



HAL
open science

Travaux de statistique

Donatien Niango

► **To cite this version:**

Donatien Niango. Travaux de statistique. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG; Université Joseph-Fourier - Grenoble I, 1978. Français. NNT: . tel-00288723

HAL Id: tel-00288723

<https://theses.hal.science/tel-00288723>

Submitted on 18 Jun 2008

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

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

THESE

présentée à

Université Scientifique et Médicale de Grenoble
Institut National Polytechnique de Grenoble

pour obtenir le grade de
DOCTEUR DE 3^{ème} CYCLE
Mathématiques Appliquées

par

Donatien NIANGO



TRAVAUX DE STATISTIQUE



Thèse soutenue le 14 septembre 1978 devant la Commission d'Examen :

Président : J. R. BARRA

Examineurs : F. BRODEAU
B. VAN CUTSEM

THESE

présentée à

Université Scientifique et Médicale de Grenoble
Institut National Polytechnique de Grenoble

pour obtenir le grade de
DOCTEUR DE 3ème CYCLE
Mathématiques Appliquées

par

Donatien NIANGO



TRAVAUX DE STATISTIQUE



Thèse soutenue le 14 septembre 1978 devant la Commission d'Examen :

Président : J. R. BARRA

Examineurs : F. BRODEAU
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
	BARNARD 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 (I.U.T. B)
	BONNET Jean-Louis	Clinique ophtalmologie
	BONNET-EYMARD Joseph	Clinique hépato-gastro-entérologie
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

MM.	CAU Gabriel	Médecine légale et toxicologie
	CAUQUIS Georges	Chimie organique
	CHABAUTY Claude	Mathématiques pures
	CHARACHON Robert	Clinique ot-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
	DEPORTES Charles	Chimie minérale
	DESRE Pierre	Métallurgie
	DODU Jacques	Mécanique appliquée (I.U.T. I)
	DOLIQUE Jean-Michel	Physique des plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	FONTAINE Jean-Marc	Mathématiques 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 (I.U.T. I)

MM.	LLIBOUTRY Louis	Géophysique
	LOISEAUX Jean-Marie	Sciences nucléaires
	LONGEQUEUE Jean-Pierre	Physique nucléaire
	LOUP Jean	Géographie
Mlle	LUTZ Elisabeth	Mathématiques pures
MM.	MALINAS Yves	Clinique obstétricale
	MARTIN-NOEL 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 (I.U.T. I)
	SOUTIF Michel	Physique générale
	TANCHE Maurice	Physiologie
	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
SUNIER Jules

CERMO
Physique

PROFESSEURS SANS CHAIRE

Mlle	AGNIUS-DELORS 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 (I.U.T. I)
	BUISSON René	Physique (I.U.T. I)
	BUTEL Jean	Orthopédie
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique (I.U.T. 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 (I.U.T. I)
	LUU DUC Cuong	Chimie organique - pharmacie
	MICHOULIER Jean	Physique (I.U.T. I)
Mme	MINIER Colette	Physique (I.U.T. I)

MM.	PELMONT Jean	Biochimie
	PERRIAUX Jean-Jacques	Géologie et minéralogie
	PFISTER Jean-Claude	Physique du solide
Mlle	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 (I.U.T. 1)
	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 (I.U.T. 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	Mathématiques pures
	CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
	CORDONNER Daniel	Néphrologie
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie

MM.	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 (I.U.T. 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
	MARECHAL Jean	Mécanique (I.U.T. I)
	MARTIN-BOUYER Michel	Chimie (CUS)
	MASSOT Christian	Médecine interne
	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique (I.U.T. I)
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	PEFFEN René	Métallurgie (I.U.T. I)
	PERRIER Guy	Géophysique-glaciologie
	PHELIP Xavier	Rhumatologie
	RACHALL Michel	Médecine interne
	RACINET Claude	Gynécologie et obstétrique
	RAMBAUD Pierre	Pédiatrie
	RAPHAEL Bernard	Stomatologie
Mme	RENAUDET Jacqueline	Bactériologie (pharmacie)
MM.	ROBERT Jean-Bernard	Chimie-physique
	ROMIER Guy	Mathématiques (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	SAKAROVITCH Michel	Mathématiques appliquées

MM.	SCHAERER René	Cancérologie
Mme	SEIGLE-MURANDI Françoise	Cryptogamie
MM.	STOEBNER Pierre	Anatomie pathologie
	STUTZ Pierre	Mécanique
	VROUSOS Constantin	Radiologie

MAITRES DE CONFERENCES ASSOCIES

MM.	DEVINE Roderick	Spectro Physique
	KANEKO Akira	Mathématiques pures
	JOHNSON Thomas	Mathématiques appliquées
	RAY Tuhina	Physique

MAITRE DE CONFERENCES DELEGUE

M.	ROCHAT Jacques	Hygiène et hydrologie (pharmacie)
----	----------------	-----------------------------------

Fait à Saint Martin d'Hères, novembre 1977

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Année universitaire 1978-1977

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

MM.	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
	SABONNADIÈRE Jean-Claude	Informatique fondamentale et appliquée
Mme	SAUCIER Gabrielle	Informatique fondamentale et appliquée
M.	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 (Métallurgie)
	DRIVER Julian	Ecole des Mines St. Etienne (Métallurgie)

MM.	KOBYLANSKI André	Ecole des Mines St. Etienne (Métallurgie)
	LE COZE Jean	Ecole des Mines St. Etienne (Métallurgie)
	LESBATS Pierre	Ecole des Mines St. Etienne (Métallurgie)
	LEVY Jacques	Ecole des Mines St. Etienne (Métallurgie)
	RIEU Jean	Ecole des Mines St. Etienne (Métallurgie)
	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'études nucléaires de Grenoble
	KAMARINOS	Centre national recherche scientifique

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

M.	BORNARD	Centre national recherche scientifique
Mme	CHERUY	Centre national recherche scientifique
MM.	DAVID	Centre national recherche scientifique
	DESCHIZEAUX	Centre national recherche scientifique

Je ne saurais trouver les mots pour exprimer ma très profonde gratitude à Monsieur F. BRODEAU, Professeur de Mathématiques à l'Institut Universitaire de Technologie II de Grenoble, pour m'avoir accordé sa confiance en acceptant de diriger mes travaux, en dépit de la grande distance qui sépare Grenoble d'Abidjan, et pour les nombreux conseils qu'il n'a cessé de me prodiguer et sans lesquels je n'aurais jamais pu soutenir cette thèse. Monsieur BRODEAU n'a pas seulement dirigé mes travaux, c'est lui qui m'a donné le goût de la statistique, lorsqu'il enseignait à l'Université Scientifique et Médicale de Grenoble.

Je tiens à remercier Monsieur J.R. BARRA, Professeur à l'Université Scientifique et Médicale de Grenoble, non seulement pour m'avoir fait l'honneur de présider le Jury de cette thèse, mais pour avoir été le principal conseiller technique et l'un de ceux qui m'ont guidé sur le chemin de la statistique, j'avais également, particulièrement apprécié ses enseignements.

Je suis très reconnaissant à Monsieur B. VAN CUTSEM, Professeur à l'Université Scientifique et Médicale de Grenoble, d'avoir accepté de faire partie du Jury et de l'aide bienveillante qu'il m'a ainsi apportée.

Je remercie également Mademoiselle G. BICAIS pour les soins particuliers qu'elle a apportés à la dactylographie de cette thèse ainsi que le Service de reprographie du Département de Mathématiques Appliquées, pour la rapidité avec laquelle le tirage a été assuré.

INTRODUCTION

-:--:--:--:--:--:

Nous avons réuni sous le titre "Travaux de Statistique Mathématique" une recherche sur la décision statistique séquentielle et une autre recherche sur les chaînes de Markov et défauts de pages en programmation.

Ces deux publications n'ont apparemment aucun point commun, la première étant très théorique et la seconde pouvant s'appliquer à un certain nombre de problèmes à caractère pratique.

Dans la première partie, nous avons voulu donner un exposé mathématique, aussi rigoureux que possible, de la statistique séquentielle telle qu'elle est exposée dans le livre de T.S. Ferguson : *Mathematic Statistics Decision Theoric*, qui constitue la première référence du sujet.

Dans le chapitre I de cette première partie, nous avons, par souci de clarté de l'exposé, regroupé les outils mathématiques, classiques pour la plupart, et auxquels il sera fait constamment appel par la suite. Nous avons également situé le problème de décision séquentielle par rapport à la décision statistique plus courante avant de le formaliser au maximum, en dégageant les deux éléments essentiels qui constituent une stratégie : la règle d'arrêt et la règle de décision terminale.

Le deuxième chapitre est consacré à la théorie séquentielle proprement dite, notre préoccupation principale étant là, de trouver des stratégies de Bayes, c'est-à-dire des stratégies qui minimisent le risque bayésien. Nous montrons qu'on peut obtenir des stratégies de Bayes en deux étapes : d'abord nous recherchons une règle de décision terminale de Bayes, puis ensuite, et c'est l'étape la plus délicate, nous donnons des conditions permettant de trouver une règle d'arrêt de Bayes en tronquant le problème en une étape fixée à l'avance.

Dans le chapitre III, nous donnons des conditions permettant d'approcher le problème général par une suite de problèmes tronqués. L'exemple que nous avons traité à ce sujet, sur les variables de Bernouilli illustre la théorie : le problème tronqué à l'étape 5 est solution du problème général. Nous avons pu noter là encore, que dans certains cas, une règle de Bayes pour le problème général est une règle de Bayes pour un problème tronqué, ce qui permet ainsi d'obtenir le minimum d'observations à faire.

Quant au chapitre IV, il est consacré à la recherche de stratégies minimax, de classes de stratégies essentiellement complètes pour des structures statistiques ayant des propriétés simples : exhaustivité et transitivité notamment.

Nous montrons d'abord comment obtenir des stratégies minimax à partir de stratégies de Bayes. Nous montrons ensuite dans ce chapitre qui est le plus développé du point de vue mathématique que, moyennant des hypothèses d'exhaustivité et de transitivité pour certaines structures statistiques utilisées, les classes de stratégies basées sur ces statistiques sont essentiellement complètes.

Dans la deuxième publication "Chaînes de Markov et défauts de pages en programmation", notre but est de simuler les instants des défauts de pages, par les instants de retour en l'un ou l'autre de deux états (choisis à l'avance) d'une chaîne de Markov.

Ainsi, après avoir rappelé rapidement ce qu'est un défaut de page en programmation et une chaîne de Markov, nous avons montré, en faisant largement appel à la méthode des fonctions génératrices, comment on peut calculer successivement la loi de retour en 0 et 1, et le coefficient de corrélation entre deux intervalles successifs ; nous avons ensuite traité un exemple dont les résultats nous permettent d'espérer que nous sommes sur la bonne voie.

Le problème de l'adéquation d'une chaîne de Markov à un exemple de défaut de page reste toutefois posé.

DECISION SEQUENTIELLE
-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-

D. NIANGO

I - INTRODUCTION

La notion de décision statistique est une notion assez récente, et de ce fait, en dépit de la rigueur des théories mathématiques élaborées à ce sujet, elle reste encore une notion controversée. La notion de décision séquentielle souffre encore plus de cette incompréhension.

