est définie par une famille $\{p_n\}_{n\geq 1}$ de semi-normes, et X un vecteur aléatoire dans $(E,\mathcal{B}(E))$. Si $p_n(X).p_m(X)$ est une v.a.r. intégrable quels que soient n et $m \geq 1$, la covariance faible Λ_X appartient au produit tensoriel symétrique projectif complété $E^{2\hat{\odot}}$ de E par lui-même.

En effet, X est intégrable d'après le THEOREME 1, puisque $[p_n(X)]^2$ et donc $p_n(X)$ est une v.a.r. intégrable pour tout $n \geq 1$; donc $E(X)$ existe dans E et on peut supposer, pour simplifier, que $E(X) = 0$ sans se restreindre.

On sait, (cf. [31], p.94, par exemple) que la topologie projective sur $E \otimes E$ peut être définie par la famille $\{p_n \otimes p_m\}_{\substack{n\geq 1\\ m\geq 1}}$ de semi-normes, où : $p_n \otimes p_m(x\otimes x') = p_n(x).p_m(x')$; $x,x' \in E$, elle fait donc de $E \otimes E$ un E.V.T.L.C.S. métrisable et séparable et le produit tensoriel projectif complété $E \hat{\otimes} E$ est donc lui-même un espace de FRECHET séparable.

On peut maintenant considérer $\chi_E(X) = X \otimes X$ comme un vecteur aléatoire à valeurs dans $(E \hat{\otimes} E,\mathcal{B}(E\otimes E)) = (E\hat{\otimes} E,\mathcal{C}(E\hat{\otimes} E,(E\hat{\otimes} E)'))$ ou encore, compte tenu de la PROPOSITION 2, CHAPITRE I, §3 dans $(E\hat{\otimes} E,\mathcal{C}(E\hat{\otimes} E,E'\otimes E'))$ c'est-à-dire comme un vecteur aléatoire dans $E \hat{\otimes} E$ relativement à la dualité entre $E \hat{\otimes} E$ et $E' \otimes E'$. En appliquant le THEOREME 1, les hypothèses impliquent que ce vecteur aléatoire est intégrable ; il existe donc un élément $E(X \otimes X)$ de $E \hat{\otimes} E$ (l'espérance mathématique de $X \otimes X$) tel que $\langle E(X \otimes X), u \rangle = E(\langle X \otimes X, u \rangle)$ quel que soit $u \in E' \otimes E'$ en notant encore $\langle .,. \rangle$ la forme bilinéaire mettant $E \hat{\otimes} E$ et $E' \otimes E'$ en dualité. $E(X \otimes X)$ est, par définition identifié à la forme linéaire : $u \longmapsto E(\langle X \otimes X, u \rangle)$ sur $E' \otimes E'$ qui n'est autre que la covariance faible Λ_X du vecteur scalairement de carré intégrable X. Enfin il est clair que le tenseur $E(X \otimes X) = \Lambda_X$ est symétrique et appartient par conséquent à l'espace $E \hat{\odot} E = E^{2\hat{\odot}}$. ∎

Nous verrons une illustration de cette proposition dans le chapitre suivant.

Comme cas particulier, si E est un espace de BANACH séparable, la condition de la proposition se ramène à $E(\|X\|^2) < \infty$. Et compte tenu de l'isomorphisme décrit au CHAPITRE I, §3, n°1,c) on retrouve un théorème classique dans le cas d'un espace de HILBERT séparable (cf. [33], p.14) :

COROLLAIRE. - Soit X un vecteur aléatoire à valeurs dans un espace de HILBERT réel séparable H tel que $E(\|X\|^2) < \infty$. Alors sa covariance faible Λ_X est un opérateur nucléaire auto-adjoint de H et $Tr(\Lambda_X) = E(\|X\|^2)$.

4. Valeurs typiques d'un échantillon.

Considérons une structure statistique vectorielle d'échantillon de taille n : $(E, \mathcal{a}, P)^n$ correspondant aux observations indépendantes $x_1, \ldots, x_n$ de X dans E.

On définit comme d'habitude, la loi de probabilité empirique associée à cet échantillon par la probabilité : $P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}$ sur $(E, \mathcal{a})$.

Pour tout $y \in (E, \mathcal{a})^m$, $\int_E \langle x, y \rangle \, dP_n(x) = \frac{1}{n} \sum_{i=1}^{n} \langle x_i, y \rangle = \langle \frac{1}{n} \sum_{i=1}^{n} x_i, y \rangle$ ce qui montre qu'un vecteur aléatoire dans $(E, \mathcal{a})$ de loi P_n est intégrable. On appelle son espérance mathématique : $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \in E$, la "moyenne empirique de l'échantillon".

Pour tous $y, y' \in (E,\mathcal{A})^m$, $\int_E \langle x-\overline{x}, y\rangle\langle x-x, y'\rangle dP_n(x)$

$$= \frac{1}{n}\sum_{i=1}^{n} \langle x_i-\overline{x}, y\rangle\langle x_i-\overline{x}, y'\rangle = \langle \frac{1}{n}\sum_{i=1}^{n} (x_i-\overline{x}) \otimes (x_i-\overline{x}) \ , \ y\otimes y'\rangle$$

ce qui montre que la covariance faible d'un vecteur aléatoire dans $(E,\mathcal{A})$ de loi P_n s'identifie à l'élément : $s^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i-\overline{x}) \otimes (x_i-\overline{x})$ de $E^{2\odot}$ qu'on appellera "<u>covariance empirique de l'échantillon</u>".

On verra au chapitre suivant que le théorème de FISHER concernant l'indépendance des statistiques $\overline{x}$ et s^2 sur une structure gaussienne s'étend en dimension quelconque.

REMARQUE 1. - Il existe une théorie récente (cf. [32]) qui définit la notion de fonction de répartition d'une loi de probabilité sur un espace vectoriel ordonné et étudie ses propriétés et notamment dans quelle mesure elle est caractéristique de cette probabilité . Partant , on peut définir la fonction de répartition empirique associée à un échantillon d'observations vectorielles très générales, et l'auteur généralise les théorèmes de KOLMOGOROV-SMIRNOV et de GLIVENKO-CANTELLI dans le cas où les observations appartiennent à un espace métrique séparable ce qui conduit à une extension du domaine d'application des tests d'adéquation classiques. ∎

DEUXIEME PARTIE

CONTRIBUTIONS A LA STATISTIQUE D'OBSERVATIONS INFINIDIMENSIONNELLES

CHAPITRE III

LA LOI DE WISHART GENERALISEE
ET AUTRES LOIS DE PROBABILITE USUELLES
DE VECTEURS ALEATOIRES DE DIMENSION QUELCONQUE.

Ce chapitre est consacré à l'étude des principaux types de lois de probabilité qui peuvent servir à la construction de structures statistiques vectorielles de dimension infinie. On construit ainsi des E.V.P. pour lesquels on illustre les notions étudiées au chapitre précédent, notamment celles qui concernent leur dual.

La plus importante, à cause de ses nombreuses applications, est la loi Normale et certaines de ses particularisations ; l'exposé se limite toutefois à la présentation, dans le langage adopté jusque-là, des définitions et propriétés les plus générales qui seront utilisées dans la suite.

L'apport le plus important de ce chapitre est constitué par une extension en dimension infinie de la loi de WISHART qui se rattache directement à la loi Normale. Elle fournit en outre des résultats utiles à l'étude de formes quadratiques de certaines f.a.r. gaussiennes décentrées ce qui permet d'améliorer certains résultats concernant des problèmes de tests d'adéquation en statistique classique et d'envisager la construction de tests pour la moyenne d'une f.a.r. gaussienne.

Les autres lois de probabilité sont plus précisément définies sur des espaces vectoriels de mesures. On étudie ainsi une extension de la loi Multinomiale, une loi de POISSON, une loi de DIRICHLET en indiquant pour chacune d'entre elles leur domaine d'application en statistique.

§1. - LOI NORMALE ET VECTEURS ALEATOIRES GAUSSIENS.

Soient E et F deux espaces vectoriels réels mis en dualité séparante par une forme bilinéaire $\langle .,. \rangle$ sur $E \times F$.

DEFINITION 1. - Un vecteur aléatoire X dans E , relativement à la dualité entre E et F est dit "gaussien" si pour tout $y \in F$, la v.a.r. $\langle X,y \rangle$, définie sur le même espace probabilisé, est gaussienne. Un vecteur aléatoire X , à valeurs dans un E.V.M.S. $(E,\mathfrak{a})$ est dit "gaussien", s'il est gaussien relativement à la dualité séparante entre E et $(E,\mathfrak{a})^m$.

L'étude des vecteurs aléatoires gaussiens, essentiellement de leur loi de probabilité, nécessite impérativement la considération de la notion de probabilité cylindrique gaussienne.

1. Probabilités cylindriques gaussiennes.

THEOREME 1. - Pour tout élément $x_o \in E$ et toute forme quadratique positive Q définie sur F , il existe une probabilité cylindrique unique sur E relativement à la dualité entre E et F , notée $\gamma_{x_o,Q}$ dont la transformée de FOURIER est :

$$y \in F \ , \ \varphi_{\gamma_{x_o,Q}}(y) = \exp\{i\langle x_o,y\rangle - \frac{1}{2}Q(y)\} \ .$$

Ce théorème résulte du THEOREME 3, CHAPITRE II, §1 . ■

DEFINITION 2. - On appelle $\gamma_{x_o,Q}$ la "probabilité cylindrique gaussienne sur E relativement à la dualité entre E et F de moyenne $x_o \in E$ et de variance, la forme quadratique positive Q sur F" . On dit qu'elle est centrée si $x_o = 0$.

Les termes de moyenne et de variance constituent des abus de langage puisqu'ils ne sont pas habituellement définis pour une probabilité cylindrique, ils sont cependant justifiés par le théorème suivant et ses conséquences :

THEOREME 2. - <u>Soient</u> E_1 <u>et</u> F_1 <u>deux autres espaces vectoriels réels en dualité séparante et</u> u , <u>une application linéaire de</u> E <u>dans</u> E_1 <u>continue pour les topologies faibles</u> $\sigma(E,F)$ <u>et</u> $\sigma(E_1,F_1)$. <u>Alors, l'image par</u> u <u>de la probabilité cylindrique gaussienne</u> $\gamma_{x_o,Q}$ <u>est la probabilité cylindrique gaussienne</u> : $\gamma_{u(x_o),\, Q\circ {}^t u}$ <u>sur</u> E_1 <u>relativement à la dualité entre</u> E_1 <u>et</u> F_1 .

Cela résulte directement du calcul de la transformée de FOURIER de la probabilité cylindrique image $\dot{u}(\gamma_{x_o,Q})$.

En prenant, en particulier pour u les applications canoniques π_M de E sur E/M quand M parcourt $\mathcal{M}(E,F)$ (notations du CHAPITRE I, §2) , on constate que le système projectif $(E/M, \mathcal{B}(E/M), (\gamma_{x_o,Q})_M)_{M\in\mathcal{M}(E,F)}$ est constitué d'E.V.M.S. de dimension finie, munis respectivement de la probabilité gaussienne de moyenne $\pi_M(x_o)$ et de variance la forme quadratique positive : $Q_M = Q \circ {}^t\pi_M$ sur M^o (en identifiant $(E/M)'$ à M^o par l'isomorphisme ${}^t\pi_M$) . Par suite, pour tout $y \in F$, l'image de $\gamma_{x_o,Q}$ par la forme linéaire faiblement continue : $x \longmapsto \langle x,y\rangle$ sur E est la loi Normale : $N(\langle x_o,y\rangle, Q(y))$ sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. Inversement, on peut montrer (cf. [8],p.78) :

THEOREME 3. - <u>Soit</u> $\mu = (\mu_M)_{M\in\mathcal{M}(E,F)}$ <u>une probabilité cylindrique sur</u> E , <u>relativement à la dualité entre</u> E <u>et</u> F . <u>Si pour tout</u> $y \in F$, <u>l'image de</u> μ <u>par la forme linéaire</u> $x \longmapsto \langle x,y\rangle$, $x \in E$, <u>est une probabilité gaussienne sur</u> $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ <u>alors</u> μ <u>est une probabilité cylindrique gaussienne</u>.

DEFINITION 3. - <u>Soit</u> H <u>un espace de</u> HILBERT <u>réel séparable</u> ; <u>on appelle "probabilité cylindrique gaussienne canonique sur</u> H" , <u>la probabilité cylindrique gaussienne centrée</u> $\gamma_{0,\|\cdot\|_H^2}$ <u>sur</u> H <u>relativement à la dualité entre</u> H <u>et</u> H' <u>identifié à</u> H , <u>de variance la forme quadratique</u> $y \longmapsto \|y\|_H^2$ <u>sur</u> H .

C'est évidemment la probabilité cylindrique donnée à l'EXEMPLE 2, CHAPITRE II, §1. Bien que n'induisant pas une probabilité sur $(H,\mathcal{B}(H))$

si H est de dimension infinie, elle joue un rôle fondamental dans la construction de probabilités gaussiennes. A titre d'exemple on peut déjà citer le théorème classique (cf. [8],p.93) :

THEOREME 4. - Soit A un opérateur linéaire continu dans H, la probabilité cylindrique gaussienne centrée : $\gamma_{0,\|{}^tA(.)\|_H^2} = A(\gamma_{0,\|.\|_H^2})$, image de la probabilité cylindrique gaussienne canonique sur H, induit une probabilité sur $(H,\mathcal{C}(H,H)) = (H,\mathcal{B}(H))$ si et seulement si H est un opérateur de HILBERT-SCHMIDT.

2. Probabilités et vecteurs gaussiens. Loi Normale.

DEFINITION 4. - Une probabilité P définie sur l'espace vectoriel faiblement mesurable $(E,\mathcal{C}(E,F))$ est dite "gaussienne" si la probabilité cylindrique $P = (P_M)_{M\in\mathcal{M}(E,F)}$ qui lui est associée est gaussienne. Une probabilité P définie sur un E.V.M.S. $(E,\mathcal{a})$ est dite "gaussienne, si elle est gaussienne sur l'E.V.M.S. affaibli $(E,\mathcal{C}(E,(E,\mathcal{a})^m))$; dans ce cas, l'E.V.P. $(E,\mathcal{a},P)$ sera appelé "espace vectoriel probabilisé gaussien" (E.V.P.G

Compte tenu de la DEFINITION 1 et du THEOREME 3, un vecteur aléatoire X dans E, relativement à la dualité entre E et F, est gaussien si sa loi de probabilité P_X est gaussienne.

DEFINITION 5. - Pour un élément $m \in E$ et une forme quadratique positive Q sur F on dira que X suit la loi Normale sur E de moyenne m et de variance Q, ce que l'on notera : $X \rightsquigarrow N_{E,F}(m,Q)$ si sa fonction caractéristique est :

$$y \in F \ , \ \varphi_X(y) = E(e^{i\langle X,y\rangle}) = \exp(i\langle m,y\rangle - \frac{1}{2}Q(y)) \ .$$

Un calcul classique montre en effet que pour tout entier $n \geq 0$ et pour tout $y \in F$ on a :

$$E(|\langle X-m,y\rangle|^n) = \int_E |\langle x-m,y\rangle|^n \, dP_X(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \, 2^{n/2} \, Q(y)^{n/2} \ ,$$

$$E(\langle X-m,y\rangle^{2n}) = \frac{(2n)!}{2^n n!} \, Q(y)^n \ ,$$

$$E(\langle X-m,y\rangle^{2n+1}) = 0 \ ;$$

ce qui implique que X est intégrable et que $E(X) = m$ d'une part, et que X est scalairement de carré intégrable et que sa variance est : $Q(y) = E(\langle X-m,y\rangle^2)$, $y \in F$. au sens des définitions du CHAPITRE II, §4, d'autre part.

EXEMPLE 1. - Soient T un ensemble et $(X_t)_{t\in T}$ une fonction aléatoire réelle gaussienne sur T , définie sur un espace probabilisé $(\Omega,\mathcal{F},P)$, de moyenne la fonction $m = \{m(t)\}_{t\in T}$ et de fonction de covariance

$$k(s,t) = E[(X(s)-m(s))(X(t)-m(t))] \ , \ (s,t) \in T\times T \ .$$

On a vu au CHAPITRE I, §4, EXEMPLE 1, qu'une f.a.r. définit un vecteur aléatoire X dans $E = \mathbb{R}^T$ relativement à la dualité entre E et $F = \sum_{t\in T} \mathbb{R}_t$, $(\mathbb{R}_t = \mathbb{R}$, $\forall t \in T)$ c'est-à-dire dans l'E.V.M.S. $(\mathbb{R}^T, \bigotimes_{t\in T} \mathcal{B}(\mathbb{R})) = (E,\mathcal{C}(E,F))$; les fonctionnelles d'évaluation $\delta_t : x \longmapsto x(t)$, $t\in T$, formant une base de F , la formule : $Q_k(\sum_{i\in I} \lambda_i \delta_{t_i}) = \sum_{I\times I} \lambda_i \lambda_j k(t_i,t_j)$ I fini, définit une forme quadratique positive Q_k sur F et il est clair que : $X \rightsquigarrow N_{E,F}(m,Q_k)$.

EXEMPLE 2. - Des hypothèses supplémentaires sur les trajectoires de la f.a.r. gaussienne précédente permettent de préciser que l'élément aléatoire X qu'elle définit est à valeurs dans un sous-espace vectoriel mesurable de $\mathbb{R}^T$: Reprenant l'EXEMPLE 2 du CHAPITRE I, §4, où T est un espace métrique σ-compact , si m est continue sur T et si k est continue sur $T\times T$ et vérifie certaines hypothèses (cf. [27]), on peut montrer que la f.a.r. $(X_t)_{t\in T}$ est p.s. à trajectoires continues sur T et, comme on l'a remarqué dans cet exemple l'élément aléatoire X qui lui est associé est à valeurs dans l'espace de FRECHET séparable $E = C(T)$ relativement à la dualité entre cet espace et l'espace $F = M_c(T)$ (notations de l'EXEMPLE 2, CHAPITRE I, §2), ou ce qui revient au même, dans l'E.V.M.S. $(C(T), \mathcal{B}[C(T)])$. La forme quadratique Q_k définie à l'EXEMPLE 1, se prolonge en une forme quadratique $\tilde{Q}_k$ sur $M_c(T)$, telle que, pour $\nu \in M_c(T)$, $\tilde{Q}_k(\nu) = \iint_{T\times T} k(s,t)\, d\nu(s)\, d\nu(t)$, de sorte que : $X \rightsquigarrow N_{C(T),M_c(T)}(m,\tilde{Q}_k)$, et a pour fonction caractéristique :

$$\varphi_X(\nu) = \exp\{i \int_T m(t)\, d\nu(t) - \frac{1}{2} \iint_{T\times T} k(s,t)\, d\nu(s)\, d\nu(t)\} \ ; \ \nu \in M_c(T) . \quad (1)$$

Inversement, si P est une probabilité gaussienne sur $(C(T), \mathcal{B}[C(T)])$, si l'on note δ_t la mesure de DIRAC au point $t \in T$, dans $M_c(T)$, en posant :

$$m(t) = \int_{C(T)} \langle x,\delta_t\rangle dP(x) = \int_{C(T)} \int_T x(s)\, d\delta_t(s)\, dP(x) = \int_{C(T)} x(t)\, dP(x) , \ t \in T ;$$

$$k(s,t) = \int_{C(T)} \langle x,\delta_s\rangle \langle x,\delta_t\rangle\, dP(x) - m(s)\, m(t) \quad (2)$$

$$= \int_{C(T)} x(s)\, x(t)\, dP(x) - m(s)\, m(t) , \ (s,t) \in T \times T ,$$

la transformée de FOURIER de P est donnée par la formule (1) ; (cf. [27]). Les deux fonctions m et k déterminent entièrement la probabilité gaussienne P ; ce sont respectivement la moyenne et la fonction de covariance de la f.a.r. gaussienne $(X_t)_{t\in T}$ sur T, définie sur l'espace probabilisé $(C(T), \mathcal{B}[C(T)], P)$ par : $X_t(x) = \langle x,\delta_t\rangle$, $x \in C(T)$, $t \in T$; le vecteur aléatoire gaussien qui lui est associé est l'application identique sur cet espace probabilisé. ■

On appellera "<u>Structure statistique gaussienne</u>", une structure statistique vectorielle $(E, \mathcal{A}, \mathcal{P})$ où $\mathcal{P}$ est une famille de lois de probabilité gaussiennes sur $(E, \mathcal{A})$. On en verra de nombreux exemples.

3. <u>Quelques résultats fondamentaux sur les probabilités gaussiennes</u>.

Compte tenu de ce qui précède, on étudie maintenant plus précisément les probabilités gaussiennes sans faire référence aux vecteurs aléatoires dont elles peuvent constituer la loi de probabilité. Cette étude consiste à rappeler quelques résultats fondamentaux de la théorie des probabilités dans ce domaine, qui sont utiles en statistique et qui concernent notamment la caractérisation du support d'une probabilité gaussienne, du dual d'un E.V.P.G. dans des cas suffisamment généraux et des propriétés d'absolue continuité de probabilités gaussiennes.

a) Support (cf. CHAPITRE II, §1, n°5)

THEOREME 5. - *Soient* T *un ensemble ayant au plus la puissance du continu et* P *une probabilité gaussienne centrée régulière sur l'E.V.M.S.* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *pour la topologie de la convergence simple. Alors, le support de* P *est la fermeture* $\overline{H(k)}$ *dans* $\mathbb{R}^T$ *de l'espace de HILBERT autoreproduisant associé au noyau* k *défini sur* $T \times T$ *par :*

$$k(s,t) = \int_{\mathbb{R}^T} x(s)\ x(t)\ dP(x)\ , \ (s,t) \in T \times T\ .$$

De plus, $H(k) \in \mathcal{B}(\mathbb{R}^T)$ *et* $P[H(k)] = 0$ *si et seulement si* $H(k)$ *est de dimension infinie.*

THEOREME 6 . - *Soient* E *un espace de FRECHET séparable et* P *une probabilité gaussienne centrée sur* $(E, \mathcal{B}(E))$ *, de variance la forme quadratique positive* Q *sur le dual topologique* E' *de* E *. Il existe un sous-espace de HILBERT réel séparable* $H(Q)$ *de* E *, uniquement déterminé par* Q *, tel que l'injection canonique* j *de* $H(Q)$ *dans* E *soit continue et transforme la probabilité cylindrique gaussienne canonique sur* $H(Q)$ *en* P *.*

Le support topologique de P *est la fermeture* $\overline{H(Q)}$ *dans* E *. De plus,* $H(Q) \in \mathcal{B}(E)$ *et* $P[H(Q)] = 0$ *si* $H(Q)$ *est de dimension infinie.*

La démonstration, (cf. [27]), établit d'abord le résultat lorsque E est un sous-espace vectoriel fermé de l'espace de FRECHET $C(T)$ de l'EXEMPLE 2 ci-dessus ; l'espace $H(Q)$ n'est autre, alors, que l'espace de HILBERT autoreproduisant $H(k)$ associé au noyau k sur $T \times T$ sousjacent à Q soit : $k(s,t) = \frac{1}{2}[Q(\delta_s+\delta_t) - Q(\delta_s) - Q(\delta_t)]$, $(s,t) \in T \times T$. Le résultat annoncé s'obtient en utilisant l'isomorphisme qui existe entre tout espace de FRECHET séparable et un sous-espace vectoriel fermé de $C(T)$. Il implique donc, que l'image par l'injection canonique j de $H(Q)$ dans $C(T)$, qui est une application linéaire continue donc continue pour les topologies faibles $\sigma(H(Q),H(Q))$ et $\sigma(C(T),M_c(T))$, de la probabilité cylindrique gaussienne canonique $\gamma_{0,\|\cdot\|^2_{H(Q)}}$ est une probabilité cylindrique gaussienne sur $C(T)$ relativement à la dualité entre $C(T)$ et $M_c(T)$ qui induit la probabilité P sur $(C(T),\mathcal{C}(C(T),M_c(T))) = (C(T),\mathcal{B}[C(T)])$.

Ces deux théorèmes rassemblent des résultats de B.S. RAJPUT [27] qui seront utilisés au § suivant. ■

REMARQUE 2. - On trouve ici une illustration de l'extension en dimension infinie de la caractérisation du support d'une probabilité au moyen de l'analyse convexe, développée au CHAPITRE II, §1, n°5 . En effet, soit P une probabilité gaussienne centrée sur $(E,\mathcal{B}(E))$ où E est un E.V.T.L.C.S. , de variance la forme quadratique positive Q sur E' , continue pour la topologie de MACKEY $\tau(E',E)$. Un calcul facile montre que la fonction cumulante de P (définie sur E') est : $c_P = \frac{1}{2}Q$, et une théorie classique (cf. [30]) montre qu'il existe un sous-espace hilbertien $H(Q)$ de E (avec injection continue), dont on peut voir, d'ailleurs, qu'il coïncide avec l'espace de HILBERT autoreproduisant associé au noyau K sur $E' \times E'$ vérifiant : $Q(y) = K(y,y)$, $y \in E'$ et tel que la polaire c_P^* de c_P soit définie par : $c_P^*(x) = \frac{1}{2}\|x\|^2_{H(Q)}$ si $x \in H(Q)$ et $c_P^*(x) = +\infty$ si $x \notin H(Q)$. Une telle fonction s'appelle "fonction convexe hilbertienne sur E" . Si E est un espace de FRECHET séparable, cet espace $H(Q)$ s'identifie évidemment avec l'espace de HILBERT séparable du THEOREME 6 .
Nous avons donc : $\mathrm{dom}(c_P^*) = H(Q)$ et ce théorème indique que $\mathrm{supp}(P) = \overline{H(Q)}$ mais comme le support de la probabilité gaussienne régulière P est un sous-espace vectoriel fermé de E , on a bien :
$\Omega = \overline{\mathrm{conv}}(\mathrm{Supp}(P)) = \mathrm{dom}(c_P^*)$ ce qui étend en dimension infinie le THEOREME 4 CHAPITRE II, §1 . En résumé :

PROPOSITION 1. - Soient E un espace de FRECHET séparable et P une probabilité gaussienne centrée sur $(E, \mathcal{B}(E))$ de variance la forme quadratique positive Q sur le dual topologique E' de E. Alors, la fonction cumulante de P est : $c_P = \frac{1}{2} Q$, sa polaire est la "fonction convexe hilbertienne" définie par : $c_P^*(x) = \frac{1}{2} \|x\|^2_{H(Q)}$ si $x \in H(Q)$ et $c_P^*(x) = +\infty$ si $x \notin H(Q)$ où $H(Q)$ est un "sous-espace hilbertien de E" uniquement déterminé par Q. De plus, le support topologique de P coïncide avec la fermeture dans E de $\mathrm{dom}(c_P^*)$. ■

b) Dual d'un espace vectoriel probabilisé gaussien.

Soit $(E, \mathcal{Q}, P_o)$ un E.V.P.G. au sens de la DEFINITION 4. Supposons pour simplifier que P_o soit centrée. La variance de P_o est donc la forme quadratique positive Q définie sur le dual $(E, \mathcal{Q})^m$ de l'E.V.M.S. $(E, \mathcal{Q})$ par : $Q(y) = \int_E \langle x, y \rangle^2 \, dP_o(x)$, $y \in (E, \mathcal{Q})^m$. Tout élément y de ce dual s'identifie à une v.a.r. gaussienne centrée Y sur $(E, \mathcal{Q}, P_o)$ et si $y, y' \in (E, \mathcal{Q})^m$ on a : $Y = Y'$, P_o-p.s $\Longleftrightarrow Q(y-y') = 0$; par conséquent, Q définit sur l'espace quotient $(E, \mathcal{Q})^m / P_o$ une forme quadratique strictement positive $\dot{Q}$ dont la forme bilinéaire symétrique associée $\dot{K}$ fait de cet espace vectoriel un espace préhilbertien séparé ; c'est un sous-espace de $L_2(E, \mathcal{Q}, P_o)$ muni de la structure préhilbertienne induite.

Notons H sa fermeture : $\overline{(E, \mathcal{Q})^m / P_o}$ dans $L_2(E, \mathcal{Q}, P_o)$ - H est, en fait, un espace gaussien sur l'espace probabilisé $(E, \mathcal{Q}, P_o)$ dans la terminologie de [22] - On peut lui associer un espace $H(Q)$ de fonctionnelles définies sur $(E, \mathcal{Q})^m$ au moyen de l'isométrie $\psi : Z \in H \longmapsto h_Z(y) = E(ZY)$.

La proposition suivante qui caractérise le dual d'un E.V.P.G. sous des conditions très générales, nous donne un premier exemple de dual d'un E.V.P. :

PROPOSITION 2. - Si $H(Q) \subset E$, $L(E, \mathcal{Q}, P_o) \equiv H$.

On sait, en effet, d'après le COROLLAIRE de la PROPOSITION 1, CHAPITRE II, § 2 , que $L(E,\mathcal{A},P_o)$ contient la fermeture dans $L_2(E,\mathcal{A},P_o)$ de l'espace vectoriel $(E,\mathcal{C})^m/P_o$ puisque tout élément de cet espace est une (classe d'équivalence de) v.a.r. de carré intégrable. L'inclusion contraire résulte, dans l'hypothèse de la proposition, d'un théorème classique (cf. [28],p.146) assurant que toute v.a.r.l. est limite en moyenne quadratique d'une suite de fonctionnelles linéaires mesurables. ∎

Sous une condition très générale, le dual d'un E.V.P.G. est donc constitué de v.a.r. gaussiennes.

Cette condition est vérifiée en particulier sous les hypothèses des THEOREMES 5 et 6 , où $H(Q)$ n'est autre que l'espace de HILBERT construit plus haut ; on en déduit par exemple :

COROLLAIRE. - <u>Si</u> E <u>est un espace de FRECHET séparable et</u> P_o <u>une probabilité gaussienne centrée</u>, <u>sur</u> $(E,\mathcal{B}(E))$, <u>le dual de l'E.V.P.G.</u> $(E,\mathcal{B}(E),P_o)$ <u>est constitué des</u> (<u>classes d'équivalence de</u>) <u>v.a.r. qui sont limites simples presque-sûrement de suites de</u> (<u>classes d'équivalence de</u>) <u>formes linéaires continues sur</u> E .

EXEMPLE 3. - En reprenant l'EXEMPLE 1 et en supposant $m = 0$, le dual de l'E.V.P.G. : $(E,C(E,F),N_{E,F}(0,Q_k))$ coïncide avec l'espace gaussien de la f.a.r. gaussienne centrée canonique $(\xi_t)_{t\in T}$, définie sur cet espace probabilisé par $\xi_t(x) = \langle x,\delta_t\rangle$, $x \in E$, $t \in T$. ∎

EXEMPLE 4. - En reprenant ce même exemple, avec ses notations, supposons de plus que T soit une partie non discrète de $\mathbf{R}$ et que la f.a.r. gaussienne $(X_t)_{t\in T}$, soit centrée et stationnaire ; on peut alors spécifier la nature des v.a.r.l. sur l'E.V.P.G. $(E,C(E,F),N_{E,F}(0,Q_k))$: dans cette hypothèse, en effet, $k(s,t) = B(s-t)$, $s,t \in T$ et cette fonction d'auto correlation admet la représentation spectrale : $B(t) = \int_{-\infty}^{+\infty} e^{it\lambda} F(d\lambda)$, $t \in T$, où F désigne la mesure spectrale de la f.a.r. stationnaire, de plus on a aussi $X(t) = \int_{-\infty}^{+\infty} e^{it\lambda}\,\Phi(d\lambda)$, $t \in T$, où Φ désigne une mesure spectrale stochastique telle que si $A,B \in \mathcal{B}(\mathbf{R})$, $E(\Phi(A)\,\overline{\Phi(B)}) = F(A\cap B)$. On montre alors (cf. [16],p.28) que toute v.a.r.l. ξ sur l'E.V.P. ci-dessus peut s'identifier à une intégrale par rapport à cette mesure stochastique de la forme : $\xi = \int_{-\infty}^{+\infty} \varphi(\lambda)\, \Phi(d\lambda)$ ou φ est un élément de la fermeture dans $L_2^{\mathbb{C}}(\mathbb{R},\mathcal{B}(\mathbb{R}),F)$ de l'espace vectoriel formé de combinaisons linéaires à coefficients réels des fonctions $\{e^{it\lambda}\}_{t\in T}$. ∎

Un autre exemple important de v.a.r.l. sera fourni par la caractérisation du dual de l'E.V.P. de WIENER au n° suivant. ∎

Ayant caractérisé le dual d'un E.V.P.G. on peut spécifier dans ce contexte les notions qui s'y rapportent introduites au CHAPITRE II, §2 . On supposera remplie la condition de la PROPOSITION 2 :

1°) - La transformée de FOURIER de l'E.V.P.G. $(E,\mathcal{Q},P_o)$ est donc, d'après ce qui précède donnée par :

$$\Phi_{P_o}(\xi) = \exp\{-\frac{1}{2}\|\xi\|_H^2\} \ , \ \xi \in L(E,\mathcal{Q},P_o) \equiv H \ .$$

Elle se déduit de la transformée de FOURIER de la probabilité cylindrique gaussienne canonique sur $H(Q)$ par l'isométrie ψ de H sur $H(Q)$ soit : $\Phi_{P_o} = \varphi_{\gamma_o,\|.\|^2_{H(Q)}} \circ \psi$.

2°) - La transformée de LAPLACE réelle est donnée par : $L_{P_o}(\xi) = \exp\{\frac{1}{2}\|\xi\|_H^2\}$; $\xi \in L(E,\mathcal{Q},P_o) = H$, par suite $D(E,\mathcal{Q},P_o) \equiv L(E,\mathcal{Q},P_o)$ et est évidemment d'intérieur algébrique non vide, par suite l'algèbre engendrée par la famille de v.a.r. $\{e^{\xi}\}_{\xi\in L(E,\mathcal{Q},P_o)}$ est dense dans $L_2(E,\mathcal{Q},P_o)$ par un argument analogue à celui de la PROPOSITION 3, CHAPITRE II, §3 . Le noyau $K^2_{L_{P_o}}$ est défini sur $L(E,\mathcal{Q},P_o) \times L(E,\mathcal{Q},P_o)$ par : $K^2_{L_{P_o}}(\xi,\xi') = L_{P_o}(\xi+\xi') = \exp\{\frac{1}{2}\|\xi+\xi'\|_H^2\}$; l'espace de HILBERT autoreproduisant associé $\mathcal{H}(K^2_{L_{P_o}})$ est engendré par la famille de fonctions $\{\exp\langle .,\xi\rangle - \frac{1}{2}\|.\|_H^2\}_{\xi\in L(E,\mathcal{Q},P_o)}$ sur $L(E,\mathcal{Q},P_o)$; l'application : $\lambda_P^2 : X \longmapsto E_{P_o}(Xe^{\xi})$ est une isométrie de $L_2(E,\mathcal{Q},P_o)$ sur $\mathcal{H}(K^2_{L_{P_o}})$ ce qui est un résultat classique (cf. [22], [14]) .