Nous nous assignons la tâche redoutable d'essayer d'y faire la lumière.

Mais avant d'en arriver là, il nous semble indispensable de rappeler certains concepts utilisés en statistique mathématique.

I.1 - RAPPELS DE QUELQUES NOTIONS DE STATISTIQUE MATHÉMATIQUE

Structure statistique

On appelle structure statistique, le triplet $[\Omega, \mathcal{A}, \mathcal{S}]$ où Ω est l'espace fondamental ou espace des observations, \mathcal{A} la tribu des parties de Ω et \mathcal{S} une famille de lois de probabilités ; la famille \mathcal{S} dépendra en général d'un paramètre θ , élément de Θ ensemble des paramètres et on écrira

$\mathcal{S} = \{P_\theta, \theta \in (\Theta, \mathcal{C})\}$ où \mathcal{C} est la tribu des parties de Θ .

Une telle structure sera notée

$$[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{C})\}]$$

Probabilité de transition

Soient $[\Omega, \mathcal{A}]$ et $[\Delta, \mathcal{D}]$ deux espaces probabilités ; on appelle probabilité de transition, une fonction $P_\omega(D)$ sur $\Omega \times \mathcal{D}$ telle que :

a) $\forall \omega \in \Omega, P_\omega(D)$ comme fonction de D est une loi de probabilité sur $[\Delta, \mathcal{D}]$

b) $\forall D \in \mathcal{D}$, $P_\omega(D)$ est une fonction \mathcal{A} -mesurable de ω .

Distribution à priori et à postérieure

Définition : On appelle distribution à priori sur $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{F})\}]$ toute loi de probabilité Q sur (Θ, \mathcal{F}) .

L'introduction d'une distribution à priori dans un problème de statistique est à la base de ce que l'on appelle le point de vue bayésien ; une distribution à priori étant destinée à représenter l'information avant toute expérience ou observation du paramètre θ ; elle est parfois concrètement justifiable parfois d'usage arbitraire.

Nous désignons par β_Q l'image de Q : elle est définie par :

$$\forall A \in \mathcal{A} \quad \beta_Q(A) = \int_{\Theta} P_\theta(A) dQ(\theta)$$

c'est comme on le voit, une loi de probabilité sur $[\Omega, \mathcal{A}]$

Distribution à postérieure :

Avant de donner une définition générale de la distribution à postérieure, nous donnons une définition de ce concept dans le cas discontinu suivant :

Théorème et définition :

Soient $C_1, \dots, C_i, \dots, C_N$, N causes incompatibles dont chacune donne un résultat R ;

$P_1, \dots, P_i, \dots, P_N$, les probabilités à priori de chaque cause et $q_1, \dots, q_i, \dots, q_N$, les probabilités respectives de R sous chacune des causes ($q_i = P\{R/C_i\}$). Alors les probabilités à postérieure des causes C_i sont :

$$P\{C_i/R\} = \frac{P_i q_i}{\sum_{i=1}^N P_i q_i}$$

En effet :

$$P\{C_i/R\} = \frac{P\{C_i \cap R\}}{P\{R\}}$$

et

$$P\{C_i \cap R\} = P\{C_i\} P\{R/C_i\} = p_i q_i$$

$$P\{R\} = \sum_{i=1}^N P\{C_i \cap R\} = \sum p_i q_i$$

et

$$P\{C_i/R\} = \frac{P\{C_i \cap R\}}{P\{R\}}$$

$$P\{C_i \cap R\} = P\{C_i\} P\{R/C_i\} = p_i q_i$$

$$P\{R\} = \sum_{i=1}^N P\{C_i \cap R\} = \sum_{i=1}^N p_i q_i$$

d'où le résultat.

Généralisons à un cas quelconque.

Définition

Considérons la structure $[\Omega, \mathcal{A}, P_\omega, \theta \in (\Theta, \mathcal{C})]$ et supposons qu'il existe une probabilité de transition $P_\omega(T)$ sur $\Omega \times \mathcal{C}$ telle que l'on ait :

$$\forall A \in \mathcal{A}, \quad \forall T \in \mathcal{C}$$

$$\int_A P_\omega(T) d\beta_Q = \int_T P_\theta(A) dQ(\theta)$$

c'est-à-dire que $P_\omega(T)$ et l'image β_Q de Q d'une part et $P_\theta(A)$ et Q d'autre part engendrent la même loi π sur $(\Omega \times \Theta, \mathcal{A} \otimes \mathcal{C})$ alors pour tout $\omega \in \Omega$, on appelle distribution à postériori, la loi de probabilité P_ω sur (Θ, \mathcal{C}) .

Dans la plupart des cas, P_θ admet une densité $p_\theta(\omega)$, $\mathcal{A} \otimes \mathcal{C}$ -mesurable ; on retrouve dans ce cas les formules de Bayes

$$\frac{dP_\omega}{dQ} = \frac{p_\theta(\omega)}{\int_{\Theta} P_\theta(\omega) dQ}$$

Par définition même de la loi π on a :

$$\forall A \in \mathcal{A} \quad \forall T \in \mathcal{C} \quad \pi(T \times A) = \int_A P_\omega(T) d\beta_Q = \int_T P_\theta(A) dQ(\theta)$$

On remarque que cette notion de distribution à postériori ne joue pas le même rôle mathématique que la distribution à priori ; en effet, la probabilité de transition $P_\omega(T)$ sur $\Omega \times \mathcal{C}$ apparaît comme l'inverse de la probabilité de transition $P_\theta(A)$ sur $\Theta \times \mathcal{A}$.

PRODUITS DE STRUCTURES

DEFINITION : Soient $[\Omega, \mathcal{A}, \mathcal{P}]$ et $[\Omega', \mathcal{A}', \mathcal{P}']$ deux structures statistiques ; on appelle produit de ces 2 structures et on note $[\Omega, \mathcal{A}, \mathcal{P}] \otimes [\Omega', \mathcal{A}', \mathcal{P}']$ la structure $(\Omega \times \Omega', \mathcal{A} \otimes \mathcal{A}', \mathcal{P} \otimes \mathcal{P}')$ où $\mathcal{P} \otimes \mathcal{P}' = \{P.P', P \in \mathcal{P}, P' \in \mathcal{P}'\}$

DEFINITION : Soient $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{C})\}]$ et $[\Omega', \mathcal{A}', \{P'_\theta, \theta \in (\Theta, \mathcal{C})\}]$ 2 structures statistiques ayant le même espace de paramètre et le même paramètre. On appelle produit restreint de ces 2 structures et on note $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{C})\}] \times [\Omega', \mathcal{A}', \{P'_\theta, \theta \in (\Theta, \mathcal{C})\}]$ la structure : $[\Omega \times \Omega', \mathcal{A} \otimes \mathcal{A}', \{P_\theta \times P'_\theta, \theta \in (\Theta, \mathcal{C})\}]$ en particulier la structure $[\Omega, \mathcal{A}, \mathcal{P}]^n = [\Omega^n, \mathcal{A}^n, \{P, P \in \mathcal{P}\}]$ est appelé échantillon empirique : c'est le produit restreint d'un nombre fini d'exemplaires d'une même structure.

NOTION D'EXHAUSTIVITE

TRIBUS ET STATISTIQUES EXHAUSTIVES

DEFINITIONS : Soit $[\Omega, \mathcal{A}, \mathcal{P}]$ une structure statistique ; une sous-tribu \mathcal{B} de \mathcal{A} est dite exhaustive si $\forall A \in \mathcal{A}$ il existe une détermination de $P\{A/\mathcal{B}\}$ commune à toutes les lois P de \mathcal{P} ; autrement dit, pour toute statistique sommable T il existe une détermination de $E_P(T/\mathcal{B})$ commune à toutes les lois P de \mathcal{P} .

DEFINITION : On dit que la statistique T définie sur la structure statistique $[\Omega, \mathcal{A}, \mathcal{P}]$ à valeurs dans (X, \mathcal{F}) est exhaustive si la tribu $T^{-1}(\mathcal{F})$ engendrée par T est exhaustive.

Si dans la pratique, la notion de statistique exhaustive est plus utilisée que la notion de tribu exhaustive, celle-ci est plus commode, en théorie, que la première. C'est pourquoi, nous rédigerons notre exposé presque exclusivement en termes de tribus ; notons d'autre part qu'il peut exister une tribu exhaustive sans pour autant qu'existe une statistique exhaustive à valeurs dans un espace mesurable donné ; concrètement la notion de statistique exhaustive ou de tribu exhaustive signifie que la statistique ou la sous-tribu considérée est aussi bon renseignement sur le paramètre que toutes les observations effectuées ou toute la tribu initiale.

Exemple : Si la structure considérée est de la forme :

$(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mathcal{P})$ dire que la statistique $(x,y) \rightarrow y$ est exhaustive, c'est dire que la loi de x conditionnelle à y ne dépend plus du paramètre ; lorsqu'on connaît y , la valeur de x est sans intérêt statistique ; y contient toute l'information contenue dans la structure et constitue de ce fait un résumé équivalent aux observations initiales.

I.2 - RELATIONS ENTRE TRIBUS ET SOUS-TRIBUS

a) D'abord quelques notations pratiques

Considérons une structure statistique $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{E})\}]$

- une fonction f définie sur Ω sera notée " f \mathcal{A} - P_θ -intégrable".
si f est mesurable par rapport à \mathcal{A} et intégrable par rapport à P_θ ,
c'est-à-dire si l'intégrale $\int_{\Omega} f dP_\theta$ existe. ($\forall P_\theta \in \mathcal{P}$)
- Pour tout $\omega \in \Omega$ soit $\pi(\omega)$ une proposition concernant ω .
Nous écrirons

$$\pi(\omega) \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

Si la proposition $\pi(\omega)$ est vraie excepté sur un élément de \mathcal{A} de P_θ -mesure nulle $\forall \theta \in \Theta$.

Relations entre tribus et sous-tribus

Soit

$$\mathcal{A}_0 \subseteq \mathcal{A}$$

Proposition 1 : une mesure m sur \mathcal{A} est aussi une mesure sur \mathcal{A}_0

Proposition 2 : une fonction \mathcal{A}_0 - m -intégrable est \mathcal{A} - m -intégrable et on a

$$(\mathcal{A}_0) \int_{\Omega} f dm = (\mathcal{A}) \int_{\Omega} f dm$$

où comme le suggèrent les notations, les intégrales de gauche et de droite sont prises sur les espaces $(\Omega, \mathcal{A}_0, m)$ et (Ω, \mathcal{A}, m) .

Soient m et n deux mesures sur \mathcal{A} :

1) si $m = n$ sur \mathcal{A} alors $m = n$ sur \mathcal{A}_0 .

2) si $m \ll n$ sur \mathcal{A} , alors $m \ll n$ sur \mathcal{A}_0 (si n est absolument continue par rapport à m sur \mathcal{A} alors n est absolument continue par rapport à m sur \mathcal{A}_0).

3) Si \mathcal{S} est complet sur \mathcal{A} alors \mathcal{S} est complet sur \mathcal{A}_0 .

4) Si n est σ -finie sur \mathcal{A}_0 alors n est σ -finie sur \mathcal{A} . La réciproque de quatre propositions précédentes est fautive en général.

Soient \mathcal{A}_1 et \mathcal{A}_2 des sous-tribus de \mathcal{A} et soit $\mathcal{S} = \{P_\theta, \theta \in (\Theta, \mathcal{E})\}$ un ensemble de mesures sur \mathcal{A} .

Nous écrirons $\mathcal{A}_1 \subseteq \mathcal{A}_2$ [\mathcal{A}, θ] $\forall \theta \in \Theta$ si correspondant à chaque $A \in \mathcal{A}_1$ il existe $B \in \mathcal{A}_2$ tels que $\chi_A(\omega) = \chi_B(\omega)$ [\mathcal{A}, θ] $\forall \theta \in \Theta$ (où χ_A est la fonction caractéristique de A).

Nous écrirons également $\mathcal{A}_1 = \mathcal{A}_2$ [\mathcal{A}, θ] $\forall \theta \in \Theta$ si on a à la fois

$$\mathcal{A}_1 \subseteq \mathcal{A}_2 \text{ et } \mathcal{A}_2 \subseteq \mathcal{A}_1 \text{ } [\mathcal{A}, \theta] \forall \theta \in \Theta$$

Rappelons enfin quelques propriétés de l'espérance mathématique conditionnelle à une sous-tribu.

○ Soit la structure statistique $[\Omega, \mathcal{A}, \mathcal{S} = \{P_\theta, \theta \in (\Theta, \mathcal{E})\}]$.

Définition

Soit \mathcal{A}' une sous-tribu quelconque de \mathcal{A} et f une fonction \mathcal{A} - p_θ -intégrable.

Alors d'après le théorème de Radon-Nykodym pour les mesures il existe une fonction g \mathcal{A}' - p_θ -intégrable, telle que, pour tout

$$A' \in \mathcal{A}' \int_{A'} g(\omega) dp_\theta = \int_{A'} f(\omega) dp_\theta$$

et g est presque partout unique, car si g^* satisfait à la même condition, on a :

$$g^*(\omega) = g(\omega) \text{ } [\mathcal{A}, \theta] \text{ } \forall \theta \in \Theta.$$

Nous écrirons $g(\omega) = E_\theta[f(\omega) / \mathcal{A}']$, et nous l'appellerons par définition, l'espérance mathématique conditionnelle de f à la sous-tribu \mathcal{A}' .

Par conséquent, pour chaque $p_\theta \in \mathcal{P}$ pour toute fonction f \mathcal{A} - p_θ -intégrable et pour toute sous-tribu \mathcal{A}' de \mathcal{A} la quantité $E_\theta [f(\omega)/\mathcal{A}']$ est une fonction définie, \mathcal{A}' - p_θ -intégrable et telle que pour chaque $A' \in \mathcal{A}'$

$$\int_{A'} E_\theta [f(\omega)/\mathcal{A}'] dp_\theta = \int_{A'} g(\omega) dp_\theta = \int_{A'} f(\omega) dp_\theta$$

Si $f(\omega) = \chi_A(\omega)$, on peut remplacer l'espérance mathématique de f par la probabilité conditionnelle de A .

Propriétés

1°) si f est \mathcal{A} - p_θ -intégrable, alors $\int_{\Omega} E_\theta [f(\omega)/\mathcal{A}'] dp_\theta = \int_{\Omega} g(\omega) dp_\theta$.

2°) si f est \mathcal{A} - p_θ -intégrable et si $c \leq f(\omega) \leq d$

$$[\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

avec

$$-\infty < c \leq d < +\infty$$

alors

$$c \leq E_\theta [f(\omega)/\mathcal{A}'] \leq d[\mathcal{A}, \theta] \quad \forall \theta \in \Theta.$$

3°) si $\{f_n\}_{n \in \mathbb{N}}$ est une suite de fonctions intégrables telle que

$$f_n(\omega) \leq f_{n+1}(\omega) \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta \quad \text{pour } n = 1, 2, \dots$$

et si $\sup_n f_n(\omega)$ est \mathcal{A} - p_θ -intégrable, alors

$$\sup_n \{E_\theta [f_n(\omega)/\mathcal{A}']\} = E_\theta [\sup_n (f_n(\omega)/\mathcal{A}')] \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

4°) si $f(\omega)$ est \mathcal{A}' - p_θ -intégrable alors

$$E_\theta [h(\omega) \cdot f(\omega)/\mathcal{A}'] = E_\theta [h(\omega)/\mathcal{A}'] f(\omega) \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

pour toute fonction h telle que $h(\omega)$ et $h(\omega) f(\omega)$ soit \mathcal{A} - p -intégrable.

En particulier

$$E_{\theta}[f(\omega)/\mathcal{G}'] = f(\omega) \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

5°) Si f est une fonction \mathcal{A} - P_{θ} -intégrable et si \mathcal{G}' et \mathcal{G}'' sont deux sous-tribus de \mathcal{A} si $\mathcal{G}' \subseteq \mathcal{G}''$ $[\mathcal{A}, \theta] \quad \forall \theta \in \Theta$.

Si n est σ -finie sur \mathcal{G}_0 alors n est σ -finie sur \mathcal{A} .

I-3 - DEFINITION DE LA DECISION STATISTIQUE SEQUENTIELLE

Afin de bien comprendre le problème de décision séquentielle, il convient de bien le situer par rapport au problème de "décision statistique", notion plus familière au statisticien.

Un problème de décision statistique, consiste en la donnée d'une structure statistique $[\Omega, \mathcal{A}, \{P_{\theta}, \theta \in (\Theta, \mathcal{C})\}]$ et d'un espace mesurable (Δ, \mathcal{D}) appelé espace des décisions. Au vu d'une observation $\omega \in \Omega$, l'expérimentateur prend une décision $d(\omega)$ de Δ suivant une stratégie $S_{\omega}(D)$, qui est une probabilité de transition sur $\Omega \times \mathcal{D}$. Une stratégie est dite déterministe si S_{ω} est pour tout $\omega \in \Omega$, la mesure de Dirac au point $s(\omega)$ de Δ . Sinon, elle est dite stochastique ou aléatoire. Une stratégie déterministe consiste donc à prendre la décision $d(\omega)$.

Une décision statistique fait intervenir d'autres éléments dont nous donnons, ci-dessous, les définitions.

1 - Image d'une stratégie S_{ω}

C'est la probabilité de transition \mathcal{P}_{θ}^S sur $\Theta \times \mathcal{D}$ composée de P_{θ} et S définie par :

$$\forall D \in \mathcal{D} \quad \mathcal{P}_{\theta}^{S(D)} = \int_{\Omega} S_{\omega}(D) dP_{\theta}(\omega) ; \text{ l'efficacité pratique}$$

d'une stratégie est définie par son image.

2 - La fonction de perte

La fonction de perte est une quantité aléatoire notée $L(\theta, d(\omega))$
application mesurable L de

$$(\Theta, \mathcal{C}) \times (\Delta, \mathcal{D}) \text{ dans } (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$$

$L(\theta, d)$ est la perte encourue quand on prend la décision d alors que la vraie valeur du paramètre est θ . La fonction de perte est une fonction qu'on suppose positive.

○

3 - La perte moyenne

C'est une fonction sur $\Theta \times \Omega$ définie par :

$$L_S(\theta, \omega) = \int_{\Delta} L(\theta, d) d(S_{\omega}(d)).$$

4 - Le risque

Le risque est une fonction sur Θ définie par

$$R_S(\theta) = \int_{\Omega} L_S(\theta, \omega) dP_{\theta}(\omega) = \int_{\Delta} L(\theta, d) d(\mathcal{S}_{\theta}^S(d)).$$

Exemple de décision statistique (très schématisée).

Dans une usine, une machine fabrique des pièces par série de 1000 dont une proportion h est composée de pièces défectueuses. Cette proportion, n'est pas connue, mais elle peut prendre six valeurs :

$$0,01, \quad 0,02, \quad 0,05, \quad 0,10, \quad 0,20, \quad 0,25$$

Chaque pièce défectueuse est à rectifier et le coût unitaire de l'opération est 300 francs CFA. Il est cependant possible d'éliminer les rebuts en procédant à un réglage complet de la machine avant mise en marche, ce réglage coûte 14000 francs CFA. Quelle décision est-il raisonnable de prendre ?

Ici l'ensemble des décisions est

$$\Delta = \{d_1, d_2\} \text{ où}$$

d_1 : ne rien faire

d_2 : procéder au réglage

Voici le tableau donnant le coût dans chaque cas.

h	décision	
	d_1	d_2
0,01	3 000	14 000
0,02	6 000	14 000
0,05	15 000	14 000
0,10	30 000	14 000
0,20	60 000	14 000
0,25	75 000	14 000

COMMENTAIRE DU TABLEAU

Le problème précédent est un problème de décision en présence d'incertitude puisque la valeur prise par h n'est pas connue, on voit, d'après ce tableau, que suivant que l'on est optimiste ou prudent, on prendra la décision d_1 ou d_2 .

Quant aux problèmes de décision séquentielle, ils diffèrent des problèmes de "décision statistique" dont nous venons de parler par le fait qu'ici, l'expérimentation s'arrête ou se poursuit en fonction du résultat obtenu.

On veut, par exemple, estimer un paramètre inconnu θ . Pour cela, le statisticien peut observer successivement une suite de variables aléatoires $X_1, X_2, \dots, X_j, \dots$ dont la loi dépend de θ , et qui sont observables en des temps notés $1, 2, \dots, j, \dots$

A chaque étape le statisticien observe les résultats obtenus ainsi au temps j , il connaît les valeurs x_1, x_2, \dots, x_j prises respectivement par X_1, X_2, \dots, X_j . Il décide alors soit de continuer et d'observer X_{j+1} , soit d'arrêter. S'il s'arrête il prend une décision terminale concernant le paramètre θ . Ce faisant, il doit envisager deux types de coût :

- un coût dû à l'observation des variables
- un coût attaché à la décision terminale

Notons également, à la différence avec les problèmes de "décision statistique classique", que l'expérimentateur est libre de choisir le nombre d'expériences et que l'ordre des variables aléatoires est donné à l'avance et ne dépend pas du statisticien.

Indiquons enfin qu'il existe des modèles plus compliqués (Wald) où le statisticien décide en outre, des variables aléatoires qu'il va observer.