3°) - Variables aléatoires réelles quadratiques sur l'E.V.P.G. $(E,\mathcal{a},P_o)$

Nous utiliserons, en statistique, la propriété suivante et son corollaire, les notations étant celles du CHAPITRE II, §2, n°5 :

PROPOSITION 3. - L'espace de HILBERT $L(E,\mathcal{a},P_o)^{2\tilde{\odot}}$ complété de l'espace préhilbertien $L(E,\mathcal{a},P_o)^{2\odot}$ est isomorphe à un sous-espace vectoriel de $L(E^{2\odot},\mathcal{a}^{2\chi},P_{o\chi_E})$.

Cette proposition, qui est de même nature que le COROLLAIRE 1 de la PROPOSITION 6 CHAPITRE II, §2 , utilise ici la structure hilbertienne de $L(E,\mathcal{a},P_o)$. En effet, $L(E,\mathcal{a},P_o)^{2\odot}$ peut-être identifié à un sous-espace vectoriel de l'espace $Q(E,\mathcal{a},P_o)$ des v.a.r.q. sur $(E,\mathcal{a},P_o)$ qui est fermé dans $L_2(E,\mathcal{a},P_o)$ d'après la PROPOSITION 5, CHAPITRE II, §2 , il en sera donc de même de l'espace de HILBERT $L(E,\mathcal{a},P_o)^{2\tilde{\odot}}$; mais on sait que $Q(E,\mathcal{a},P_o)$ est isomorphe à $L(E^{2\odot},\mathcal{a}^{2\chi},P_{o\chi_E})$ d'où le résultat. ∎

COROLLAIRE. - Tout élément $\zeta \in L(E,\mathcal{a},P_o)^{2\tilde{\odot}}$ peut s'écrire $\zeta = \xi \circ \chi_E$, où ξ est une v.a.r.l. sur l'E.V.P. $(E^{2\odot},\mathcal{a}^{2\chi},P_{o\chi_E})$ et χ_E est l'application quadratique sur E .

Il est probable que l'espace $Q(E,\mathcal{a},P_o)$ des v.a.r.q. au sens de la DEFINITION 6 CHAPITRE II, §2 , coïncide avec l'espace $L(E,\mathcal{a},P_o)^{2\tilde{\odot}}$, la condition $H(Q) \subset E$ étant remplie. Cependant le résultat du COROLLAIRE ci-dessus qui consiste en une représentation des v.a.r. appartenant à l'espace de HILBERT $L(E,\mathcal{a},P_o)^{2\tilde{\odot}}$ est suffisant pour notre propos. ∎

c) Théorèmes d'absolue continuité.

Les résultats concernant la propriété d'absolue continuité d'une probabilité gaussienne par rapport à une autre sont aussi nombreux que variés suivant le contexte dans lequel se placent leurs auteurs, les plus généraux sont dans [22] ou [28] . Nous envisagerons, en statistique, le cas où ces probabilités diffèrent par leur moyenne et le cas où elles diffèrent par leur variance. Dans la présentation que nous en avons faite, ces théorèmes peuvent s'exprimer de la façon suivante :

Soient $(E,\mathfrak{a})$ un E.V.M.S. et P_o la probabilité gaussienne centrée de variance la forme quadratique positive Q_o sur $(E,\mathfrak{a})^m$ qui servira de probabilité de référence. Avec les notations du b), on supposera que la condition de la PROPOSITION 2 est remplie, soit : $H(Q_o) \subset E$ de sorte que $L(E,\mathfrak{a},P_o) \equiv H$ et que $\mathfrak{a}$ coïncide avec la tribu linéaire $\mathfrak{a}_L$ (cf. DEFINITION 4, CHAPITRE II, §2) .

THEOREME 7. - <u>Pour tout</u> $m \in E$ <u>la probabilité gaussienne de moyenne</u> m P_m <u>sur</u> $(E,\mathfrak{a})$ <u>est soit équivalente soit étrangère à</u> P_o . <u>Pour qu'elle soit équivalente</u>, <u>il faut et il suffit que</u> $m \in H(Q_o)$, <u>alors</u> $L(E,\mathfrak{a},P_m) = L(E,\mathfrak{a},P_o)$ <u>et</u> : $\frac{dPm}{dP_o} = \exp\{\psi^{-1}(m) - \frac{1}{2}\|\psi^{-1}(m)\|^2_H\} = \exp\{\psi^{-1}(m) - \frac{1}{2}\|m\|^2_{H(Q_o)}\}$.
<u>où</u> ψ^{-1} <u>est l'isométrie inverse de l'isométrie</u> ψ <u>de</u> H <u>sur</u> $H(Q_o)$ <u>du b)</u> .

Remarquons que, sachant que $L(E,\mathfrak{a},P_m) = L(E,\mathfrak{a},P_o)$, la densité peut s'obtenir par application de la PROPOSITION 2, CHAPITRE II, §3 .

THEOREME 8. - <u>Soit</u> P <u>une probabilité gaussienne centrée sur</u> $(E,\mathfrak{a})$, <u>de variance la forme quadratique positive</u> Q <u>sur</u> $(E,\mathfrak{a})^m$. <u>Alors</u>, P <u>est soit étrangère à</u> P_o <u>soit équivalente à</u> P_o <u>et dans ce cas</u> $L(E,\mathfrak{a},P) = L(E,\mathfrak{a},P_o)$ <u>et</u> ,

$$\frac{dP}{dP_o} = \frac{e^{\zeta_Q}}{E_{P_o}(e^{\zeta_Q})}$$

<u>où</u> ζ_Q <u>est un élément uniquement déterminé par</u> Q <u>de l'espace de HILBERT</u> $L(E,\mathfrak{a},P_o)^{2\widetilde{\odot}}$.

Les conditions d'équivalence sont données dans [22] , nous n'utiliserons ici que la forme de la densité.

On trouve dans [28] ou [16] des spécialisations de ces théorèmes lorsque les probabilités gaussiennes considérées sont les lois de vecteurs aléatoires associés à divers types de f.a.r. gaussiennes notamment stationnaires. ■

Structures Statistiques gaussiennes.

Les THEOREMES 7 et 8 impliquent que dans une structure statistique gaussienne, la famille $\mathcal{P}$ est nécessairement constituée de lois de probabilités équivalentes. Quant aux THEOREMES 5 et 6 , ils impliquent que l'observation x n'appartient presque-sûrement pas à l'espace de HILBERT $H(Q)$ quand il est de dimension infinie, ce qui aura pour conséquence l'impossibilité de l'extension en dimension infinie de la plupart des méthodes classiques de la statistique gaussienne.

4. - La loi de WIENER.

A titre d'illustration de ce qui précède, nous décrivons maintenant une loi de probabilité gaussienne particulière qui intervient très souvent dans les applications. Ces rappels, loin d'être exhaustifs se limitent aux seuls résultats utilisés dans la suite.

De plus nous nous restreindrons à un intervalle de temps borné, $[0,1]$ par exemple, ce qui pour les applications statistiques correspond aux conditions d'observation.

Soit $C_o[0,1]$ l'espace de BANACH séparable des fonctions numériques continues et nulles en 0 sur l'intervalle $[0,1]$ de $\mathbb{R}$. Son dual topologique peut s'identifier à l'espace $M[0,1]$ des mesures boréliennes bornées sur $[0,1]$; la forme bilinéaire canonique mettant ces deux espaces en dualité séparante étant définie par :

$$x \in C_o[0,1] \ , \ \nu \in M[0,1] \ , \ \langle x,\nu\rangle = \int_0^1 x(t)\, d\nu(t) \ .$$

D'après le THEOREME 3 et le THEOREME 4 du CHAPITRE I, §2, on a :

$$\mathcal{C}(C_o[0,1],M[0,1]) = \mathcal{B}(C_o[0,1]) \quad \text{et} \quad (C_o[0,1],\mathcal{B}(C_o[0,1]))^m \equiv M[0,1] \ .$$

DEFINITION 6. - On appelle "loi de WIENER centrée" la probabilité gaussienne centrée P_W sur l'E.V.M.S. $(C_o[0,1], \mathcal{B}(C_o[0,1]))$ de variance la forme quadratique positive Q_W définie sur $M[0,1]$ par :

$$\nu \in M[0,1] \;,\; Q_W(\nu) = \int_0^1 \int_0^1 \inf(s,t)\, d\nu(s)\, d\nu(t) \;.$$

On appelle "mouvement brownien standard", toute f.a.r. $(X_t)_{t\in[0,1]}$ définie sur un espace probabilisé quelconque telle que le vecteur aléatoire X qui lui est associé suive la loi de WIENER centrée. On appelle "mouvement brownien standard canonique", le mouvement brownien standard $(\xi_t)_{t\in[0,1]}$ défini sur l'espace probabilisé $(C_o[0,1], \mathcal{B}(C_o[0,1]), P_W)$ par : $\xi_t(x) = \langle x, \delta_t \rangle \;,\; x \in C_o[0,1], t \in [0,1]$.

Cette probabilité gaussienne, s'inscrit bien dans le cadre de la PROPOSITION 2 et l'on montre :

THEOREME 9. - L'espace de HILBERT $H(Q_W)$ est constitué des fonctions x appartenant à $C_o[0,1]$, admettant une dérivée x' de carré intégrable par rapport à la mesure de LEBESGUE sur $[0,1]$ et $\|x\|^2_{H(Q_W)} = \int_0^1 x'^2(t)\,dt$. Le dual de l'E.V.P.G. $(C_o[0,1], \mathcal{B}(C_o[0,1]), P_W)$, soit H , est constitué des intégrales stochastiques de la forme : $\xi = \int_0^1 f(t)\, d\xi_t$, $f \in L_2([0,1], \mathcal{B}([0,1]), dt)$, par rapport au mouvement brownien standard canonique. L'isométrie ψ^{-1} de $H(Q_W)$ sur H étant définie par : $x \longmapsto \int_0^1 x'(t)\, d\xi_t$, $x \in H(Q_W)$.

Ce théorème fournit un exemple supplémentaire de dual d'un E.V.P. il permet aussi de présenter ces intégrales stochastiques ξ comme des v.a.r.l. sur l'E.V.P. $(C_o[0,1], \mathcal{B}(C_o[0,1], P_W)$ et donc il existe une version de ξ coïncidant avec une fonctionnelle mesurable, définie et linéaire sur un sous-espace vectoriel de P_W-probabilité 1 de $C_o[0,1]$ et d'après le COROLLAIRE de la PROPOSITION 2, on peut considérer une telle version comme une limite simple P_W-p.-p. de formes linéaires continues sur $C_o[0,1]$ (donc d'éléments de $M[0,1]$) . ∎

La version simplifiée suivante d'un théorème de GIRSANOV est d'un usage fréquent dans les applications concernant certains modèles différentiels stochastiques (cf. [23]) .

THEOREME 10. - Il existe une correspondance biunivoque entre l'ensemble des fonctions numériques θ définies sur $[0,1] \times C_o[0,1]$ telles que :

a) θ soit mesurable par rapport à la tribu des parties de $[0,1] \times C_o[0,1]$ engendrée par les ensembles de la forme : $\{(s,x) : x \in A \ , \ s \geq T(x)\}$ où T est un temps d'arrêt du mouvement brownien standard canonique $(\xi_t)_{t\in[0,1]}$ et A un ensemble de la tribu $\mathcal{B}_T$ du passé par rapport à T - On dit que θ est "bien mesurable" - ,

b) $\int_0^1 \theta^2(t,x)\,dt < \infty$, pour P_W-presque tout $x \in C_o[0,1]$,

c) $E\{\exp(\int_0^1 \theta\, d\xi_t - \frac{1}{2}\int_0^1 \theta^2\, dt)\} = E(M_1^\theta) = 1$;

et l'ensemble des probabilités sur $(C_o[0,1], \mathcal{B}(C_o[0,1]))$ absolument continues par rapport à P_W . Cette correspondance est définie par l'application $\theta \longmapsto P_\theta$ où $P_\theta = M_1^\theta \, . \, P_W$.

En outre, sur l'E.V.P. $(C_o[0,1], \mathcal{B}(C_o[0,1]), P_\theta)$ la f.a.r. : $(\xi_t - \int_0^t \theta\, ds)_{t\in[0,1]}$ est un mouvement brownien standard.

COROLLAIRE. - La solution d'une équation différentielle stochastique de la forme :

$dX_t = \alpha(t)\, X_t\, dt + dX_t^o$, $X_o = 0$, $t \in [0,1]$ où $\alpha(.) \in L^2([0,1], \mathcal{B}([0,1]), dt)$ et $(X_t^o)_{t\in[0,1]}$ est un mouvement brownien standard sur un espace probabilisé quelconque, définit un vecteur aléatoire gaussien (*) sur cet espace

---

(*) au sens des EXEMPLES 1 et 2 du CHAPITRE I, §4 et de ce § .

probabilisé, à valeurs dans l'E.V.M.S. $(C_o[0,1], \mathcal{B}(C_o[0,1]))$ dont la loi de probabilité P_α est équivalente à la loi de WIENER et vérifie :

$$\frac{dP_\alpha}{dP_W} = \exp\left(\int_0^1 \alpha(t)\, \xi_t \, d\xi_t - \frac{1}{2}\int_0^1 \alpha^2(t)\, \xi_t^2 \, dt\right).$$

Des hypothèses supplémentaires sur la fonction α, permettent, le cas échéant dans les applications, de calculer une version de cette densité, on retrouve alors d'anciens théorèmes de CAMERON et MARTIN (cf. [39], par ex.) . ∎

Dans cet esprit, donnons un exemple de probabilité gaussienne équivalente à la loi de WIENER pour laquelle la densité de RADON-NIKODYM se calcule explicitement :

Rappelons qu'une fonction de covariance réelle k sur $[0,1] \times [0,1]$ est dite "triangulaire" si elle est de la forme :
$k_{a,b}(s,t) = a[\min(s,t)] \, . \, b[\max(s,t)]$, $s,t \in [0,1]$ où a et b sont deux fonctions numériques sur $[0,1]$ telles que la fonction numérique $\frac{a}{b}$ soit positive et non décroissante sur cet intervalle. Les f.a.r. gaussiennes centrées à covariance triangulaire sont en particulier markoviennes. (cf. [22],p.39) .

THEOREME 11. - (cf. [38],p.272) Soit $(X_t)_{t\in[0,1]}$ une f.a.r. gaussienne centrée à covariance triangulaire $k_{a,b}$ où les fonctions a et b vérifient de plus : $a(0) \geq 0$; $b(t) > 0$ pour tout $t \in [0,1]$; a et b sont de classe C^2 et $b(t)\,a'(t) - a(t)\,b'(t) > 0$ pour tout $t \in [0,1]$. Alors cette f.a.r. détermine un vecteur aléatoire gaussien X à valeurs dans l'E.V.M.S. $(C_o[0,1], \mathcal{B}(C_o[0,1]))$ dont la loi de probabilité $P_{a,b}$ est équivalente à à la loi de WIENER si et seulement si : $b(t)\,a'(t) - a(t)\,b'(t) = 1$, pour tout $t \in [0,1]$ et $a(0) = 0$ auquel cas, si $x \in C_o[0,1]$,

$$\frac{dP_{a,b}}{dP_W}(x) = \sqrt{\frac{b(0)}{b(1)}}\ \exp\left(\frac{1}{2}\int_0^1 b'(t)\, d\left[\frac{x^2(t)}{b(t)}\right]\right).$$

L'intégrale intervenant dans la densité s'entend au sens de

RIEMANN-STIELTJES soit, en l'intégrant par parties, on obtient :

$$\frac{d P_{a,b}}{d P_W}(x) = \sqrt{\frac{b(0)}{b(1)}}\ \exp\left\{\frac{1}{2}\left(\frac{b'(1)}{b(1)}\ x^2(1) - \int_0^1 \frac{b''(t)}{b(t)}\ x^2(t)\, dt\right)\right\}.$$

On remarque que cette densité ne dépend que de la fonction b . ∎

§2. - LA LOI DE WISHART

DEFINITION 1. - Soit X un vecteur aléatoire gaussien. On appellera Loi de WISHART associée à X, la loi de probabilité du vecteur aléatoire $X \otimes X$. On dira qu'elle est "centrale" ou "non centrale" suivant que X est centré ou décentré, dans l'espace vectoriel où il prend ses valeurs.

La spécification de l'E.V.M.S. $(E,\mathfrak{a})$ dans lequel le vecteur gaussien X prend ses valeurs, détermine l'ensemble des valeurs possibles du vecteur aléatoire $X \otimes X$ composée de X et de l'application quadratique χ_E sur E. La théorie qui suit repose donc sur les résultats exposés au §3 du CHAPITRE I et au §1, n°3 du CHAPITRE II, dont on gardera les notations.

La loi de WISHART sera entièrement déterminée par sa fonctionnelle caractéristique, dont le calcul est le résultat de ce § . Comme pour la loi Normale l'étude de cette loi passe par l'étude de la notion de probabilité cylindrique de WISHART, les résultats de ce § ont été annoncés dans [35] et [37] .

1. Probabilités cylindriques de WISHART.

DEFINITION 2. - Soient E et F deux espaces vectoriels réels en dualité séparante. On appellera "probabilité cylindrique de WISHART relativement à cette dualité", la probabilité cylindrique sur $E^{2\odot}$ relativement à la dualité séparante entre $E^{2\odot}$ et $F^{2\odot}$, image par l'application quadratique χ_E d'une probabilité cylindrique gaussienne sur E relativement à la dualité entre E et F.

Le bien-fondé de cette définition résulte de la PROPOSITION 1, CHAPITRE II, §1, n°3 qui assure l'existence et l'unicité d'une telle probabilité cylindrique. En particulier, compte tenu de la REMARQUE 4, loc.cit., on posera :

DEFINITION 3. - Soient H un espace de HILBERT réel séparable et m un élément quelconque de H . On appellera "probabilité cylindrique de WISHART canonique relative à H , à un degré de liberté, de paramètre de décentrage m" , la probabilité cylindrique sur $H^{2\odot}$ relativement à la dualité séparante entre cet espace vectoriel et lui-même : $\chi_H(\gamma_{m,\|.\|_H^2})$, image de la probabilité cylindrique gaussienne canonique sur H , décentrée de m , i.e. de transformée de FOURIER $\varphi\gamma_{m,\|.\|_H^2} = \exp\{i\langle m,.\rangle - \frac{1}{2}\|.\|_H^2\}$ sur H . On la notera : $\mathfrak{w}_H(1,m;\|.\|_H^2)$:

Dans la suite, on notera H_c l'espace de HILBERT complexe : $H + iH$, compléxifié de H .

PROPOSITION 1. - La probabilité cylindrique $\mathfrak{w}_H(1,m;\|.\|_H^2)$ a pour transformée de FOURIER :

$$(1) \qquad u \in H^{2\odot} \; , \; \varphi_{\mathfrak{w}(1,m;\|.\|_H^2)}(u) = \text{dét}(\mathbb{I}-2iu)^{-\frac{1}{2}} \exp\{i\langle m,(\mathbb{I}-2iu)^{-1}\circ u(m)\rangle_{H_c}\} \; ,$$

où $\text{dét}(\mathbb{I}-2iu)$ désigne la valeur au point $2i$ du déterminant de FREDHOLM de l'opérateur auto-adjoint de rang fini u (identifié à un opérateur de H_c) et $(\mathbb{I}-2iu)^{-1}$ est l'inverse de l'opérateur $(\mathbb{I}-2iu)$ de H_c .

En effet, puisque $H^{2\odot} = \bigcup_{K\in\mathcal{F}(H^{2\odot})} K$ (notations de la REMARQUE 4, CHAPITRE II, §1, n°3) il suffit de montrer que l'égalité (1) est vraie sur tout sous-espace de dimension finie de H . Mais d'après la construction de l'image quadratique d'une probabilité cylindrique définie sur un espace de HILBERT séparable, qui est donnée dans cette remarque, il suffit de vérifier l'égalité (1) seulement sur tout sous-espace de $H^{2\odot}$ de la forme : $V^{2\odot}$ où $V \in \mathcal{F}(H)$. Or, si $V \in \mathcal{F}(H)$, en notant P_V le projecteur orthogonal de H sur V , $(\gamma_{m,\|.\|_H^2})_V = P_V(\gamma_{m,\|.\|_H^2})$ est la probabilité gaussienne sur $(V,\mathcal{B}(V))$ de moyenne $P_V(m)$ et de variance la forme quadratique unité sur V . Le calcul de la transformée de FOURIER $\chi_V[(\gamma_{m,\|.\|_H^2})_V]$ repose alors sur le théorème classique :

THEOREME 1. - (WISHART-JAMES, cf. [5]) . Soit X un vecteur aléatoire gaussien dans $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ de moyenne $a \in \mathbb{R}^k$ de matrice de covariance Λ. La matrice aléatoire $X.{}^tX$ à valeurs dans l'espace de HILBERT S_k des matrices carrées symétriques d'ordre k, pour le produit scalaire $\langle A,B\rangle_{S_k} = \text{Trace}(AB)$, a pour fonction caractéristique :

$$U \in S_k \ , \ \varphi_{X^tX}(U) = \text{dét}(I-2iU\Lambda)^{-\frac{1}{2}} \exp\{i.\text{Trace}(a{}^ta.U.(I-2iU\Lambda)^{-1}\}$$

et pour espérance mathématique : $(a{}^ta + \Lambda) \in S_k^+$ (cône des matrices ≥ 0) .

Supposons maintenant V de dimension k ; en choisissant une base orthonormale dans V, cet espace devient isomorphe à $\mathbb{R}^k$ et l'espace de HILBERT $V^{2\odot}$ (ou l'espace de HILBERT des opérateurs continus auto-adjoints dans V, muni du produit scalaire $\langle u,v\rangle = \text{Trace}(u \circ v)$ est isométrique à l'espace de HILBERT S_k .

En prenant, dans le THEOREME 1 , $a = P_V(m)$ et $\Lambda = I$ on obtient :

$$u \in V^{2\odot} \ , \ \varphi_{X_V[(\gamma_{m,\|.\|_H^2})_V]}(u) = \text{dét}(\mathbb{I}-2iu)^{-\frac{1}{2}} \exp\{i\ \text{Trace}(P_V(m)\otimes P_V(m) \circ u \circ (\mathbb{I}-2iu)^{-1}\}$$

où $\text{dét}(\mathbb{I}-2iu)$ désigne la valeur au point $2i$ du déterminant de FREDHOLM de l'opérateur (de rang fini) u dans V, associé à la matrice U, qui coïncide avec le déterminant de la matrice $(I-2iU) \in S_k + i\,S_k$ (cf. CHAPITRE I, §3, n°1, DEFINITION 5). Comme $P_V(m) \otimes P_V(m) = P_V^{2\odot}(m\otimes m)$ (projection orthogonale de $m \otimes m \in H^{2\odot}$ sur $V^{2\odot}$) , nous avons :

$$\begin{aligned}
\text{Trace}(P_V(m) \otimes P_V(m) \circ u \circ (\mathbb{I}-2iu)^{-1}) &= \langle P_V(m) \otimes P_V(m)\ ,\ u \circ (\mathbb{I}-2iu)^{-1}\rangle_{H_c^{2\odot}} \\
&= \langle m \otimes m\ ,\ u \circ (\mathbb{I}-2iu)^{-1}\rangle_{H_c^{2\odot}} \\
&= \langle m,\ (\mathbb{I}-2iu)^{-1} \circ u(m)\rangle_{H_c} .
\end{aligned}$$

Ceci montre que la formule (1) est valable sur tout sous-espace de dimension finie de $H^{2\odot}$ de la forme $V^{2\odot}$ avec $V \in \mathcal{F}(H)$ puisque si $u \in V^{2\odot}$, $\varphi_{X_H(\gamma_{m,\|.\|_H^2})}(u) = \varphi_{X_V[(\gamma_{m,\|.\|_H^2})_V]}(u)$. ■

PROPOSITION 2. - Dans les mêmes conditions qu'à la proposition précédente, il existe une probabilité cylindrique unique $\hat{\mathfrak{w}}_H(1,m,\|.\|_H^2)$, sur $H^{2\odot}$ relativement à la dualité entre $H^{2\odot}$ et $H^{2\hat{\odot}}$ - restriction à ces deux espaces de la dualité entre l'espace de HILBERT $H^{2\hat{\odot}}$ et lui-même induite par son produit scalaire - dont la transformée de FOURIER est :

$$(2) \quad u \in H^{2\hat{\odot}} \;,\; \varphi_{\hat{\mathfrak{w}}_H(1,m,\|.\|_H^2)}(u) = \text{dét}(\mathbb{I}-2iu)^{\frac{-1}{2}} \exp\{i\langle m,(\mathbb{I}-2iu)^{-1}\circ u(m)\rangle_{H_c}\}$$

où dét$(\mathbb{I}-2iu)$ est la valeur au point $2i$ du déterminant de FREDHOLM de l'opérateur nucléaire auto-adjoint u de H (identifié à un opérateur de H_c) .

En effet, on constate que la fonction à valeurs complexes (1) définie sur $H^{2\odot}$ est continue pour la topologie projective (induite par celle de $H^{2\otimes}$) associé à la topologie de la norme hilbertienne de H ; il résulte alors du THEOREME 7 , CHAPITRE I, §3 que cette fonction se prolonge en une fonction continue sur le produit tensoriel symétrique projectif complété $H^{2\hat{\odot}}$ et que ce prolongement est défini par l'équation (2) . Il suffit, pour achever la démonstration, de remarquer que ce prolongement est encore une fonction complexe, de type positif, continue sur les sous-espaces vectoriels de dimension finie de $H^{2\hat{\odot}}$, prenant la valeur 1 à l'origine, et d'appliquer le THEOREME 3, CHAPITRE II, §1. Il est naturel de noter $\hat{\mathfrak{w}}_H(1,m,\|.\|_H^2)$ la probabilité cylindrique ainsi définie puisque sa transformée de FOURIER est le prolongement de celle de la probabilité cylindrique $\mathfrak{w}_H(1,m,\|.\|_H^2)$ et que la dualité entre $H^{2\odot}$ et lui-même, n'est autre que la restriction de la dualité entre $H^{2\odot}$ et $H^{2\hat{\odot}}$. ∎

COROLLAIRE. - Soient H un espace de HILBERT réel séparable, n un entier positif et $m_1,\ldots,m_n$ des éléments de H . Il existe une probabilité cylindrique unique $\mathfrak{w}_H(n,m_1,\ldots,m_n;\|.\|_H^2)$ [resp. $\hat{\mathfrak{w}}_H(n,m_1,\ldots,m_n;\|.\|_H^2)$] sur $H^{2\odot}$ relativement à la dualité entre $H^{2\odot}$ et lui-même [resp. entre $H^{2\odot}$ et $H^{2\hat{\odot}}$] que l'on appellera "probabilité cylindrique de WISHART canonique relative à H , à n degrés de liberté, de paramètres de décentrage $m_1,\ldots,m_n$" dont la transformée de FOURIER est définie par :

$u \in H^{2\odot}$ [resp. $u \in H^{2\hat{\odot}}$] , $\varphi_{\mathfrak{W}_H(n;m_1,\ldots,m_n,\|.\|^2_H)}(u) = \text{dét}(\mathbb{I}-2iu)^{-\frac{n}{2}}$.

$\exp\{i \sum_{j=1}^{n} \langle m_j, (\mathbb{I}-2iu)^{-1} \circ u(m_j)\rangle_{H_c}\}$ [resp. $= \varphi_{\hat{\mathfrak{W}}_H(n,m_1,\ldots,m_n,\|.\|^2_H)}(u)$]

Cette probabilité cylindrique n'est autre en effet, que la convolution des probabilités cylindriques à un degré de liberté $\mathfrak{W}_H(1,m_j;\|.\|^2_H)$, $j = 1,\ldots,n$ [resp. $\hat{\mathfrak{W}}_H(1,m_j;\|.\|^2_H)$] . ∎

Il est maintenant facile de construire une probabilité cylindrique de WISHART relativement à la dualité entre E et F comme celle de la DEFINITION 2 . Soit π une application linéaire de H dans E continue pour les topologies faibles $\sigma(H,H)$ et $\sigma(E,F)$; nous sommes alors dans les conditions de la PROPOSITION 2, CHAPITRE II, §1 : $\pi(\gamma_{m,\|.\|^2_H})$ est une probabilité cylindrique gaussienne sur E , relativement à la dualité entre E et F , d'après le THEOREME 2, §1 , $\chi_E \circ \pi(\gamma_{m,\|.\|^2_H})$ est une probabilité cylindrique de WISHART relativement à cette dualité, elle peut donc s'écrire aussi : $\pi^{2\odot} \circ \chi_H(\gamma_{m,\|.\|^2_H})$ ou encore : $\pi^{2\odot}[\mathfrak{W}_H(1,m,\|.\|^2_H)]$, ce qui permet de calculer sa transformée de FOURIER, laquelle est définie sur $F^{2\odot}$ par :

(3) $$\varphi_{\chi_E \circ \pi(\gamma_{m,\|.\|^2_H})} = \varphi_{\mathfrak{W}_H(1,m,\|.\|^2_H)} \circ {}^t(\pi^{2\odot}) ,$$

où ${}^t(\pi^{2\odot})$ désigne la transposée de l'application linéaire $\pi^{2\odot}$ de $H^{2\odot}$ dans $E^{2\odot}$ continue pour les topologies faibles $\sigma(H^{2\odot},H^{2\odot})$ et $\sigma(E^{2\odot},F^{2\odot})$.

De même que précédemment, la fonctionnelle (3) définie sur $F^{2\odot}$ peut se prolonger à tout espace vectoriel G contenant $F^{2\odot}$ qui soit en dualité avec $E^{2\odot}$ pourvu que cette dualité, restreinte à $E^{2\odot}$ et $F^{2\odot}$ coïncide avec la précédente et que l'application ${}^t(\pi^{2\odot})$ admette un prolongement continu de G dans $H^{2\hat{\odot}}$ pour les topologies faibles $\sigma(G,E^{2\odot})$ et $\sigma(H^{2\hat{\odot}},H^{2\odot})$.

Enfin, si π transforme la probabilité cylindrique gaussienne canonique décentrée sur H , $\gamma_{m,\|.\|^2_H}$ en une probabilité cylindrique induisant une probabilité sur l'E.V.M.S. $(E,\mathcal{C}(E,F))$, comme on l'a remarqué à la proposition citée plus haut, la probabilité cylindrique de WISHART $\chi_E \circ \pi(\gamma_{m,\|\cdot\|_H^2}) = \pi^{2\odot}[w_H(1,u,\|\cdot\|_H^2)]$ induit une probabilité sur l'E.V.M.S. $(E^{2\odot},\mathcal{C}(E^{2\odot},F^{2\odot}))$ que l'on appellera "probabilité de WISHART".

On saura donc calculer la transformée de FOURIER de l'image quadratique d'une probabilité gaussienne définie sur un E.V.M.S. pourvu que celle-ci soit elle-même l'image linéaire d'une probabilité cylindrique gaussienne canonique décentrée définie sur un espace de HILBERT séparable. En particulier on saura calculer la fonction caractéristique du vecteur aléatoire $X \otimes X$ de loi de WISHART, pourvu que la loi du vecteur gaussien décentré X vérifie cette hypothèse. C'est ce que nous allons appliquer au n° suivant.

2. - La loi de WISHART associée à un vecteur aléatoire gaussien décentré à valeurs dans un espace de FRECHET séparable.

PROPOSITION 3. - Soient E un espace de FRECHET séparable, $\mathcal{B}(E)$ sa tribu borélienne E' son dual topologique, $E^{2\odot}$ [resp. $E'^{2\odot}$] le produit tensoriel symétrique projectif de E [resp. de E' muni de la topologie forte] par lui-même, $\mathcal{B}(E^{2\odot})$ la tribu borélienne de l'espace topologique $E^{2\odot}$ et $E'^{2\hat{\odot}}$ le complété de l'espace topologique $E'^{2\odot}$.

Soient $N_E(0,Q)$ la probabilité gaussienne centrée sur $(E,\mathcal{B}(E))$ de variance la forme quadratique positive Q sur E' et $H(Q)$ le sous-espace hilbertien de E qui lui correspond.

Pour tout $\theta \in H(Q)$ il existe une probabilité unique sur $(E^{2\odot},\mathcal{B}(E^{2\odot}))$ dont la transformée de FOURIER φ soit telle que si $t \in E'^{2\hat{\odot}}$,

$$\varphi(t) = \text{dét}(\mathbb{1} - 2i\eta_Q(t))^{-\frac{1}{2}} \quad \exp\{i\langle \theta, (\mathbb{1}-2i\eta_Q(t))^{-1} \circ \eta_Q(t)(\theta)\rangle_{H(Q)_c}\}$$

où η_Q est une application linéaire continue de $E'^{2\hat{\odot}}$ dans $H(Q)^{2\hat{\odot}}$ uniquement déterminée par Q .

Si X est un vecteur aléatoire de la forme : $\theta + X_o$ où X_o est un vecteur aléatoire gaussien de loi $N_E(0,Q)$ dans $(E,\mathcal{B}(E))$, la relation ci-dessus détermine la fonction caractéristique du vecteur aléatoire $X \otimes X$ dans $(E^{2\odot},\mathcal{B}(E^{2\odot}))$.

On notera $W_E(1,\theta;Q)$ la loi de probabilité de $X \otimes X$ ainsi caractérisée, que l'on appellera loi de WISHART associée à X (ou à un degré de liberté associée à la probabilité gaussienne décentrée $N_E(\theta,Q)$.

On définit aisément, par convolution, la loi de WISHART à n degrés de liberté $W_E(n;\theta_1,\dots,\theta_n,Q)$.

D'après la DEFINITION 4, §1, le THEOREME 3, CHAPITRE I, §2, et le THEOREME 4, CHAPITRE I, §2 la probabilité gaussienne centrée $N_E(0,Q)$ sur $(E,\mathcal{B}(E))$, peut s'écrire aussi $N_{E,E'}(0,Q)$ et sa transformée de FOURIER est définie sur E' par : $x' \in E'$, $\varphi_{N_E(0,Q)}(x') = \exp\{-\frac{1}{2}Q(x')\}$.

D'après le THEOREME 6, §1, il existe un sous-espace hilbertien séparable $H(Q)$ de E uniquement déterminé par Q, tel que l'injection canonique j de $H(Q)$ dans E soit continue et transforme la probabilité cylindrique gaussienne canonique $\gamma_{0,\|\cdot\|^2_{H(Q)}}$ sur $H(Q)$ en la probabilité gaussienne $N_E(0,Q)$ sur $(E,\mathcal{B}(E))$. Soit $\theta \in H(Q)$; la probabilité cylindrique gaussienne canonique décentrée de θ, $\gamma_{\theta,\|\cdot\|^2_{H(Q)}}$ sur $H(Q)$ est définie par sa transformée de FOURIER sur $H(Q)$:

$$h \in H(Q)\ ,\ \varphi_{\gamma_{\theta,\|\cdot\|^2_{H(Q)}}}(h) = \exp\{i\langle\theta,h\rangle_{H(Q)} - \frac{1}{2}\|h\|^2_{H(Q)}\}\ .$$

L'application linéaire j étant continue de $H(Q)$ dans E, elle est continue pour les topologies faibles $\sigma(H(Q),H(Q))$ et $\sigma(E,E')$, (cf. [6]), et l'image par j de $\gamma_{\theta,\|\cdot\|^2_{H(Q)}}$ induit la probabilité gaussienne décentrée $N_E(\theta,Q)$ sur $(E,\mathcal{B}(E))$ dont la transformée de FOURIER est :

$$x' \in E' \ , \ \varphi_{j(\gamma_{\theta,\|.\|^2_{H(Q)}})}(x') = \exp\{i\langle\theta, {}^t j(x')\rangle_{H(Q)} - \frac{1}{2} \ \|{}^t j(x')\|^2_{H(Q)}\}$$
$$= \exp\{i\langle\theta,x'\rangle - \frac{1}{2} \ Q(x')\} \ , \text{ car } \ j(\theta) = \theta \ .$$

L'image quadratique $\chi_E \circ j(\gamma_{\theta,\|.\|^2_{H(Q)}}) = \chi_E[N_E(\theta,Q)]$ est donc une probabilité sur l'E.V.M.S. $(E^{2\odot},\mathcal{C}(E^{2\odot},E'^{2\odot}))$ d'après les raisonnements du n°1 et sa transformée de FOURIER est :

$$\varphi_{\chi_E[N_E(\theta,Q)]} = \varphi_{\mathfrak{w}_{H(Q)}(1,\theta,\|.\|^2_{H(Q)})} \circ {}^t j^{2\odot}$$

soit :

$$(4) \ t \in E'^{2\odot}, \quad \varphi_{\chi_E[N_E(\theta,Q)]}(t) = \text{dét}(\mathbb{I} - 2i\ {}^t j^{2\odot}(t))^{-\frac{1}{2}} \ .$$
$$\exp\{i\langle\theta,(\mathbb{I}-2i\,{}^t j^{2\odot}(t))^{-1} \circ {}^t j^{2\odot}(t)(\theta)\rangle_{H(Q)_c}\} \ .$$

Mais d'après la PROPOSITION 2, CHAPITRE I, §3 la tribu faible $\mathcal{C}(E^{2\odot},E'^{2\odot})$ coïncide avec la tribu des boréliens $\mathcal{B}(E^{2\odot})$ du produit tensoriel symétrique projectif $E^{2\odot}$. Donc la probabilité $\chi_E[N_E(\theta,Q)]$ est borélienne c'est-à-dire qu'elle est définie sur l'E.V.M.S. $(E^{2\odot},\mathcal{B}(E^{2\odot}))$, et elle est entièrement déterminée par (4) .

Cependant, il est intéressant pour les applications de prolonger la transformée de FOURIER de cette probabilité afin d'en avoir une expression explicite sur un sous-espace le plus grand possible de $(E^{2\odot})'$, c'est ce qui est annoncé dans la proposition ; en effet, ${}^t j^{2\odot} = ({}^t j)^{2\odot}$; de plus, ${}^t j$ est une application linéaire faiblement et fortement continue de E' dans $H(Q)$, (cf. [6]), par suite ${}^t j^{2\odot}$ est continue du produit tensoriel symétrique projectif $E'^{2\odot}$ dans le produit tensoriel symétrique $H(Q)^{2\odot}$ associés aux topologies fortes respectivement sur E' et sur $H(Q)$. (cette dernière étant la topologie hilbertienne de $H(Q)$) . Soit η_Q le prolongement canonique continu de ${}^t j^{2\odot}$ du produit tensoriel symétrique projectif complété $E'^{2\hat{\odot}}$ dans le produit tensoriel symétrique projectif complété $H(Q)^{2\hat{\odot}}$;

alors, la fonction φ définie sur $E'^{2\hat{\odot}}$ par :

$$(5) \quad t \in E'^{2\hat{\odot}} \,, \ \varphi(t) = \varphi_{\mathfrak{w}_{H(Q)}(1,\theta,\|.\|^2_{H(Q)})} \circ \eta_Q(t)$$

$$= \text{dét}(\mathbb{I} - 2i\eta_Q(t))^{\frac{-1}{2}} \ \exp\{i\langle\theta,(\mathbb{I}-2i\eta_Q(t))^{-1} \circ \eta_Q(t)(\theta)\rangle_{H(Q)_c}\}$$

est la transformée de FOURIER d'une loi de probabilité unique sur l'E.V.M.S. $(E^{2\odot}, \mathcal{C}(E^{2\odot}, E'^{2\hat{\odot}}))$ mais comme, d'après le COROLLAIRE de la PROPOSITION 2, CHAPITRE I, §3, $\mathcal{C}(E^{2\odot}, E'^{2\hat{\odot}}) = \mathcal{B}(E^{2\odot})$, c'est celle de $\chi_E[N_E(\theta,Q)]$. ∎

Enfin, la dernière partie de la proposition est évidente, puisque (4) et (5) caractérisent la même loi de probabilité, celle du vecteur aléatoire $X \otimes X$ quand $X \rightsquigarrow N_E(\theta,Q)$ dans $(E,\mathcal{B}(E))$. Si $X_1,\dots,X_n$ sont n vecteurs aléatoires indépendants de lois respectives $N_E(\theta_j,Q); j = 1,\dots,n$ dans $(E,\mathcal{B}(E))$ on aura $\sum_{j=1}^{n} X_j \otimes X_j \rightsquigarrow W_E(n,\theta_1,\dots,\theta_n,Q)$.

REMARQUE 1. - Si $X \otimes X \rightsquigarrow W_E(1,\theta;Q)$, d'après la PROPOSITION 1, CHAPITRE II, §4, $X \otimes X$ est scalairement intégrable et l'application bilinéaire symétrique sur $E' \times E'$:

$$: (y,y') \longmapsto E(\langle X\otimes X, y\otimes y'\rangle) = E(\langle X,y\rangle\langle X,y'\rangle) \quad \text{coïncide avec :}$$

$$(y,y') \longmapsto K(y,y') + \langle \theta\otimes\theta, y\otimes y'\rangle \,,$$

où $K(.,.)$ est la forme bilinéaire symétrique associée à la forme quadratique positive Q sur E'.

Sauf si E est de dimension finie, cette forme bilinéaire qui peut s'identifier à un élément du dual algébrique $(E'^{2\odot})^*$ de $E'^{2\odot}$ n'appartient pas à $E^{2\odot}$ et donc $X \otimes X$ n'a pas d'espérance mathématique dans cet espace. On sait cependant, d'après cette même proposition qu'elle appartient à l'espace $E^{2\hat{\odot}}$, elle est donc de la forme : $\theta \otimes \theta + \Lambda_K$ où $\Lambda_K \in E^{2\hat{\odot}}$ est la covariance faible de X. ∎

On peut également généraliser en dimension infinie le théorème de WISHART-BARTLET, connu en dimension 1 sous le nom de théorème de FISHER (cf. [5], p.96) :

PROPOSITION 4. - <u>Avec les mêmes hypothèses qu'à la PROPOSITION</u> 2, <u>soient</u> $X_1, \ldots X_n$ <u>des vecteurs aléatoires gaussiens indépendants de même loi que</u> X c.à.d. $N_E(\theta, Q)$ <u>dans</u> $(E, \mathcal{B}(E))$. <u>Alors les vecteurs aléatoires</u> $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ <u>et</u> $S = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}) \otimes (X_j - \bar{X})$ <u>sont indépendants et suivent respectivement la loi</u> $N_E(\theta, \frac{Q}{n})$ <u>et la loi de WISHART centrale à</u> $n-1$ <u>degrés de liberté</u> $W_E(n-1, 0, Q)$.

La démonstration est calquée sur celle qui est donnée en dimension finie dans [16] . Remarquons que $S = \frac{1}{n}\sum_{j=1}^{n} X_j \otimes X_j - n\bar{X} \otimes \bar{X}$ et que l'on peut se ramener au cas où $\theta = 0$. Soit $C = (c_{j,k})$, $j = 1, \ldots, n$; $k = 1, \ldots$ une matrice carrée orthogonale d'ordre n de première ligne :

$(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$ et posons $Y_k = \sum_{j=1}^{n} c_{j,k} X_j$, $k = 1, \ldots, n$.

Les vecteurs aléatoires $Y_1, \ldots, Y_n$ dans $(E, \mathcal{B}(E))$ sont également indépendants et de loi $N_E(0, Q)$ et l'on a : $\sum_{j=1}^{n} X_j \otimes X_j = \sum_{k=1}^{n} Y_k \otimes Y_k$. En effet, le vecteur aléatoire $(Y_1, \ldots, Y_n)$ à valeurs dans $(E^n, \mathcal{B}(E^n))$ a pour fonction caractéristique :

$$x' = (x'_1, \ldots, x'_n) \in E'^n \ , \ \varphi_{Y_1, \ldots, Y_n}(x') = E(e^{i \sum_{k=1}^{n} \langle Y_k, x'_k \rangle})$$

$$= E(e^{i \sum_{k=1}^{n}\sum_{j=1}^{n} c_{j,k} \langle X_j, x'_k \rangle})$$

$$= E(e^{i \sum_{j=1}^{n} \langle X_j, \sum_{k=1}^{n} c_{j,k} x'_k \rangle})$$

$$= \prod_{j=1}^{n} \varphi_{X_j}(\sum_{k=1}^{n} c_{j,k} x'_k)$$

$$= \exp\{-\frac{1}{2} \sum_{j=1}^{n} Q(\sum_{k=1}^{n} c_{j,k}\, x'_k)\}$$

$$= \exp\{-\frac{1}{2} \sum_{k=1}^{n} Q(x'_k)\}$$

$$= \prod_{k=1}^{n} \varphi_{Y_k}(x'_k) \ ,$$

compte tenu des propriétés de la matrice C . Ceci montre à la fois que les Y_k , $k = 1,\dots,n$ sont indépendants et sont de loi $N_E(0,Q)$. De plus, on a $X = \frac{1}{\sqrt{n}} Y_1$ et $S = \sum_{k=2}^{n} Y_k \otimes Y_k$ d'où les résultats annoncés. ∎

3. Application aux fonctions aléatoires réelles gaussiennes.

La connaissance de la transformée de FOURIER de la loi de WISHART généralisée permet d'apporter une contribution à l'analyse quadratique de certaines f.a.r. gaussiennes :

PROPOSITION 5. - Soient T un espace métrique σ-compact et $(X_t)_{t\in T}$ une fonction aléatoire réelle gaussienne, presque-sûrement à trajectoires continues sur T telle que :

a) $E(X_t) = \theta(t)$, $t \in T$,

b) $E[(X_s - \theta(s))(X_t - \theta(t))] = k(s,t)$, $(s,t) \in T \times T$,

c) la covariance k soit continue sur $T \times T$,

d) la fonction θ appartienne à l'espace de HILBERT autoreproduisant $H(k)$.

Alors, quelle que soit la mesure ν borélienne régulière, à support compact sur $T \times T$ telle que $\nu(A \times B) = \nu(B \times A)$ si A et B sont des boréliens de T , la variable aléatoire réelle :

$$Z_\nu = \iint_{T\times T} X_s\, X_t \, d\nu(s,t)$$

a pour fonction caractéristique :

$$\tau \in \mathbb{R} \ , \ \varphi_{Z_\nu}(\tau) = \text{dét}(\mathbb{I} - 2i\tau A_\nu^k)^{-\frac{1}{2}} \exp\{i\tau \langle \theta, (\mathbb{I} - 2i\tau A_\nu^k)^{-1} \circ A_\nu^k(\theta) \rangle_{H(k)_C}\}$$

où A_ν^k est l'opérateur nucléaire autoadjoint dans $H(k)$ défini par :

$$f \in H(k) \ , \ A_\nu^k(f) = \iint_{T \times T} f(s)\ k(t,.)\ d\nu(s,t) \ ;$$

et pour espérance mathématique :

$$E(Z_\nu) = \langle \theta, A_\nu^k(\theta) \rangle_{H(k)} + \text{Trace}\,(A_\nu^k) = \iint_{T \times T} [\theta(s)\,\theta(t) + k(s,t)]\ d\nu(s,t) \ .$$

En particulier, quelle que soit la mesure μ borélienne régulière à support compact sur T , la variable aléatoire réelle :

$$Y_\mu = \int_T X_t^2\ d\mu(t)$$

a pour fonction caractéristique :

$$\tau \in \mathbb{R} \ , \ \varphi_{Y_\mu}(\tau) = \text{dét}(\mathbb{I} - 2i\tau B_\mu^k)^{-\frac{1}{2}} \exp\{i\tau \langle \theta, (\mathbb{I} - 2i\tau B_\mu^k)^{-1} \circ B_\mu^k(\theta) \rangle_{H(k)_C}\}$$

où B_μ^k est l'opérateur nucléaire auto-adjoint dans $H(k)$ défini par :

$$f \in H(k) \ , \ B_\mu^k(f) = \int_T f(t)\ k(t,.)\ d\mu(t)$$

et pour espérance mathématique :

$$E(Y_\mu) = \langle \theta, B_\mu^k(\theta) \rangle_{H(k)} + \text{Trace}(B_\mu^k) = \int_T [\theta^2(t) + k(t,t)]\ d\mu(t) \ .$$

Plus généralement, si $\nu_1, \ldots, \nu_n$ [resp. $\mu_1, \ldots, \mu_n$] sont n mesures définies comme ν [resp. comme μ] , les vecteurs aléatoires dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $Z = (Z_{\nu_1}, \ldots, Z_{\nu_n})$ et $Y = (Y_{\mu_1}, \ldots, Y_{\mu_n})$ ont respectivement pour fonction caractéristique :

$$(\tau_1,\ldots,\tau_n) \in \mathbb{R}^n \ , \ \varphi_Z(\tau_1,\ldots,\tau_n) = \varphi_{Z_{\sum_{j=1}^{n} \tau_j \nu_j}}(1) \ ,$$

$$\varphi_Y(t_1,\ldots,t_n) = \varphi_{Y_{\sum_{j=1}^{n} \tau_j \mu_j}}(1)$$

La démonstration résulte du calcul de la transformée de FOURIER de la loi de WISHART associée au vecteur aléatoire gaussien X défini par la f.a.r. gaussienne $(X_t)_{t\in T}$.

Avec les conditions imposées à cette f.a.r., le vecteur X est exactement celui qui est décrit à l'EXEMPLE 2, §1, il prend ses valeurs dans l'espace de FRECHET séparable $C(T)$ des fonctions réelles continues sur T , muni de sa tribu borélienne dont le dual (topologique ou mesurable) est identifié à l'espace $M_c(T)$ des mesures boréliennes régulières à support compact sur T . Sa loi de probabilité P_X est la loi normale $N_{C(T)}(\theta, Q_k)$ dont la transformée de FOURIER est :

$$\mu \in M_c(T) \ , \ \varphi_X(\mu) = \exp\{ i\langle \theta,\mu\rangle - \frac{1}{2} \ Q_k(\mu)\} \ , \text{ soit :}$$

$$\mu \in M_c(T) \ , \ \varphi_X(\mu) = \exp\{ i\int_T \theta(t)\,d\mu(t) - \frac{1}{2} \iint_{T\times T} k(s,t)\ d\mu(s)\ d\mu(t)\} \ .$$

Nous sommes donc exactement dans les conditions de la PROPOSITION 3, en prenant pour espace E l'espace $C(T)$ et pour E' l'espace $M_c(T)$. Le vecteur aléatoire $X \otimes X$ à valeurs dans $(C(T)^{2\odot}, \mathcal{B}(C(T)^{2\odot}))$ est en fait le vecteur aléatoire associé à la f.a.r. $(X_s.X_t)_{(s,t)\in T\times T}$ sur $T\times T$.

On sait que l'espace hilbertien $H(Q_k)$ n'est autre que l'espace de HILBERT séparable autoreproduisant associé au noyau k sur $T\times T$, que l'on note plutôt $H(k)$; il est engendré par les fonctions sur T : $\{k(t,.), t\in T\}$, contenu dans $C(T)$ et l'injection canonique j de $H(k)$ dans $C(T)$ est continue.

La fonction caractéristique de $X \otimes X$ est entièrement déterminée par ses valeurs sur $M_c(T)^{2\odot}$:

$$\nu \in M_c(T)^{2\odot}, \quad \varphi_{X\otimes X}(\nu) = E(e^{i\langle X\otimes X,\nu\rangle}) = E(\exp\{i\iint_{T\times T} X_s X_t \, d\nu(s,t)\})$$

$$= \text{dét}(\mathbb{I}-2i\,{}^tj^{2\odot}(\nu))^{-\frac{1}{2}} \exp\{i\langle \theta, (\mathbb{I}-2i\,{}^tj^{2\odot}(\nu))^{-1} \circ {}^tj^{2\odot}(\nu)(\theta)\rangle_{H(k)}$$

Il ne reste plus qu'à expliciter l'application linéaire ${}^tj^{2\odot}$ de $M_c(T)^{2\odot}$ dans $H(k)^{2\odot}$ d'une part, et à montrer, d'autre part, qu'elle se prolonge en une application linéaire de l'espace $M_c^s(T\times T)$ des mesures boréliennes régulières ν à support compact sur $T \times T$ et symétriques (c'est-à-dire, telles que $\nu(A\times B) = \nu(B\times A)$, $A,B \in \mathcal{B}(T)$) dans l'espace $H(k)^{2\hat{\odot}}$ identifié à l'espace des opérateurs nucléaires auto-adjoints dans $H(k)$.

Explicitons d'abord l'application : ${}^tj^{2\otimes}$ de $M_c(T)^{2\otimes}$ dans $H(k)^{2\otimes}$; on obtiendra l'application "symétrique" ${}^tj^{2\odot}$ associée. ${}^tj^{2\otimes} = ({}^tj)^{2\otimes}$ et ${}^tj$, transposée de l'injection canonique j de $H(k)$ dans $C(T)$ est définie par :

$$t \in T \ , \ f \in H(k) \ \langle {}^tj(\delta_t), f\rangle_{H(k)} = \langle f,\delta_t\rangle = f(t)$$

ce qui implique, d'après la propriété du produit scalaire de $H(k)$ que ${}^tj(\delta_t) = k(t,.)$ pour $t \in T$. Par conséquent, on a :

$$(s,t) \in T\times T \ , \ ({}^tj)^{2\otimes}(\delta_s \otimes \delta_t) = {}^tj(\delta_s) \otimes {}^tj(\delta_t) = k(s,.) \otimes k(t,.) = k(s,.).k(t,.)$$

de plus, l'opérateur $A^k_{\delta_s \otimes \delta_t}$ de rang 1 , dans $H(k)$, associé à l'élément ${}^tj^{2\otimes}(\delta_s \otimes \delta_t)$ de $H(k)^{2\otimes}$ est défini par :

$$f \in H(k) \ , \ A^k_{\delta_s \otimes \delta_t}(f) = \langle f, k(s,.)\rangle_{H(k)} \, . \, k(t,.) \ ,$$

soit :

$$A^k_{\delta_s \otimes \delta_t}(f)(.) = f(s).k(t,.) = \iint_{T\times T} f(u)\, k(v,.)\, d\delta_s \otimes \delta_t(u,v) \ ,$$

et :

$$\mathrm{Trace}(A^k_{\delta_s \otimes \delta_t}) = \langle k(s,.),k(t,.)\rangle_{H(k)} = k(s,t) = \iint_{T\times T} k(u,v)\, d\delta_s \otimes \delta_t(u,v) \ .$$

On en déduit immédiatement l'opérateur A^k_ν associé à ${}^t j^{2\otimes}(\nu) \in H(k)^{2\otimes}$ lorsque ν appartient à l'espace vectoriel engendré par la famille $\{\delta_s \otimes \delta_t ; (s,t) \in T\times T\}$; comme il est séquentiellement dense dans $M_c(T\times T)$ pour la topologie faible $\sigma(M_c(T\times T), C(T\times T))$, on en déduit que l'application linéaire ${}^t j^{2\otimes}$ se prolonge en une application linéaire ${}^t j^{2\otimes}$ de $M_c(T\times T)$ dans $H(k)^{2\hat{\otimes}}$ telle que si $\nu \in M_c(T\times T)$, l'opérateur A^k_ν dans $H(k)$ associé à ${}^t j^{2\otimes}(\nu)$ soit défini par :

$$f \in H(k) \ , \ A^k_\nu (f)(.) = \iint_{T\times T} f(s)\, k(t,.)\, d\nu(s,t) \ ,$$

avec :

$$\mathrm{Trace}\ (A^k_\nu) = \iint_{T\times T} k(s,t)\, d\nu(s,t) \ .$$

Enfin, l'opérateur A^k_ν est autoadjoint, donc identifié à un élément de $H(k)^{2\odot}$ si et seulement si pour tous f et g dans $H(k)$ on a : $\iint_{T\times T} f(s)\, g(t)\, d\nu(s,t) = \iint_{T\times T} f(t)\, g(s)\, d\nu(s,t)$ ce qui implique que la mesure ν soit symétrique, donc appartienne à l'espace $M^s_c(T\times T)$ défini plus haut. La restriction de ${}^t j^{2\hat{\otimes}}$ à cet espace, est l'application linéaire cherchée de $M^s_c(T\times T)$ dans $H(k)^{2\hat{\odot}}$; par conséquent, l'élément aléatoire $X\otimes X$ dans $(C(T)^{2\odot}, \mathcal{B}(C(T)^{2\odot}))$ est tel que sa fonction caractéristique vérifie :

$$\nu \in M^s_c(T\times T) \ , \ \varphi_{X\otimes X}(\nu) = \det(\mathbb{I} - 2iA^k_\nu)^{-\frac{1}{2}} \exp[i\langle \theta, (\mathbb{I} - 2iA^k_\nu)^{-1} \circ A^k_\nu(\theta)\rangle_{H(k)_c}]$$

et comme $\varphi_{X\otimes X}(\nu) = E(e^{i\langle X\otimes X,\nu\rangle})$, la v.a.r. $Z_\nu = \langle X\otimes X, \nu\rangle = \iint_{T\times T} X_s X_t\, d\nu(s,t)$ a pour fonction caractéristique : $\varphi_{Z_\nu}(t) = \varphi_{X\otimes X}(\tau\nu)$, $t \in \mathbb{R}$; d'où la formule de la proposition.

Pour le calcul de $E(Z_\nu)$, nous savons que pour tout $y \in H(k)^{2\hat{\odot}}$, la v.a.r. $x \longmapsto \langle x,y\rangle$ sur $H(k)^{2\odot}$ muni de la probabilité cylindrique de

WISHART canonique à un degré de liberté, de paramètre de décentrage θ, définie à la PROPOSITION 2 a pour espérance mathématique : $\langle \theta \otimes \theta, y\rangle + \mathrm{Trace}(y)$ (pour la dualité entre $H^{2\odot}$ et $H^{2\hat{\odot}}$. Par suite, le vecteur aléatoire X est tel que :

$$\nu \in M_c^s(T\times T) \ , \ E(\langle X\otimes X, \nu\rangle) = \langle \theta \otimes \theta, A_\nu^k\rangle + \mathrm{Trace}(A_\nu^k)$$

$$= \langle \theta, A_\nu^k(\theta)\rangle_{H(k)} + \mathrm{Trace}(A_\nu^k)$$

$$= \iint_{T\times T} [\theta(s)\,\theta(t) + k(s,t)]\, d\nu(s,t) \ .$$

A partir de la loi de $X\otimes X$, nous pouvons déduire la loi de probabilité du vecteur aléatoire G associé au processus : $(X_t^2)_{t\in T}$. En effet, G est à valeurs dans $(C(T), \mathcal{B}(C(T)))$ et sa fonction caractéristique sera définie par :

$$\mu \in M_c(T) \ , \ \varphi_G(\mu) = E(e^{i\langle G,\mu\rangle}) = E(\exp\{i\int_T X_t^2\, d\mu(t)\}) \ .$$

Soit, alors, Δ l'application mesurable de $(T, \mathcal{B}(T))$ dans $(T\times T, \mathcal{B}(T)\otimes \mathcal{B}(T))$ définie par : $\Delta(t) = (t,t)$, $t \in T$; on a : $\int_T X_t^2\, d\mu(t) = \int_{T\times T} X\otimes X\, d\,\mu\circ\Delta^{-1}$ où $\mu\circ\Delta^{-1}$ est la mesure image de μ par Δ, portée par la diagonale de $T\times T$ et appartenant donc à $M_c^s(T\times T)$. D'où $\varphi_G(\mu) = \varphi_{X\otimes X}(\mu\circ\Delta^{-1})$. Il suffit maintenant, de remarquer que $A^k_{\mu\circ\Delta^{-1}} = B_\mu^k$ est défini par :

$$f \in H(k) \ , \ B_\mu^k(f) = \iint_{T\times T} f(s)\; k(t,.)\; d\,\mu\circ\Delta^{-1}(s,t)$$

$$= \int_T f(t)\; k(t,.)\; d\mu(t) \ .$$

on obtient ainsi le deuxième groupe de formules de la proposition. Enfin, les fonctions caractéristiques des vecteurs aléatoires Z et Y résultent d'un calcul évident. ∎

Cette proposition constitue déjà dans le cas centré une extension à une classe assez large de f.a.r. gaussiennes des résultats connus concernant la fonction caractéristique de certaines v.a.r. quadratiques, associées à ces f.a.r.

On retrouve en effet, comme cas particuliers les résultats de HIDA (cf. [15]) pour le mouvement brownien centré, en considérant l'isométrie entre les espaces de HILBERT $H(Q_W)$ et $L_2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), dt)$ qui s'expriment ainsi avec les notations de la théorie de FREDHOLM classique sur ce dernier espace.

L'origine du problème remonte au calcul de la fonction caractéristique de la v.a.r. $\int_0^1 X^2(t)\,dt$ quand $[X(t)]_{t \in [0,1]}$ est un "pont brownien" $(X(1) = 0)$ dans le but d'obtenir des propriétés asymptotiques du célèbre test d'adéquation de CRAMÉR-Von MISES en statistique non paramétrique (cf. [2]) .

4. <u>Contribution à l'étude asymptotique de certains tests basés sur le critère de CRAMÉR-Von MISES</u>.

Cette extension des formules connues, à des f.a.r. gaussiennes centrées autres que celle du mouvement (ou du pont) brownien, ainsi que l'extension au cas décentré, peuvent s'appliquer avec profit à des problèmes de statistique classique :

En effet, depuis l'article [2] , de nombreux auteurs se sont attachés à utiliser un critère analogue à celui de CRAMÉR-Von MISES pour divers tests d'hypothèses et à étudier les propriétés asymptotiques de la fonction puissance du test utilisé sous l'hypothèse à tester ainsi que sous certaines de ses alternatives. La bibliographie dans ce domaine est très importante. De façon générale, étant donnée une hypothèse à tester H_o concernant la loi de probabilité d'un élément aléatoire X , au vu d'un échantillon de n observations indépendantes de X , on utilise un test basé sur une statistique T_n qui, sous l'hypothèse H_o converge en loi vers une v.a.r. de la forme $\int_0^1 X^2(t)\,\psi(t)\,dt$ où $(X_t)_{t\in[0,1]}$ est une f.a.r. gaussienne centrée de covariance connue et ψ une fonction mesurable positive connue et vérifiant certaines conditions, définies sur l'intervalle $[0,1]$.

Très souvent, on considère alors la limite en loi de cette statistique sous certaines hypothèses alternatives $H' \in H_0^c$ et l'on montre qu'elle se comporte comme une v.a.r. de la forme $Z' = \int_0^1 [\delta(t)+X(t)]^2\,\psi(t)\,dt$ où δ est une fonction donnée sur $[0,1]$. (cf. [9], par ex.) . Ces limites en loi sont sensées décrire les propriétés asymptotiques des tests utilisés car sauf dans des cas très particuliers concernant la covariance de la f.a.r. $(X(t))_{t\in[0,1]}$ ainsi que les fonctions ψ et δ la fonction caractéristique et à fortiori la loi de probabilité des v.a.r. Z et Z' n'est pas connue. Par conséquent, les résultats précédents sont susceptibles d'apporter une contribution à la solution de ce problème.

Citons très brièvement quelques exemples récents de problèmes ainsi traités :

- Tests d'adéquation à une loi de probabilité quand les paramètres importuns sont préalablement estimés (cf. [10], par ex.)

- Tests de changement de la valeur du paramètre au cours d'une série d'observations de structure exponentielle scalaire (cf. [18]) .

- Tests de symétrie concernant une loi de probabilité sur la droite (cf. [29]) où les auteurs se limitent à des fonctions ψ de la forme : $\psi(t) = \alpha t^k$, $k>-2$) .

Enfin les théorèmes de KOLMOGOROV-SMIRNOV s'étendant à des observations non indépendantes (processus mélangeants) certains auteurs utilisent des tests basés sur un critère quadratique qui conduisent à des problèmes semblables à celui évoqué plus haut. ∎

5. - Application à un test d'hypothèses concernant la moyenne d'une f.a.r. gaussienne - Tests quadratiques.

Les résultats nouveaux obtenus au n°3 nous permettent d'envisager la construction de certains tests concernant la moyenne inconnue d'une f.a.r. gaussienne dans la structure statistique gaussienne correspondant à l'observation d'une trajectoire de cette f.a.r. L'intérêt de ces tests que nous appellerons "tests quadratiques" est que nous disposerons d'une expression de la fonction puissance. Cependant nous ne faisons qu'esquisser ici une telle théorie, en fournissant quelques lemmes préliminaires à une étude plus approfondie de ces tests et de leur fonction puissance.

a) Tests d'hypothèses linéaires.

On dispose d'une observation $x = (x_t)_{t\in T}$ d'une f.a.r. $(X_t)_{t\in T}$ sur un ensemble T , de la forme : $X_t = \theta(t) + X_t^o$; $t \in T$ où $\theta = [\theta(t)]_{t\in T}$ est une fonction déterministe, inconnue et $(X_t^o)_{t\in T}$ une f.a.r. gaussienne centrée sur T de covariance k connue sur $T \times T$.

En se reportant à l'EXEMPLE 1, §1, pour les notations, on en déduit que la structure statistique associée à cette observation est de la forme :

$$(\mathbb{R}^T, \underset{t\in T}{\otimes} \mathcal{B}(\mathbb{R}) , \{P_\theta = N_{\mathbb{R}^T}(\theta, Q_k) \; ; \; \theta \in \Theta \subset H(Q_k)\})$$

qui est une structure statistique gaussienne, à moyenne inconnue, celle-ci variant dans une partie Θ de l'espace de HILBERT autoreproduisant $H(Q_k)$ $(=H(k))$ associé au noyau k . L'ensemble Θ , qui traduit une hypothèse à priori faite sur la moyenne inconnue θ , doit être contenu dans $H(k)$, afin que les lois de probabilités P_θ soient équivalentes, compte tenu du THEOREME 7 §1, et donc que la structure statistique ait un sens. On supposera, dans la suite, que Θ est un sous-espace vectoriel fermé de $H(k)$.

<u>On appellera</u> "<u>test d'hypothèse linéaire concernant</u> θ" , <u>le test de l'hypothèse</u> H_V : "$m \in V$" <u>contre l'hypothèse</u> "$m \notin V$" <u>où</u> V <u>désigne</u>

un sous-espace vectoriel de Θ .

Lorsque T est fini, ce problème de test concernant la moyenne d'un vecteur aléatoire gaussien est bien connu pour son application à l'analyse de variance (cf. [5]) et admet une solution satisfaisante (même lorsque k n'est connue qu'à un facteur constant près). Le test porte sur la statistique quadratique : $\|x_{V^\perp}\|^2$, où $x_{V^\perp}$ désigne la projection orthogonale de l'observation x sur le sous-espace vectoriel de Θ orthogonal à V dans cet espace ; ce choix étant motivé par le fait que la statistique $x_{V^\perp}$ est exhaustive pour le vecteur $\theta_{V^\perp}$ et que l'hypothèse à tester est équivalente à $\|\theta_{V^\perp}\|^2 = 0$.

Mais lorsque T est infini, comme on l'a vu au THEOREME 5, §1, quelle que soit l'hypothèse θ , l'espace de HILBERT $H(k)$ est de probabilité nulle dès qu'il est de dimension infinie ; et donc l'observation x , qui presque-sûrement n'appartient pas à cet espace, ne saurait être projetée sur un de ses sous-espaces, et la méthode ne peut être malheureusement pas étendue en dimension quelconque. On peut penser, en effet, que les tests d'hypothèses linéaires dans le cas général trouvent leur application dans des problèmes d'analyse de variance avec variation continue du niveau des facteurs.

On peut cependant conserver l'idée de construire des tests portant sur des formes quadratiques de l'observation. C'est ce que nous ferons dans le cas particulier important suivant :

b) Tests quadratiques.

Supposons que l'ensemble T et la f.a.r. $(X_t)_{t\in T}$ satisfassent les hypothèses de la PROPOSITION 5, et que θ ne soit à priori soumis à d'autre condition que d'appartenir à $H(k)$.

Considérons le test de l'hypothèse simple H_o : "$\theta = 0$" contre l'hypothèse : "$\theta \neq 0$" dans la structure statistique gaussienne :

(1) $$(C(T),\mathcal{B}(C(T)),\ \{N_{C(T)}(\theta,Q_k)\ ;\ \theta \in H(k)\})\ .$$

DEFINITION 4. - On appellera "test quadratique" de "$\theta = 0$" contre "$\theta \neq 0$", dans la structure statistique (1), tout test de la forme :

$$x \in C(T) \; , \; \Phi_{\mu,l}(x) = \mathbb{1}_{\{[q_\mu(x)]^{\frac{1}{2}} > l\}}$$

où $q_\mu(x) = \int_T x^2(t)\,d\mu(t)$, $\mu \in M_c^+(T)$, cône des mesures boréliennes régulières à support compact, positives sur T, $l \in \mathbb{R}$.

Le choix de tels tests peut être justifié de plusieurs façons :

1°) Pour tout $\mu \in M_c^+(T)$, $x \longmapsto [q_\mu(x)]^{1/2}$ définit une semi-norme sur $C(T)$ et bien que $C(T)$ soit métrisable, on ne connaît pas la loi de probabilité d'une v.a.r. de la forme $d(X,0)$ où d est une distance quelconque sur $C(T)$.

2°) On connaît la fonction caractéristique $\varphi^\theta_{Y_\mu}$ de la v.a.r. $Y_\mu = q_\mu(X)$ pour tout $\theta \in H(k)$, elle est donnée par la PROPOSITION 5 ; on a donc une expression de la fonction puissance de ces tests sur $H(k)$. En effet, d'après un théorème classique, sous réserve que $\int_{-\infty}^{+\infty} |\varphi^\theta_{Y_\mu}(\tau)|\,d\tau < \infty$, la fonction de répartition $F^\theta_{Y_\mu}$ est partout dérivable et vaut :

$$u \in \mathbb{R}_+ \; , \; F^\theta_{Y_\mu}(u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-i\tau u}-1}{-i\tau} \, \varphi^\theta_{Y_\mu}(\tau)\,d\tau$$

Par conséquent, la fonction puissance du test $\Phi_{\mu,l}$ se calcule comme :

$$\theta \in H(k) \; , \; \beta_{\Phi_{\mu,l}}(\theta) = 1 - F^\theta_{Y_\mu}(l^2)$$

$$= 1 - \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-i\tau l^2}-1}{-i\tau} \, \varphi^\theta_{Y_\mu}(\tau)\,d\tau \; .$$

3°) $E_\theta(Y_\mu) = \int_T [\theta^2(t) + k(t,t)]\, d\mu(t) = q_\mu(\theta) + \int_T k(t,t)\, d\mu(t)$. donc la statistique utilisée croît en moyenne comme $q_\mu(\theta)$.

c) <u>Etude de la Fonction puissance des tests quadratiques</u>.