I.4 - ELEMENTS MATHEMATIQUES DU PROBLEME DE DECISION SEQUENTIEL

Un problème de décision séquentielle comporte les éléments suivants :

- a) (Θ, \mathcal{C}) : espace des paramètres
- b) (Δ, \mathcal{D}) : espace des décisions terminales ; nous noterons $\delta \in \Delta$ une décision à prendre.
- c) Une fonction réelle L définie sur $(\Theta, \mathcal{C}) \times (\Delta, \mathcal{D})$ à valeurs dans $(\bar{\mathbb{R}}, \mathcal{B}_{\bar{\mathbb{R}}})$; cette fonction sera appelée fonction de perte ; $L(\theta, y)$ sera la perte encourue par l'expérimentateur si θ est la vraie valeur du paramètre et y la décision terminale prise par l'observateur.

d) Nous avons désigné par $X_1, X_2, \dots, X_j, \dots$ les variables aléatoires associées aux observations $x_1, x_2, \dots, x_j, \dots$ respectivement ; ces variables aléatoires seront dans toute la suite des variables indépendantes, de sorte que nous avons les structures statistiques :

$$[\Omega_1, \mathcal{A}_1, \{p_1(\theta; dx_j), \theta \in (\Theta, \mathcal{E})\}], \dots, [\Omega_j, \mathcal{A}_j, \{p_j(\theta; dx_j), \theta \in (\Theta, \mathcal{E})\}], \dots$$

Posant $\Omega = \Omega_1 \times \dots \times \Omega_j \times \dots$

$$\mathcal{A} = \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_j \otimes \dots$$

nous pouvons considérer la structure statistique produit, unique :

$$[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{E})\}] \text{ où } P_\theta = \prod_j p_j(\theta; dx_j)$$

Le vecteur aléatoire associé à l'observation

$$x = (x_1, x_2, \dots, x_j, \dots) \text{ sera noté}$$

$$X = (X_1, X_2, \dots, X_j, \dots)$$

et le vecteur aléatoire associé à l'observation (x_1, \dots, x_j)

$$\mathcal{X}_j = (x_1, \dots, x_j) \in \Omega_1 \times \dots \times \Omega_j;$$

sa loi sera définie par

$$p_j(\theta; d\mathcal{X}_j) = \prod_{i=1}^j p_i(\theta; dx_i).$$

e) $C_j(\theta, x_1, \dots, x_j)$ $j = 1, 2, \dots$ le coût relatif à l'observation de x_1, x_2, \dots, x_j , une suite de fonctions réelles définies sur

$$(\Theta, \mathcal{E}) \times (\Omega_1, \mathcal{A}_1) \times \dots \times (\Omega_j, \mathcal{A}_j)$$

Le nombre $C_j(\theta, x_1, \dots, x_j)$ représente le coût de l'expérimentateur en faisant les observations $X_1 = x_1, X_2 = x_2, \dots, X_j = x_j$ et en arrêtant l'expéri-

mentation lorsque θ est la vraie valeur du paramètre.

Avec ces cinq éléments le statisticien doit observer la suite des valeurs x_j et décider à quelle étape il doit arrêter l'expérimentation et quelle décision (décision terminale) il doit prendre en fonction du résultat obtenu.

Son but est aussi de minimiser la valeur moyenne du coût ainsi que la fonction de perte, sachant que si une expérimentation plus courte coûte moins cher, elle entraîne une perte moyenne plus grande. Ce qui pose d'énormes problèmes.

○

Nous allons décrire maintenant la stratégie que l'on doit utiliser en décision séquentielle.

Examinons d'abord la stratégie aléatoire, la stratégie déterministe étant un cas particulier.

La stratégie aléatoire comporte deux parties :

- * la règle d'arrêt
- * la règle de décision terminale.

1°) La règle d'arrêt, notée ψ

Une règle d'arrêt est une suite de fonctions mesurables

$$\psi(X) = (\psi_0, \psi_1(X_1), \dots, \psi_j(X_1, X_2, \dots, X_j), \dots)$$

où

$$\psi_j : (\Omega_1 \times \dots \times \Omega_j, \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_j) \rightarrow [0,1]$$

La fonction mesurable $\psi_j(x_1, \dots, x_j)$ représente la probabilité conditionnelle d'arrêter l'expérimentation sachant que l'on a fait j observations :

$X_1 = x_1, \dots, X_j = x_j$ (ψ_0 étant une constante représentant la probabilité de faire les observations ou non) ; parallèlement, nous devons définir une règle d'arrêt par la suite des fonctions :

$$\Psi(X) = (\Psi_0, \Psi_1(X_1), \dots, \Psi_j(X_1, \dots, X_j), \dots)$$

où $\Psi_j(X_1, \dots, X_j)$ représente la probabilité conditionnelle de ne pas s'arrêter après les $j-1$ premières observations et d'arrêter l'expérimentation à la j ème étape sachant que l'on a : $X_1 = x_1, \dots, X_j = x_j$.

Il découle des définitions de φ_j et ψ_j que l'on a :

$$\Psi_0 = \varphi_0 \quad \Psi_j(X_1, \dots, X_j) = (1 - \varphi_0)(1 - \varphi_1) \dots (1 - \varphi_{j-1}) \varphi_j(X_1, \dots, X_j) \quad (1)$$

$$(j = 1, 2, \dots).$$

La probabilité d'arrêter à l'étape n est alors

$$E_\theta[\Psi_n(X_1, \dots, X_n)]$$

Par conséquent la probabilité d'arrêter éventuellement (ou ce qui est équivalent, d'un nombre fini d'observations) est :

$$\sum_{n=0}^{\infty} E_\theta[\Psi_n(X_1, \dots, X_n)]$$

On a

$$\sum_{n=0}^{\infty} E_\theta[\Psi_n(X_1, \dots, X_n)] \leq 1 \text{ car}$$

$$\sum_{n=0}^{\infty} \Psi_n(X_1, \dots, X_n) \leq 1 \quad \forall \theta \in (\Theta, \mathcal{E})$$

(cela se démontre pour une somme finie, en utilisant des propriétés de barycentre).

Nous ferons l'hypothèse que $\sum_{n=0}^{\infty} E_\theta[\Psi_n(X_1, \dots, X_n)] = 1 \quad \forall \theta$
ceci afin que le coût moyen soit de valeur finie.

Le temps d'arrêter peut être défini comme une variable aléatoire N :

$$P_{\theta}\{N = n / X = x\} = \psi_n(X_1, \dots, X_n) \quad n = 0, 1, 2, \dots$$

et

$$P_{\theta}\{N = \infty / X = x\} = 1 - \sum_{j=0}^{\infty} \psi_j(X_1, \dots, X_j)$$

Remarquons que si

$$\sum_{j=0}^{\infty} E_{\theta}[\psi_j(X_1, \dots, X_j)] = 1$$

on a

$$\sum_{j=0}^{\infty} \psi_j(X_1, \dots, X_j) = 1 \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

C'est une remarque essentielle et nous supposons cette condition réalisée.

2°) La règle de décision terminale

Une règle de décision terminale est une suite de fonctions

$$\delta(X) = (\delta_0, \delta_1(X_1), \dots, \delta_j(X_1, \dots, X_j))$$

où pour tout j $\delta_j(x_1, \dots, x_j)$ est une mesure de transition définie sur $(\Omega_1, \mathcal{A}_1) \times \dots \times (\Omega_j, \mathcal{A}_j) \times (\Delta, \mathcal{D})$ et est telle que l'espérance mathématique de la fonction de perte correspondante $E_{\theta}[L(\theta, Y)]$ existe et est finie pour tout $\theta \in (\Theta, \mathcal{C})$; concrètement δ_j à la signification suivante: si l'on décide d'arrêter l'expérimentation après avoir observé $X_1 = x_1, \dots, X_j = x_j$, on choisira un point y de (Δ, \mathcal{D}) dont la variable aléatoire associée est Y de loi de probabilité $\delta_j(x_1, \dots, x_j; D)$ $D \in \mathcal{D}$. Pour une règle d'arrêt donnée φ_j , les δ_j sont définies uniquement pour (x_1, \dots, x_j) pour lesquels $\psi_j(x_1, \dots, x_j) > 0$. Il sera cependant utile de définir une règle de décision terminale qui soit indépendante, comme nous le verrons, de la règle d'arrêt.

Enfin pour simplifier nous supposons que Ω_j est pour tout j , \mathbb{R} muni de sa tribu borélienne que nous notons encore \mathcal{A}_j .

Une stratégie séquentielle sera donc, un couple (φ, δ) où φ est la règle d'arrêt et δ la règle de décision terminale.

La fonction de risque d'une stratégie séquentielle est :

$$R[\theta, (\delta, \psi)] = \sum_{j=0}^{\infty} \int_{\mathbb{R}^j} \psi_j(X_1, \dots, X_j) \left\{ \int_{\Delta} L(\theta, Y) \delta_j(X_1, \dots, X_j, Y) dy + C_j(\theta, X_1, \dots, X_j) \right\} P_j(\theta; dX_j)$$

ou encore

$$R[\theta, (\delta, \psi)] = \sum_{j=0}^{\infty} E_{\theta} \{ \psi_j(X_1, \dots, X_j) [E_{\theta} (L(\theta, Y) + C_j(\theta, X_1, \dots, X_j))] \}$$

Nous noterons $d = (d_0, d_1, \dots, d_j, \dots)$ avec d_j définie sur $(\mathbb{R}, \mathcal{A}_1) \times \dots \times (\mathbb{R}, \mathcal{A}_j) \times (\Delta, \mathcal{D})$ toute règle de décision terminale qui ne sera pas aléatoire ; c'est un cas particulier du cas général.

Notons enfin, que par définition, nous dirons qu'une règle d'arrêt ψ est déterministe si $\psi_j(X_1, \dots, X_j)$ est égale à 0 ou à 1 pour tout j -uple (X_1, \dots, X_j) de variables aléatoires. Nous n'adopterons aucune notation particulière pour désigner les règles d'arrêt déterministes.

L'étude d'une décision séquentielle requiert des hypothèses supplémentaires sur certains des concepts que nous venons de définir.

Hypothèses supplémentaires

$L(\theta, y)$ sera une fonction bornée inférieurement, de plus on supposera $L(\theta, y) \geq 0$.

Ferguson suppose que

enfin $C_j(\theta, X_1, \dots, X_j) \geq 0 \quad \forall j, \theta, X_1, \dots, X_j$

et $C_j(\theta, X_1, \dots, X_j) \leq C_{j+1}(\theta, X_1, \dots, X_{j+1})$

$$C_j(\theta, X_1, \dots, X_j) \xrightarrow{j \rightarrow \infty} \infty \quad \forall X, \forall \theta$$

En particulier, le coût d'un nombre fini d'observations est infini.
Si donc pour une stratégie, le coût est fini c'est que

$$\sum_{j=0}^{\infty} \psi_j(X_1, \dots, X_j) = 1 \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

Exemple :

Dans la plupart des cas on prendra, $C_j(\theta, X_1, \dots, X_j) = jC$ ou C est une constante indépendante de X_1, \dots, X_j .

I.5 - COMPARAISON DES REGLES DE DECISION

Soit \mathcal{J} l'ensemble des stratégies, supposé muni d'une relation de préordre

- (δ, φ) est aussi bonne que (δ', φ')

si

$$R[\theta, (\delta, \varphi)] \leq R[\theta, (\delta', \varphi')] \quad \forall \theta \in \Theta$$

que l'on peut noter simplement $(\delta, \varphi) \geq (\delta', \varphi')$.

- (δ, φ) est strictement meilleure que (δ', φ') , si elle est aussi bonne que (δ', φ') et si de plus, il existe $\theta \in \Theta$ tel que $R[\theta, (\delta, \varphi)] < R[\theta, (\delta', \varphi')]$

- (δ, φ) est équivalente à (δ', φ') si (δ, φ) est aussi bonne que (δ', φ') et réciproquement (δ', φ') est admissible s'il n'existe aucune règle de décision (δ', φ') strictement meilleure.