1°) <u>Calcul</u>.

D'après ce qui précède, et compte tenu de la PROPOSITION 5, et avec ses notations, on a :

$$(2) \quad \theta \in H(k),\ \beta_{\Phi_{\mu,l}}(\theta) = 1 - \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-i\tau l^2} - 1}{-i\tau} \cdot \text{dét}(\mathbb{I} - 2i\tau B_\mu^k)^{-\frac{1}{2}} \cdot$$

$$\exp\{i\tau \langle \theta, (\mathbb{I} - 2i\tau B_\mu^k)^{-1} \circ B_\mu^k(\theta)\rangle_{H(k)_C}\} d\tau$$

et son niveau de signification est donné par :

$$\alpha = \beta_{\Phi_{\mu,l}}(0) = 1 - \int_{-\infty}^{+\infty} \frac{e^{-i\tau l^2} - 1}{-i\tau} \ \text{dét}(\mathbb{I} - 2i\tau B_\mu^k)^{-\frac{1}{2}} \, d\tau \ .$$

ce qui permet, au moins théoriquement, de calculer l en fonction de μ et de α .

2°) <u>Continuité</u>.

PROPOSITION 6. - <u>Pour tous</u> l <u>et</u> μ <u>fixés la fonctionnelle</u> : $\theta \longmapsto \beta_{\Phi_{\mu,l}}(\theta)$ <u>est continue sur</u> $H(k)$. <u>Pour tous</u> l <u>et</u> θ <u>fixés la fonctionnelle</u> $\mu \longmapsto \beta_{\Phi_{\mu,l}}(\theta)$ <u>est continue sur</u> $M_C^+(T)$ <u>muni de la topologie de la convergence vague</u>.

En effet, la continuité de la fonction puissance $\beta_{\Phi_{\mu,1}}(.)$ sur $H(k)$ est évidente. Pour la continuité de l'application $\mu \longmapsto \beta_{\Phi_{\mu,1}}(\theta)$, θ fixé, il suffit de remarquer que si une suite $\{\mu_n\}_{n\geq 1}$ de mesures converge vaguement vers la mesure μ dans $M_c^+(T)$, $\int_T x^2(t)\,d\mu_n(t) \underset{n\to\infty}{\longmapsto} \int_T x^2(t)\,d\mu(t)$ quel que soit $x \in C(T)$ ce qui implique que la suite de v.a.r. $\{q_{\mu_n}(.)\}_{n\geq 1}$ converge (presque) sûrement vers la v.a.r. $q_\mu(.)$ sur chaque espace probabilisé : $(C(T),\mathcal{B}(C(T))$, $N_{C(T)}(\theta,Q_k))$, $\theta \in H(k)$; elle converge aussi en loi donc $\varphi^{\theta}_{Y_{\mu_n}}(\tau) \underset{n\to\infty}{\longmapsto} \varphi^{\theta}_{Y_\mu}(t)$, $\forall\tau \in \mathbb{R}$ quel que soit $\theta \in H(k)$ et donc $\beta_{\phi_{\mu_n,1}}(\theta) \underset{n\to\infty}{\longmapsto} \beta_{\Phi_{\mu,1}}(\theta)$ quel que soit θ .

On pourrait aussi démontrer ce résultat en utilisant la continuité de l'application $\mu \longmapsto B^k_\mu$ de $M_c^+(T)$ dans $H(k)^{2\odot}$. ■

d) Propriétés des tests quadratiques.

PROPOSITION 7. - Pour tout réel α $(0\leq\alpha\leq 1)$, il existe au moins un test quadratique sans biais de niveau de signification α .

Il suffit de considérer le test $\phi_{\delta_{t_o},1}$ pour t_o quelconque fixé dans T ; ce test, qui s'écrit $\phi_{\delta_{t_o},1}(x) = \mathbb{1}_{\{|x(t_o)|\geq 1\}}$, $x \in C(T)$ est libre et sans biais pour tester l'hypothèse linéaire H_{t_o} : "$\theta(t_o) = 0$" contre l'hypothèse "$\theta(t_o)\neq 0$" ; en effet, il coïncide avec le test uniformément le plus puissant parmi les tests sans biais (U.M.P.B.) de l'hypothèse simple "$\theta(t_o) = 0$" contre l'hypothèse "$\theta(t_o)\neq 0$" sur la structure image de la structure (1) par la statistique réelle linéaire $x \longmapsto \langle x,\delta_{t_o}\rangle$, qui se trouve être la structure gaussienne scalaire :

$$(\mathbb{R},\mathcal{B}(\mathbb{R}),\ \{N(\theta(t_o),k(t_o,t_o))\ ;\ \theta(t_o) \in \mathbb{R}\}).$$

Par conséquent, pour tout $\theta \in H(k)$ la fonction puissance du test $\phi_{\delta_{t_o},1(t_o,\alpha)}$ de seuil α , vérifie : $\beta_{\phi_{\delta_{t_o},1(t_o,\alpha)}}(\theta) = \alpha$ si $\theta(t_o) = 0$

et $\beta_{\phi_{\delta_{t_o},1(t_o,\alpha)}}(\theta) > \alpha$ si $\theta(t_o) \neq 0$ comme le montre un calcul facile.

Il s'ensuit que le test $\phi_{\delta_{t_o},1(t_o,\alpha)}$ est <u>sans biais et de niveau de signification</u> α <u>pour tester</u> "$\theta = 0$" <u>contre</u> "$\theta \neq 0$" puisque pour tout $\theta \neq 0$ dans $H(k)$ on a $\beta_{\phi_{\delta_{t_o},1(t_o,\alpha)}}(\theta) \geq \alpha$ suivant que $\theta(t_o) \neq 0$ ou que $\theta(t_o) = 0$.

On ne sait pas si les tests quadratiques sont sans biais en général.

Un autre problème intéressant apparaît dans le choix optimal (suivant un critère à fixer) de la mesure μ dans $M_c^+(T)$, par exemple, on dira qu'un test ϕ_μ est uniformément le plus puissant parmi les tests quadratiques (U.M.P.Q.), à son niveau de signification $\alpha = \alpha(\phi_\mu)$ si pour toute mesure μ' telle que $\alpha(\phi_{\mu'}) \leq \alpha(\phi_\mu)$ on a :

$\beta_{\phi_{\mu'}}(\theta) \leq \beta_{\phi_\mu}(\theta)$, $\forall \theta \neq 0$.

Le lemme suivant montre que ce critère impose que le choix de la mesure μ doit tenir compte de k.

LEMME - <u>Une condition nécessaire pour qu'un test quadratique</u> ϕ_μ <u>soit U.M.P.Q. à son niveau de signification et que pour tout</u> $\theta \in H(k)$, $\mu\{t \in T : \theta(t) \neq 0\} > 0$.

En effet, s'il existait $\theta_o \in H(k)$ tel que : $\mu\{t \in T : \theta_o(t) \neq 0\} = 0$ on aurait $B_\mu^k(\theta_o) = 0$ et donc $\beta_{\phi_\mu}(\theta_o) = \beta_{\phi_\mu}(0) = \alpha(\phi_\mu) = \alpha$. Alors d'après la PROPOSITION 7 le test $\phi_{\delta_t,1(\alpha(\phi_\mu),t)}$ où $t \in \{t \in T : \theta_o(t) \neq 0\}$ serait de même niveau de signification, et sa puissance en θ_o strictement supérieure à α soit :

$$\alpha(\phi_{\delta_t}) = \alpha(\phi_\mu) \quad \text{et} \quad \beta_{\phi_{\delta_t}}(\theta_o) > \beta_{\phi_\mu}(\theta_o) = \alpha$$

donc ϕ_μ ne serait pas U.M.P.Q. Remarquons que la condition implique que : $\forall s \in T$, $\mu\{t \in T : k(s,t) = 0\} = 0$ en particulier, $\mu\{t \in T : k(t,t) = 0\} = 0$.

Remarquons enfin que l'on peut espérer des résultats plus significatifs dans ce domaine en supposant T compact ce qui d'ailleurs est plus compatible avec les conditions d'observation d'une f.a.r. . Il est possible, dans le cas où T est un intervalle borné de la droite que les meilleurs tests quadratiques s'obtiennent à partir de mesures diffuses sur T . ∎

REMARQUE 3 . - On peut envisager d'utiliser un critère quadratique pour tester l'égalité des moyennes de deux processus gaussiens indépendants X et X' de même fonction de covariance connue, au vu d'observations x et x' respectives de leurs trajectoires. En effet, tout test de la forme

$$\Phi_{\mu,1}(x,x') = \begin{cases} 1 & \text{si} \quad (\int_T [x(t) - x'(t)]^2 \, d\mu(t))^{1/2} > 1 \\ 0 & \text{si} \quad (\int_T [x(t) - x'(t)]^2 \, d\mu(t))^{1/2} < 1 \end{cases} \quad ; \; x, x' \in C(T)$$

est un test libre de "$\theta = \theta'$" contre "$\theta \neq \theta'$" dans la structure statistique produit :

$$(C(T), \mathcal{B}(C(T)), \{N_{C(T)}(\theta, Q_k) \; ; \; \theta \in H(k)\}) \; . \; (C(T), \mathcal{B}(C(T)), \{N_{C(T)}(\theta', Q_k) ; \theta' \in H(k)\})$$

puisque $X - X' \rightsquigarrow N_{C(T)} \; (\theta - \theta', Q_{2k})$. ∎

§3. - LA LOI MULTINOMIALE.

1. Définition et propriétés.

Soit $(\Omega, \mathcal{F})$ un espace mesurable quelconque. Notons E l'espace vectoriel des mesures bornées à support fini sur Ω et $\mathcal{A}$ la plus grande tribu sur E rendant mesurable l'application $T : \omega \longmapsto \delta_\omega$, de $(\Omega, \mathcal{F})$ dans E, où δ_ω désigne la mesure de DIRAC au point $\omega \in \Omega$.

LEMME 1. - Il existe un isomorphisme entre le dual de l'E.V.M. $(E, \mathcal{A})$ et l'espace vectoriel $\mathcal{L}_o(\Omega, \mathcal{F})$ de toutes les fonctions réelles mesurables définies sur $(\Omega, \mathcal{F})$.

En effet, soit ψ l'application de $\mathcal{L}_o(\Omega, \mathcal{F})$ dans $E^* = \mathbb{R}^E$ définie par :

$$f \in \mathcal{L}_o(\Omega, \mathcal{F}) \ , \ \mu = \sum_{j=1}^{p} \lambda_j \ \delta_{\omega_j} \in E \ , \ \psi(f)(\mu) = \sum_{j=1}^{p} \lambda_j \ f(\omega_j) \ ;$$

- ψ est linéaire et à valeurs dans $(E, \mathcal{A})^m \subset \mathbb{R}^E$, puisque $\psi(f)(\mu) = \langle \mu, \psi(f) \rangle$ est la valeur en $(\mu, \psi(f))$ de la forme bilinéaire canonique sur $E \times E^*$ et que $\psi(f)$ est $\mathcal{A}$-mesurable ; en effet, $f = \psi(f) \circ T$ est $\mathcal{F}$-mesurable et $\mathcal{A}$ est la plus grande tribu rendant T mesurable.

- ψ est injective, puisque $\psi(f) = 0 \Rightarrow \psi(f)(\delta_\omega) = 0$ quelque soit $\omega \in \Omega$ donc $f = 0$.

- ψ est surjective, puisque pour tout $u \in (E, \mathcal{A})^m$, il existe $f \in \mathcal{L}_o(\Omega, \mathcal{F})$ telle que $\psi(f) = u$. (Il suffit de prendre $f = u \circ T$, alors $\psi(u \circ T)(\mu) = \sum_{j=1}^{p} \lambda_j (u \circ T)(\omega_j) = \sum_{j=1}^{p} \lambda_j \ u(\delta_{\omega_j}) = u(\sum_{j=1}^{p} \lambda_j \ \delta_{\omega_j}) = u(\mu)$ dès que $\mu = \sum_{j=1}^{p} \lambda_j \ \delta_{\omega_j}$).

ψ est donc bien une application linéaire bijective, qui permet donc d'identifier le dual $(E, \mathcal{A})^m$ à $\mathcal{L}_o(\Omega, \mathcal{F})$. ∎

La forme bilinéaire canonique sur $E \times (E,\mathcal{a})^m$ étant, compte tenu de cette identification, définie par :

$$\mu = \sum_{j=1}^{p} \lambda_j \delta_{\omega_j} \in E \ , \ f \in \mathcal{L}_o(\Omega,\mathcal{F}) \equiv (E,\mathcal{a})^m \ ; \ \langle \mu, f \rangle = \sum_{j=1}^{p} \lambda_j \ f(\omega_j) \ .$$

La dualité entre E et $(E,\mathcal{a})^m$ est séparante et donc $(E,\mathcal{a})$ est un E.V.M.S.

Soit maintenant P une probabilité sur $(\Omega,\mathcal{F})$.

LEMME 2. - <u>Quel que soit l'entier positif</u> n , <u>la fonctionnelle</u> φ <u>définie sur</u> $(E,\mathcal{a})^m$ <u>par</u> : $f \in \mathcal{L}_o(\Omega,\mathcal{F}) \equiv (E,\mathcal{a})^m$, $\varphi(f) = [\int_\Omega e^{if(\omega)} dP(\omega)]^n$ <u>est la transformée de FOURIER d'une probabilité unique sur l'E.V.M.S.</u> $(E,\mathcal{a})$.

Il suffit de remarquer, en effet, que φ est la fonction caractéristique du vecteur aléatoire X à valeurs dans $(E,\mathcal{a})$ et défini sur l'espace probabilisé produit $(\Omega^n, \overset{n}{\underset{1}{\otimes}}\mathcal{F}, P^{n\otimes})$ par :
$X(\omega_1,\ldots,\omega_n) = \sum_{i=1}^{n} T(\omega_i)$, $(\omega_1,\ldots,\omega_n) \in \Omega^n$, c'est-à-dire, la transformée de FOURIER de la loi de probabilité de X ; de fait, pour $n = 1$,

$$E(e^{i\langle T,f\rangle}) = \varphi_{P \circ T^{-1}}(f) = \int_\Omega e^{if(\omega)} dP(\omega) \ , \ f \in \mathcal{L}_o(\Omega,\mathcal{F}) \ . \ \blacksquare$$

DEFINITION 1. - <u>On appelle</u> "<u>loi multinomiale de paramètres l'entier</u> n , <u>et la probabilité</u> P <u>sur</u> $(\Omega,\mathcal{F})$" , <u>la loi de probabilité du vecteur aléatoire</u> X <u>dans</u> $(E,\mathcal{a})$, <u>ainsi défini</u>, <u>et on la note</u> : $\mathcal{M}_{(\Omega,\mathcal{F})}(P;n)$.

Si Ω est fini, de cardinal $k \geq 1$,
$\mathcal{F} = \mathfrak{P}(\Omega)$, $P = \{p_1,\ldots,p_k\}$, $\sum_{j=1}^{k} p_j = 1$ on peut identifier X à un vecteur aléatoire dans $\mathbb{R}^k$ de loi multinomiale classique $\mathcal{M}(p_1,\ldots,p_k;n)$, ce qui justifie la généralisation.

PROPOSITION 1. - <u>Si</u> X $\mathcal{M}_{(\Omega,\mathcal{F})}(P;n)$, <u>pour toute partition mesurable</u> $\{F_1,\dots,F_k\}$ <u>de</u> Ω <u>le vecteur aléatoire</u> : $(\langle X,1_{F_1}\rangle,\dots,\langle X,1_{F_k}\rangle)$ <u>suit la loi multinomiale dans</u> $\mathbb{R}^k$: $\mathcal{M}(P(F_1),\dots,P(F_k);n)$.

En effet, soit $t=(t_1,\dots,t_k)\in\mathbb{R}^k$, alors,

$$\varphi_{\langle X,1_{F_1}\rangle,\dots,\langle X,1_{F_k}\rangle}(t_1,\dots,t_k) = \varphi_X(\sum_{j=1}^k t_j\, 1_{A_j}) = (\sum_{j=1}^k P(A_j)\, e^{it_j})^n \ . \blacksquare$$

REMARQUE 1. - D'après cette proposition, pour tout $F\in\mathcal{F}$, la variable aléatoire $\langle X,1_F\rangle$ suit la loi binomiale : $\mathcal{B}(P(F)\ ;\ n)$ et donc $E(\langle X,1_F\rangle) = n\,P(F)$ ce qui par linéarité implique que pour toute fonction $f\in\mathcal{L}_o(\Omega,\mathcal{F})$, $E(\langle X,f\rangle) = n(\int_\Omega f\,dP)$ dès que f est P-intégrable ; par conséquent, compte tenu de la DEFINITION 2 du CHAPITRE II, §4 , X n'est pas scalairement intégrable relativement à la dualité entre E et $(E,\mathcal{a})^m = \mathcal{L}_o(\Omega,\mathcal{F})$. Pour obtenir l'intégrabilité, il suffirait de considérer X comme un vecteur aléatoire à valeurs dans l'espace vectoriel $\mathcal{M}(\Omega,\mathcal{F})$ des mesures bornées sur $(\Omega,\mathcal{F})$ relativement à la dualité entre cet espace et l'espace vectoriel $\mathcal{L}_\infty(\Omega,\mathcal{F})$ des fonctions numériques mesurables bornées sur $(\Omega,\mathcal{F})$ alors, comme on l'a vu à l'EXEMPLE 3, CHAPITRE I, §2, la tribu faible $\mathcal{C}(\mathcal{M}(\Omega,\mathcal{F}),\mathcal{L}_\infty(\Omega,\mathcal{F}))$ est engendrée par les applications $\mu\longmapsto\mu(F) = \langle\mu,1_F\rangle$, $F\in\mathcal{F}$. La restriction à $\mathcal{L}_\infty(\Omega,\mathcal{F})$ de la fonctionnelle φ du LEMME 2 caractérise entièrement la loi de probabilité de X qui dans ce cas est un vecteur aléatoire scalairement intégrable et même intégrable puisque la formule ci-dessus montre que $E(X) = nP\in\mathcal{M}(\Omega,\mathcal{F})$. $\blacksquare$

A titre d'illustration du §2, du CHAPITRE II, nous pouvons caractériser le dual de l'E.V.P. $(E,\mathcal{a},\mathcal{M}_{(\Omega,\mathcal{F})}(P;n))$:

PROPOSITION 2. - <u>Il existe un isomorphisme entre le dual</u> $L(E,\mathcal{a},\mathcal{M}_{(\Omega,F)}(P;n))$ <u>et l'espace vectoriel</u> $L_o(\Omega,\mathcal{F},P)$ <u>des(classes d'équivalence de) v.a.r. sur</u> $(\Omega,\mathcal{F},P)$.

En effet, il suffit de supposer $n = 1$; nous savons que $(E,\mathcal{A})^m \subset \mathcal{L}(E,\mathcal{A},P\circ T^{-1})$ et d'après le LEMME 1, que $(E,\mathcal{A})^m$ est isomorphe à $\mathcal{L}_o(\Omega,\mathcal{F})$. Comme il existe une correspondance biunivoque entre les sous-ensembles mesurables de P-probabilité 1 de Ω et les sous-espaces vectoriels de E de $P\circ T^{-1}$ - probabilité 1, on peut donc établir une bijection entre l'ensemble des fonctions numériques mesurables définies P-presque partout sur Ω et l'ensemble $\mathcal{L}(E,\mathcal{A},P\circ T^{-1})$; mais comme on peut prolonger arbitrairement de telles fonctions à l'ensemble Ω tout entier sans changer leur classe d'équivalence, on peut en déduire un isomorphisme $\tilde{\psi}$ de l'espace vectoriel $L_o(\Omega,\mathcal{F},P)$ des classes d'équivalence de v.a.r. sur l'espace probabilisé $(\Omega,\mathcal{F},P)$ sur l'espace $L(E,\mathcal{A},P\circ T^{-1})$ des classes d'équivalence de v.a.r.l. sur l'E.V.P. $(E,\mathcal{A},P\circ T^{-1})$. Il est défini par $\tilde{\psi}^{-1}(\xi) = j(\xi) = \xi\circ T$ pour $\xi \in L(E,\mathcal{A},P\circ T^{-1})$. ∎

REMARQUE 2. - Dans cet exemple, $(E,\mathcal{A})^m/P\circ T^{-1} = L(E,\mathcal{A},P\circ T^{-1})$.

La transformée de FOURIER Φ de l'E.V.P. $(E,\mathcal{A},\mathcal{M}_{(\Omega,\mathcal{F})}(P;n))$ se déduit facilement du LEMME 2 et de la PROPOSITION 2, elle est définie par :

$$\xi \in L(E,\mathcal{A},\mathcal{M}_{(\Omega,\mathcal{A})}(P;n)) \equiv L_o(\Omega,\mathcal{F},P) \ , \ \Phi(\xi) = \left(\int_\Omega e^{i\xi}\,dP\right)^n .$$

2. Application.

Nous remarquons que : $\mathbb{P}_n = \frac{1}{n} X$ n'est autre que la "probabilité empirique d'ordre n associée à P" qui est à la base de l'étude statistique des échantillons empiriques.

Si l'on considère le vecteur X comme un vecteur aléatoire dans $\mathcal{M}(\Omega,\mathcal{F})$ relativement à la dualité entre $\mathcal{M}(\Omega,\mathcal{F})$ et $\mathcal{L}_\infty(\Omega,\mathcal{F})$, la REMARQUE 1 exprime que $\mathbb{P}_n$ est un estimateur sans biais de P sur la structure statistique d'échantillon la plus générale : $(\Omega,\mathcal{F},\mathcal{P})^n$, $\mathcal{P}$ désignant la famille de toutes les lois de probabilité sur $(\Omega,\mathcal{F})$. Lorsque $\Omega = \mathbf{R}$ ou même $\mathbf{R}^k$, $k > 1$, les théorèmes concernant la convergence de la fonction de répartition(empirique) F_n de $\mathbb{P}_n$ vers la fonction de répartition F de P sont anciens et extrêmement classiques de même que les théorèmes concernant la limite du processus stochastique $\sqrt{n}(F_n(x) - F(x))_{x\in\mathbb{R}}$ dans le cas réel par exemple, comme processus gaussien.

Dans le cas où Ω est quelconque, on peut cependant étudier la convergence en loi du vecteur aléatoire $D_n = \sqrt{n}(\mathbb{P}_n - P)$ dans $(\mathcal{M}(\Omega,\mathcal{F}), \mathcal{C}(\mathcal{M}(\Omega,\mathcal{F}), \mathcal{L}_\infty(\Omega,\mathcal{F})))$ quand $n \to \infty$. :

En effet, pour tout $f \in \mathcal{L}_\infty(\Omega,\mathcal{F})$, $\varphi_{\sqrt{n}(\mathbb{P}_n - P)}(f) = e^{-i\langle P,f\rangle\sqrt{n}}[E(e^{i\frac{f}{\sqrt{n}}})]^n$,

et en faisant un développement limité des exponentielles on constate que pour tout $f \in \mathcal{L}_\infty(\Omega,\mathcal{F})$,

$$\varphi_{\sqrt{n}(\mathbb{P}_n - P)}(f) \underset{n\to\infty}{\longmapsto} \exp\left\{-\frac{1}{2}[\langle f^2,P\rangle - (\langle f,P\rangle^2)]\right\}$$

où $\langle f^2,P\rangle - (\langle f,P\rangle^2) = \int_\Omega f^2\,dP - (\int_\Omega f\,dP)^2 = \sigma_P^2(f)$ est une forme quadratique positive sur $\mathcal{L}_\infty(\Omega,\mathcal{F})$.

Ce qui montre que la fonction caractéristique de D_n converge en tout point vers la transformée de FOURIER d'une probabilité cylindrique gaussienne centrée sur $\mathcal{M}(\Omega,\mathcal{F})$ relativement à la dualité entre $\mathcal{M}(\Omega,\mathcal{F})$ et $\mathcal{L}_\infty(\Omega,\mathcal{F})$, qui, d'après le THEOREME 2, CHAPITRE II, §1, détermine une probabilité gaussienne centrée unique sur l'E.V.M.S. $(\mathcal{L}_\infty^*(\Omega,\mathcal{F}), \mathcal{C}(\mathcal{L}_\infty^*, \mathcal{L}_\infty))$. <u>On peut donc en conclure que le vecteur aléatoire</u> D_n <u>converge en loi cylindri lorsque</u> $n \to \infty$, <u>vers un vecteur aléatoire gaussien centré</u> D, <u>à valeurs dans le dual algébrique</u> $\mathcal{L}_\infty^*(\Omega,\mathcal{F})$ <u>de</u> $\mathcal{L}_\infty(\Omega,\mathcal{F})$ <u>relativement à la dualité canonique entre ces deux espaces</u>, <u>de variance la forme quadratique positive</u> σ_P^2.

Il serait alors intéressant de trouver un sous-espace vectoriel $\tilde{E}$ de $\mathcal{L}_\infty^*(\Omega,\mathcal{F})$, le plus petit possible, contenant presque sûrement D ; mais ceci est un problème ouvert. Cependant, dans le but d'utiliser les théorèmes modernes concernant le support d'une probabilité gaussienne, on peut avoir intérêt à étudier la f.a.r. gaussienne sur $\mathcal{F}$: $(D(F))_{F \in \mathcal{F}}$ dont la fonction de covariance est le noyau symétrique de type positif k_P défini sur F par : $k_P(F,G) = \frac{1}{2}[\sigma_P^2(1_F + 1_G) - \sigma_P^2(1_F) - \sigma_P^2(1_G)]$

$$= P(F \cap G) - P(F)\ P(G)\ .$$

Montrons que l'espace de HILBERT autoreproduisant associé à k_P est le sous-espace vectoriel $H(k_P)$ des fonctions d'ensembles sur $(\Omega,\mathcal{F})$ constitué des mesures μ absolument continues par rapport à P telles que $\mu = f.P$, $f \in L^2(\Omega,\mathcal{F},P)$ et $E_P(f) = 0$, muni du produit scalaire :

$$\langle \mu_1, \mu_2 \rangle_{H(k_P)} = \langle f_1, f_2 \rangle_{L_2(\Omega,\mathcal{F},P)} ; \mu_i = f_i.P\ ,\ i = 1,2\ .$$

En effet, pour tout $F \in \mathcal{F}$, $k_P(F,.) \in H(k_P)$ puisque $k_P(F,.) \ll P$ et $k_P(F,.) = [1_F - P(F)].P$; de plus, si $\mu = f.P \in H(k_P)$ on a $\langle \mu, k_P(F,.)\rangle_{H(k_P)} = \langle f, 1_F - P(F)\rangle_{L_2(\Omega,\mathcal{F},P)} = \mu(F)$ ce qui, d'après l'unicité de l'espace autoreproduisant, montre que $H(k_P)$ est bien celui qui est associé au noyau k_P . ∎

§4. - LA LOI DE POISSON.

Nous définissons ici une loi de probabilité sur un espace vectoriel de mesures qui se présente comme une généralisation de la loi de POISSON. Tout en fournissant une illustration du CHAPITRE II, les hypothèses simplificatrices que nous faisons, font que cette loi n'est pas la plus générale possible, qui puisse servir à l'étude de tous les processus ponctuels de POISSON (cf. [24]), mais elle correspond au processus réellement observé dans la plupart des cas pratiques.

Soit $(T,\mathcal{T})$ un espace mesurable quelconque. Comme au § précédent, considérons la dualité séparante entre l'espace vectoriel $\mathcal{M}(T,\mathcal{T})$ des mesures bornées sur $(T,\mathcal{T})$ et l'espace vectoriel $\mathcal{L}_\infty(T,\mathcal{T})$ des fonctions numériques mesurables bornées sur $(T,\mathcal{T})$ définie par la forme bilinéaire : $\langle\mu,f\rangle = \int_T f\,d\mu$, $\mu \in \mathcal{M}(T,\mathcal{T})$, $f \in \mathcal{L}_\infty(T,\mathcal{T})$. La tribu faible : $C(\mathcal{M}(T,\mathcal{T}),\mathcal{L}_\infty(T,\mathcal{T}))$ est engendrée par les formes linéaires $\mu \longmapsto \langle\mu,1_A\rangle = \mu(A)$ lorsque A parcourt $\mathcal{T}$. (cf. EXEMPLE 3, CHAPITRE I, §2).

THEOREME 1. - <u>Pour toute mesure bornée positive</u> λ <u>fixée dans</u> $\mathcal{M}(T,\mathcal{T})$, <u>la fonctionnelle complexe</u> φ_λ <u>définie sur</u> $\mathcal{L}_\infty(T,\mathcal{T})$ <u>par</u> :

$$f \in \mathcal{L}_\infty(T,\mathcal{T}) \ , \ \varphi_\lambda(f) = \exp\{\int_T (e^{if(t)} - 1)\, d\lambda(t)\}$$

<u>est la transformée de FOURIER d'une probabilité unique</u> $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$ sur $(\mathcal{M}(T,\mathcal{T}),C(\mathcal{M},\mathcal{L}_\infty))$ <u>appelée loi de POISSON sur</u> $\mathcal{M}(T,\mathcal{T})$ <u>de paramètre</u> "<u>d'intensité</u>" λ .

En effet, φ_λ a bien les propriétés d'une transformée de FOURIER et définit donc une probabilité cylindrique unique sur $\mathcal{M}(T,\mathcal{T})$ relativement à la dualité considérée. Il suffit de montrer qu'elle induit une probabilité $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$ sur la tribu cylindrique $C(\mathcal{M},\mathcal{L}_\infty)$. Pour cela, montrons qu'il existe un espace probabilisé $(\Omega,\mathcal{F},P_\lambda)$ et un vecteur aléatoire X dans $\mathcal{M}(T,\mathcal{T})$ relativement à cette dualité, défini sur cet espace probabilisé, dont φ_λ soit la fonction caractéristique (DEFINITION 1, CHAPITRE II, §4).

Prenons pour Ω , l'ensemble des suites finies ordonnées d'éléments de T et notons ω_o la suite vide. Il est clair que pour tout $n \geq 0$, $T^n \subset \Omega$, en posant $T^o = \{\omega_o\}$.

Posons $\mathcal{F} = \{F \subset \Omega : F \cap T^n \in \mathcal{T}^{n\otimes}, \forall n \geq 0\}$ qui est bien une tribu de parties de Ω . Enfin définissons la probabilité P_λ sur $(\Omega,\mathcal{F})$ par :

$$F \in \mathcal{F} \ , \ P_\lambda(F) = e^{-\lambda(T)} \sum_{n\geq 0} \frac{\lambda^{n\otimes}(F\cap T^n)}{n!} \quad \text{avec} \quad \lambda^{0\otimes} = \delta_{\omega_o} \ .$$

Si $\omega \in \Omega$ et $\omega \neq \omega_o$, il existe donc une suite finie de points de T soit : $\{t_1(\omega),\ldots,t_{n(\omega)}(\omega)\}$ pour un $n(\omega)$ unique. Posons alors pour un tel $\omega \in \Omega$: $X(\omega) = \sum_{j=1}^{n(\omega)} \delta_{t_j(\omega)}$ et $X(\omega_o) = 0$ et montrons que l'application X de Ω dans $\mathcal{M}(T,\mathcal{T})$ est $\mathcal{F} - \mathcal{C}(\mathcal{M},\mathcal{L}_\infty)$-mesurable ; pour cela il suffit de montrer que pour tout $A \in \mathcal{T}$, l'application $\omega \longrightarrow \langle X(\omega),1_A\rangle$ est une v.a.r..
Mais, $\langle X(\omega),1_A\rangle = \int_A d[\sum_{j=1}^{n(\omega)} \delta_{t_j(\omega)}] =$ nombre de $t_j(\omega)$ dans $A = N_A(\omega)$
or pour tous entiers k et $n \geq 0$,

$$N_A^{-1}(\{k\}) \cap T^n = \begin{cases} \emptyset & \text{si } n < k \ , \\ A^k(A^c)^{n-k} & \text{sinon ;} \end{cases}$$

c'est donc un ensemble $\mathcal{F}$-mesurable et donc $N_A(\omega)$ est une v.a.r. sur $(\Omega,\mathcal{F})$.

Enfin, calculons la fonction caractéristique de X :

$$f \in \mathcal{L}_\infty(T,\mathcal{T}) \ , \ \varphi_X(f) = E(e^{i\langle X,f\rangle}) = \int_\Omega e^{i\langle X(\omega),f\rangle} \, dP_\lambda$$

comme $\langle X(\omega),f\rangle = \sum_{j=1}^{n(\omega)} f(t_j(\omega))$, cette intégrale vaut, d'après la définition de P_λ :

$$\varphi_X(f) = \sum_{n\geq 0} e^{-\lambda(T)} \frac{1}{n!} \int_{T^n} e^{i\sum_{j=1}^{n} f(t_j)} \, d\lambda^{n\otimes}(t_1,\ldots,t_n)$$

$$= e^{-\lambda(T)} \sum_{n\geq 0} \frac{1}{n!} [\int_T e^{if(t)} \, d\lambda(t)]^n$$

$$\varphi_X(f) = e^{-\lambda(T)}\, e^{\int_T e^{if} d\lambda} = \exp \int_T (e^{if} - 1)\, d\lambda = \varphi_\lambda \ .$$

On a donc : $\mathfrak{P}_{(T,\mathcal{T})}(\lambda) = P_\lambda \circ X^{-1}$. on dira que X suit la loi de POISSON de paramètre λ . ∎

THEOREME 2. - <u>Pour tout</u> $A \in \mathcal{T}$, <u>la v.a.r.</u> $\langle X, 1_A \rangle$ <u>suit la loi de POISSON de paramètre</u> $\lambda(A)$ - <u>Si</u> $f_1, \ldots, f_n$ <u>sont des fonctions appartenant à</u> $\mathcal{L}_\infty(T,\mathcal{T})$ <u>de supports disjoints, les v.a.r.</u> $\langle X, f_1 \rangle, \ldots, \langle X, f_n \rangle$ <u>sont indépendantes. Le vecteur</u> X <u>est scalairement intégrable et son espérance mathématique est</u> $E(X) = \lambda$.

Calculons la fonction caractéristique de la v.a.r. $\langle X, 1_A \rangle$:

$$u \in \mathbb{R} \ , \ \varphi_{\langle X, 1_A \rangle}(u) = E(e^{iu\langle X, 1_A \rangle}) = \varphi_X(u.1_A)$$

$$= \exp \int_T (e^{iu\,1_A} - 1)\, d\lambda$$

$$= e^{\lambda(A)(e^{iu} - 1)}$$

C'est bien la fonction caractéristique d'une v.a.r. de loi de POISSON de paramètre $\lambda(A)$.

La fonction caractéristique du vecteur aléatoire $Z = (\langle X, f_1 \rangle, \ldots, \langle X, f_n \rangle)$ est $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$, $\varphi_Z(u) = E(e^{i\langle u, Z \rangle}) = E(e^{i\langle X, \sum_{j=1}^n u_j f_j \rangle})$

$$= \exp \int_T (e^{i \sum_{j=1}^n u_j f_j} - 1)\, d\lambda$$

or, si les f_j ont des supports disjoints, $e^{i\sum_{j=1}^n u_j f_j} = \sum_{j=1}^n e^{iu_j f_j}$ sur T ,

donc $\varphi_Z(u) = \prod_{j=1}^n \varphi_{\langle X, f_j \rangle}(u_j)$, d'où l'indépendance des composantes de Z .

Enfin, pour tout $f \in \mathcal{L}_\infty(T,\mathcal{T})$, $\dfrac{d}{du} \varphi_{\langle X, f \rangle}(u)\Big|_{u=0} = i \int_T f\, d\lambda$

donc $E(\langle X, f \rangle) = \int_T f\, d\lambda = \langle \lambda, f \rangle = \langle E(X), f \rangle$, si l'on prend $E(X) = \lambda \in \mathcal{M}(T,\mathcal{T})$, ce qui démontre la dernière assertion. ∎

REMARQUE 1. - La loi de POISSON est reliée à la loi Multinomiale décrite au § précédent par la propriété suivante qui généralise un théorème bien connu concernant les processus de POISSON (cf. [24] pour une démonstration).

"Soit X un vecteur aléatoire de loi $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$. Pour tout ensemble $A \in \mathcal{T}$ tel que $\lambda(A) > 0$ la loi de probabilité du vecteur aléatoire : $X_A = 1_A.X$, conditionnelle à la v.a.r. $\langle X,1_A\rangle$ est la loi Multinomiale : $\mathcal{M}_{(A,A\cap\mathcal{T})}(\frac{1_A.\lambda}{\lambda(A)} ; \langle X,1_A\rangle)$."

REMARQUE 2. - Il peut être utile (comme on le verra au chapitre suivant) de présenter la loi de POISSON de façon analogue à celle utilisée pour la loi multinomiale.

Soient $(T,\mathcal{T})$ un espace mesurable quelconque, $(\Omega,\mathcal{F})$ l'espace mesurable "somme directe" des espaces mesurables $(T^n,\mathcal{T}^{n\otimes})$ $n \geq 0$ déjà défini et X l'application définie sur Ω par : $X(\omega) = \sum_{j=1}^{n(\omega)} \delta_{t_j(\omega)}$, $X(\omega_o) = 0$.

Soit E l'espace vectoriel engendré par la famille de mesures $\{\sum_{j=1}^{n(\omega)} \delta_{t_j}(\omega)\}_{\omega\in\Omega}$ et $\mathcal{A}$ la plus grande tribu sur E rendant mesurable l'application X de $(\Omega,\mathcal{F})$ dans E.

LEMME 1. - Il existe un isomorphisme entre le dual de l'E.V.M. $(E,\mathcal{A})$ et l'espace vectoriel $\mathcal{L}_o(T,\mathcal{T})$ de toutes les fonctions mesurables réelles définies sur $(T,\mathcal{T})$.

En effet, on utilise une démonstration analogue à celle du LEMME 1, §3, qui consiste à montrer que l'application ψ définie sur $\mathcal{L}_o(T,\mathcal{T})$ par :

$$f \in \mathcal{L}_o(T,\mathcal{T}) \;,\; \psi(f)\,(\sum_{j=1}^{n(\omega)} \delta_{t_j(\omega)}) = \sum_{j=1}^{n(\omega)} f(t_j(\omega)) = \psi(f)(\mu) \;;\; \mu \in E$$

est une application linéaire, bijective de $\mathcal{L}_o(T,\mathcal{T})$ sur $(E,\mathcal{A})^m$.

THEOREME 3. - <u>Pour toute mesure positive bornée</u> λ <u>sur</u> $(T,\mathcal{T})$, <u>la fonctionnelle</u> φ_λ <u>définie sur</u> $(E,\mathcal{A})^m$ (<u>identifié à</u> $\mathcal{L}_o(T,\mathcal{T})$) <u>par</u> :

$$f \in \mathcal{L}_o(T,\mathcal{T}) \ , \ \varphi_\lambda(f) = \exp\{\int_T (e^{if(t)}-1)\,d\lambda(t)\}$$

<u>est la transformée de FOURIER d'une probabilité unique</u> $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$ <u>sur l'E.V.M.S.</u> $(E,\mathcal{A})$, <u>appelée loi de POISSON d'intensité</u> λ <u>relative à l'espace mesurable</u> $(T,\mathcal{T})$.

La démonstration consiste à montrer que φ_λ est la fonction caractéristique du vecteur aléatoire X à valeurs dans $(E,\mathcal{A})$, défini sur l'espace probabilisé $(\Omega,\mathcal{F},P_\lambda)$, comme pour le THEOREME 1.

De la même façon qu'au §3, on peut maintenant caractériser le dual de l'E.V.P. $(E,\mathcal{A},\mathfrak{P}_{(T,\mathcal{T})}(\lambda))$:

PROPOSITION 1 . - <u>Il existe un isomorphisme</u> ψ <u>entre l'espace vectoriel</u> $L_o(T,\mathcal{T},\lambda)$ <u>des classes d'équivalence</u> (<u>par rapport à la mesure positive bornée</u> λ) <u>de fonctions mesurables réelles définies sur</u> $(T,\mathcal{T})$ <u>et le dual</u> $L(E,\mathcal{A},\mathfrak{P}_{(T,\mathcal{T})}(\lambda))$. <u>Compte tenu de cette identification</u>, <u>la transformée de FOURIER</u> Φ_λ <u>de l'E.V.P.</u> $(E,\mathcal{A},\mathfrak{P}_{(T,\mathcal{T})}(\lambda))$ <u>est définie par</u> :

$$\xi \in L(E,\mathcal{A},\mathfrak{P}_{(T,\mathcal{T})}(\lambda)) \equiv L_o(T,\mathcal{T},\lambda) \ , \ \Phi_\lambda(\xi) = \exp\{\int_T (e^{i\xi}-1)\,d\lambda\} \ .$$

Comme pour la PROPOSITION 2, §3, la démonstration consiste à utiliser le LEMME 1 et à remarquer qu'il y a une correspondance biunivoque entre les sous-espaces vectoriels mesurables de $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$-probabilité 1 et les sous-ensembles $\mathcal{F}$-mesurables de Ω de P_λ probabilité 1 , ceux-ci étant eux-mêmes en correspondance avec les sous-ensembles $\mathcal{T}$-mesurables de T de mesure égale à $\lambda(T)$. Cet isomorphisme est tel que : $\psi(\xi)(\delta_t) = \xi(t)$, $t \in T$, $\xi \in L_o(T,\mathcal{T},\lambda)$. ∎

Nous utiliserons en statistique la propriété d'absolue continuité suivante :

PROPOSITION 2 . - <u>Soient</u> λ_o <u>et</u> λ <u>deux mesures positives bornées sur</u> $(T,\mathcal{T})$. <u>Une condition nécessaire et suffisante pour que les lois de POISSON</u> $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$ <u>et</u> $\mathfrak{P}_{(T,\mathcal{T})}(\lambda_o)$ <u>soient équivalentes sur</u> $(E,\mathcal{A})$ <u>est que les mesures</u> λ <u>et</u> λ_o <u>le soient sur</u> $(T,\mathcal{T})$. <u>Dans ce cas, en posant</u> : $\xi_\lambda = \dfrac{d\lambda}{d\lambda_o}$ <u>on a</u> :

$$\frac{d\,\mathfrak{P}_{(T,\mathcal{T})}(\lambda)}{d\,\mathfrak{P}_{(T,\mathcal{T})}(\lambda)} = e^{\psi(\mathrm{Log}\,\xi_\lambda) - \int_T (\xi_\lambda - 1)\, d\lambda_o} ,$$

<u>où</u> ψ <u>est l'isomorphisme de la PROPOSITION 1 restreint à</u> $L_1(T,\mathcal{T},\lambda_o)$.

En effet, une C.N.S. pour que $\mathfrak{P}_{(T,\mathcal{T})}(\lambda)$ et $\mathfrak{P}_{(T,\mathcal{T})}(\lambda_o)$ soient équivalentes sur $(E,\mathcal{A})$ est que P_λ et P_{λ_o} le soient sur $(\Omega,\mathcal{F})$ ou encore que $\lambda^{n\otimes}$ et $\lambda_o^{n\otimes}$ le soient sur $(T^n,\mathcal{T}^{n\otimes})$ pour tout $n \geq 1$ ce qui revient à dire que les mesures λ et λ_o sont équivalentes sur $(T,\mathcal{T})$.

Pour tout $F \in \mathcal{F}$ on a :

$$P_\lambda(F) = \int_F dP_\lambda = e^{-\lambda(T)} \sum_{n\geq 1} \frac{1}{n!} \int_{F\cap T^n} d\lambda^{n\otimes}$$

soit $\xi_\lambda^{n\otimes} = \dfrac{d\lambda^{n\otimes}}{d\lambda_o^{n\otimes}}$, alors

$$P_\lambda(F) = e^{-\int_T \xi_\lambda\, d\lambda_o} \sum_{n\geq 1} \frac{1}{n!} \int_{F\cap T^n} \xi_\lambda^{n\otimes}\, d\lambda_o^{n\otimes}$$

$$= e^{-\int_T (\xi_\lambda - 1)\, d\lambda_o} \int_F \prod_{j=1}^{n(\omega)} \xi_\lambda(t_j(\omega))\, dP_{\lambda_o}(\omega) .$$

Comme $\lambda_o\{\xi_\lambda = 0\} = 0$, cette formule montre que

$$\frac{dP_\lambda}{dP_{\lambda_o}} = e^{-\int_T (\xi_\lambda - 1)\, d\lambda_o + \psi(\mathrm{Log}\,\xi_\lambda)\circ X} , \text{ p.p sur } \Omega ,$$

d'où la formule annoncée. ∎

REMARQUE 3. - La restriction à des lois de POISSON d'intensité bornée correspond généralement aux conditions d'observation, dans les problèmes de statistique concernant les processus ponctuels de POISSON ; en effet le temps (ou le champ) d'observation est en pratique une partie bornée de $\mathbf{R}$ (ou de $\mathbf{R}^n$) soit T, dont la mesure (pour l'intensité du processus) est finie. ∎

§ 5. - LA LOI DE DIRICHLET.

Dans [11] FERGUSON a construit une loi de probabilité sur l'ensemble des probabilités définies sur un espace mesurable quelconque. Cette loi de probabilité, comme celles des deux § précédents peut être interprétée comme une généralisation en dimension infinie d'une loi classique, la loi de DIRICHLET pour les vecteurs aléatoires ; elle est utilisée comme loi de probabilité à priori pour l'étude bayésienne de structures statistiques non paramétriques. Elle nous fournit un exemple supplémentaire très intéressant de loi de probabilité en dimension infinie. Les résultats suivants sont empruntés de [11] .

DEFINITION 1. - Soient $\alpha_1,\dots,\alpha_k$ des réels non négatifs. Un vecteur aléatoire $X = (X_1,\dots,X_k)$ dans $(\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k))$ suit la loi de DIRICHLET de paramètres $\alpha_1,\dots,\alpha_k$ notée : $\mathcal{D}(\alpha_1,\dots,\alpha_k)$ si et seulement si pour tout $j = 1,2,\dots,k$, X_j est une v.a.r. de la forme : $Z_j/\sum_{i=1}^{k} Z_i$ où $Z_1,\dots,Z_k$ sont des v.a.r. de loi $\gamma(\alpha_j,1)$ respectivement, de densité par rapport à la mesure de LEBESGUE $f_{\gamma(\alpha_j,1)}(x) = e^{-x}\, x^{\alpha_j - 1}\, 1\!\!1_{[0,\infty[}$ si $\alpha_j > 0$, ou identique à la mesure de DIRAC au point 0 si $\alpha_j = 0$.

Le vecteur aléatoire X n'a pas de densité par rapport à la mesure de LEBESGUE sur $(\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k))$ puisque $\sum_{j=1}^{k} X_j = 1$ mais si $\alpha_j > 0$ pour tous les $j = 1,\dots,k$ le vecteur aléatoire $(X_1,\dots,X_{k-1})$ est absolument continu par rapport à la mesure de LEBESGUE sur $(\mathbb{R}^{k-1},\mathcal{B}(\mathbb{R}^{k-1}))$ et a pour densité :

$$f(x_1,\dots x_{k-1}) = \frac{\Gamma(\alpha_1+\dots+\alpha_k)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_k)}\; x_1^{\alpha_1 - 1}\dots x_{k-1}^{\alpha_{k-1}-1}\,(1 - x_1 - \dots - x_{k-1})^{\alpha_k - 1}\, 1\!\!1_S$$

où S désigne le simplexe de $\mathbb{R}^{k-1}$: $\{x = (x_1\dots x_{k-1}) \in \mathbb{R}^{k-1} : x_j \geq 0, \sum_{j=1}^{k-1} x_j \leq 1\}$.

En particulier pour $k = 2$, X_1 suit la loi $\beta(\alpha_1,\alpha_2)$ sur l'intervalle $[0,1]$.

Toutes les lois marginales du vecteur X sont aussi des lois de DIRICHLET.

DEFINITION 2. - Soient $(\Omega,\mathcal{F})$ un espace mesurable quelconque et α une mesure positive bornée sur $(\Omega,\mathcal{F})$. Une f.a.r. $\{X(F)\}_{F\in\mathcal{F}}$ sur $\mathcal{F}$, est appelée "processus de DIRICHLET de paramètre α" , si pour toute partition mesurable $\{F_1,\ldots,F_k\}$, $k \geq 1$ de Ω , le vecteur aléatoire $(X(F_1),\ldots,X(F_k))$ suit la loi $\mathcal{D}(\alpha(F_1),\ldots,\alpha(F_k))$, dans $(\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k))$.

THEOREME 1. - (FERGUSON) Un processus de DIRICHLET de paramètre α , induit une loi de probabilité unique $\mathcal{D}_{(\Omega,\mathcal{F})}(\alpha)$ sur $(\mathbb{R}^{\mathcal{F}},\bigotimes_{\mathcal{F}}\mathcal{B}(\mathbb{R}))$ - c'est la loi de probabilité du vecteur aléatoire X à valeurs dans $\mathbb{R}^{\mathcal{F}}$ associé à la f.a.r. $\{X(F)\}_{F\in\mathcal{F}}$. - De plus, le support de cette probabilité lorsque $\mathbb{R}^{\mathcal{F}}$ est muni de la topologie de la convergence ponctuelle est contenu dans l'ensemble des probabilités sur $(\Omega,\mathcal{F})$, il contient l'ensemble des probabilités absolument continues par rapport à α sur $(\Omega,\mathcal{F})$. ∎

La première partie de ce théorème résulte du théorème de KOLMOGOROV dont les hypothèses de cohérence sont vérifiées grâce aux propriétés de la loi de DIRICHLET. La deuxième partie (cf.[11]) montre que le vecteur X de la loi de DIRICHLET généralisée : $\mathcal{D}_{(\Omega,\mathcal{F})}(\alpha)$ définit bien une probabilité aléatoire.

L'application de cette loi de probabilité à la statistique bayésienne, s'exprime dans le théorème suivant :

THEOREME 2. - (cf. [11]). Soit $(\Omega,\mathcal{F},\mathcal{P})^n$ une structure statistique d'échantillon où $\mathcal{P}$ désigne la famille de toutes lois de probabilité sur $(\Omega,\mathcal{F})$. Si l'on prend pour loi de probabilité à priori sur $\mathcal{P}$ la loi de DIRICHLET $\mathcal{D}_{(\Omega,\mathcal{F})}(\alpha)$, la loi de probabilité à postériori est la loi de DIRICHLET $\mathcal{D}_{(\Omega,\mathcal{F})}(\alpha + n\,\mathbb{P}_n)$ où $\mathbb{P}_n$ désigne la loi de probabilité empirique de l'échantillon. (cf. §3) .

CHAPITRE IV

STRUCTURES STATISTIQUES EXPONENTIELLES GENERALISEES.

On dit habituellement qu'une structure statistique est exponentielle si elle correspond à l'observation d'un vecteur aléatoire dans $\mathbb{R}^k$, dont la loi de probabilité appartient à une famille de lois $\{P_\theta\}_{\theta\in\Theta}$ absolument continues par rapport à l'une d'entre elles : P_{θ_o}, $\theta_o \in \Theta$ et si leur densité par rapport à cette loi admet une version de la forme :

$$\theta \in \Theta \ , \ x \in \mathbb{R}^k \ , \ \frac{d P_\theta}{d P_{\theta_o}}(x) = C(\theta) \ \exp\{\sum_{j=1}^{s} T_j(x) \ Q_j(\theta)\} \ . \ h(x) \ ; \ s \geq 1 \ .$$

Cette structure statistique admet un résumé exhaustif de l'observation x : $T(x) = \{T_1(x), \ldots, T_s(x)\}$ qui ramène l'étude du problème statistique à la structure image par la statistique T - dite exponentielle canonique - dans laquelle, les lois de probabilité, moyennant une reparamétrisation, admettent une densité par rapport à une mesure dominante m, de la forme :

$$(1) \qquad t \in \mathbb{R}^s \ , \ \frac{d P_{\tilde{\theta},T}}{d m}(t) = \tilde{C}(\tilde{\theta}) \ \exp\{\langle t, \tilde{\theta}\rangle\} \ .$$

La théorie de ces structures, qui sont un champ d'étude très riche, notamment pour leurs propriétés analytiques, et qui se trouvent être le modèle mathématique utilisé pour de nombreux problèmes de statistique paramétrique a été introduite par DARMOIS, KOOPMAN, PITMAN dans les années 1935-36, et abordée par de nombreux auteurs dont LEHMANN, LINNIK, CHENTZOV, etc. ; elle fait l'objet d'un exposé systématique et moderne dans [5] et connaît des développements récents dûs à O.BARNDORFF-NIELSEN [4] et ses élèves.

Il est naturel d'essayer d'étendre cette notion de structure statistique exponentielle, en dimension quelconque ; c'est-à-dire au cas où l'observation et le paramètre peuvent ne pas être de dimension finie. C'est ce que nous faisons dans ce chapitre, en justifiant le bien fondé de cette démarche par de nombreux exemples d'application et en essayant de plus, de dégager les propriétés intrinsèques de ces structures statistiques vectorielles. A cet effet, on fera très souvent référence aux chapitres précédents puisque les notions qui y sont étudiées sont toutes utilisées ici.

Le point de départ de cette généralisation consiste à remarquer que dans (1) , le paramètre $\tilde{\theta}$ peut être identifié à une forme linéaire de l'observation t , ou encore, que la densité (1) est à une constante près, une exponentielle d'une variable aléatoire réelle linéaire.

Il est aisé de construire artificiellement des structures exponentielles généralisées en dimension quelconque ; sommairement : il suffit de considérer une famille de probabilités $\{P_\theta ; \theta \in \Theta\}$, absolument continues par rapport à une probabilité de référence P_o sur un espace vectoriel E , le paramètre θ parcourant une partie Θ adéquate, contenant l'origine d'un espace vectoriel F mis en dualité séparante avec E par une forme bilinéaire ; ces probabilités sont définies sur la tribu faible $\mathcal{C}(E,F)$ de E , par leur densité par rapport à P_o , à laquelle on impose d'admettre une version de la forme (1) , en remplaçant le produit scalaire par la forme bilinéaire.

C'est ce que nous avons proposé en 1973 dans [34] . Il semble d'ailleurs que d'autres auteurs aient eu indépendamment cette idée, S. JOHANSEN [17] et aussi H.D. BRUNK dans un rapport technique en 1975, non publié, et cité en référence dans [17] .

Mais les limites d'un tel modèle, pour plaisant qu'il soit à étudier, apparaissent très rapidement dès que l'on s'intéresse à ses applications ; ainsi, par exemple, une structure statistique gaussienne à moyenne inconnue n'est pas de ce type en dimension infinie, à moins de restreindre arbitrairement le domaine de variation de la moyenne ! C'est ce qui nous a amené à utiliser la notion plus générale de v.a.r. linéaire du CHAPITRE II, §2, c'est-à-dire à faire varier le paramètre de la famille de probabilités dans le dual de l'espace vectoriel probabilisé de référence. Le modèle précédent s'insère alors naturellement, comme cas particulier de cette théorie générale ([36]) qui elle, répond aux besoins des applications.

§ 1. - DEFINITIONS ET GENERALITES.

1. Structures exponentielles canoniques.

Soit $(E,\mathcal{A},\mathcal{P})$ une structure statistique vectorielle , où $\mathcal{P}$ est une famille de lois de probabilité équivalentes sur l'E.V.M.S. $(E,\mathcal{A})$; il s'ensuit que le dual $L(E,\mathcal{A},P)$ de l'E.V.P. $(E,\mathcal{A},P)$ [resp. l'espace quotient $(E,\mathcal{A})^m/P$] est le même quel que soit $P \in \mathcal{P}$, on le notera donc $L(E,\mathcal{A},\mathcal{P})$ [resp. $(E,\mathcal{A})^m/\mathcal{P}$] et on l'appellera "dual de la structure statistique vectorielle $(E,\mathcal{A},\mathcal{P})$" , ses éléments étant des (classes de $\mathcal{P}$-équivalence de) "<u>statistiques réelles linéaires</u>" (s.r.l.) définies sur cette structure.

On peut supposer sans restreindre la généralité, que $\mathcal{A}$ coïncide avec la tribu linéaire $\mathcal{A}_L$ relativement à $\mathcal{P}$ (cf. CHAPITRE II, §2, DEFINITION 4) puisque le dual ne change pas.

De plus, puisque l'espace des (classes d'équivalence de) v.a.r. [resp. de plus, P-essentiellement bornées définies sur $(E,\mathcal{A},P)$ ne dépend pas de P dans $\mathcal{P}$, on le notera $L_o(E,\mathcal{A},\mathcal{P})$ [resp. $L_\infty(E,\mathcal{A},\mathcal{P})$] en l'assimilant à l'espace des (classes d'équivalence de) statistiques réelles [resp. de plus $\mathcal{P}$-essentiellement bornées] définies sur cette structure.

Enfin, on appellera "statistique vectorielle linéaire" tout vecteur aléatoire linéaire défini sur chaque E.V.P. $(E,\mathcal{A},P)$, $P \in \mathcal{P}$ comme au CHAPITRE II, §2, n°3.

DEFINITION 1. - <u>On dira que la structure statistique vectorielle</u> $(E,\mathcal{A},\mathcal{P})$ <u>est une "structure exponentielle canonique"</u>, <u>si pour une loi</u> $P^* \in \mathcal{P}$ <u>fixée, on a</u> :

$$\{\mathrm{Log}\ \frac{dP}{dP^*}\}_{P \in \mathcal{P}} \subset L(E,\mathcal{A},\mathcal{P}) \oplus \mathbb{R}\ .$$

Remarquons que la v.a.r. $\mathrm{Log}\ \dfrac{dP}{dP^*}$ est bien définie pour tout

$P \in \mathcal{P}$, puisque l'équivalence des lois de la famille $\mathcal{P}$ implique que $P^*\{\frac{dP}{dP^*} > 0\} = 1$.

De plus, si la relation de la définition est vérifiée pour une loi P^* , elle l'est aussi pour toute autre loi fixée dans $\mathcal{P}$, ce qui montre que cette caractérisation ne dépend pas du choix de P^* .

Cette définition implique que pour tout $P \in \mathcal{P}$, il existe un élément unique $\xi_P \in L(E,\mathcal{a},\mathcal{P})$ (une s.r.l.), et un réel k_P tels que : $\frac{dP}{dP^*} = \exp\{\xi_P + k_P\}$, le couple (ξ_P, k_P) étant lié par la relation : $\int_E e^{\xi_P + k_P}\, dP^* = 1$ - Ce qui permet d'identifier la famille $\{\mathrm{Log}\frac{dP}{dP^*}\}_{P\in\mathcal{P}}$ a une variété dans l'espace vectoriel $L_o(E,\mathcal{a},\mathcal{P})$ -

On a donc : $k_P = -\mathrm{Log}\int_E e^{\xi_P}\, dP^* = -\mathrm{Log}\, L_{P^*}(\xi_P) = -C_{P^*}(\xi_P)$.

où L_{P^*} et C_{P^*} sont respectivement la transformée de LAPLACE réelle et la fonction cumulante de l'E.V.P. $(E,\mathcal{a},P^*)$.

Soit u l'application : $P \longmapsto \xi_P$ de $\mathcal{P}$ dans $L(E,\mathcal{a},\mathcal{P})$; elle est injective puisque : $P \longmapsto \mathrm{Log}\frac{dP}{dP^*}$ l'est, et $u(P^*) = 0$. Posons $\Theta = u(\mathcal{P})$; u est une bijection de $\mathcal{P}$ sur Θ et pour $\theta \in \Theta$, posons $P_\theta = u^{-1}(\theta)$. On peut donc paramétrer la famille $\mathcal{P}$ par θ , de sorte que la DEFINITION 1 est équivalente à la suivante :

DEFINITION 2. - <u>Une structure statistique vectorielle</u> $(E,\mathcal{a},\mathcal{P} = \{P_\theta;\theta\in\Theta\})$ <u>est dite exponentielle canonique si</u> :

i) - Θ <u>est un sous-ensemble, contenant l'origine, du convexe</u> $D(E,\mathcal{a},P_o) \subset L(E,\mathcal{a},P_o)$, <u>domaine de définition de la transformée de LAPLACE réelle</u> L_{P_o} <u>de l'E.V.P.</u> $(E,\mathcal{a},P_o)$.

ii) - <u>Les lois de probabilité</u> P_θ , $\theta \in \Theta$ <u>sont équivalentes</u> .

iii) - <u>Pour tout</u> $\theta \in \Theta$, $P_\theta = \dfrac{e^\theta}{L_{P_o}(\theta)} \circ P_o$.

DEFINITION 3. - <u>On appellera</u> θ <u>le</u> "<u>paramètre naturel</u>" <u>de la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P})$ <u>et</u> "<u>codage</u>" (<u>relatif à</u> $P^* \in \mathcal{P}$) , <u>la bijection de</u> $\mathcal{P}$ <u>sur</u> Θ <u>qui permet de rendre équivalentes les DEFINITIONS 1 et 2</u>.

On appellera parfois le convexe $D(E,\mathcal{A},\mathcal{P})$, "domaine du paramètre naturel" relatif à l'E.V.P. $(E,\mathcal{A},P_o)$, à cause de la définition ci-dessus.

DEFINITION 4. - <u>On dira que la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P} = \{P_\theta;\theta\in\Theta\})$ <u>est de</u> "<u>type convexe</u>" [<u>resp.</u> "<u>saturée</u>"] <u>si</u> Θ <u>est une partie convexe de</u> $L(E,\mathcal{A},\mathcal{P})$ [<u>resp. si de plus</u> $\Theta = D(E,\mathcal{A},P_o)$] ; <u>qu'elle est à</u> "<u>paramètre de dimension finie</u>" <u>si l'espace vectoriel engendré par</u> Θ <u>dans</u> $L(E,\mathcal{A},\mathcal{P})$ <u>est de dimension finie</u> ; <u>enfin qu'elle est</u> "<u>réduite</u>" <u>si</u> $\mathcal{A}$ <u>coïncide avec la tribu exhaustive minimale de la structure statistique</u> $(E,\mathcal{A},\mathcal{P})$, <u>c'est-à-dire</u>, <u>si elle est engendrée par la famille</u> $\{\theta;\theta\in\Theta\}$ <u>de s.r.l</u>.

DEFINITION 5. - <u>Si</u> Θ' <u>est une partie non vide contenant l'origine</u>, <u>de</u> Θ , <u>nous dirons que la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P}' = \{P_\theta;\theta\in\Theta'\})$ <u>est une sous-structure de la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P} = \{P_\theta;\theta\in\Theta\})$.

La définition suivante permet de retrouver, comme cas particulier, le modèle initialement proposé dans [34] évoqué dans l'introduction ci-dessus.

DEFINITION 6. - <u>Avec les notations de la DEFINITION 1</u>, <u>on dira que la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P})$ <u>est</u> "<u>EXPLICITE</u>" <u>si de plus</u>,

$$\{\mathrm{Log}\,\frac{dP}{dP^*}\}_{P\in\mathcal{P}} \subset (E,\mathcal{A})^m/\mathcal{P} \oplus R\ .$$

En utilisant le codage u (relatif à $P^* \in \mathcal{P}$) de cette structure, cela revient à dire qu'elle peut s'écrire $(E,\mathcal{A},\{P_\theta;\theta\in\Theta\})$ avec $\Theta \subset (E,\mathcal{A})^m/\mathcal{P} \subset L(E,\mathcal{A},\mathcal{P})$. Mais il existe, alors, une version de toute v.a.r. $\theta \in \Theta$ de la forme : $x \longmapsto \theta(x) = \langle x,\theta\rangle$ pour un élément noté encore θ du dual $(E,\mathcal{A})^m$ de l'E.V.M.S. $(E,\mathcal{A})$ où $\langle .,. \rangle$ désigne la forme bilinéaire canonique mettant E et $(E,\mathcal{A})^m$ en dualité séparante. En utilisant cet abus de notation, la définition ci-dessus est équivalente à la suivante :

DEFINITION 7. - <u>Avec les notations de la DEFINITION 2</u>, <u>la structure exponentielle canonique</u> $(E,\mathcal{A},\{P_\theta;\theta\in\Theta\})$ <u>est dite</u> "<u>EXPLICITE</u>" <u>si pour tout</u> $\theta \in \Theta$, <u>il existe une version de la densité de</u> P_θ <u>par rapport à</u> P_o <u>de la forme</u> :

$$\frac{d P_\theta}{d P_\theta}(x) = \frac{e^{\langle x,\theta\rangle}}{l_{P_o}(\theta)} \quad , \ x \in E \ , \ \theta \in (E,\mathcal{A})^m \ .$$

On peut toujours supposer que $\mathcal{A}$ coïncide avec la tribu affaiblie $C(E,(E,\mathcal{A})^m)$.

Plus généralement, si E et F sont deux espaces vectoriels réels en dualité séparante et mesurable (i.e. $(E,C(E,F))^m = F$, cf. CHAPITRE I, §2, n°3) - ce que l'on peut toujours supposer - pour toute loi de probabilité P_o sur l'espace vectoriel faiblement mesurable $(E,C(E,F))$ et toute partie contenant l'origine du sous-ensemble convexe de F : $\{y\in F : l_{P_o}(y) = \int_E e^{\langle x,y\rangle} d P_o(y) < \infty\}$ on peut construire une structure statistique exponentielle explicite :

$$(E,C(E,F) \ , \ \{P_\theta;\theta\in\Theta\})$$

où P_θ est définie par sa densité par rapport à P_o sur $(E,C(E,F))$ admettant une version de la forme :

$$\frac{d P_\theta}{d P_o}(x) = \frac{e^{\langle x,\theta\rangle}}{l_{P_o}(\theta)} \ ; x \in E \ , \ \theta \in \Theta \subset F$$

C'est cette définition que nous avions proposée, dans un premier temps, dans [34] .

Evidemment, si E est de dimension finie, toute structure exponentielle canonique est explicite. Les structures exponentielles canoniques explicites admettent par définition une fonction de vraisemblance explicite dont la forme généralise celle des structures exponentielles canoniques classiques.

Mais la véritable généralisation en dimension quelconque de cette notion - et les exemples d'application le prouveront - est celle donnée par les DEFINITIONS 1 et 2 que l'on pourrait appeler aussi "structures exponentielles canoniques IMPLICITES" .

2. Structures exponentielles.

Il est rare qu'en pratique, une structure statistique se trouve être une structure exponentielle canonique, mais il est fréquent, de rencontrer des structures statistiques dont l'image par une statistique exhaustive soit une structure de ce type.

Il est important de les reconnaître puisque les décisions statistiques interviennent dans cette dernière, après réduction par exhaustivité, et que le modèle canonique décrit plus haut a, comme on le verra, des propriétés remarquables.

DEFINITION 8. - Une structure statistique $(\mathcal{X},\mathcal{F},\mathcal{P})$ est appelée "structure exponentielle" [resp. "structure exponentielle explicite"], s'il existe une statistique exhaustive T , définie sur cette structure, à valeurs dans un E.V.M.S. $(E,\mathcal{a})$, telle que la structure image par T soit une structure exponentielle canonique [resp. une structure exponentielle canonique explicite].

Nous n'imposons pas que $\mathcal{X}$ soit un espace vectoriel comme on le fait en théorie classique (cf. [5]) . Nous envisagerons les implications de cette définition dans le cas général et dans le cas "explicite".

1°) Pour tout $P \in \mathcal{P}$, notons P_T (plutôt que $P \circ T^{-1}$), la loi de probabilité de T sur $(E,\mathcal{A})$ et posons $\mathcal{P}_T = \{P_T, P \in \mathcal{P}\}$. La structure image $(E,\mathcal{A},\mathcal{P}_T)$ doit donc, d'après les définitions du n°1, être telle que :

a) les lois P_T, $P \in \mathcal{P}$ soient équivalentes sur $(E,\mathcal{A})$, ce qui implique que les lois $P \in \mathcal{P}$ le soient sur la sous-tribu $T^{-1}(\mathcal{A})$ de $\mathcal{F}$ engendrée par T ; notons $P_{T^{-1}(\mathcal{A})}$ la restriction de P à cette tribu, $P \in \mathcal{P}$.

b) Pour une loi $P^* \in \mathcal{P}$ on ait :

$$\{\mathrm{Log}\ \frac{d P_T}{d P_T^*}\}_{P\in\mathcal{P}} \subset L(E,\mathcal{A},\mathcal{P}_T) \oplus \mathbb{R}$$

ce qui implique que pour tout $P \in \mathcal{P}$, il existe $\zeta_P \in L(E,\mathcal{A},\mathcal{P}_T)$ et $k_P \in \mathbb{R}$ tels que, puisque T est exhaustive :

$$E_{P^*}\left(\frac{d P}{d P^*}\,\middle|\,T\right) = \frac{d P}{d P^*} = \exp\{\zeta_P \circ T + k_P\}\ ,\ P \in \mathcal{P}\ .$$

avec $k_P = -\mathrm{Log} \int_E e^{\zeta_P}\, d P_T^* = -\mathrm{Log} \int_{\mathcal{X}} e^{\zeta_P \circ T}\, d P^*$; soit, en passant au paramétrage naturel de la structure (canonique) image, c'est-à-dire en utilisant le codage de celle-ci relatif à P_T^*, que l'on puisse écrire :

$$\frac{d P}{d P^*} = \frac{e^{v(P) \circ T}}{\tilde{c}(P)}\ ,\ P \in \mathcal{P}$$

en notant v l'application injective de $\mathcal{P}$ dans $L(E,\mathcal{A},\mathcal{P}_T)$ définie par :

$v(P) = u(P_T)$ avec $v(P^*) = u(P_T^*) = 0$. Alors, $\tilde{c}(P) = L_{P^*}(v(P)) = \int_E e^{v(P)} d P_T^*$.