Classe complète : soit \mathcal{C} une partie de \mathcal{J} ; \mathcal{C} est complète si

$$\forall (\varphi', \delta') \in \mathcal{J} \quad \exists (\delta, \varphi) \in \mathcal{C} \text{ tel que } (\delta, \varphi) > (\delta', \varphi')$$

Classe essentiellement complète

Une partie \mathcal{C} de \mathcal{J} est dite essentiellement complète si

$$\forall (\varphi', \delta') \in \bigcup_{\mathcal{J}} \mathcal{C}, \exists (\delta, \varphi) \in \mathcal{C} \text{ tel que } (\delta, \varphi) \geq (\delta', \varphi')$$

Règle de Bayes

Soit τ une probabilité à priori sur (θ, \mathcal{C}) . On appelle risque de Bayes correspondant, la quantité

$$R[\tau, (\delta, \varphi)] = \int_{\Theta} R[\theta, (\delta, \varphi)] d(\tau(\theta))$$

Une règle (δ, φ) est dite de Bayes relativement à τ si

$$R[\tau, (\delta, \varphi)] = \inf_{(\delta', \varphi') \in \mathcal{J}} R[\tau, (\delta', \varphi')]$$

ϵ -Stratégie de Bayes ○

Une stratégie (δ, φ) est dite ϵ - stratégie de Bayes relativement à τ si

$$R[\tau, (\delta, \varphi)] \leq \inf_{(\delta', \varphi') \in \mathcal{J}} R[\tau, (\delta', \varphi')] + \epsilon .$$

II - THEORIE DE LA DECISION SEQUENTIELLE

II.1 - DETERMINATION DES REGLES DE BAYES ; DETERMINATION DE LA REGLE DE DECISION TERMINALE

Soit τ une mesure sur (Θ, \mathcal{E}) .

On suppose que les risques Bayésiens considérés sont finis. La recherche d'une stratégie relativement à τ se divise en deux parties :

- on fixe φ et on détermine δ minimisant $R[\tau, (\delta, \varphi)]$
- ensuite on cherche φ minimisant $\inf_{\delta} R[\tau, (\delta, \varphi)]$.

Cette deuxième partie, comme nous le verrons, étant plus difficile à réaliser. Soit donc φ fixé ; minimiser $R[\tau, (\delta, \varphi)]$ par rapport à δ revient à minimiser

$$\int_{\Theta} R[\theta, (\delta, \varphi)] d(\tau(\theta)) = \int_{\Theta} \sum_{j=0}^{\infty} E_{\theta}\{\psi_j(X_1, \dots, X_j)\} [E_{\theta}(L(T, Y) + C_j(T, X_j))] d(\tau(\theta))$$

○

$$= E_{\theta}\{R[T, (\delta, \varphi)]\}$$

où T est une variable aléatoire sur (Θ, \mathcal{E}) de loi de probabilité τ .

Comme $\psi_j \geq 0$ ainsi que $L(\theta, y) \forall j$, il suffit, pour chaque valeur de j , de minimiser l'expression :

$$\int_{\Theta} \sum_{j=0}^{\infty} E_{\theta}\{\psi_j(X_1, \dots, X_j)\} E_{\theta}(L(T, Y)) d\tau(\theta)$$

$$= \int_{\Theta} \sum_{j=0}^{\infty} \int_{\mathbb{R}^j} \psi_j(X_1, \dots, X_j) \int_{\Delta} L(\theta, y) \delta_j(X_1, \dots, X_j, dy) P_j(\theta, X_j) d(\tau(\theta))$$

$$\int_{\Theta} \sum_{j=0}^{\infty} E_{\theta}[\psi_j(X_1, \dots, X_j) \int_{\Delta} L(\theta, y) \delta_j(X_1, \dots, X_j; dy)] d\tau(\theta)$$

$C_j(T, X_1, \dots, X_j)$ étant indépendant de δ_j ; on minimisera simplement

$$\begin{aligned} & E_{\theta}[\psi_j(X_1, \dots, X_j) \int_{\Delta} L(T, y) \delta_j(X_1, \dots, X_j ; dy)] \\ &= E_{\theta}\{\psi_j(\mathfrak{X}_j) E_{\theta}[\int_{\Delta} L(T, y) \delta_j(\mathfrak{X}_j ; dy) / X_1, \dots, X_j]\} \end{aligned}$$

Par conséquent, s'il existe, pour chaque valeur de j un δ_j noté δ_j^0 minimisant

$$E_{\theta}[\int_{\Delta} L(T, y) \delta_j(X_1, \dots, X_j ; dy) / (X_1, \dots, X_j)]$$

la suite $\{\delta_j^0\}$ minimisera le risque bayésien (pour φ fixé).

D'où le lemme 1

Si $\delta_j^0(X_1, \dots, X_j)$ est une règle de Bayes basée sur X_1, \dots, X_j avec la loi de probabilité à priori τ alors pour toute règle d'arrêt donnée φ , $r(\tau, (\delta, \varphi))$ est minimisé par $\delta^0 = (\delta_0^0, \delta_1^0, \dots)$.

Remarque : Dans les conditions du lemme précédent, $\{\delta_j^0\}$ ne dépend pas de φ , ce qui facilitera les calculs. Si la règle d'arrêt a déjà été choisie, l'étude du problème revient à l'étude d'un échantillon donné.

Une règle δ_j^0 , si elle existe (Ferguson l'appelle règle de Bayes, mais c'est un abus de langage) est une règle de Bayes pour l'échantillon fixe X_1, \dots, X_j , la réciproque étant fausse.

Cas particulier important dans la pratique

Si $P_j(\theta; d\mathfrak{X}_j)$ et τ définissent une mesure μ sur $R^j \times \Theta$ et si μ se désintègre en $P(d\mathfrak{X}_j)$ et $P(\mathfrak{X}_j ; d\theta)$ alors, on appliquera la formule de Fubini à l'intégrale

$$\int_{\Theta} \sum_{j=0}^{\infty} E_{\theta}\{\psi_j(X_j, \dots, X_j) \int_{\Delta} [L(T, y) \delta_j(\mathfrak{X}_j ; dy)]\} d(\tau(\theta))$$

et on est ramené à essayer de minimiser l'intégrale

$$\int_{\Theta} P(\mathbf{x}_j ; d\theta) \int_{\Delta} L(T,y) \delta_j(\mathbf{x}_j ; dy)$$

donc si la distribution à postérieure de θ existe, cette méthode doit être assez efficace. Nous le verrons dans un exemple.

Si de plus $\Theta = \mathbb{R}$ d'après le théorème de Jirina (livre de Hennequin Tortrat sur la théorie de probabilités et quelques applications p. 233 à 235), la distribution à postérieure existe et

$$\forall \epsilon > 0 \quad \forall j \exists \delta_j^0 :$$

$$E_{\theta} \left\{ \int_{\Delta} L(T,y) \delta_j(\mathbf{x}_j ; dy) / X_1, \dots, X_j \right\} \leq \epsilon + \inf_{\delta_j} E_{\theta} \left\{ \int_{\Delta} L(T,y) \delta_j(\mathbf{x}_j ; dy) / X_1, \dots, X_j \right\}$$

d'où en multipliant par $\psi_j(X_1, \dots, X_j)$ et en intégrant

$$E_{\theta} [\psi_j(\mathbf{x}_j) \int_{\Delta} L(T,y) \delta_j^0] \leq \epsilon(\psi_j(\mathbf{x}_j)) + \inf_{\delta_j} E_{\theta} \left\{ \psi_j \int_{\Delta} L(T,y) \delta_j \right\}$$

Soit en additionnant :

$$R[\tau, (\delta^0, \psi)] \leq \epsilon + \inf_{\delta} R[\tau, (\delta, \psi)] \quad \text{où } \delta^0 = \{\delta_j^0\}$$

On peut ainsi déterminer des ϵ -stratégies de Bayes.

II.2 - DETERMINATION DE LA REGLE D'ARRET POUR UN PROBLEME TRONQUE

Nous abordons ainsi la deuxième étape de la procédure.

On tronque le problème à l'étape J et on cherche une stratégie de Bayes pour le problème tronqué. La troisième étape consistera à approcher le problème général par un problème tronqué.

On s'arrête donc systématiquement en J.

$$\varphi_J(X_1, \dots, X_J) = 1 \quad \forall (X_1, \dots, X_J)$$

ce qui entraîne d'ailleurs

$$\sum_{j=0}^J \psi_j(X_1, \dots, X_j) = 1$$

Supposons qu'existe $\{\delta_j^0\}_{j=0,1,\dots,J}$ déterminée dans la première étape. Si on s'arrête en j , on utilise la règle δ_j^0 .

Montrons comment on peut déterminer une suite $\{\varphi_j^0\}$ telle que $\{\delta_j^0, \varphi_j^0\}$ soit une stratégie de Bayes pour le problème tronqué.

Posons :

$$\forall j=0, 1, 2, \dots, J$$

$$U_j(X_1, \dots, X_j) = \inf_{\delta_j} \{ E \left[\int_{\Delta} L(T, y) \delta_j(\mathfrak{X}_j; dy) / X_1, \dots, X_j \right] \}$$

$$+ E_{\theta} [C_j(T, X_1, \dots, X_j) / (X_1, \dots, X_j)]$$

$$= E_{\theta} \left[\int_{\Delta} L(T, y) \delta_j^0(\mathfrak{X}_j; dy) / X_1, \dots, X_j \right] + E_{\theta} [C_j(T, \mathfrak{X}_j) / X_1, \dots, X_j]$$

Pour toute stratégie $(\delta, \varphi) = \{\delta_j, \varphi_j\}_j$, on a :

$$R[\tau, (\delta, \varphi)] = \sum_{j=0}^J E_{\theta} \{ \psi_j(\mathfrak{X}_j) E_{\theta} \left[\int_{\Delta} L(T, y) \delta_j(\mathfrak{X}_j; dy) / \mathfrak{X}_j \right] \}$$

$$+ \sum_{j=0}^J E_{\theta} \{ \psi_j(X_1, \dots, X_j) E_{\theta} [C_j(T, X_1, \dots, X_j) / (X_1, \dots, X_j)] \}$$

$$R[\tau, (\delta, \varphi)] = \sum_{j=0}^J E_{\theta} \{ \psi_j(X_1, \dots, X_j) [E_{\theta} \{ \int_{\Delta} L(T, y) \delta_j(x_j; dy) \} + E_{\theta} \{ C_j(T, X_1, \dots, X_j) \}] \}$$

$$R[\tau, (\delta, \varphi)] \geq \sum_{j=0}^J E_{\theta} [\psi_j(X_1, \dots, X_j) U_j(X_1, \dots, X_j)] = R[\tau, (\delta^0, \varphi)]$$

Cette inégalité est valable pour toute stratégie (δ, φ) mais elle n'a de sens que si $U_j(X_1, \dots, X_j)$ est mesurable. (D'ailleurs, le problème n'aura d'intérêt que si $U_j(X_1, \dots, X_j)$ est intégrable.