Posons $\Theta_T = v(\mathcal{P}) \subset L(E,\mathcal{a},\mathcal{P}_T)$ et pour $\theta \in \Theta_T$, $P_\theta = v^{-1}(\theta)$; on peut donc paramétrer la famille $\mathcal{P}$ par $\theta \in \Theta_T$, si bien que la structure statistique est de la forme :

$$(\mathcal{X},\mathcal{F},\mathcal{P} = \{P_\theta ; \theta \in \Theta_T\})$$

avec :

$$P_\theta = \frac{e^{\theta \circ T}}{L_{P_{o,T}}(\theta)} \circ P_o \quad \text{et} \quad L_{P_{o,T}}(\theta) = \int_E e^\theta \, dP_{o,T} \, , \; \theta \in \Theta_T$$

<u>On appellera</u> v , <u>le</u> "<u>codage naturel</u>" <u>de la structure exponentielle</u> $(\mathcal{X},\mathcal{F},\mathcal{P})$ <u>relativement à la loi</u> $P^* \in \mathcal{P}$ <u>et à la statistique exhaustive</u> T , <u>et</u> θ <u>le paramètre naturel</u>.

2°) Dans le cas explicite, il existe une version de θ de la forme : $t \longmapsto \langle t,\theta\rangle$, $t \in E$, en notant encore θ un élément de $(E,\mathcal{a})^m$; ce qui implique que la structure exponentielle explicite puisse s'écrire : $(\mathcal{X},\mathcal{F},\mathcal{P} = \{P_\theta ; \theta \in \Theta_T\})$ avec :

$$\frac{dP_\theta}{dP_o}(x) = \frac{e^{\langle T(x),\theta\rangle}}{l_{P_{o,T}}(\theta)} \, , \; x \in E$$

où $l_{P_{o,T}}(\theta) = \int_E e^{\langle t,\theta\rangle} \, dP_{o,T}(t)$, $\theta \in \Theta_T \subset \{y \in (E,\mathcal{a})^m : \int_E e^{\langle t,y\rangle} dP_{o,T}(y) < \infty\}$.

On peut résumer ces caractérisations en une définition équivalente :

DEFINITION 9. - <u>Une structure statistique</u> $(\mathcal{X},\mathcal{F},\{P_\theta , \theta \in \Theta\})$ <u>est appelée</u> "<u>structure exponentielle</u>" [<u>resp</u>. "<u>structure exponentielle explicite</u>"] <u>s'il existe une statistique exhaustive</u> T <u>à valeurs dans un E.V.M.S.</u> $(E,\mathcal{a})$ <u>et telle que</u> :

i) - Θ soit un sous-ensemble, contenant l'origine, du convexe $D(E,\mathcal{a},P_{o,T}) \subset L(E,\mathcal{a},P_{o,T})$ - domaine de définition de la transformée de LAPLACE réelle de l'E.V.P. $\check{}(E,\mathcal{a},P_{o,T})$ - [resp. du convexe $\{y \in (E,\mathcal{a})^m : l_{P_{o,T}}(y) = \int_E e^{\langle t,y\rangle} dP_{o,T}(t) < \infty\}$] .

ii) - Les lois de probabilité $P_\theta ; \theta \in \Theta$ soient équivalentes sur la sous-tribu de $\mathcal{F}$ engendrée par T .

iii) - Pour tout $\theta \in \Theta$, $P_\theta = \dfrac{e^{\theta \circ T}}{L_{P_{o,T}}(\theta)} \cdot P_o$.

$$[\text{resp.}\ \frac{dP_\theta}{dP_o}(x) = \frac{e^{\langle T(x),\,\theta\rangle}}{l_{P_{o,T}}(\theta)} \quad , x \in \mathcal{X}] .$$

REMARQUE 1. - Une structure exponentielle canonique est une structure exponentielle relativement à la statistique exhaustive triviale (identité).

REMARQUE 2. - Si $(\mathcal{X},\mathcal{F},\mathcal{P})$ est une structure exponentielle relativement à la statistique exhaustive T , la structure d'échantillon de taille $n \geq 1$ correspondante : $(\mathcal{X},\mathcal{F},\mathcal{P})^n$ l'est aussi relativement à la statistique exhaustive $T^{(n)} : \mathcal{X}^n \longmapsto E$ définie par : $T^{(n)}(x_1,\ldots,x_n) = \sum_{i=1}^{n} T(x_i)$, $x_i \in \mathcal{X}$, $i = 1,\ldots,n$; en effet, pour tout $P \in \mathcal{P}$, $\dfrac{dP^{n\otimes}}{dP^{*n\otimes}} = \exp\{\zeta_P \circ T^{(n)} + n k_P\}$ car ζ_P est une v.a.r.l. P^*_T - p.p. sur $(E,\mathcal{a},P^*_{T^{(n)}})$.

REMARQUE 3. - Il est clair que si $(\mathcal{X},\mathcal{F},\mathcal{P})$ est une structure statistique exponentielle classique en dimension finie, elle est une structure exponentielle au sens ci-dessus compte tenu de la caractérisation du dual d'un E.V.P. de dimension finie (cf. REMARQUE 4 , CHAPITRE II, §2) .

REMARQUE 4. - Toute structure statistique d'échantillon, dont les hypothèses sont des lois de probabilité équivalentes est une structure exponentielle (relativement à la statistique exhaustive triviale constituée par la loi de probabilité empirique de l'échantillon).

Compte tenu de la REMARQUE 2, il suffit de le démontrer pour la structure statistique correspondant à une seule observation. Soit donc $(\Omega,\mathcal{F},\mathcal{P})$ une structure statistique arbitraire telle que toutefois, les lois $P \in \mathcal{P}$ soient équivalentes.

Considérons la statistique $T : \omega \longmapsto \delta_\omega$ de $(\Omega,\mathcal{F})$ dans l'espace vectoriel E des mesures bornées à support fini sur Ω, muni de la plus grande tribu rendant l'application T mesurable. La démonstration repose alors sur la caractérisation du dual $L(E,\mathcal{a},\mathcal{P}_T)$ donnée par la PROPOSITION 2, CHAPITRE III, §3, puisque les lois de $\mathcal{P}$ sont équivalentes : "Il existe un isomorphisme ψ de $L_o(\Omega,\mathcal{F},\mathcal{P})$ sur $L(E,\mathcal{a},\mathcal{P}_T)$ tel que $\psi(X) = (u \longmapsto \int_\Omega X d\mu)$, $X \in L_o(\Omega,\mathcal{F},\mathcal{P}), \mu \in E$ et $\psi^{-1}(\zeta) = \zeta \circ T$, $\zeta \in L(E,\mathcal{a},\mathcal{P}_T)$".

Pour toute probabilité $P^* \in \mathcal{P}$ fixée, et pour tout $P \in \mathcal{P}$, $\dfrac{dP}{dP^*} = e^{\zeta_P \circ T}$, (1), à condition de prendre $\zeta_P = \psi(\text{Log}\,\dfrac{dP}{dP^*}) \in L(E,\mathcal{a},\mathcal{P}_T)$.

Ceci implique que la structure $(\Omega,\mathcal{F},\mathcal{P})$ est, d'après les DEFINITIONS 8 et 9, une structure exponentielle relativement à la statistique exhaustive T, avec pour codage relatif à P^* et à T, l'application v de $\mathcal{P}$ dans $L(E,\mathcal{a},\mathcal{P}_T)$ définie par $v(P) = \psi(\text{Log}\,\dfrac{dP}{dP^*})$ la constante k_P étant nulle pour tout $P \in \mathcal{P}$.

Si $(\Omega,\mathcal{F},\mathcal{P})^n$, $n \geq 1$ est la structure statistique correspondant à un échantillon de taille n : $(\omega_1,\dots,\omega_n)$, elle est exponentielle relativement à la statistique exhaustive $T^{(n)}$ définie par $T^{(n)}(\omega_1,\dots,\omega_n) = \sum_{i=1}^n T(\omega_i) = \sum_{i=1}^n \delta_{\omega_i}$ ou de façon équivalente à la statistique : $\frac{1}{n} T^{(n)} = \mathbb{P}_n$ qui est la loi de probabilité empirique de l'échantillon. La structure exponentielle canonique, image par $T^{(n)}$ de la structure d'échantillon $(\Omega,\mathcal{F},\mathcal{P})^n$, qui sera dite "structure exponentielle canonique empirique", est donc :

$$(E,\mathcal{a},\ \{\mathcal{M}_{(\Omega,\mathcal{F})}(P;n)\ ;\ P \in \mathcal{P}\})$$

où $\mathcal{M}_{(\Omega,\mathcal{F})}(P;n)$ est la loi <u>multinomiale (généralisée)</u> de paramètres l'entier n et la probabilité P sur $(\Omega,\mathcal{F})$ décrite au CHAPITRE III, §3 .

Cette structure exponentielle est même explicite puisque, comme l'indique la REMARQUE 2, CHAPITRE III, §3, $(E,\mathcal{a})^m/\mathcal{P}_T = L(E,\mathcal{a},\mathcal{P}_T)$. De fait, l'équation (1) s'écrit aussi :

$$\frac{dP}{dP^*}(\omega) = e^{\langle T(\omega),\mathrm{Log}\,\frac{dP}{dP^*}\rangle}\ ,\ \omega \in \Omega$$

en identifiant $\mathrm{Log}\,\frac{dP}{dP^*}$ à un élément de $(E,\mathcal{a})^m/\mathcal{P}_T$, où $\langle \mu,X\rangle = \int_\Omega X\,d\mu\ ;\ \mu \in E\ ,\ X \in L_o(\Omega,\mathcal{F},\mathcal{P})$.

Il n'est pas surprenant d'obtenir un tel résultat si l'on admet de paramétrer une structure statistique par un paramètre de dimension infinie (en particulier par la densité des lois de probabilité hypothèses). Ce genre de résultat est à rapprocher formellement de celui que l'on obtient en théorie de l'exhaustivité qui conclut à l'exhaustivité de la statistique d'ordre d'un échantillon, résultat peu utile mais qui n'enlève rien à cette théorie. De même, ici, les structures exponentielles généralisées seront intéressantes lorsqu'elles seront non triviales comme nous allons le voir dans les exemples du § suivant.

§ 2. - EXEMPLES DE STRUCTURES EXPONENTIELLES.

1. Structures statistiques gaussiennes à moyenne inconnue.

Les résultats de ce n° reposent sur ceux du § 1 du CHAPITRE III dont on reprendra les notations.

Soit $(E,\mathcal{A},P_o)$ un E.V.P.G. (au sens de la DEFINITION 4, CHAPITRE III, §1) où P_o est la probabilité gaussienne centrée sur l'E.V.M.S. $(E,\mathcal{A})$ de variance, la forme quadratique positive Q sur $(E,\mathcal{A})^m$; supposons vérifiée la condition de la PROPOSITION 2, CHAPITRE III, §1 , c'est-à-dire que l'espace de HILBERT $H(Q)$ associé à Q soit contenu dans E .

Pour tout $m \in E$, notons P_m la probabilité gaussienne de moyenne m et de variance Q c'est-à-dire l'image de P_o par l'application mesurable $x \longmapsto x + m$ de $(E,\mathcal{A})$ dans lui-même.

PROPOSITION 1. - <u>La structure statistique gaussienne à moyenne inconnue</u> :

$$(E,\mathcal{A},\{P_m, m \in H(Q)\})$$

<u>est une structure exponentielle canonique saturée</u>.

En effet, le THEOREME 7, CHAPITRE III, §1, indique que les lois $\{P_m, m \in H(Q)\}$ sont équivalentes et que pour tout $m \in H(Q)$, $\frac{dP_m}{dP_o} = \exp\{\psi^{-1}(m) - \frac{1}{2}\|m\|^2_{H(Q)}\}$ où ψ^{-1} est un isomorphisme de $H(Q)$ sur l'espace $L(E,\mathcal{A},P_o)$ qui se trouve être l'espace de HILBERT, fermeture dans $L_2(E,\mathcal{A},P_o)$ de l'espace $(E,\mathcal{A})^m/P_o$, d'après la PROPOSITION 2, CHAPITRE III, §1 . Nous reconnaissons donc une structure exponentielle canonique, le codage relatif à P_o étant défini par :

$u(P_m) = \psi^{-1}(m) = \theta$. Elle est saturée puisque $\Theta = u\{P_m; m \in H(Q)\} = \psi^{-1}(H(Q)) = D(E,\mathcal{A},P_o) = L(E,\mathcal{A},P_o)$. ■

APPLICATION : Détection d'un signal dans un bruit brownien.

Supposons que l'on observe un processus $(X_t)_{t\in T}$ sur l'intervalle de temps $T = [0,1]$ par exemple, défini sur un espace probabilisé quelconque $(\Omega,\mathcal{F},P)$ et tel que : $X_t = m(t) + W_t$, $t \in T$, où $\{W_t\}_{t\in T}$ est un mouvement brownien standard (centré) défini sur le même espace probabilisé (le bruit) et $m = \{m(t)\}_{t\in T}$ est une fonction réelle inconnue définie sur T (le signal), supposée cependant continue et nulle à l'instant 0 .

Tout problème de décision concernant la fonction inconnue m est un problème de statistique à partir de l'observation x de l'élément aléatoire X associé à la f.a.r. $(X_t)_{t\in T}$, et il est facile de constater, en se reportant au CHAPITRE III, §1, n°4 pour les notations et les résultats concernant la loi de WIENER, que la structure statistique associée à x est :

$$(C_o[0,1]\ ,\ \mathcal{B}(C_o[0,1])\ ,\ \{P_m, m\in\mathcal{M}\})$$

où $\mathcal{M}$ est l'ensemble des hypothèses possibles pour le signal inconnu et P_m la loi du vecteur aléatoire X pour la valeur m de ce signal, P_o désignant donc la loi de WIENER.

Mais ce problème n'a de sens que si les lois P_m ne sont pas étrangères auquel cas, comme elles sont gaussiennes elles sont équivalentes et pour qu'il en soit ainsi, il faut et il suffit de restreindre l'ensemble des hypothèses à l'espace de HILBERT $H(Q_W)$ associé à la loi de WIENER, c'est-à-dire de supposer que m est de plus dérivable et que sa dérivée est de carré intégrable par rapport à la mesure de LEBESGUE sur $[0,1]$. (THEOREME 9, CHAPITRE III, §4) , la structure statistique de notre problème est donc la structure exponentielle canonique (gaussienne).

(1) $$(C_o[0,1],\mathcal{B}(C_o[0,1])\ ,\ \mathcal{P} = \{P_m, m \in H(Q_W)\})$$

avec,

$$\frac{dP_m}{dP_o} = \exp\left(\int_0^1 m'(t)\, d\xi_t - \frac{1}{2}\int_0^1 m'^2(t)\, dt\right)\ ,\ m \in H(Q_W)$$

le codage étant défini par : $u(P_m) = \theta = \int_0^1 m'(t)\, d\xi_t \in \Theta = H$, le paramètre naturel θ est une intégrale stochastique par rapport au mouvement brownien standard canonique, H étant l'espace gaussien associé à ce dernier, en fait le dual de l'E.V.P. $(C_o[0,1], \mathcal{B}(C_o[0,1]), P_o)$ et $L_{P_o}(\theta) = \exp(\frac{1}{2}\|\theta\|_H^2)$.

<u>Cette structure exponentielle devient explicite</u>, <u>moyennant des hypothèses supplémentaires sur</u> m (<u>donc sur</u> θ) :

On sait en effet que si m' est de plus à variation bornée sur $[0,1]$ l'intégrale stochastique $\int_0^1 m'(t)\, d\xi_t$ admet pour version :

$$(\int_0^1 m'(t)\, d\xi_t)(x) = x(1)m(1) - \int_0^1 x(t)\, dm'(t) \quad , \; x \in C_o[0,1] \; .$$

En remplaçant donc $\mathcal{P}$ par la sous-famille $\mathcal{P}'$ obtenue en restreignant de la sorte le domaine de variation de m , la sous-structure exponentielle canonique :

$$(C_o[0,1], \mathcal{B}(C_o[0,1]), \mathcal{P}')$$

devient explicite, en posant pour $\theta = u(P_m)$ la mesure sur $[0,1]$ définie par :

$$A \in \mathcal{B}([0,1]) \; , \; \theta(A) = m'(1)\, \delta_1(A) - \int_A dm'(t)$$

où δ_1 désigne la mesure de DIRAC au point 1 .

Alors $\theta \in (C_o[0,1], \mathcal{B}(C_o[0,1]))^m / \mathcal{P}'$ (cf. CHAPITRE III, §1, n°4) et la famille $\mathcal{P}'$ peut être paramétrée par $\theta \in \Theta' = u(\mathcal{P}')$ c'est-à-dire que l'on a :

$$\mathcal{P}' = \{P_\theta ; \theta \in \Theta'\} \quad \text{avec} \quad \frac{dP_\theta}{dP_o}(x) = \frac{e^{\langle x, \theta \rangle}}{l_{P_o}(\theta)} \quad , \; x \in C_o[0,1] \; , \; \theta \in \Theta' \; ,$$

où $\langle x, \theta \rangle = \int_0^1 x(t)\, d\theta(t)$ est la valeur en (x, θ) de la forme bilinéaire canonique mettant $C_o[0,1]$ et $M[0,1] = (C_o[0,1], \mathcal{B}(C_o[0,1]))^m$ en dualité. On vérifie que

$$1_{P_o}(\theta) = 1_{P_o}(u(P_m)) = \exp\{\frac{1}{2}\int_0^1 m'^2(t)\,dt\} = \exp\{\frac{1}{2}\iint_{T\times T} \inf(s,t)\,d\theta(s)\,d\theta(t)\} \;,\; \theta\in\Theta$$

soit en notant Q_W la variance de la loi de WIENER P_o que

$$1_{P_o}(\theta) = \exp\{\frac{1}{2}\,Q_W(\theta)\} \;.$$

En effet, comme $\theta = \theta_1 + \theta_2$ avec $\theta_1 = m'(1)\,\delta_1$ et $\theta_2 = \int dm'$, on doit avoir : $Q_W(\theta_1+\theta_2) = Q_W(\theta_1) + Q_W(\theta_2) + 2K_W(\theta_1,\theta_2)$; calculons chacun de ces termes :

$$Q_W(\theta_1) = \int_0^1\int_0^1 \inf(s,t)\;[m'(1)]^2\,d\delta_1(s)\,d\delta_1(t) = [m'(1)]^2$$

$$Q_W(\theta_2) = \int_0^1\int_0^1 \inf(s,t)\,dm'(s)\,dm'(t) = \int_0^1\int_0^1\int_0^1 1\!\!1_{[0,1]}(u)\;1\!\!1_{[0,1]}(u)\,du\,dm'(s)\,dm'(t)$$

$$= \int_0^1 [\int_u^1\int_u^1 dm'(s)\,dm'(t)]\,du = \int_0^1 [m'(1)-m'(u)]^2\,du$$

$$= [m'(1)]^2 - 2m'(1)\;[m(1)-m(0)] + \int_0^1 [m'(u)]^2\,du \;,$$

$$2K(\theta_1,\theta_2) = -2\int_0^1\int_0^1 \inf(s,t)\,m'(1)\,d\delta_1(s)\,dm'(t) = -2\int_0^1 [\int_u^1 m'(1)\,d\delta_1(s)\int_0^1 dm'(t)]du$$

$$= -2\int_0^1 m'(1)\;[m'(1)-m'(u)]\,du = -2\,[m'(1)]^2 + 2m'(1)\;[m(1)-m(0)]$$

en additionnant les trois termes on trouve bien : $Q_W(\theta) = \int_0^1 m'(t)^2\,dt$.

REMARQUE 1. - On pourrait étudier de façon analogue le modèle obtenu en remplaçant $(W_t)_{t\in T}$ par une f.a.r. gaussienne centrée quelconque sur un ensemble d'indices T , l'espace des observations serait spécifié par les connaissances supplémentaires concernant les trajectoires de cette f.a.r. sur T .

REMARQUE 2. - Supposons que dans la structure statistique (1) , on admette que le paramètre varie dans un sous-espace de dimension finie K de $H(Q_W)$, elle devient :

(2) $$(C_o[0,1],\mathcal{B}(C_o[0,1]),\ \mathcal{P}'' = \{P_m, m\in K\subset H(Q_W)\})$$

et si $\{m_1,\dots,m_N\}$ est une base de K on a :

$$\frac{dP_m}{dP_o} = \exp\{\sum_{i=1}^{N} \lambda_i(m) \int_0^1 m_i'(t)\, d\xi_t - \frac{1}{2}\int_0^1 m'(t)^2\, dt\}$$

soit $T = (T_1,\dots,T_N)$ le vecteur aléatoire défini sur $(C_o[0,1],\mathcal{B}(C_o[0,1]))$ par $T_i = \int_0^1 m_i'(t)\, d\xi_t$; $i = 1,\dots,N$ et posons $\theta = (\theta_1,\dots,\theta_N) = v(P_m)$ avec $\theta_i = \lambda_i(m); i = 1,\dots,N$. On peut paramétrer la famille $\mathcal{P}''$ par $\theta \in \mathbf{R}^N$ de sorte que $\mathcal{P}'' = \{P_\theta;\theta\in\mathbf{R}^N\}$ avec

$$\frac{dP_\theta}{dP_o} = \frac{e^{\sum_{i=1}^{N} T_i\theta_i}}{l_{P_{o,T}}(\theta)} \quad ; \theta\in\mathbf{R}^N$$

ce qui exprime que la structure statistique (2) est exponentielle, relativement à la statistique exhaustive T et que la structure exponentielle canonique, image de celle-ci par T est une structure exponentielle classique de dimension finie :

$$(\mathbf{R}^N,\mathcal{B}(\mathbf{R}^N)\ ,\ \mathcal{P}_T'' = \{P_{\theta,T};\theta\in\mathbf{R}^N\})$$

avec

$$\frac{dP_{\theta,T}}{dP_{o,T}}(t) = \frac{e^{\langle t,\theta\rangle}}{l_{P_{o,T}}(\theta)} \quad , t\in\mathbf{R}^N,\ \theta\in\mathbf{R}^N .$$

Nous remarquons, que dès que la paramètre est de dimension finie, il existe une statistique exhaustive de dimension finie. Ce fait caractéristique des structures exponentielles sera étudié au § suivant.

2. Structures statistiques gaussiennes à variance inconnue.

Dans les mêmes conditions qu'au n°1, soit P_{Q_o} la probabilité gaussienne centrée sur $(E,\mathcal{a})$ de variance la forme quadratique positive Q_o sur $(E,\mathcal{a})^m$.

PROPOSITION 2. - <u>La structure statistique gaussienne à variance inconnue :</u>

$$(E,\mathcal{a},\mathcal{P}_{Q_o})$$

<u>où $\mathcal{P}_{Q_o}$ désigne la famille des lois de probabilité gaussiennes centrées sur $(E,\mathcal{a})$ équivalentes à la loi P_{Q_o} est une structure exponentielle relativement à la statistique exhaustive</u> : $x \longrightarrow \chi_E(x) = x \otimes x$, <u>de</u> E <u>dans</u> $E \otimes E$.

On sait en effet, d'après le THEOREME 8, CHAPITRE III, §1, que si $P_Q \in \mathcal{P}_{Q_o}$, $\dfrac{dP_Q}{dP_{Q_o}} = \dfrac{e^{\zeta_Q}}{E_{P_{Q_o}}(\zeta_Q)}$ où ζ_Q est un élément uniquement déterminé de l'espace de HILBERT $L(E,\mathcal{a},P_{Q_o})^{2\tilde{\odot}}$. Il suffit alors d'appliquer le COROLLAIRE de la PROPOSITION 3, CHAPITRE III, §1 qui exprime qu'un tel élément ζ_Q peut s'écrire $\zeta_Q = \xi_Q \circ \chi$ où ξ_Q est une v.a.r.l. unique, appartenant au dual $L(E^{2\odot},\mathcal{a}^{2\chi},P_{Q_o,\chi_E})$, où $E^{2\odot}$ est l'espace vectoriel engendré par l'ensemble $\chi_E(E)$ dans $E \otimes E$, muni de la plus grande tribu $\mathcal{a}^{2\chi}$ rendant mesurable l'application quadratique χ_E , P_{Q_o,χ_E} étant la loi de χ_E . Alors

$$\frac{dP_Q}{dP_{Q_o}} = \frac{e^{\xi_Q \circ \chi_E}}{C(Q)} \quad , \; P_Q \in \mathcal{P}_{Q_o} \; . \blacksquare$$

APPLICATIONS : a) <u>Structure statistique associée à l'observation d'une solution d'une équation différentielle stochastique linéaire</u>.

Soit $x = (x(t))_{t\in[0,1]}$ une observation faite pendant un intervalle de temps fini, fixé à $[0,1]$ pour simplifier, d'un processus stochastique $(X_t)_{t\in[0,1]}$ défini sur un espace probabilisé quelconque et régi par une équation différentielle stochastique de la forme :

$$\begin{cases} dX_t = \alpha(t)\, X_t\, dt + dW_t \,, \\ X_o = 0 \,; \end{cases}$$

où $(W_t)_{t\in[0,1]}$ est un mouvement brownien standard, défini sur le même espace probabilisé et $\alpha(t)$, un coefficient réel inconnu, dépendant du temps.

La théorie des équations différentielles stochastiques (cf. COROLLAIRE, du THEOREME 10, CHAPITRE III, §1) permet d'affirmer que si la fonction inconnue $\alpha = (\alpha(t))_{t\in[0,1]}$ appartient cependant à l'espace $L^2([0,1],\mathcal{B}([0,1]),dt)$, x peut être considéré comme l'observation d'un élément aléatoire X défini, comme on l'a déjà fait précédemment à partir de la fonction aléatoire $(X_t)_{t\in[0,1]}$, à valeurs dans l'E.V.M.S. $(C_o[0,1],\mathcal{B}(C_o[0,1]))$ dont la loi de probabilité P_α est équivalente, quelque soit α, à la loi P_o qui coïncide alors avec la loi de WIENER sur ce dernier espace mesurable et vérifie :

$$\frac{dP_\alpha}{dP_o} = \exp\Big(\int_0^1 \alpha(t)\,\xi_t\, d\xi_t - \frac{1}{2}\int_0^1 \alpha^2(t)\,\xi_t^2\, dt\Big) \tag{3}$$

cette densité s'exprime au moyen d'une intégrale stochastique par rapport au processus de WIENER standard canonique $(\xi_t)_{t\in[0,1]}$.

Quant à la variable aléatoire définie par la deuxième intégrale elle est déterminée partout sur $C_o[0,1]$ par :

$$x \in C_o[0,1]\,,\quad \Big(\int_0^1 \alpha^2(t)\,\xi_t^2\, dt\Big)(x) = \int_0^1 \alpha^2(t)\, x^2(t)\, dt\,.$$

Nous allons montrer que si l'on restreint α à appartenir à un certain sous-espace A de $L^2([0,1],\mathcal{B}([0,1]),dt)$, la structure statistique :

$$(C_o[0,1],\mathcal{B}(C_o[0,1]))\,,\ \mathcal{P} = \{P_\alpha ; \alpha\in A\}) \tag{4}$$

correspondant à l'observation x de X est une structure exponentielle explicite relativement à la statistique exhaustive T définie par $T(x) = x^2(.)$

à valeurs dans le cône positif de $C_o[0,1]$.

On peut montrer, en utilisant un théorème de CAMERON et MARTIN par exemple (cf. [39], pour les détails) que si α est une fonction continue et à variation bornée sur $[0,1]$, l'intégrale stochastique dans (3) admet pour version :

$$\left(\int_0^1 \alpha(t)\, \xi_t\, d\xi_t\right)(x) = \frac{1}{2}\left(\int_0^1 \alpha(t)\, dx^2(t) - \int_0^1 \alpha(t)\, dt\right),$$

pour P_o-presque tout x dans $C_o[0,1]$.

L'intégrale $\int_0^1 \alpha(t)\, dx^2(t)$ est alors bien définie et entendue comme une intégrale de RIEMANN-STIELJES, c'est-à-dire comme :

$$\lim_{n\to\infty} \sum_{i=1}^{n} \alpha\left(\frac{i}{n}\right) \left[x\left(\frac{i}{n}\right) + x\left(\frac{i-1}{n}\right)\right] \left[x\left(\frac{i}{n}\right) - x\left(\frac{i-1}{n}\right)\right].$$

La fonction de vraisemblance de la structure statistique ci-dessus, s'écrit alors,

$$\forall \alpha \in A,\ \frac{dP_\alpha}{dP_o}(x) = \exp\left\{\frac{1}{2}\left(\int_0^1 \alpha(t)\, dx^2(t) - \int_0^1 \alpha^2(t)\, x^2(t)\, dt - \int_0^1 \alpha(t)\, dt\right)\right\},$$

P_o presque partout sur $C_o[0,1]$, ou encore en intégrant la première intégrale par parties :

$$\forall \alpha \in A,\ \frac{dP_\alpha}{dP_o}(x) = \exp\left\{\frac{1}{2}\left(\alpha(1)\, x^2(1) - \int_0^1 x^2(t)\, d\alpha(t) - \int_0^1 \alpha^2(t)\, x^2(t)\, dt - \int_0^1 \alpha(t)\, dt\right)\right\}$$

P_o-presque partout.

Le théorème de factorisation, montre bien que $T(x) = x^2(.)$ et une statistique exhaustive.

Pour montrer que la structure statistique (4) est une structure exponentielle explicite relativement à la statistique T, il suffit de constater que cette fonction de vraisemblance s'écrit :

$$\forall\alpha \in A \ , \ \frac{d P_\alpha}{d P_o}(x) = \frac{e^{\langle T(x), u(\alpha)\rangle}}{C(\alpha)} \quad , \quad x \in C_o[0,1] \ , \ P_o\text{-p.p.}$$

en prenant pour paramètre naturel la mesure bornée $\theta = u(\alpha)$ définie sur $([0,1], \mathcal{B}([0,1]))$ par : $B \in \mathcal{B}([0,1])$, $\theta(B) = \frac{1}{2}[\alpha(1)\delta_1(A) - \int_B d\alpha(t) - \int_B \alpha^2(t)\, dt]$ alors $\langle T(x), u(\alpha)\rangle = \int_0^1 x^2(t)\, d\theta(t)$, qui est la valeur en $(T(x), \theta)$ de la forme bilinéaire mettant $C_o[0,1]$ et l'espace $M[0,1]$ des mesures de RADON bornées sur $[0,1]$ en dualité.

Il s'ensuit que l'on peut paramétrer la famille $\mathcal{P}$ par $\theta \in u(A) = \Theta \subset M[0,1]$, la structure statistique devient :

$$(C_o([0,1]) \ , \ \mathcal{B}(C_o[0,1]) \ , \ \mathcal{P} = \{P_\theta ; \theta \in \Theta\})$$

avec

$$\frac{d P_\theta}{d P_o}(x) = \frac{e^{\langle T(x), \theta\rangle}}{l_{P_{o,T}}(\theta)} \quad ; \quad x \in C_o([0,1]) \ P_o\text{-p.p.}$$

où,

(5) $$l_{P_{o,T}}(\theta) = l_{P_{o,T}}(u(\alpha)) = C(\alpha) = \exp\left(\frac{1}{2}\int_0^1 \alpha(t)\, dt\right) .$$

La structure exponentielle canonique explicite image de cette dernière par T est donc :

$$(C_o([0,1]), \mathcal{B}(C_o[0,1]), \ \mathcal{P}_T = \{P_{\theta,T} \ ; \ \theta \in \Theta\})$$

où $\Theta \subset D(C_o[0,1], \mathcal{B}(C_o[0,1]), P_o \circ T^{-1})$.

Pour étudier cette structure exponentielle canonique, il convient donc de connaître la loi de probabilité $P_o \circ T^{-1}$ de T quand $\theta = 0$, c'est-à-dire la loi de l'élément aléatoire : $T(X) = X^2(.)$ dans $(C_o[0,1], \mathcal{B}(C_o[0,1]))$

quand $X = (X(.))$ est l'élément aléatoire associé à la f.a.r. gaussienne centrée du mouvement brownien.

On dispose en particulier de sa fonction caractéristique :

$$\nu \in M[0,1] \; ; \; \varphi_T(\nu) = E(e^{i\langle T,\nu\rangle}) ,$$

comme cas particulier des résultats obtenus au CHAPITRE III, §2 concernant la loi de probabilité de WISHART associée à un élément aléatoire gaussien X dans un espace de FRECHET séparable :

Dans ce cas très simple on obtient :

$$\varphi_T(\nu) = \text{dét}(\mathbb{I} - 2i B_\nu)^{-\frac{1}{2}} \quad ; \nu \in M[0,1] .$$

où $\text{dét}(\mathbb{I} - 2i B_\nu)$ est la valeur au point $2i$ du déterminant de FREDHOLM de l'opérateur nucléaire autoadjoint dans $H(Q_W)$ défini par :

$$f \in H(Q_W) \; , \; B_\nu(f) = \int_0^1 f(t) \; \inf(.,t) \; d\nu(t) .$$

En utilisant l'isométrie entre $H(Q_W)$ et $L^2([0,1],\mathcal{B}([0,1]),dt)$ on peut d'ailleurs montrer que :

$$\text{dét}(\mathbb{I} - 2i B_\nu) = \delta(2i, \tilde{B}_\nu)$$

où avec des notations plus classiques $\delta(.,\tilde{B}_\nu)$ est le déterminant de FREDHOLM de l'opérateur de $L^2([0,1],\mathcal{B}([0,1]),dt)$ défini par le noyau :

$$s,t \in [0,1] \; , \; \tilde{B}_\nu(s,t) = \nu[\max(s,t),1] .$$

De ce fait, $$l_{P_{o,T}}(\theta) = \int_{C_o[0,1]} e^{\langle t,\theta\rangle} \, dP_{o,T}(t) = \varphi_T(-i\,\theta)$$

$$= \text{dét}(\mathbb{I} + 2B_\theta)^{-\frac{1}{2}}$$

D'ou la description du convexe : $D(C_o[0,1],\mathcal{B}(C_o[0,1]), P_o \circ T^{-1})$ dans

$M[0,1]$ et l'identité : $\text{dét}(\mathbb{I} + 2B_{u(\alpha)}) = \exp(-\int_0 \alpha(t)\,dt)$, d'après (5) .

Les résultats du CHAPITRE III, §2 permettent d'envisager la construction de structures exponentielles canoniques explicites plus générales à partir de la loi de WISHART généralisée.

b) <u>Structure statistique associée à l'observation d'une f.a.r. gaussienne centrée à covariance triangulaire</u>.