Il reste à minimiser par rapport à φ_j la quantité

$$\sum_{j=0}^J E_{\theta} [\psi_j(X_1, \dots, X_j) U_j(X_1, \dots, X_j)].$$

On déterminera les φ_j dans le sens de J vers 0 car on remarque que φ_{J-1} ne figure en fait que dans deux termes de la somme, les deux derniers

○

$$\psi_{J-1}(X_1, \dots, X_{j-1}) = (1 - \varphi) (1 - \varphi_1(X_1)) \dots (1 - \varphi_{J-1}) \varphi_{J-1}(X_1, \dots, X_{J-1})$$

$$\psi_J(X_1, \dots, X_J) = (1 - \varphi_0) \dots (1 - \varphi_{J-1}) \quad (\varphi_J(X_1, \dots, X_J) = 1)$$

La somme des deux derniers termes s'écrit :

$$E_{\theta} \{ (1 - \varphi_0) \dots (1 - \varphi_{J-1}) [\varphi_{J-1}(X_1, \dots, X_{J-1}) U_{J-1}(X_1, \dots, X_{J-1})$$

$$+ (1 - \varphi_{J-1}) E_{\theta} (U_J(X_1, \dots, X_J) / (X_1, \dots, X_{J-1})) \}.$$

On minimise donc le terme relatif à l'étape J-1 si on détermine φ_{J-1} minimisant

$$\varphi_{J-1}(X_1, \dots, X_{J-1}) U_{J-1}(X_1, \dots, X_{J-1}) + (1 - \varphi_{J-1}) E_{\theta} [U_J(X_1, \dots, X_J) / (X_1, \dots, X_{J-1})]$$

on prendra alors la règle optimale, évidente suivante :

$$\varphi_{J-1}^{\circ}(X_1, \dots, X_{J-1}) = \begin{cases} 1 & \text{si } U_{J-1}(X_1, \dots, X_{J-1}) < E_{\theta} [U_J / (X_1, \dots, X_{J-1})] \\ 0 & \text{si } U_{J-1}(X_1, \dots, X_{J-1}) > E_{\theta} [U_J / (X_1, \dots, X_{J-1})] \\ \lambda \text{ (quelconque)} & \text{si } U_{J-1}(X_1, \dots, X_{J-1}) = E_{\theta} [U_J / (X_1, \dots, X_{J-1})] \end{cases}$$

Commentaire :

On arrêtera à l'étape J-1 si l'espérance conditionnelle du coût à l'étape J-1, après avoir observé $X_1 = x_1 \dots X_{J-1} = x_{J-1}$ est inférieure à l'espérance conditionnelle du coût de faire une expérience supplémentaire après observation de $X_1 = x_1, \dots, X_{J-1} = x_{J-1}$; sinon, on continue l'expérimentation.

Posons alors :

$$V_{J-1}^J(X_1, \dots, X_{J-1}) = \text{Min} \begin{cases} U_{J-1}(X_1, \dots, X_{J-1}) \\ E_{\theta} \{U_J(X_1, \dots, X_J) / X_1=x_1, \dots, X_{J-1}=x_{J-1}\} \end{cases}$$

On a immédiatement :

$$\begin{aligned} & \varphi_{J-1}(X_1, \dots, X_{J-1}) U_{J-1}(X_1, \dots, X_{J-1}) + (1 - \varphi_{J-1}(X_1, \dots, X_{J-1})) \\ & \quad E_{\theta} \{U_J(X_1, \dots, X_J) / (X_1, \dots, X_{J-1})\} \\ & \geq V_{J-1}^J(X_1, \dots, X_{J-1}) \end{aligned}$$

On peut alors écrire :

$$R[\tau, (\delta, \varphi)] \geq \sum_{j=0}^{J-2} E_{\theta} [\psi_j(X_1, \dots, X_j) U_j(X_1, \dots, X_j)] \\ + E_{\theta} [(1 - \varphi_0) \dots (1 - \varphi_{J-2}(X_1, \dots, X_{J-2})) V_{J-1}^J(X_1, \dots, X_{J-1})]$$

et on recommence avec les deux derniers termes qui sont les seuls où figure $\varphi_{J-2}(X_1, \dots, X_{J-2})$, soit :

$$(1 - \varphi_0) (1 - \varphi_{J-3}(X_1, \dots, X_{J-3})) \{ \varphi_{J-2}(X_1, \dots, X_{J-2}) U_{J-2}(X_1, \dots, X_{J-2}) \\ + (1 - \varphi_{J-2}) E_{\theta} [V_{J-1}^J(X_1, \dots, X_{J-1}) / (X_1, \dots, X_{J-2})] \}$$

ce qui entraîne

$$\varphi_{J-2}(X_1, \dots, X_{J-2}) = \begin{cases} 1 & \text{si } U_{J-2} < E_{\theta} [V_{J-1}^J / (X_1, \dots, X_{J-2})] \\ 0 & \text{si } U_{J-2} > E_{\theta} [V_{J-2}^J / (X_1, \dots, X_{J-2})] \\ \mu & \text{(quelconque) s'il y a égalité} \end{cases}$$

et ainsi de suite.

On définira la suite V_k^J par :

$$(I) \left\{ \begin{array}{l} V_J^J(X_1, \dots, X_J) = U_J(X_1, \dots, X_J) \\ V_{k-1}^J(X_1, \dots, X_{k-1}) = \text{Min} \left\{ \begin{array}{l} U_{k-1}(X_1, \dots, X_{k-1}) \\ E_{\theta} [V_k^J(X_1, \dots, X_k) / (X_1, \dots, X_{k-1})] \end{array} \right. \\ \varphi_{k-1}^0(X_1, \dots, X_{k-1}) = \left\{ \begin{array}{l} 1 \text{ si } U_{k-1} < E_{\theta} [V_k^J(X_1, \dots, X_k) / (X_1, \dots, X_{k-1})] \\ 0 \text{ si } U_{k-1} > E_{\theta} [V_k^J(X_1, \dots, X_k) / (X_1, \dots, X_{k-1})] \\ \text{quelconque s'il y a égalité.} \end{array} \right. \end{array} \right.$$

Ce procédé épuise le problème et conduit d'ailleurs à :

d'où $\varphi_0 U_0 + (1 - \varphi_0) E_{\theta} [V_1^{(J)}(X_1)]$ dont le minimum est V_0^J par définition

$$R[\tau, (\delta^0, \varphi^0)] \geq V_0^J, \text{ minimum qui est atteint.}$$

La constante V_0^J désigne le minimum du risque de Bayes pour le problème tronqué en J. La règle d'arrêt de Bayes est alors

$$\varphi^0 = (\varphi_0^0, \varphi_1^0, \dots, \varphi_j^0) \text{ où } \varphi_k^0 \text{ est définie par la dernière des relations (I)}$$

on arrête l'expérimentation à l'étape j-1 si l'espérance mathématique conditionnelle du coût après avoir observé $X_1 = x_1, \dots, X_{j-1} = x_{j-1}$ est inférieure à l'espérance mathématique conditionnelle du coût de faire une observation supplémentaire et d'utiliser la règle $\varphi_j^0, \dots, \varphi_j^0$.

Théorème

Si $\delta^0 = (\delta_0^0, \delta_1^0, \dots, \delta_j^0(X_1, \dots, X_j), \dots)$ est une règle de décision terminale de Bayes (au sens du lemme 1) et si φ^0 est une règle d'arrêt, déterminée par les relations (I), alors $((\varphi^0, \delta^0))$ est une stratégie de Bayes pour le problème tronqué en J.

Démonstration :

Il suffira de prouver que

$$\forall (\varphi, \delta) \quad r(\tau, (\varphi^0, \delta^0)) \leq r(\tau, (\varphi, \delta))$$

pour toute mesure à priori donnée τ sur (Θ, \mathcal{G}) .

On a, par définition

$$\begin{aligned} r(\tau, (\varphi^0, \delta^0)) &= \sum_{j=0}^{\infty} E_{\theta} \left\{ \psi_j(\mathfrak{X}_j) \int_{\Delta} L(T, y) \delta_j^0(\mathfrak{X}_j; dy) \right\} \\ &+ \sum_{j=0}^{\infty} E_{\theta} [\psi_j^0(X_1, \dots, X_j) C_j(T, \mathfrak{X}_j)] \end{aligned}$$

Posons :

$$\begin{aligned} \rho_j(X_1, \dots, X_j) &= \inf_{\delta_j} E_{\theta} \left\{ \int_{\Delta} L(T, y) \delta_j(\mathbf{X}_j, dy) / X_1=x_1, \dots, X_j=x_j \right\} \\ &= \inf_{\delta \in \Delta} E_{\theta} \left\{ \int_{\Delta} L(T, y) \delta_j(\mathbf{X}_j; dy) / X_1=x_1, \dots, X_j=x_j \right\}. \end{aligned}$$

Nous supposons que $\rho_j(X_1, \dots, X_j)$ est intégrable.

On a donc :

$$U_j(X_1, \dots, X_j) = \rho_j(X_1, \dots, X_j) + E_{\theta} [C_j(T, X_1, \dots, X_j) / X_1=x_1, \dots, X_j=x_j]$$

Or

$$\begin{aligned} E_{\theta} [\Psi_j^0(X_1, \dots, X_j) \int_{\Delta} [L(T, y) \delta_j(\mathbf{X}_j, dy)]] \\ = E_{\theta} \{ \Psi_j^0(X_1, \dots, X_j) E_{\theta} \left[\int_{\Delta} L(T, y) \delta_j^0(\mathbf{X}_j, dy) / X_1, \dots, X_j \right] \} \end{aligned}$$

de même

$$\begin{aligned} E_{\theta} [\Psi_j^0(X_1, \dots, X_j) C_j(T, X_1, \dots, X_j)] = \\ E_{\theta} \{ \Psi_j(X_1, \dots, X_j) E_{\theta} [C_j(T, X_1, \dots, X_j) / (X_1, \dots, X_j)] \} \end{aligned}$$

Le second membre de l'avant-dernière égalité n'est autre que :

$$E_{\theta} [\Psi_j^0(X_1, \dots, X_j) \rho_j(X_1, \dots, X_j)] \text{ et la somme donne}$$

$$E_{\theta} [\Psi_j^0(X_1, \dots, X_j) U_j(X_1, \dots, X_j)]$$

Mais par définition même de $U_j(X_1, \dots, X_j)$ on a :

$$\begin{aligned} U_j(X_1, \dots, X_j) \leq E_{\theta} \left[\int_{\Delta} L(T, y) \delta_j(X_1, \dots, X_j; dy) / X_1, \dots, X_j \right] + \\ E_{\theta} [C_j(T, X_1, \dots, X_j) / X_1, \dots, X_j] \end{aligned}$$

en multipliant les 2 membres par $E_\theta[\psi_j^0(X_1, \dots, X_j)]$ et en additionnant :

$$\sum_{j=0}^{\infty} E_\theta[\psi_j^0(X_1, \dots, X_j)U_j(X_1, \dots, X_j)] \leq \sum_{j=0}^{\infty} E_\theta[\psi_j^0(X_1, \dots, X_j)] E_\theta \int_{\Delta} [L(T, y) \delta_j(\mathcal{X}_j; dy) / (X_1, \dots, X_j)] \\ + \sum_{j=0}^{\infty} E_\theta[\psi_j^0(X_1, \dots, X_j)] E_\theta [C_j(T, X_1, \dots, X_j) / X_1, \dots, X_j]$$

Le second membre de cette inégalité est encore égale à :

$$\sum_{j=0}^{\infty} E_\theta[\psi_j^0(X_1, \dots, X_j)] \int_{\Delta} L(T, \delta_j(X_1, \dots, X_j)) + \sum_{j=0}^{\infty} E_\theta[\psi_j^0(X_1, \dots, X_j)] C_j(T, X_1, \dots, X_j) \\ = r(\tau, (\varphi^0, \delta)).$$

Le premier membre est égal à $r(\tau, (\varphi^0, \delta^0))$.

On a donc montré que

$$r(\tau, (\varphi^0, \delta^0)) \leq r(\tau, (\varphi^0, \delta))$$

Ceci achève la démonstration si l'on tient compte du lemme 1.

Définition

Nous dirons qu'une règle d'arrêt $\varphi^0 = (\varphi_0^0, \varphi_1^0, \dots)$ donnée par les relations (I) est une règle de Bayes (suivant une loi π) pour le problème tronqué en J, s'il existe une règle terminale de loi π . En général les règles d'arrêt de Bayes pour les problèmes de décision tronqués existent.

III- APPROXIMATION DU PROBLEME GENERAL PAR UNE SUITE DE PROBLEMES TRONQUES

Pour compléter l'étude, il reste à dégager des conditions sous lesquelles, on peut approcher le problème général par la suite des problèmes tronqués.

Supposons qu'à partir d'un certain rang V_0^J soit fini ($V_0^J < \infty$).

Lemme 2

La suite $\{V_0^J\}$ est monotone décroissante, donc convergente.

Démonstration :

De

$$V_{J-1}^J(X_1, \dots, X_{J-1}) = \min \begin{cases} U_{J-1}(X_1, \dots, X_{J-1}) \\ E_{\theta} [V_J^J(X_1, \dots, X_J) / X_1, \dots, X_{J-1}] \end{cases}$$

et

$$V_{J-1}^{J-1} = U_{J-1}(X_1, \dots, X_{J-1}) \text{ on déduit}$$

$$V_{J-1}^J(X_1, \dots, X_{J-1}) \leq V_{J-1}^{J-1}(X_1, \dots, X_{J-1})$$

Puis de

$$V_{J-2}^J = \min \begin{cases} U_{J-2}(X_1, \dots, X_{J-2}) \\ E_{\theta} [V_{J-1}^J(X_1, \dots, X_{J-1}) / X_1, \dots, X_{J-2}] \end{cases}$$

et

$$V_{J-2}^{J-1}(X_1, \dots, X_{J-2}) = \min \begin{cases} U_{J-2}(X_1, \dots, X_{J-2}) \\ E_{\theta} [V_{J-1}^{J-1}(X_1, \dots, X_{J-1}) / X_1, \dots, X_{J-2}] \end{cases}$$

Il résulte que :

$$V_{J-2}^J(X_2, \dots, X_{J-2}) \leq V_{J-2}^{J-1}(X_1, \dots, X_{J-2}) \quad \forall X = (X_1, X_2, \dots, X_J, \dots).$$

en continuant le raisonnement, on obtient :

$$V_0^J \leq V_0^{J-1} \quad \forall X$$

et puisque $V_0^J \geq 0 \quad \forall X$ le lemme est établi.

Il reste à donner les conditions sous lesquelles cette limite est égale à V_0^∞ ; minimum du risque Bayésien.

On a de façon évidente

$$V_0^J \geq V_0^\infty \quad \forall J$$

Théorème

Si $V_0^\infty < \infty$ et si $E_\theta[\rho_J(X_1, \dots, X_j)]$ tend vers 0 quand $j \rightarrow \infty$ ou si L est bornée la suite $\{V_0^J\}$ converge vers V_0^∞ .

Démonstration

Si $V_0^\infty < \infty$, on a $\sum_{j=0}^{\infty} \psi_j(X_0, \dots, X_j) = 1 \quad [a, \theta] \quad \forall \theta$

Désignons par $(\varphi^\varepsilon, \delta^\varepsilon)$ une stratégie de Bayes pour le problème général (une telle stratégie existe toujours) et modifions si nécessaire, δ^ε de manière que δ_j^ε soit une règle de Bayes pour le problème basé sur X_1, \dots, X_j au sens du lemme 1.

Désignons par $\varphi^{\varepsilon, J}$ la règle φ^ε pour le problème tronqué en J ; autrement dit

$$\varphi_j^{\varepsilon, J} = \varphi_j^\varepsilon \quad \forall j \leq J-1 \quad \text{et} \quad \varphi_J^{\varepsilon, J} = 1$$

alors

$$\begin{aligned}
 & r(\tau, (\varphi^{\varepsilon, J}, \delta^{\varepsilon})) - r(\theta, (\varphi^{\varepsilon}, \delta^{\varepsilon})) \\
 &= \sum_{j=J}^{\infty} E_{\theta} \{ \psi_j^{\varepsilon}(X_1, \dots, X_j) \int_{\Delta} [L(T, y) \delta_j^{\varepsilon}(X_1, \dots, X_j; dy) - L(T, y) \delta_j^{\varepsilon}(X_1, \dots, X_j, dy)] \} \\
 &+ \sum_{j=J}^{\infty} E_{\theta} \{ \psi_j^{\varepsilon}(X_1, \dots, X_j) [C_j(T, X_1, \dots, X_j) - \mathcal{E}_j(T, X_1, \dots, X_j)] \}
 \end{aligned}$$

d'où

$$r(\tau, (\varphi^{\varepsilon, J}, \delta^{\varepsilon})) - r(\tau, (\varphi^{\varepsilon}, \delta^{\varepsilon})) \leq E_{\theta} [\psi_J^{\varepsilon, J}(X_1, \dots, X_J) \int_{\Delta} L(T, y) \delta_J^{\varepsilon}(X_1, \dots, X_J; dy)]$$

si $E_{\theta} [\rho_J(X_1, \dots, X_J)] \rightarrow 0$ quand J tend vers ∞ ou si L est bornée, le second membre de cette inégalité peut être rendu inférieur à 2ε pour J suffisamment grand (puisque

$$\int_{\Delta} [L(T, y) \delta_J^{\varepsilon}(X_1, \dots, X_J; dy)] \leq E_{\theta} [\rho_J(X_1, \dots, X_J) + \varepsilon]$$

Alors puisque :

$$r(\tau, (\varphi^{\varepsilon}, \delta^{\varepsilon})) \leq V_0^{\infty} + \varepsilon$$

nous avons pour J suffisamment grand

$$V_0^J \leq r(\tau, (\varphi^{\varepsilon, J}, \delta^{\varepsilon})) \leq V_0^{\infty} + 3\varepsilon .$$

Comme ε est arbitraire ceci achève la démonstration.

Exemple : Nous supposons que $\Theta = [0, 1] = \Delta$

$$L(\theta, y) = K(\theta - y)^2, \quad C_j(\theta, X_1, \dots, X_j) = jC$$

où C est une constante.

Si θ est la valeur du paramètre X_1, \dots, X_j , sont des variables aléatoires de Bernouilli indépendantes : $P(X_i = 1/\theta) = \theta$.

On veut trouver une stratégie de Bayes pour la distribution uniforme sur $[0,1]$.

Nous utilisons la procédure exposée.
 Déterminons la loi de $X = (X_1, X_2, \dots)$ et de T .

Si $P(X_1 = x_1 \cup X_2 = x_2 \cup \dots \cup X_j = x_j / T = \theta) = f(\theta)$

et si la densité de T est $g(\theta)$ on a :

$$P(X_1=x_1 \cup X_2=x_2 \cup \dots \cup X_j=x_j \cup \{0 \leq T \leq \theta\}) = \int_0^\theta f(t)g(t)dt$$

Or $P(X_1=x_1 \cup X_2=x_2 \dots \cup X_j=x_j / T = \theta)$

$$= P(X_1=x_1/T = \theta) \cdot P(X_2=x_2/T = \theta) \dots P(X_j=x_j/T = \theta)$$

$$= \theta^{\bar{x}_j} (1-\theta)^{j-\bar{x}_j} \text{ où } \bar{x}_j = \sum_{i=1}^j x_i = x_1+x_2+\dots+x_j$$

Ici, on aura :

$$P(X_1=x_1 \cup \dots \cup X_j=x_j) = \int_0^1 t^{\bar{x}_j} (1-t)^{j-\bar{x}_j} dt = B(\bar{x}_j+1, j-\bar{x}_j+1)$$

Il s'agit là d'une fonction eulérienne, de deuxième espèce

On a donc :

$$\int_0^1 t^{\bar{x}_j} (1-t)^{j-\bar{x}_j} dt = \frac{\Gamma(\bar{x}_j+1)\Gamma(j-\bar{x}_j+1)}{\Gamma(j+2)}$$

Déterminons la loi à postérieure

$$P[T \leq \theta / X_1=x_1 \cup \dots \cup X_j=x_j] = \frac{\Gamma(j+2) \int_0^\theta t^{\bar{x}_j} (1-t)^{j-\bar{x}_j} dt}{\Gamma(\bar{x}_j+1)\Gamma(j-\bar{x}_j+1)}$$

La loi à postérieure a donc pour densité

$$\frac{\Gamma(j+2)}{\Gamma(\bar{x}_j+1)\Gamma(j-\bar{x}_j+1)} \theta^{\bar{x}_j} (1-\theta)^{j-\bar{x}_j}$$

C'est donc la loi $\beta(\bar{x}_j+1, j-\bar{x}_j+1)$ puisque $\beta(a,b)$ a pour densité

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \text{ sur } [0,1]$$

On sait qu'une telle loi a pour moyenne :

$$\frac{\bar{x}_j+1}{j+2} = E(T)$$

variance :

$$\frac{(\bar{x}_j+1)(j-\bar{x}_j+1)}{(j+2)^2(j+3)} = \sigma^2(T)$$

Détermination d'une règle de décision terminale

On essaie de minimiser $\forall X_1, \dots, X_j$

$$\rho_j(X_1, \dots, X_j) = \inf_y \frac{\Gamma(j+2) k}{\Gamma(\bar{x}_j+1)\Gamma(j-\bar{x}_j+1)} \int_0^1 \theta^{\bar{x}_j} (1-\theta)^{j-\bar{x}_j} (\theta-y)^2 d\theta$$

$$(\theta-y)^2 = \theta^2 - 2y\theta + y^2 \text{ et } E(T^2) - E^2(T) = \sigma^2(T)$$

d'où

$$\rho_j(X_1, \dots, X_j) = \frac{\inf_y (\bar{x}_j+1)(j-\bar{x}_j+1)}{(j+2)^3(j+3)} + \frac{(\bar{x}_j+1)^2}{(j+2)^2} - 2y \frac{\bar{x}_j+1}{j+2} + y^2$$

il suffit donc de minimiser

$$y^2 - 2y \frac{\bar{x}_j+1}{j+2} = h(y)$$

on aura :

$$h'(y) = 2y - 2 \frac{\bar{x}_j+1}{j+2} = 0 \Rightarrow y = \frac{\bar{x}_j+1}{j+2}$$

et $y = \frac{\bar{x}_j+1}{j+2}$ est un élément de $[0,1]$

autrement dit, la règle terminale est une règle non aléatoire

$$d^0 = (d_0^0, d_1^0, \dots) \quad \text{où} \quad d_j^0(X_1, \dots, X_j) = \frac{\bar{x}_j + 1}{j+2}$$