Soit $x = (x(t))_{t \in [0,1]}$ l'observation faite pendant l'intervalle de temps fini $[0,1]$ d'un processus réel gaussien $(X_t)_{t \in [0,1]}$ centré, de fonction de covariance inconnue, mais supposée triangulaire c'est-à-dire de la forme $k_{a,b}(s,t) = a[\min(s,t)].b[\max(s,t)]$, $s,t \in [0,1]$ où a et b sont deux fonctions numériques définies sur $[0,1]$ telles que $\frac{a}{b}$ soit positive et non décroissante sur cet intervalle.

Le THEOREME 11, CHAPITRE III, §1 exprime que si l'on suppose de plus que les fonctions a et b sont de classe C^2 et telles que : $a(0) = 0$; $b(t) > 0$, $t \in [0,1]$; $b(t)\,a'(t) - a(t)\,b'(t) = 1$, $t \in [0,1]$, la stucture statistique associée à cette observation x est de la forme :

$$(C_o[0,1] \; , \; \mathcal{B}(C_o[0,1]) \; , \; \mathcal{P}_U = \{P_{a,b} ; (a,b) \in U\})$$

où U est l'ensemble des couples de fonctions vérifiant les conditions imposées à a et b . Les lois de probabilité $P_{a,b}$ sont équivalentes, en particulier la densité de $P_{a,b}$ par rapport à l'une d'entre elles, la loi de WIENER, obtenue pour : $a_o(t) = t$ et $b_o(t) = 1$, $\forall t \in [0,1]$, est donnée par :

$$x \in C_o[0,1] \; , \; \frac{dP_{a,b}}{dP_{a_o,b_o}}(x) = \sqrt{\frac{b(0)}{b(1)}} \; \exp\{-\frac{1}{2}\,(\frac{b'(1)}{b(1)}\,x^2(1) - \int_0^1 x^2(t)\frac{b''(t)}{b(t)}\,dt)\}$$

Ce qui montre que cette structure statistique est exponentielle explicite relativement à la statistique exhaustive $T : x \longmapsto x^2(.)$.

En posant : $\theta = \theta(a,b) = \frac{1}{2}\left(\frac{b'(1)}{b(1)}\delta_1 - \int \frac{b''(t)}{b(t)}\,dt\right)$

on définit une mesure borélienne bornée sur $[0,1]$, et donc un codage relatif à P_{a_o, b_o} de $\mathcal{P}_U$, ce qui fait que la structure statistique considérée est de la forme :

$$(C_o[0,1], \mathcal{B}(C_o[0,1]), \mathcal{P} = \{P_\theta ; \theta \in \Theta = \theta(U) \subset M[0,1]\})$$

avec : $\dfrac{dP_\theta}{dP_o}(x) = \dfrac{e^{\langle T(x), \theta\rangle}}{l_{P_{o,T}}(\theta)}$, $x \in C_o[0,1]$, $\theta \in \Theta$

où $l_{P_{o,T}}(\theta(a,b)) = \left[\dfrac{b(0)}{b(1)}\right]^{-\frac{1}{2}}$; $a,b \in U$.

ce qui, au passage, fournit l'identité : $\det(\mathbb{I} + 2B_{\theta(a,b)}) = \dfrac{b(0)}{b(1)}$; $(a,b) \in U$

lorsque B_θ est l'opérateur défini comme dans l'application a) précédente, relativement à la mesure θ . ∎

3. Structures statistiques de POISSON. Application.

PROPOSITION 3. - La structure statistique vectorielle :

(1) $$(E, \mathcal{A}, \mathcal{P} = \{\mathfrak{P}_{(T,\mathcal{T})}(\lambda) , \lambda \in \Lambda(\lambda_o)\})$$

où E désigne l'espace vectoriel des mesures bornées à support fini sur un espace mesurable $(T,\mathcal{T})$, $\mathcal{A}$ la plus grande tribu rendant mesurable l'application X de l'espace mesurable somme directe $(\Omega,\mathcal{F}) = \sum\limits_{n\geq 0} (T^n, \mathcal{T}^{n\otimes})$ définie par : $X(\omega) = \sum\limits_{j=1}^{n(\omega)} \delta_{t_j(\omega)}$, $\omega \in \Omega$, $X(\omega_o) = 0$, $\{\omega_o\} = T^o$ et $\mathcal{P}$ la famille des lois de POISSON d'intensité la mesure positive bornée λ équivalente à la mesure positive bornée λ_o sur $(T,\mathcal{T})$, est une structure statistique exponentielle canonique.

Cela résulte directement de la PROPOSITION 2, CHAPITRE III, §4, avec les notations de laquelle, le codage relatif à $\mathfrak{P}_{(T,\mathcal{T})}(\lambda_o)$ de cette structure s'écrit : $u(\mathfrak{P}_{(T,\mathcal{T})}(\lambda)) = \psi(\mathrm{Log}\,\frac{d\lambda}{d\lambda_o}) = \theta$, où ψ est un isomorphisme de $L_o(T,\mathcal{T},\lambda_o)$ sur $L(E,\mathcal{A},\mathcal{P})$. Cette structure est bien de la forme :

$$(E,\mathcal{A},\mathcal{P} = \{P_\theta\,;\theta \in \Theta\})$$

avec
$$\frac{dP_\theta}{dP_o} = \frac{e^\theta}{L_{P_o}(\theta)}, \ \theta \in \Theta \ ;$$

où,
$$L_{P_o}(\theta) = \exp\{\int_T (e^{\psi^{-1}(\theta)} - 1)\, d\lambda_o\} \ .$$

Remarquons, que d'après la démonstration de la proposition citée, il revient au même de dire que la structure statistique :

$$(\Omega,\mathcal{F},\{P_\lambda\,;\lambda \in \Lambda(\lambda_o)\}) = \sum_{n\geq 0} (T^n, \mathcal{T}^{n\otimes}, \{e^{-\lambda(T)}\frac{\lambda^{n\otimes}}{n!}, \ \lambda \in \Lambda(\lambda_o)\})$$

est une structure exponentielle généralisée relativement à la statistique exhaustive constituée par le vecteur X , puisque la structure exponentielle canonique (1) en est l'image par X . ∎

<u>Application</u> : <u>Structure statistique associée à l'observation d'un processus de POISSON d'intensité équivalente à la mesure de LEBESGUE sur la droite</u>.

Supposons que l'on observe pendant un intervalle de temps fini $T = [0,1]$ par exemple un processus de POISSON N d'intensité inconnue λ définie par sa densité par rapport à la mesure de LEBESGUE sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ f_λ que l'on suppose être une fonction positive continue, soit Λ l'ensemble de ces mesures d'intensité.

L'observation réelle se résume à la donnée des instants successifs $t_1,\ldots,t_n$ de réalisation d'un événement donné, elle est équivalente à la donnée de la mesure $\sum_{j=1}^{n} \delta_{t_j}$ qui est l'observation x d'un vecteur aléatoire X de loi de POISSON : $\mathfrak{P}_{(T,\mathcal{T})}(\lambda_T)$, où λ_T est la mesure bornée, restriction de λ à l'espace mesurable $(T,\mathcal{B}(T))$, puisque cette dernière est équivalente à l'observation de la famille de v.a.r. $\{N(F)\}_{F\in T\cap\mathcal{B}(\mathbb{R})}$ où $N(F)$

est le nombre de t_j dans F . On peut poser $X = 1_T . N$.

La structure statistique associée à ces observations est donc une structure de POISSON :

$$(E, \mathcal{A}, \{ \mathbb{P}_{(T,\mathcal{T})}(\lambda) ; \lambda \in \Lambda \})$$

dont la fonction de vraisemblance vaut :

$$\text{si } x = \sum_{j=1}^{n} \delta_{t_j} , \quad \mathcal{L}(\lambda ; x) = \frac{d\, \mathbb{P}_{(T,\mathcal{T})}(\lambda)}{d\, \mathbb{P}_{(T,\mathcal{T})}(dt)} (x) = \prod_{j=1}^{n} f_\lambda(t_j)\, e^{-\int_0^1 f_\lambda(t)\, dt + 1} \quad . \blacksquare$$

§ 3. - PROPRIETES DES STRUCTURES EXPONENTIELLES CANONIQUES.

Etant donné ce qui précède, la notion de structure exponentielle canonique a une importance particulière, puisque c'est la structure statistique obtenue après "réduction de l'observation" par une statistique exhaustive, et donc dans laquelle sont prises les décisions statistiques. Il est alors souhaitable d'en connaître les propriétés générales caractéristiques.

Dans tout ce § on considère donc une structure exponentielle canonique, munie de son paramètre naturel, c'est-à-dire une structure statistique vectorielle $(E,\mathcal{A},\mathcal{P} = \{P_\theta ; \theta \in \Theta\})$ vérifiant les conditions i) à iii) de la DEFINITION 2, §1 .

1. Intégrabilité.

PROPOSITION 1. - Avec les définitions du CHAPITRE II, §2, n° 2, pour tout $\theta \in \Theta$, l'E.V.P. $(E,\mathcal{A},P_\theta)$ a pour transformée de FOURIER :

$$\xi \in L(E,\mathcal{A},P_\theta) \ , \ \Phi_{P_\theta}(\xi) = \int_E e^{i\xi}\, dP_\theta = \frac{\tilde{L}_{P_o}(\theta + i\xi)}{L_{P_o}(\theta)}$$

(où $\tilde{L}_{P_o}$ désigne l'extension de la transformée de LAPLACE réelle L_{P_o} de l'E.V.P. $(E,\mathcal{A},P_o)$ à la bande $D(E,\mathcal{A},P_o) + i\, L(E,\mathcal{A},\mathcal{P})$ de variables aléatoires complexes).

De plus, en notant $\overset{o}{D}(E,\mathcal{A},P_o)$ l'intérieur algébrique du convexe $D(E,\mathcal{A},P_o)$, on a :

$$L(E,\mathcal{A},\mathcal{P}) = \bigcap_{\theta \in \Theta \cap \overset{o}{D}(E,\mathcal{A},P_o)} L_1(E,\mathcal{A},P_\theta) \ .$$

En effet, la première partie de la proposition est évidente puisque $P_\theta = \dfrac{e^\theta}{L_{P_o}(\theta)} . P_o$. Pour la deuxième, on remarque que

$\theta \in \overset{o}{D}(E,\mathcal{A},P_o) \Leftrightarrow 0 \in D(E,\mathcal{A},P_\theta)$, par suite, en appliquant la

PROPOSITION 2, CHAPITRE II, §2, on a : $L(E,\mathcal{a},P_\theta) \subset L_1(E,\mathcal{a},P_\theta)$ pour tout $\theta \in \overset{\circ}{\Theta} \cap D(E,\mathcal{a},P_o)$ ce qui implique la relation proposée puisque $L(E,\mathcal{a},P_\theta) = L(E,\mathcal{a},\mathcal{P}), \forall\, \theta \in \Theta$. Elle exprime que toute statistique réelle linéaire définie sur la structure est P_θ-intégrable dès que θ est un point interne de $D(E,\mathcal{a},P_o)$. De plus, d'après la REMARQUE 6, CHAPITRE II, §2, si $\theta \in D(E,\mathcal{a},P_o)$, pour tout $\xi \in L(E,\mathcal{a},\mathcal{P})$, $E_{P_\theta}(\xi) = \dfrac{L'_{P_o}(\theta,\xi)}{L_{P_o}(\theta)} = C'_{P_o}(\theta,\xi)$. ∎

2. Structure image par une statistique linéaire et caractérisation.

Soient $(E_1,\mathcal{a}_1)$ un E.V.M. et U une statistique vectorielle linéaire définie sur la structure exponentielle canonique et à valeurs dans $(E_1,\mathcal{a}_1)$.

La structure statistique image est $(E_1,\mathcal{a}_1,\mathcal{P}_U = \{P_\theta \circ U^{-1} ; \theta \in \Theta\})$.

Compte tenu de la PROPOSITION 4, CHAPITRE II, §2, définissant l'application transposée de U de $L(E_1,\mathcal{a}_1,P_o \circ U^{-1})$ dans $L(E,\mathcal{a},P_o)$, un élément θ_1 de $L(E_1,\mathcal{a}_1,P_o \circ U^{-1})$ appartient à $D(E_1,\mathcal{a}_1,P_o \circ U^{-1})$ si et seulement si, $\int_{E_1} e^{\theta_1}\, dP_o \circ U^{-1} = \int_E e^{\theta_1 \circ U}\, dP_o < \infty$ c'est-à-dire, si $\theta_1 \circ U = {}^tU(\theta_1) \in D(E,\mathcal{a},P_o)$, donc $D(E_1,\mathcal{a}_1,P_o \circ U^{-1}) = {}^tU^{-1}(D(E,\mathcal{a},P_o))$ et $L_{P_o \circ U^{-1}}(\theta_1) = L_{P_o}({}^tU(\theta_1))$.

Posons $\Theta_1 = {}^tU^{-1}(\Theta)$, $\mathcal{P}_1 = \{P_{\theta_1} : \theta_1 \in \dot{\Theta}_1\}$ et

$$P_{\theta_1} = \frac{e^{\theta_1}}{L_{P_o \circ U^{-1}}(\theta_1)} \cdot P_o \circ U^{-1}$$

, de sorte que $(E_1,\mathcal{a}_1,\mathcal{P}_1 = \{P_{\theta_1} ; \theta_1 \in \Theta_1\})$ est une structure exponentielle canonique.

LEMME 1. - $\mathcal{P}_1 \subset \mathcal{P}_U$.

En effet, soit $\theta_1 \in \Theta_1$ et montrons que $P_{\theta_1} \in \mathcal{P}_U$, c'est-à-dire qu'il existe $\theta \in \Theta$ tel que $P_{\theta_1} = P_\theta \circ U^{-1}$. D'après la définition de P_{θ_1} on a :

$$\frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \circ U = \frac{e^{\theta_1 \circ U}}{L_{P_o \circ U^{-1}}(\theta_1)} = \frac{e^{{}^tU(\theta_1)}}{L_{P_o}({}^tU(\theta_1))}$$

posons $\theta = {}^tU(\theta_1) \in \Theta$, alors

(1) $$\frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \circ U = \frac{d P_\theta}{d P_o}$$

donc,

$$E_{P_o}\left(\frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \circ U \middle| U \right) = E_{P_o}\left(\frac{d P_\theta}{d P_o} \middle| U \right) , \quad P_o\text{-p.s.}$$

soit,

$$\frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \circ U = \frac{d P_\theta \circ U^{-1}}{d P_o \circ U^{-1}} , \ P_o\text{-ps.}$$

ce qui implique que :

$$\forall A_1 \in \mathcal{a}_1 , \ P_{\theta_1}(A_1) = \int_{A_1} \frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \, d P_o \circ U^{-1} = \int_{U^{-1}(A_1)} \left(\frac{d P_{\theta_1}}{d P_o \circ U^{-1}} \circ U \right) d P_o$$

$$= \int_{U^{-1}(A_1)} \left(\frac{d P_\theta \circ U^{-1}}{d P_o \circ U^{-1}} \circ U \right) d P_o = \int_{A_1} d P_\theta \circ U^{-1} ;$$

donc $P_{\theta_1} = P_\theta \circ U^{-1}$.

Ce lemme revient donc à montrer que la structure image $(E_1, \mathcal{a}_1, \mathcal{P}_U)$ "contient" la structure exponentielle canonique $(E_1, \mathcal{a}_1, \mathcal{P}_1)$.

CONSEQUENCES :

1°) - Exhaustivité

La relation (1) implique, en utilisant le théorème de factorisation, que la statistique U est exhaustive sur la sous-structure $(E,\mathcal{A},\mathcal{P}' = \{P_{{}^tU(\theta_1)} \; ; \; \theta_1 \in \Theta_1\})$ de $(E,\mathcal{A},\mathcal{P})$, dont l'image par U , est la structure exponentielle canonique $(E_1,\mathcal{A}_1,\mathcal{P}_1 \subset \mathcal{P}_U)$ ce qui peut se schématiser par le diagramme :

$$\begin{array}{ccc} (E,\mathcal{A},\mathcal{P}) & \supset & (E,\mathcal{A},\mathcal{P}') \\ U \downarrow & & \downarrow U \text{ exhaustive} \\ (E_1,\mathcal{A}_1,\mathcal{P}_U) & \supset & (E_1,\mathcal{A}_1,\mathcal{P}_1) \end{array}$$

on obtient comme corollaire :

PROPOSITION 2. - Si ${}^tU$ est surjective de Θ_1 sur Θ , U est exhaustive et la structure image est exponentielle canonique.

2°) - Liberté.

Si U est une statistique libre, alors $\mathcal{P}_U = \{P_o \circ U^{-1}\}$ et le LEMME 1 implique que $\mathcal{P}_1 = \{P_o \circ U^{-1}\}$ c'est-à-dire $\Theta_1 = \{0\} = {}^tU^{-1}(\Theta)$ ce qui est équivalent à $\theta_1 \neq 0 \Rightarrow {}^tU(\theta_1) \notin \Theta$, ou bien :
${}^tU(L(E_1,\mathcal{A}_1,P_o \circ U^{-1})) \cap \Theta = \{0\}$ qui implique que si 0 est un point interne de Θ (ce qui entraîne que Θ est d'intérieur algébrique non vide) ${}^tU(L(E_1,\mathcal{A}_1,P_o \circ U^{-1}) = 0$ soit ${}^tU \equiv 0$ et donc U est la statistique linéaire libre triviale $U \equiv 0$ d'où :

PROPOSITION 3. - Si l'origine de $L(E,\mathcal{A},\mathcal{P})$ et un point interne de Θ , la seule statistique vectorielle linéaire libre est la statistique triviale identiquement nulle.

3°) - Caractérisation.

Reprenant les notations et les résultats du CHAPITRE II, §2, n°4, soit N un élément de la famille $\mathcal{F}(L)$ des sous-espaces vectoriels de dimension finie de $L(E,\mathcal{A},\mathcal{P})$ de dimension n par exemple. En prenant pour U la statistique linéaire vectorielle $\xi_N = (\xi_1,\dots,\xi_n)$ à valeurs dans $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ où $\{\xi_i ; i=1,\dots,n\}$ est une base de N, la structure statistique image est $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n), \{P_\theta \circ \xi_N^{-1} ; \theta \in \Theta\})$ et comme on l'a vu l'application ${}^t\xi_N$ est un isomorphisme de $L(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n), P_o \circ \xi_N^{-1})$ sur N ; par suite $\Theta_N = {}^t\xi_N^{-1}(\Theta)$ est tel que ${}^t\xi_N(\Theta_N) = \Theta \cap N$. Si $\Theta \subset N$, ${}^t\xi_N$ est surjective de Θ_N sur Θ et d'après la PROPOSITION 2 ξ_N est exhaustive, d'où :

PROPOSITION 4. - <u>Si la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P})$ <u>est à paramètre de dimension finie</u> <u>(DEFINITION 4, §1)</u>, <u>il existe au moins une statistique vectorielle linéaire exhaustive à valeurs dans un espace vectoriel mesurable de dimension finie et la structure image et exponentielle canonique</u>.

Cette proposition ramène donc l'étude d'un problème de type exponentiel, dont le paramètre est de dimension finie, dans une structure exponentielle classique en dimension finie, même si l'observation est de dimension infinie comme on l'a déjà vu au §2 . Elle caractérise déjà en partie les structures exponentielles canoniques, car on peut énoncer :

PROPOSITION 5. - <u>Une condition nécessaire et suffisante pour qu'une structure statistique de la forme</u> $(E,\mathcal{A},\mathcal{P} = \{P_\theta, \theta \in \Theta\})$ <u>vérifiant les conditions</u> i) <u>et</u> ii) <u>de la DEFINITION</u> 2, §1 <u>soit une structure exponentielle canonique est que le système projectif</u> :

(2) $$\left(E_N, \mathcal{B}(E_N)\, ,\, \{P_\theta \circ \xi_N^{-1}\, ;\, \theta \in \Theta \cap N\}\right)_{N \in \mathcal{F}(L)}$$

<u>soit constitué de structures statistiques exponentielles canoniques de dimension finie classiques</u>.

En effet, l'élément générique du système projectif (2) est la structure image de la sous-structure exponentielle canonique $(E, \mathcal{a}, \mathcal{P}_N = \{P_\theta\, ;\, \theta \in \Theta \cap N\})$ par la statistique linéaire (qui est alors exhaustive) ξ_N et la condition nécessaire découle de la PROPOSITION 4.

Inversement, comme $\Theta = \bigcup_{N \in \mathcal{F}(L)} \Theta \cap N$, par un argument semblable à celui de la REMARQUE 9, CHAPITRE II, §2,

si $\theta \in \Theta \cap N$, $L_{P_o}(\theta) = \int_E e^\theta\, dP_o = \int_{E_N} e^{t_{\xi_N^{-1}}(\theta)}\, dP_o \circ \xi_N^{-1}$;

par conséquent, la fonctionnelle L_{P_o} est parfaitement déterminée par ses valeurs sur les parties de Θ qui sont intersection de Θ avec les sous-espaces vectoriels de dimension finie de $L(E, \mathcal{a}, \mathcal{P})$ et donc :

$$\forall \theta \in \Theta \cap N\, ,\ \frac{dP_\theta}{dP_o} = \frac{dP_\theta \circ \xi_N^{-1}}{dP_o \circ \xi_N^{-1}} \circ \xi_N = \frac{e^\theta}{L_{P_o}(\theta)}$$

puisque par hypothèse, ξ_N est exhaustive sur la structure $(E, \mathcal{a}, \mathcal{P}_N)$ ci-dessus. ∎

Cette proposition a été énoncée dans le cas explicite dans [34], elle constitue une caractérisation des structures exponentielles canoniques.

3. <u>Espaces de statistiques réelles et complétion</u>.

Soit $V_1(E, \mathcal{a}, \mathcal{P})$ le sous-espace vectoriel de $L_1(E, \mathcal{a}, P_o)$ engendré par la famille de v.a.r. $\{\frac{dP_\theta}{dP_o} = \frac{e^\theta}{L_{P_o}(\theta)}\, ,\ \theta \in \Theta\}$ ou encore par la famille $\{e^\theta\, ;\, \theta \in \Theta\}$.

Posons $U_1(E,\mathcal{A},\mathcal{P}) = \bigcap_{\theta\in\Theta} L_1(E,\mathcal{A},P_\theta)$ qui est l'espace vectoriel des statistiques réelles $\mathcal{P}$-intégrables sur la structure exponentielle canonique $(E,\mathcal{A},\mathcal{P} = \{P_\theta ; \theta\in\Theta\})$.

On a évidemment $L_\infty(E,\mathcal{A},\mathcal{P}) \subset U_1(E,\mathcal{A},\mathcal{P})$; nous savons également, d'après la PROPOSITION 1, que si $\Theta \subset \overset{o}{D}(E,\mathcal{A},P_o)$ supposé non vide, on a $L(E,\mathcal{A},\mathcal{P}) \subset U_1(E,\mathcal{A},\mathcal{P})$.

Pour $\xi \in U_1(E,\mathcal{A},\mathcal{P})$ [resp. $\xi \in L_\infty(E,\mathcal{A},\mathcal{P})$] et $\eta \in V_1(E,\mathcal{A},\mathcal{P})$, l'application $(\xi,\eta) \longmapsto E_{P_o}(\xi.\eta)$ est une forme bilinéaire (notée $\langle .,. \rangle_{\mathcal{P}}$) sur $U_1(E,\mathcal{A},\mathcal{P}) \times V_1(E,\mathcal{A},\mathcal{P})$ [resp. $L_\infty(E,\mathcal{A},\mathcal{P}) \times V_1(E,\mathcal{A},\mathcal{P})$] qui met donc ces deux couples d'espaces respectivement en dualité .

Ces dualités sont séparantes en $V_1(E,\mathcal{A},\mathcal{P})$ c'est-à-dire que :
$\langle \xi,\eta \rangle_{\mathcal{P}} = 0$, $\forall \xi \in U_1(E,\mathcal{A},\mathcal{P})$ [resp. $\forall \xi \in L_\infty(E,\mathcal{A},\mathcal{P})$] $\Rightarrow \eta = 0$ dans $V_1(E,\mathcal{A},\mathcal{P})$ (en effet, en posant $\eta = \sum_{i=1}^{n} \lambda_i \dfrac{dP_{\theta_i}}{dP_o}$, on a pour tout $A \in \mathcal{A}$, dans cette hypothèse, $\sum_{i=1}^{n} \lambda_i \; P_{\theta_i}(A) = 0$ ce qui implique que $\lambda_i = 0$, $i = 1,\dots,n$ si la structure statistique n'est pas dégénérée). Par contre, l'hypothèse $\langle \xi,\eta \rangle_{\mathcal{P}} = 0$, $\forall_\eta \in V_1(E,\mathcal{A},\mathcal{P})$ n'entraîne pas nécessairement $\xi = 0$, ce qui exprime que la dualité considérée n'est pas toujours séparante en $U_1(E,\mathcal{A},\mathcal{P})$. [resp. $L_\infty(E,\mathcal{A},\mathcal{P})$] ; ce fait se traduit d'ailleurs en termes de complétion de la structure statistique :

PROPOSITION 6. - <u>Une condition nécessaire et suffisante pour que la dualité entre</u> $U_1(E,\mathcal{A},\mathcal{P})$ [<u>resp</u>. $L_\infty(E,\mathcal{A},\mathcal{P})$] <u>et</u> $V_1(E,\mathcal{A},\mathcal{P})$ <u>soit séparante est que la structure statistique</u> $(E,\mathcal{A},\mathcal{P})$ <u>soit complète</u> [<u>resp</u>. <u>quasi-complète</u>].

Cette proposition se démontre facilement. ∎

REMARQUE 1. - La dualité entre $L_\infty(E,\mathcal{A},\mathcal{P})$ et $V_1(E,\mathcal{A},\mathcal{P})$ n'est autre que la restriction à ces deux espaces de la dualité classique entre $L_\infty(E,\mathcal{A},P_o)$ et $L_1(E,\mathcal{A},P_o)$; dire que la structure statistique est quasi-complète est donc équivalent à dire que l'espace vectoriel $V_1(E,\mathcal{A},\mathcal{P})$ est dense dans $L_1(E,\mathcal{A},P_o)$

pour la topologie faible $\sigma(L_1, L_\infty)$ ou encore que l'ensemble $\{e^\theta ; \theta \in \Theta\}$ est total dans L_1 pour cette topologie.

REMARQUE 2. - Soient $1 \leq p < \infty$ et q le réel conjugué de p. Si $V_1(E,\mathcal{A},\mathcal{P}) \subset L_p(E,\mathcal{A},P_o)$, ce qui est équivalent à $p\Theta \subset D(E,\mathcal{A},P_o)$, alors $U_1(E,\mathcal{A},\mathcal{P}) \supset L_q(E,\mathcal{A},P_o)$ puisque $\langle \xi,\eta \rangle_{\mathcal{P}} = \langle \xi,\eta \rangle_{L_q,L_p}$ pour tout $\xi \in L_q(E,\mathcal{A},P_o)$; en particulier, si $p \geq 2$ alors $V_1(E,\mathcal{A},\mathcal{P}) \subset L_p(E,\mathcal{A},P_o) \subset L_q(E,\mathcal{A},P_o) \subset U_1(E,\mathcal{A},\mathcal{P}) \subset L_1(E,\mathcal{A},P_o)$ et donc $V_1(E,\mathcal{A},\mathcal{P}) \subset U_1(E,\mathcal{A},\mathcal{P})$.

LEMME 2. - <u>Si la structure exponentielle canonique</u> $(E,\mathcal{A},\mathcal{P})$ <u>est réduite</u> (DEFINITION 4, §1) <u>une condition suffisante pour qu'elle soit complète</u> [<u>resp. quasi-complète</u>] <u>est que pour tout sous-espace vectoriel</u> N <u>de dimension finie de</u> $L(E,\mathcal{A},\mathcal{P})$ <u>la structure exponentielle canonique de dimension finie</u> :

$$(E_N, \mathcal{B}(E_N), \{P_\theta \circ \xi_N^{-1} ; \theta \in \Theta \cap N\})$$

<u>soit complète</u> [<u>resp. quasi-complète</u>].

En effet, avec les notations déjà employées dans la PROPOSITION 5, cela revient à dire que pour tout $N \in \mathcal{F}(L)$, la statistique vectorielle linéaire ξ_N est complète [resp. quasi complète] sur la sous-structure $(E,\mathcal{A},\mathcal{P}_N = \{P_\theta ; \theta \in \Theta \cap N\})$ sur laquelle elle est exhaustive ; ou encore que la structure restreinte $(E,\mathcal{A}_N,\mathcal{P}_N)$ l'est aussi, où $\mathcal{A}_N$ est la tribu (ne dépendant que de N) engendrée par ξ_N définie à la REMARQUE 7, CHAPITRE II, §2, et $\mathcal{P}_N$ la famille des restrictions à cette tribu des $P_\theta ; \theta \in \Theta \cap N$.

Soit $V_1(E,\mathcal{A}_N,\mathcal{P}_N)$ l'espace vectoriel engendré par la famille de v.a.r. $\{e^\theta ; \theta \in \Theta \cap N\}$ c'est bien un sous-espace de $L_1(E,\mathcal{A}_N,P_o)$ puisque si $\theta \in N$, θ est $\mathcal{A}_N$-mesurable.

Comme $\Theta = \bigcup_{N \in \mathcal{F}(L)} \Theta \cap N$ il est facile de montrer que :

(3) $$V_1(E,\mathcal{A},\mathcal{P}) = \bigcup_{N \in \mathcal{F}(L)} V_1(E,\mathcal{A}_N,\mathcal{P}_N) \ .$$

Notons alors $U_1(E,\mathcal{A}_N,\mathcal{P}_N) = \bigcap_{\theta \in \Theta \cap N} L_1(E,\mathcal{A}_N,P_\theta)$.

Supposons que pour tout $N \in \mathcal{F}(L)$, la structure $(E,\mathcal{A}_N,\mathcal{P}_N)$ soit complète [resp. quasi-complète] c'est-à-dire, d'après la PROPOSITION 6, que :

$$(\forall \xi \in U_1(E,\mathcal{A}_N,\mathcal{P}_N) \ [\text{resp. } L_\infty(E,\mathcal{A}_N,\mathcal{P}_N)] \ , \int_E \xi \eta \, dP_o = 0, \forall \eta \in V_1(E,\mathcal{A},\mathcal{P}_N)) \Rightarrow \xi$$

Soit, maintenant $\xi \in U_1(E,\mathcal{A},\mathcal{P})$ [resp. $L_\infty(E,\mathcal{A},\mathcal{P})$] tel que $\int_E \xi\eta \, dP_o = 0$ quel que soit $\eta \in V_1(E,\mathcal{A},\mathcal{P})$; cela implique d'après (3) que quel que soit $N \in \mathcal{F}(L)$, $\int \xi \eta \, dP_o = 0$ quel que soit $\eta \in V_1(E,\mathcal{A}_N,\mathcal{P}_N)$ et donc que $\int_E E^{\mathcal{A}_N}(\xi) \, \eta \, dP_o^E = 0$ quel que soit $\eta \in V_1(E,\mathcal{A}_N,\mathcal{P}_N)$; comme $E^{\mathcal{A}_N}(\xi) \in U_1(E,\mathcal{A}_N,\mathcal{P}_N)$ [resp. $L_\infty(E,\mathcal{A}_N,\mathcal{P}_N)$] , l'hypothèse faite entraîne donc que $E^{\mathcal{A}_N}(\xi) = 0$, $\forall N \in \mathcal{F}(L)$ et par suite $\xi = 0$ dès que $\mathcal{A} \subseteq \bigvee_{N \in \mathcal{F}(L)} \mathcal{A}_N$.

Ce qui est le cas puisque, d'après la REMARQUE 7, CHAPITRE II, §2, $\mathcal{A}_L = \bigvee_{N \in \mathcal{F}(L)} \mathcal{A}_N$. ∎

PROPOSITION 7. - <u>Si l'origine de</u> $L(E,\mathcal{A},\mathcal{P})$ <u>est un point interne de</u> Θ , <u>la structure exponentielle canonique réduite</u> $(E,\mathcal{A},\mathcal{P} = \{P_\theta ; \theta \in \Theta\})$ <u>est complète</u>.

En effet, dire que 0 est un point interne de Θ est équivalent à dire que pour tout sous-espace vectoriel de dimension finie N de $L(E,\mathcal{A},\mathcal{P})$, $\Theta \cap N$ est d'intérieur non vide dans N . La proposition découle alors directement du LEMME 2 et de la condition suffisante très connue pour qu'une structure exponentielle canonique classique (c'est-à-dire en dimension finie) soit complète (cf. [5] Théorème 1 p. 167). ∎

Il est remarquable que cette condition suffisante, soit purement algébrique, puisque la notion du point interne d'un ensemble dans un espace vectoriel, n'est pas une notion topologique, elle implique que Θ est bien dispersé dans toutes les directions de $L(E,\mathcal{A},P)$, (ce qui implique que $D(E,\mathcal{A},P_o)$ l'est aussi), en fait que Θ engendre $L(E,\mathcal{A},P)$, $\mathcal{A}$ est donc égale à $\mathcal{A}_L$.

4. Propriétés de conditionnement.

Soient N_1 et N_2 deux sous-espaces vectoriels supplémentaires de $L(E,\mathcal{A},P)$ de sorte que tout élément $\xi \in L(E,\mathcal{A},P)$ s'écrive de façon unique $\xi = \xi_1(\xi) + \xi_2(\xi)$ avec $\xi_1(\xi) \in N_1$ et $\xi_2(\xi) \in N_2$.

On pose $\Theta_1 = pr_1(\Theta) = \{\theta_1(\theta) \; ; \; \theta \in \Theta\}$.

On note $\mathcal{A}_1$ et $\mathcal{A}_2$ les sous-tribus de $\mathcal{A}$ engendrées respectivement par les éléments de N_1 et N_2 , $P_{\mathcal{A}_1}$ la restriction d'une probabilité P sur $(E,\mathcal{A})$ à la sous-tribu $\mathcal{A}_1$ et $P_{\mathcal{A}_1}^{\mathcal{A}_2}$ la loi de probabilité conditionnelle à $\mathcal{A}_2$, si elle existe, sur $\mathcal{A}_1$.

PROPOSITION 8. - Si $\Theta_1 \subset D(E,\mathcal{A},P_o)$, il existe une probabilité unique P_2 sur $(E,\mathcal{A}_2)$ telle que pour tout $\theta \in \Theta$, en posant $\theta = \theta_1 + \theta_2, (\theta_1 \in N_1 , \theta_2 \in N_2)$ on ait :

$$\frac{d P_{\theta,\mathcal{A}_2}}{d P_2} = \frac{e^{\theta_2}}{L_{P_o}(\theta)} \; ; \; \theta \in \Theta \; .$$

En effet, pour tout $A_2 \in \mathcal{A}_2$, $P_{\theta,\mathcal{A}_2}(A_2) = \frac{1}{L_{P_o}(\theta)} \int_E 1_{A_2} \; e^{\theta_1+\theta_2} \, d P_o$

$$= \frac{1}{L_{P_o}(\theta)} \int_{A_2} e^{\theta_2} \, E_{P_o}^{\mathcal{A}_2}(e^{\theta_1}) d P_{o,\mathcal{A}_2}$$

puisque θ_2 est $\mathcal{A}_2$-mesurable et que $E_{P_o}^{\mathcal{A}_2}(e^{\theta_1})$ est bien définie dès que

$\theta_1 \in D(E,\mathcal{A},P_o)$; il suffit alors de poser : $P_2 = E_{P_o}^{\mathcal{A}_2}(e^{\theta_1}) \cdot P_{o,\mathcal{A}_2}$. ■

PROPOSITION 9. - <u>Si</u> $\Theta_1 \subset D(E,\mathcal{A},\mathcal{P})$ <u>et si la probabilité conditionnelle</u> $P_{o,\mathcal{A}_1}^{\mathcal{A}_2}$ <u>est régulière on a</u>, <u>en posant</u> : $\theta = \theta_1 + \theta_2$:

$$\frac{d P_{\theta,\mathcal{A}_1}^{\mathcal{A}_2}}{d P_{o,\mathcal{A}_1}^{\mathcal{A}_2}} = \frac{e^{\theta_1}}{L_{P_{o,\mathcal{A}_1}}^{\mathcal{A}_2}(\theta_1)} \text{ ,p-s. } ; \quad \theta \in \Theta .$$

En effet, soit $A_1 \in \mathcal{A}_1$; pour tout $A_2 \in \mathcal{A}_2$ on doit avoir :

$$P_\theta(A_1 \cap A_2) = \int_{A_2} P_{\theta,\mathcal{A}_1}^{\mathcal{A}_2}(A_1)\, d P_{\theta,\mathcal{A}_2} .$$

Comme $P_\theta \ll P_o$, $P_{\theta,\mathcal{A}_1}^{\mathcal{A}_2} \ll P_{o,\mathcal{A}_1}^{\mathcal{A}_2}$ p.s , car si $A_1 \in \mathcal{A}_1$ est tel que

$P_{o,\mathcal{A}_1}^{\mathcal{A}_2}(A_1) = 0$ p.s. alors $P_o(A_1) = E_{P_o}(P_o^{\mathcal{A}_2}(A_1)) = 0$

mais on a :

$$P_{\theta,\mathcal{A}_1}^{\mathcal{A}_2}(A_1) = \frac{E_{P_o}^{\mathcal{A}_2}\left(\frac{d P_\theta}{d P_o} \cdot 1_{A_1}\right)}{E_{P_o}^{\mathcal{A}_2}\left(\frac{d P_\theta}{d P_o}\right)}$$

et comme $\frac{d P_\theta}{d P_o} \cdot 1_{A_1} = 0$, P_o-p.s. , le numérateur est une v.a.r. nulle.

soit f_θ la densité conditionnelle : $\frac{d P_{\theta,\mathcal{A}_1}^{\mathcal{A}_2}}{d P_{o,\mathcal{A}_1}^{\mathcal{A}_2}}$; en tenant compte de la

PROPOSITION 8,

$$(4) \qquad P_\theta(A_1 \cap A_2) = \int_{A_2} \left[\int_{A_1} f_\theta \, dP^{\mathcal{A}_2}_{o,\mathcal{A}_1}\right] \frac{e^{\theta_2} E^{\mathcal{A}_2}_{P_o}(e^{\theta_1})}{L_{P_o}(\theta)} \, dP_{o,\mathcal{A}_2} .$$

D'autre part, on a :

$$(5) \qquad \int_{A_1 \cap A_2} \frac{e^{\theta_1 + \theta_2}}{L_{P_o}(\theta)} \, dP_o = \int_{A_2} \frac{e^{\theta_2} E^{\mathcal{A}_2}_{P_o}(e^{\theta_1} . 1_{A_1})}{L_{P_o}(\theta)} \, dP_{o,\mathcal{A}_2} .$$

En identifiant (4) et (5), on obtient :

$$E^{\mathcal{A}_2}_{P_o}(e^{\theta_1} . 1_{A_1}) = E^{\mathcal{A}_2}_{P_o}(e^{\theta_1}) \int_{A_1} f_\theta \, dP^{\mathcal{A}_2}_{o,\mathcal{A}_1}$$

on constate que cette équation est vérifiée dès que l'on prend pour f_θ

$$f_\theta = \frac{e^{\theta_1}}{L_{P^{\mathcal{A}_2}_{o,\mathcal{A}_1}}(\theta_1)}$$

car le deuxième membre vaut alors : $\dfrac{E^{\mathcal{A}_2}_{P_o}(1_{A_1} e^{\theta_1})}{L_{P^{\mathcal{A}_2}_{o,\mathcal{A}_1}}(\theta_1)} E^{\mathcal{A}_2}_{P_o}(e^{\theta_1}) = E^{\mathcal{A}_2}_{P_o}(e^{\theta_1} . 1_{A_1})$

car $L_{P^{\mathcal{A}_2}_{o,\mathcal{A}_1}}(\theta_1) = \int_E e^{\theta_1} \, dP^{\mathcal{A}_2}_{o,\mathcal{A}_1} = E^{\mathcal{A}_2}_{P_o}(e^{\theta_1}) < \infty$ p.s puisque $\Theta_1 \subset D(E,\mathcal{A},P_o)$

la proposition résulte alors de l'unicité de f_θ. ■

CORROLLAIRE. - <u>Si</u> N_1 <u>est de dimension finie et si</u> $\Theta_1 \subset D(E,\mathcal{A},P_o)$, <u>la sous-tribu</u> $\mathcal{A}_1$ <u>est libre conditionnellement à</u> $\mathcal{A}_2$, <u>par rapport au paramètre</u> θ_2 <u>et pour tout</u> $x \in E$ <u>la structure statistique conditionnelle à</u> $\mathcal{A}_2$, <u>restreinte à</u> $\mathcal{A}_1$

$$(E,\mathcal{A}_1, \{P^{\mathcal{A}_2}_{\theta,\mathcal{A}_1} ; \theta \in \Theta\})$$

<u>est de type exponentiel canonique et admet pour paramètre naturel</u> $\theta_1 \in \Theta_1$.

Il suffit de constater que si N_1 est de dimension finie, $\mathcal{a}_1$ est la tribu engendrée sur E par un vecteur aléatoire réel de dimension finie (par exemple, une base de N_1) et les conditions du théorème de JIRINA sont remplies, qui font que la probabilité conditionnelle $P_{\theta,\mathcal{a}_1}^{\mathcal{a}_2}$ est régulière pour tout $\theta \in \Theta$. Il suffit alors d'appliquer la PROPOSITION 9 qui assure que :

$$\frac{d P_{\theta,\mathcal{a}_1}^{\mathcal{a}_2}}{d P_{o,\mathcal{a}_1}^{\mathcal{a}_2}} = \frac{e^{\theta_1}}{L_{P_{o,\mathcal{a}_1}^{\mathcal{a}_2}}(\theta_1)}$$

ce qui montre à la fois que la loi de probabilité conditionnelle $P_{\theta,\mathcal{a}_1}^{\mathcal{a}_2}$ ne dépend pas de θ_2 et que la structure conditionnelle ainsi construite est de type exponentiel canonique. ■

<u>Conséquences</u> : Si dans un problème de statistique sur une structure exponentielle canonique, le paramètre θ_2 tient lieu de paramètre nuisible, et s'il existe une statistique engendrant $\mathcal{a}_2$, conditionnellement à celle-ci, la structure obtenue ne dépend plus de θ_2 et de plus est une structure exponentielle de dimension finie classique.

En particulier si N_1 est de dimension 1, on peut essayer de généraliser la méthode des tests conditionnels de LEHMANN qui permettrait de trouver des tests optimaux pour un paramètre scalaire en présence d'une infinité de paramètres nuisibles.

§ 4. - ESTIMATIONS CONCERNANT LE PARAMETRE NATUREL D'UNE STRUCTURE EXPONENTIELLE.

Dans les problèmes pratiques, il peut être utile d'estimer le paramètre naturel (ou certaines de ses fonctionnelles) d'une structure exponentielle ; comme celui-ci n'apparaît qu'après réduction par la statistique vectorielle exhaustive de référence, on supposera donnée dans tout ce § , une structure exponentielle canonique $(E,\mathcal{a},\mathcal{P} = \{P_\theta ; \theta \in \Theta\})$ munie de son paramètre naturel, telle qu'elle est décrite dans la DEFINITION 2, §1 (ou dans la DEFINITION 7, §1 si elle est explicite). On étudiera successivement l'estimation sans biais des fonctionnelles de ce paramètre et l'estimation par la méthode du maximum de vraisemblance du paramètre lui-même, cette dernière n'ayant de sens que lorsque l'on dispose d'une version de la fonction de vraisemblance, c'est-à-dire dans le cas d'une structure explicite.

1. Estimation sans biais.

DEFINITION 1. - Une fonction réelle f définie sur Θ sera dite "admissible par rapport à P_{θ_o}", $\theta_o \in \Theta$, s'il existe un estimateur sans biais de f, de variance finie pour la loi P_{θ_o}, c'est-à-dire, une statistique $\xi \in U_1(E,\mathcal{a},\mathcal{P})$ telle que :

(1) $$\text{i - } \int_E \xi \, dP_\theta = f(\theta) \ , \ \forall \theta \in \Theta \ ,$$

$$\text{ii - } \sigma^2_{P_{\theta_o}}(\xi) < \infty \ .$$

La fonction f sera dite "uniformément admissible" si de plus $\sigma^2_{P_\theta}(\xi) < \infty$ pour tout $\theta \in \Theta$.

Choisissons par exemple $\theta_o = 0$:

PROPOSITION 1. - Si la structure exponentielle canonique $(E,\mathcal{a},\mathcal{P} = \{P_\theta ; \theta \in \Theta\})$ est réduite, telle que $\Theta + \Theta \subset D(E,\mathcal{a},P_o)$ et telle que l'origine de $L(E,\mathcal{a},\mathcal{P})$ soit un point interne de Θ, une condition nécessaire et suffisante pour qu'une fonctionnelle f soit admissible par rapport à P_o est que la fonctionnelle :

$f.L_{P_o}$ appartienne à l'espace de HILBERT autoreproduisant $\mathcal{H}(K_\Theta)$ de fonctionnelles définies sur Θ, associé au noyau K_Θ défini sur $\Theta \times \Theta$ par $K_\Theta(\theta,\theta') = L_{P_o}(\theta+\theta')$; $\theta,\theta' \in \Theta$. Dans ce cas, l'estimateur sans biais de f est unique et donné par la relation : $\xi = \lambda_\Theta^{-1}(f.L_{P_o})$ où λ_Θ est une isométrie de $L_2(E,\mathcal{a},P_o)$ sur $\mathcal{H}(K_\Theta)$, tandis que sa variance par rapport à P_o vaut :

$$\sigma^2_{P_o}(\xi) = \|f.L_{P_o}\|^2_{\mathcal{H}(K_\Theta)} - \langle f L_{P_o}, L_{P_o}\rangle^2_{\mathcal{H}(K_\Theta)} .$$

Si de plus, $\xi^2 \in L_2(E,\mathcal{a},P_o)$, f est uniformément admissible, ξ est l'estimateur sans biais de f, de variance minimum :

$$\sigma^2_{P_\theta}(\xi) = \frac{[\lambda_\Theta(\xi^2)](\theta)}{L_{P_o}(\theta)} - [f(\theta)]^2 .$$

En effet, d'après la REMARQUE 2, §3, la condition $\Theta+\Theta \subset D(E,\mathcal{a},P_o)$ implique que $V_1(E,\mathcal{a},\mathcal{P}) \subset L_2(E,\mathcal{a},P_o) \subset U_1(E,\mathcal{a},P_o)$ ce qui permet d'envisager l'existence de fonctionnelles admissibles par rapport à P_o, de plus elle entraîne que la fonction K_Θ définie sur $\Theta \times \Theta$ par $K_\Theta(\theta,\theta') = L_{P_o}(\theta+\theta')$, $\theta,\theta' \in \Theta$ est un noyau symétrique de type positif sur Θ, comme on l'a vu à la REMARQUE 1 du CHAPITRE II, §3. Par un argument analogue à celui de la PROPOSITION 4 qui la précède, il existe donc un espace de HILBERT réel autoreproduisant $\mathcal{H}(K_\Theta)$ associé à ce noyau, constitué de fonctionnelles définies sur Θ et tel que :

1°) les fonctionnelles $K_\Theta(.,\theta) = L_{P_o}(.+\theta)$ appartiennent à $\mathcal{H}(K_\Theta)$ et l'engendrent quand θ parcourt Θ.

2°) Si $f \in \mathcal{H}(K_\Theta)$, $f(\theta) = \langle f, K_\Theta(.,\theta)\rangle_{\mathcal{H}(K_\Theta)}$, $\theta \in \Theta$.

Si l'origine de $L(E,\mathcal{a},\mathcal{P})$ est un point interne de Θ, un argument analogue à celui de la PROPOSITION 3, CHAPITRE II, §3, implique que $V_1(E,\mathcal{a},\mathcal{P})$ est dense dans $L_2(E,\mathcal{a},P_o)$. La transformation λ_Θ définie par :

$$\xi \in L_2(E,\mathcal{a},P_o) , \quad [\lambda_\Theta(\xi)](\theta) = \int_E \xi e^\theta \, dP_o , \quad \theta \in \Theta$$

est une isométrie de l'espace de HILBERT $L_2(E,\mathcal{a},P_o)$ sur l'espace de HILBERT $\mathcal{H}(K_\Theta)$ puisque $\lambda_\Theta(e^\theta)(.) = K_\Theta(.,\theta)$. Remarquons en outre que cette condition implique, d'après la PROPOSITION 7, §3, que la structure est complète.

La relation (1) de la DEFINITION 1, s'écrit alors : $[\lambda(\xi)](\theta) = f(\theta)\, L_{P_o}(\theta),\ \theta\in\Theta$; elle implique que pour que la fonctionnelle f soit admissible par rapport à P_o il faut et il suffit que le produit $f.L_{P_o}$ soit dans $\mathcal{H}(K_\Theta)$, auquel cas, $\xi = \lambda_\Theta^{-1}(f\, L_{P_o})$ est uniquement déterminé par f ; on a alors,

$$\sigma^2_{P_o}(\xi) = \int_E \xi^2\, dP_o - [E_{P_o}(\xi)]^2 = \|\xi\|^2_{L_2} - [f(0)]^2$$

$$= \|f.L_{P_o}\|^2_{\mathcal{H}(K_\Theta)} - \langle f.L_{P_o}, L_{P_o}\rangle^2_{\mathcal{H}(K_\Theta)} .$$

Si de plus, $\xi^2 \in L_2(E,\mathcal{a},P_o)$, $E_{P_\theta}(\xi^2) = \int_E \xi^2 \frac{e^\theta}{L_{P_o}(\theta)}\, dP_o < \infty,\ \forall\theta\in\Theta$

donc $$\sigma^2_{P_\theta}(\xi) = \int_E \xi^2 \frac{e^\theta}{L_{P_o}(\theta)}\, dP_o - [f(\theta)]^2 = \frac{\lambda_\Theta(\xi^2)(\theta)}{L_{P_o}(\theta)} - [f(\theta)]^2 . \blacksquare$$

2. <u>Estimation par la méthode du maximum de vraisemblance</u>.

Comme on l'a dit au début de ce §, il est nécessaire de supposer que la structure est explicite, c'est-à-dire qu'elle admet pour fonction de vraisemblance :

$$\mathcal{L}(\theta;x) = \frac{e^{\langle x,\theta\rangle}}{l_{P_o}(\theta)} \quad , \quad x\in E \ , \ \theta\in\Theta\subset(E,\mathcal{a})^m$$

ou encore :

(1) $$\mathcal{L}(\theta;x) = \exp\{\langle x,\theta\rangle - c_{P_o}(\theta)\} \ , \ x\in E \ , \ \theta\in\Theta ,$$

où, avec les notations du CHAPITRE II, §1, n°4,5, c_{P_o} désigne la fonction cumulante de la probabilité de référence P_o sur l'E.V.M.S. $(E,\mathcal{a})$.

Nous supposerons, pour simplifier que cette structure exponentielle canonique est saturée c'est-à-dire que Θ est identique au convexe D_{P_o}

de $(E,\mathcal{A})^m$ domaine de définition de la transformée de LAPLACE réelle (ou de la fonction cumulante) de P_o .

L'estimation par la méthode du maximum de vraisemblance consiste à déterminer la ou les valeurs du paramètre θ , si elles existent, pour lesquelles cette fonction atteint son maximum sur D_{P_o} . L'ensemble $\hat{\theta}$ (éventuellement vide) de ces valeurs dépend donc de l'observation x .

DEFINITION 2. - On appelle statistique de maximum de vraisemblance, la multiapplication : $x \longmapsto \hat{\theta}(x)$ définie sur E , à valeurs dans $\mathfrak{P}(D_{P_o})$. Si pour tout $x \in E$, l'ensemble $\theta(x)$ est réduit à un seul point, on appelle cette application de E dans $D_{P_o} = \Theta$, estimateur de maximum de vraisemblance de θ .

La méthode de BARNDORFF-NIELSEN (cf. [4]) , qui consiste à remarquer que cette estimation se ramène à un problème d'optimisation convexe, s'étend sans difficulté en dimension infinie :

On constate en effet, qu'une valeur $\hat{\theta}$ du paramètre rend maximum la fonction (1) si et seulement si : $\langle x,\hat{\theta}\rangle - c_{P_o}(\hat{\theta}) = \operatorname*{Sup}_{\theta \in D_{P_o}} \{\langle x,\theta\rangle - c_{P_o}(\theta)\}$

c'est-à-dire si $c^*_{P_o}(x) + c_{P_o}(\hat{\theta}) = \langle x,\hat{\theta}\rangle$. Ce qui exprime que les points $x \in E$ et $\hat{\theta} \in D_{P_o}$ sont conjugués par rapport au couple de fonctions convexes duales $(c_{P_o}, c^*_{P_o})$.

Rappelons brièvement (cf. [20]) qu'une fonction $h \in \overline{\mathbb{R}}^E$ est dite sous-différentiable au point $x_o \in E$, s'il existe une fonction affine continue, prenant la même valeur qu'elle au point x_o et la minorant sur E ; une telle minorante peut s'écrire : $x \longmapsto \langle x - x_o , y_o\rangle + h(x_o)$. On dit que la pente $y_o \in (E,\mathcal{A})^m$ de cette fonction est un sous-gradient de h au point x_o ; l'ensemble (éventuellement vide) des sous-gradients de h au point x_o est appelé "sous différentiel de la fonction en ce point et noté $\partial h(x_o)$" .

Si h est convexe et si elle est finie et continue au point $x_o \in E$

pour une topologie compatible avec la dualité, l'ensemble $\partial h(x_o)$ est une partie faiblement compacte non vide de $(E,\mathfrak{a})^m$. Si h et g sont deux fonctions duales, pour un couple de points $x \in E$, $y \in (E,\mathfrak{a})^m$ les trois propriétés suivantes sont équivalentes $y \in \partial h(x)$; $x \in \partial g(y)$; $h(x)+g(y) = \langle x,y \rangle$.

En appliquant cette dernière propriété à notre problème, compte tenu de ce qui précède, on en déduit :

PROPOSITION 2. - <u>L'ensemble des valeurs du paramètre qui sont de maximum de vraisemblance</u>, <u>au vu de l'observation</u> $x \in E$ <u>est le sous-différentiel</u> $\partial c^*_{P_o}(x)$ <u>de la polaire de la fonction cumulante de</u> P_o <u>au point</u> x , <u>et donc la statistique de maximum de vraisemblance est la multiapplication</u> : $x \longmapsto \partial c^*_{P_o}(x)$.

Il est important de savoir dans quels cas, cette statistique est un estimateur ponctuel de maximum de vraisemblance, c'est-à-dire quand $\partial c^*_{P_o}(x)$ contient au plus un point.

Rappelons encore que si $h \in \overline{\mathbb{R}}^E$ et $x_o \in E$ est tel que $h(x_o) < \infty$, en notant (comme au CHAPITRE II, §2, REMARQUE 6) $h'(x_o,x)$ la dérivée de h , au point x_o selon la direction $x \in E$, la fonction $h_o : x \longmapsto h(x_o,x)$ est définie pour tout x et convexe si h est convexe ; si elle est linéaire et faiblement continue, c'est-à-dire si elle peut s'écrire $h_o(x) = \langle x,y_o \rangle$, on dit que h est <u>faiblement différentiable</u> en x_o , l'élément $y_o \in (E,\mathfrak{a})^m$ est son <u>gradient</u>. Alors si h est convexe et faiblement différentiable au point x_o , le sous-différentiel $\partial h(x_o)$ consiste en l'unique élément y_o , gradient de h en ce point. Inversement, si h est convexe, finie et continue au point x_o et si le sous-différentiel $\partial h(x_o)$ contient un seul élément $y_o \in (E,\mathfrak{a})^m$ la fonction h est faiblement différentiable au point x_o et y_o est son gradient ; d'où :

PROPOSITION 5. - <u>Une condition suffisante pour qu'il existe une estimation de maximum de vraisemblance unique du paramètre</u> θ , <u>au vu de l'observation</u> x <u>est que</u> $c^*_{P_o}$ <u>soit faiblement différentiable au point</u> x <u>et cette estimation est le gradient de cette fonction en ce point</u>.

On constate l'importance du lien qui existe entre ce problème d'estimation et les problèmes développés dans le n°5 du CHAPITRE II, §2 concernant la caractérisation du support d'une probabilité borélienne sur un E.V.T. au moyen de l'analyse convexe. Dans le cas particulier d'une probabilité gaussienne, il suffit de se reporter au CHAPITRE III, §1, n°3 , notamment à la PROPOSITION 1, pour construire une illustration de la théorie ci-dessus :

EXEMPLE 1. - Soit E un espace de FRECHET séparable et P_o la probabilité gaussienne centrée sur $(E,\mathcal{B}(E))$ de variance la forme quadratique positive Q sur le dual topologique E' de E $[(E,\mathcal{B}(E))^m = E']$. Considérons la structure exponentielle canonique explicite saturée :

$$(E,\mathcal{B}(E) \;,\; \{P_\theta \,;\, \theta \in \Theta = E'\})$$

où :

$$\mathcal{L}(\theta,x) = \frac{dP_\theta}{dP_o}(x) = \frac{e^{\langle x,\theta\rangle}}{l_{P_o}(\theta)} = e^{\langle x,\theta\rangle - \frac{1}{2}Q(\theta)} \quad ; x\in E \;,\; \theta\in E'$$

qui est la forme canonique d'une structure gaussienne à moyenne inconnue de la forme $m_\theta = K(.,\theta) \in E$, où K est la forme bilinéaire symétrique associée à Q puisque d'après la PROPOSITION 1, §3, la transformée de FOURIER de P_θ est :

$$\varphi_{P_\theta}(y) = \frac{\tilde{l}_{P_o}(\theta+iy)}{l_{P_o}(\theta)} \quad , \; y \in E'$$

soit ici

$$\varphi_{P_\theta}(y) = e^{iK(\theta,y) - \frac{1}{2}Q(y)}$$

qui exprime que P_θ est la loi de probabilité gaussienne dont la moyenne m_θ et de la forme indiquée et parcourt l'espace vectoriel image de E' par l'application : $y \longmapsto K(.,y)$ de E' , qui est un sous-espace dense dans le sous-espace hilbertien $H(Q)$ de E associé à Q . La proposition citée ci-dessus indique que : $c^*_{P_o}(x) = \frac{1}{2}\|x\|^2_{H(Q)}$, si $x \in H(Q)$ et

$c^*_{P_o}(x) = +\infty$ sinon. D'autre part l'espace $H(Q)$ est de P_θ-probabilité nulle pour tout $\theta \in E'$ s'il est de dimension infinie, il s'ensuit que la fonction $c^*_{P_o}$ est presque-sûrement infinie pour l'observation x , et donc il n'existe pas d'estimateur de maximum de vraisemblance pour ce problème, ce qui a été déjà démontré de façon différente par divers auteurs. (cf. [28], par ex.) ∎

Si on suppose que la structure n'est pas saturée mais reste convexe, c'est-à-dire que Θ est un convexe contenu dans D_{P_o} , en notant ψ_Θ l'indicatrice convexe de Θ , le problème du maximum de vraisemblance revient à calculer :

$$\sup_{\theta \in (E,\mathcal{A})^m} \{\langle x,\theta\rangle - c_{P_o}(\theta) - \psi_\Theta(\theta)\} = (c_{P_o} + \psi_\Theta)^* = c^*_{P_o} \nabla \psi^*_\Theta$$

où ∇ est l'opération d'inf-convolution ; il suffit donc de remplacer $c^*_{P_o}$ par $c^*_{P_o} \nabla \psi^*_\Theta$ dans ce qui précède. ∎

§ 5. - CONCLUSIONS SUR LES STRUCTURES EXPONENTIELLES GENERALISEES.

Cette théorie des structures exponentielles généralisées présente à notre avis deux avantages :

D'une part, elle fournit un modèle mathématique précis pour des problèmes de statistique portant sur des observations vectorielles de dimension infinie, qui garde les propriétés analytiques essentielles des structures exponentielles classiques. Les nombreux exemples d'application que nous en avons donnés le prouvent. C'était d'ailleurs, à l'origine, le but de cette généralisation.

D'autre part, elle peut contribuer à éclairer de façon nouvelle certains concepts fondamentaux de la statistique. En effet, en admettant de paramétrer une famille de lois de probabilité par un vecteur de dimension infinie, on supprime le clivage habituel entre la statistique dite paramétrique et la statistique non paramétrique. La REMARQUE 4, §1 le prouve bien, qui exprime que l'on peut considérer toute structure statistique d'échantillon, dont les hypothèses sont équivalentes, comme une structure exponentielle et que l'étude de la loi de probabilité empirique (qui est la statistique exhaustive) a pour cadre naturel une structure exponentielle canonique, à savoir, la structure multinomiale généralisée.

De plus, la propriété caractéristique de cohérence des structures exponentielles, à savoir que si le paramètre est de dimension finie il existe nécessairement une statistique vectorielle exhaustive de dimension finie est un argument supplémentaire en faveur de l'utilisation de cet outil analytique en statistique et va dans le sens d'une unification de la théorie des structures statistiques.

On peut envisager, en particulier, d'aborder certains problèmes réputés "non paramétriques" (estimation de densité, par exemple), avec des méthodes paramétriques notamment celles qui sont développées au §4, ce qui est dans la ligne des tendances actuelles de la statistique mathématique (cf. [11], par exemple). ■

CONCLUSION.

Comme on le constate, ce travail constitue une introduction à une nouvelle façon d'aborder certains problèmes de statistique, en particulier ceux concernant les processus stochastiques de toutes sortes, et qui consiste à les situer dans le cadre d'une généralisation en dimension infinie de l'analyse multidimensionnelle classique. Il se situe, en effet, en amont de tout problème de décision à partir de l'observation d'un vecteur aléatoire de dimension infinie.

Toutefois, dans son souci de généralité, il contribue à apporter une unité dans la présentation des structures statistiques attachées à ces problèmes, qui nous semble constituer une étape indispensable, préalable à leur résolution.

Il permet ensuite de constater que cette généralisation n'est pas toujours immédiate qui, outre les problèmes difficiles auxquels nous confronte l'étude des lois de probabilité de tels vecteurs ou autres problèmes d'analyse mathématique, nécessite bien souvent l'investigation d'outils nouveaux : Par exemple, la volonté d'éviter toute hypothèse topologique à priori sur l'espace des observations (ce qui est naturel pour un statisticien) nous conduit à mettre en évidence la nécessité d'une étude intrinsèque de la notion d'espace vectoriel mesurable et révèle l'utilité de l'outil que constitue le dual d'un tel espace. Il en est de même de la notion de dual d'un espace vectoriel probabilisé qui s'avère être l'outil fondamental dans la théorie des structures exponentielles généralisées.

Il se produit le même phénomène en ce qui concerne les problèmes de décision. Parmi ceux que nous avons abordés et pour ne parler que des structures statistiques gaussiennes à moyenne inconnue on aura pu constater la faillite d'une extension directe des méthodes classiques utilisées en dimension finie. C'est le cas par exemple de l'estimation de maximum de vraisemblance ou encore de la théorie des tests d'hypothèses linéaires. Il nous semble donc que ces problèmes de décision concernant un paramètre de dimension infinie doivent être également abordés avec des méthodes nouvelles qui restent à concevoir, ce qui constituerait d'ailleurs un prolongement naturel au présent travail.

L'objectif que nous nous sommes fixé nous restreint, dans ses applications à la statistique des processus stochastiques, à une conception globale de leurs observations : ainsi, nous n'avons pas abordé les méthodes de discrétisation ou de variation du domaine d'observation et donc les problèmes asymptotiques auxquels elles conduisent.

Enfin, parmi les aspects positifs, on peut souligner que paradoxalement, certains résultats obtenus ont quelques répercussions en statistique classique. Il en est ainsi des formules obtenues pour la fonction caractéristique de formes quadratiques de certaines f.a.r. gaussiennes décentrées qui contribuent à l'amélioration de l'étude asymptotique de certains tests d'adéquation. Dans un domaine plus théorique, la notion de structure exponentielle généralisée conduit à une réflexion sur les fondements de la statistique non paramétrique, qui dans l'optique du prolongement à ce travail indiqué plus haut, permet d'envisager la résolution de certains de ses problèmes par des méthodes nouvelles. ∎

BIBLIOGRAPHIE

[1] AHMAD, S. - "Eléments aléatoires dans les espaces vectoriels topologiques". Ann. Inst. H. POINCARE, Sec.B, Vol II, n°2, 1965, p.95-135.

[2] ANDERSON, T.W. DARLING, D.A. - "Asymptotic theory of certain goodness of fit criteria based on stochastic processes". Ann. Math. Stat., 23, n°2, 1965, p.193-212.

[3] BADRIKIAN, A. - "Séminaire sur les fonctions aléatoires linéaires et les mesures cylindriques". Lecture Notes in Math. n°139, 1970, Springer V.

[4] BARNDORFF-NIELSEN, O. - "Exponential families. Exact theory". Various publications Séries n°19, 1970, Aarhus Univ.

[5] BARRA, J.R. - " Notions fondamentales de statistique mathématique". Dunod, Paris, 1971.

[6] BOURBAKI, N. - "Eléments de mathématique. Espaces vectoriels topologiques, chap. III, IV, V". Hermann, Paris.

[7] BOURBAKI, N. - "Eléments de mathématique. Algèbre, chap. III". Hermann, Paris, 1958.

[8] BOURBAKI, N. - "Eléments de mathématique. Intégration, chap. IX". Hermann, Paris, 1970.

[9] CHIBISOV, D.M. - "An investigation of the asymptotic power of the tests of fit". Theory of Prob. and appl. Vol X, n°3, 1965, p.421-437.

[10] DURBIN, J. - "Weak convergence of the sample distribution function when parameters are estimated". Ann. Stat. Vol.1, n°2, 1973, p.279-290.

[11] FERGUSON, T.S. - "A bayesian analysis of some non parametric problems". Ann. Stat. Vol.1, n°2, p.209-230.

[12] GRENANDER, U. - "Stochastic processes and statistical inference". Arkiv für matematik, Band 1, n°17, 1950.

[13] GROTHENDIECK, A. - "La théorie de Fredholm". Bull. Soc.math. France, 84, 1956, p.319-384.

[14] HIDA, T. - "Stationary stochastic processes". Princeton Univ. Press., Princeton, 1970.

[15] HIDA, T. - "Quadratic functionals of brownian motion". Journ. Multivariate Analysis, 1, 1971, p.58-69.

[16] IBRAHIMOV, I.Z.
ROZANOV, Y.V. - "Processus aléatoires gaussiens". Editions Mir, Moscou, 1974.

[17] JOHANSEN, S. - "Homomorphisms and general exponential families". Congrès Européen des Statisticiens, Grenoble 1976 in Recent developments in Statistics, North-Holl. Publ.,1977.

[18] MAC NEILL, I.B. - "Tests for change of parameter at unknown times and distributions of some related functionals of Brownian motion". Ann. Stat. Vol.2, n°5, 1974, p.950-962.

[19] MEYER, P.A. - "Probabilités et potentiel". Hermann, Paris, 1966.

[20] MOREAU, J.J. - "Séminaire sur les équations aux dérivées partielles, II, Fonctionnelles convexes". Collège de France, multigraphié 1966-67.

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

[22] NEVEU, J. - "Processus aléatoires gaussiens". Séminaire de mathématiques supérieures, Montréal 1968.

[23] NEVEU, J. - "Intégrales stochastiques". Cours de 3ème cycle, notes Université de Paris VI, 1972.

[24] NEVEU, J. - "Processus ponctuels". Ecole d'été de St Flour, 1976.

[25] PADGETT, W.J. Z TAYLOR, R.L. - "Laws of large numbers for normed linear spaces and certain Fréchet spaces". Lecture Notes in Math. n°360, 1973, Springer V.

[26] PARTHASARATHY, K.R. - "Probability measures on metric spaces". Academic Press, N.Y., London, 1967.

[27] RAJPUT, B.S. - " On gaussian measures on certain locally convex spaces". Journ. Multivariate Analysis, 2, 1972, p.282-306.

[28] ROZANOV, Ju.A. - "Infinite-dimensional Gaussian distributions". Amer. Math. Soc., Providence, 1971.

[29] ROTHMAN, E.D. Z WOODROOFE, M. - "A CRAMÉR-Von MISES type statistic for testing symmetry". Ann. Math. Stat. Vol.43, n°6, 1972, p.2035-2038.

[30] SCHWARTZ, L. - "Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés". Journal d'analyse math. Jerusalem 1964.

[31] SCHAEFER, H.H. - "Topological vector spaces". Springer V., 1970.

[32] SIMONS, G. - "Generalized distribution functions : the linearly ordered case with applications to non parametric statistics". Ann.Prob. Vol.2, n°3, 1974, p.501-508.

[33] SKOROHOD, A.V. - "Integration in Hilbert space". Springer V., 1974.

[34] SOLER, J.L. - "m-dual d'un espace vectoriel mesurable et structures statistiques de type exponentiel". C.R. Acad. Sc. Paris t.277, Ser.A 1973, p.49-52.

[35] SOLER, J.L. - "La loi de probabilité de WISHART associée à un élément aléatoire gaussien décentré à valeurs dans un espace de Fréchet séparable. Application aux fonctions aléatoires gaussiennes". C.R. Acad. Sc. Paris t.281, Ser.A, 1975, p.471-474.

[36] SOLER, J.L. - "Infinite-dimensional exponential type statistical spaces. (Generalized exponential families)". Congrès européen des statisticiens, Grenoble 1976 in Recent developments in Statistics, North-Holland Publ. co. 1977.

[37] SOLER, J.L. - "Some results for the quadratic analysis of gaussian processes and applications". International conference on mathematical statistics, WISLA, POLOGNE, Déc. 1976 (à paraître).

[38] VARBERG, D. - "On gaussian measures equivalent to WIENER Measure". Trans. AM5 113, 1964, p.262-273.

[39] YEH, J. - "Stochastic processes and WIENER integrals". M. Dekker, inc. N.Y., 1973. ■