Recherche de la règle d'arrêt

On aura alors :

$$\begin{aligned} U_j(X_1, \dots, X_j) &= K \left[-\frac{(\bar{x}_j + 1)}{(j+2)^2} + \frac{(\bar{x}_j + 1)(j - \bar{x}_j + 1)}{(j+2)^2(j+3)} + \frac{(\bar{x}_j + 1)^2}{(j+2)^2} \right] + jC \\ &= K \frac{(\bar{x}_j + 1)(j - \bar{x}_j + 1)}{(j+2)^2(j+3)} + jC \end{aligned}$$

Pour déterminer la règle d'arrêt pour le problème tronqué en J il faut calculer

$$E [U_j(X_1, \dots, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}]$$

Il faut donc connaître la loi de X_j conditionnelle à $X_1 = x_1, \dots, X_{j-1} = x_{j-1}$;
On a :

$$P\{X_j = x_j / X_1 = x_1, \dots, X_{j-1} = x_{j-1}\} = \frac{\Gamma(\bar{x}_j + 1) \Gamma(j - \bar{x}_j + 1) \Gamma(j+1)}{\Gamma(j+2) \Gamma(\bar{x}_{j-1} + 1) \Gamma(j - \bar{x}_{j-1})}$$

si $x_j = 0$ $\bar{x}_j = \bar{x}_{j-1}$ et $P\{X_j = x_j / X_1 = x_1, \dots, X_{j-1} = x_{j-1}\} = \frac{j - \bar{x}_{j-1}}{j+1}$

si $x_j = 1$ $\bar{x}_j = 1 + \bar{x}_{j-1}$ et $P\{X_j = x_j / X_1 = x_1, \dots, X_{j-1} = x_{j-1}\} = \frac{\bar{x}_{j-1} + 1}{j+1}$

en utilisant cette loi, nous devons calculer l'espérance conditionnelle :

$$E_0 [V_j^{(J)}(x_1, \dots, x_{j-1}, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}]$$

nécessaire pour le calcul de V_j^J .

Nous explicitons les calculs dans un exemple précis où $J = 5$, $K = 200$ et $C = 1$, c'est-à-dire $L(\theta, y) = 200 (\theta - y)^2$ et le coût constant égal à 1. Le tableau suivant résume les calculs.

Enfin, puisque $L(\theta, y)$ est bornée, les solutions des problèmes tronqués convergent vers la solution du problème général.

Il arrive fréquemment qu'une règle de Bayes pour le problème général soit une règle de Bayes pour un problème tronqué. Le théorème suivant donne une condition suffisante pour ce cas et aussi pour un nombre minimum d'observations requises par la règle de Bayes.

Théorème

Si $V_0^J \rightarrow V_0^\infty$ quand $J \rightarrow \infty$ et si pour tout $j > J_0$

$$\rho_{j-1}(X_1, \dots, X_{j-1}) - E_\theta \{ \rho_j(X_1, \dots, X_{j-1}, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1} \} \leq$$

$$E_\theta \{ C_j(T, X_1, \dots, X_j) - C_{j-1}(T, X_1, \dots, X_{j-1}) / X_1 = x_1, \dots, X_{j-1} = x_{j-1} \} [\underline{\alpha}, \theta] \forall \theta \in \Theta$$

alors

$$V_0^J = V_0^\infty.$$

Démonstration

Soit $J > J_0$ et considérons le problème de décision séquentielle tronqué en J . L'inégalité précédente s'écrit encore :

$$U_{j-1}(X_1, \dots, X_{j-1}) \leq E_\theta [U_j(X_2, \dots, X_j) / X_2 = x_1, \dots, X_{j-1} = x_{j-1}]$$

pour $j > J_0$ avec la probabilité 1. Donc $V_{J-1}^J = U_{J-1}$ (avec la probabilité 1) et par conséquent $V_j^J = U_j$ pour $j = J_0, \dots, J$ (avec la probabilité 1) ; cela entraîne que $V_{J_0}^J = V_{J_0}^{J_0}$ pour tout $J > J_0$ (avec la probabilité 1) donc $V_0^J = V_0^\infty$.

\bar{X}_5	d_5^0	coût C_5	ρ_5	$U_5 = V_5^5$			
0	1/7	5	3,06 12	8,06 12			
1	2/7	5	5,10 20	10,10 20			
2	3/7	5	6,12 24	11,12 24			
3	4/7	5	6,12 24	11,12 24			
4	5/7	5	5,10 20	10,10 20			
5	6/7	5	3,06 12	8,00 12			
\bar{X}_4	d_4^0	coût C_4	ρ_4	U_4	$E[V_5^5/X_1, \dots, X_4]$	V_4^5	φ_4^0
0	1/6	4	3,96 89	7,96 83	8,40 13	7,96 83	1
1	2/6 ou 1/3	4	6,34 92	10,34 92	10,44 21	10,34 92	1
2	3/6 ou 1/2	4	7,14 29	11,14 29	11,12 24	11,12 24	0
3	4/6 ou 2/3	4	6,34 92	10,34 92	10,44 21	10,34 92	1
4	5/6	4	3,96 83	7,96 83	8,40 13	7,96 83	1
\bar{X}_3	d_3^0	coût C_3	ρ_3	U_3	$E[V_4^5/X_1, \dots, X_3]$	V_3^5	φ_3^0
0	1/5	5	5,33 33	8,33 33	8,44 45	8,33 33	1
1	2/5	3	8,00 00	11,00 00	10,65 85	10,65 85	0
2	3/5	3	8,00 00	11,00 00	10,65 85	10,65 85	0
3	4/5	3	5,33 33	8,33 33	8,44 45	8,33 33	1
\bar{X}_2	d_2^0	coût C_2	ρ_2	U_2	$E[V_3^5/X_1, X_2]$	V_2^5	φ_2^0
0	1/4	2	7,50 00	9,50 00	9,00 00	9,00 00	0
1	1/2	2	10,00 00	12,00 00	10,00 00	10,00 00	0
2	3/4	2	7,50 00	9,50 00	9,00 00	9,00 00	0
\bar{X}_1	d_1^0	coût C_1	ρ_1	U_1	$E[V_2^5 / X_1]$	V_1^5	φ_1^0
0	1/3	1	11,11 11	12,11 11	10,33 33	10,33 33	0
1	2/3	1	11,11 11	12,11 11	10,33 33	10,33 33	0
\bar{X}_0	d_0^0	coût C_0	ρ_0	U_0	$E[V_1^5]$	V_0^5	φ_0^0
0	1/2	0	16,66 66	16,66 66	11,33 33	11,33 33	0

5

Remarque : La quantité

$$E_{\theta} \{C_j(T, X_1, \dots, X_j) - C_{j-1}(T, X_1, \dots, X_{j-1}) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}\}$$

est l'espérance conditionnelle du coût de faire une observation supplémentaire sachant que l'on a observé $X_1 = x_1, \dots, X_{j-1} = x_{j-1}$. La quantité

$$\rho_{j-1}(X_1, \dots, X_{j-1}) - E_{\theta} \{\rho_j(X_1, \dots, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}\}$$

représente l'espérance conditionnelle de la réduction de la perte réalisée en faisant une observation supplémentaire sachant que $X_1 = x_1, \dots, X_{j-1} = x_{j-1}$ a été observé. La conclusion $V_0^J = V_0^{\infty}$ signifie que nous pouvons obtenir les mêmes résultats en utilisant des règles tronquées en J_0 qu'en utilisant des règles non tronquées, en ce sens que le risque de Bayes est le même. Autrement dit, le théorème signifie ceci si $V_0^J \rightarrow V_0^{\infty}$ quand $J \rightarrow \infty$ et si pour tout $j \geq J_0$ l'espérance conditionnelle en "économisant la perte" est inférieure à l'espérance conditionnelle du coût en faisant une observation supplémentaire sachant que l'on a observé $X_1 = x_1, \dots, X_{j-1} = x_{j-1}$, alors une règle de Bayes pour le problème tronqué en J_0 est aussi une règle de Bayes pour le problème général.

Revenons à l'exemple. Ici

$$\begin{aligned} E_{\theta} [\rho_j(X_1, \dots, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}] &= \frac{K}{(j+2)^2(j+3)} E_{\theta} [(\bar{X}_j + 1)(j - \bar{X}_j + 1) / X_1, \dots, X_{j-1}] \\ &= \frac{K(1 + \bar{x}_{j-1})}{(j+1)(j+2)^2(j+3)} (j+1)(j - \bar{x}_{j-1} + 1)(j - 2\bar{x}_{j-1} - 1) = \frac{K(\bar{x}_{j-1} + 1)(j - \bar{x}_{j-1})}{(j+1)(j+2)} \end{aligned}$$

Donc

$$\begin{aligned} \rho_{j-1}(X_1, \dots, X_{j-1}) - E_{\theta} [\rho_j(X_1, \dots, X_{j-1}, X_j) / X_1 = x_1, \dots, X_{j-1} = x_{j-1}] \\ = \frac{K(\bar{x}_{j-1} + 1)(j - \bar{x}_{j-1})}{(j+1)^2(j+2)^2} \end{aligned}$$

Cette quantité est maximum pour $\bar{x}_{j-1} = \frac{j-1}{2}$. Donc

$$\frac{K(1 + \bar{x}_{j-1})(j - \bar{x}_{j-1})}{(j+1)^2(j+2)^2} \leq \frac{K \frac{j+1}{2} \frac{j+1}{2}}{(j+1)^2(j+2)^2} = \frac{K}{4(j+2)^2}$$

fonction décroissante de $j = 1, 2, \dots$.

Par conséquent, il existe J_0 (le plus petit entier j) tel que

$$\frac{K}{4(j+2)^2} \leq C$$

Pour $K = 200$ et $C = 1$ nous trouvons $J_0 = 5$. Alors la règle de Bayes que nous calculons pour le problème tronqué en $J_0 = 5$ est la règle de Bayes pour le problème général.

Approximation de V_0^∞

Avant de considérer la règle de Bayes pour le problème tronqué comme une méthode de calcul pour le problème général, il nous reste à éclairir un point ; en effet pour savoir si un entier J est, pour le problème tronqué assez grand pour réaliser l'approximation désirée, il est nécessaire de trouver le nombre V_0^∞ ou tout au moins de l'approcher aussi près que possible. Nous disposons déjà d'une suite décroissante $\{V_0^J\}$ convergeant vers V_0^∞ . Montrons maintenant que nous pouvons construire une suite croissante $\{W_0^J\}$ qui converge vers V_0^∞ .

Considérons une suite de problèmes tronqués identiques avec la seule modification suivante : si le problème est tronqué en J et si J est atteint on ne tient pas compte de la perte due au choix d'une décision terminale. Les W_0^J représentent le risque de Bayes minimum dans ce problème.

Théorème

Les W_0^J ainsi construits vérifient :

$$W_0^1 \leq W_0^2 \leq \dots \leq V_0^\infty$$

Si de plus, $E_\theta[\rho_J(X_2, \dots, X_j)] \rightarrow 0$ quand $J \rightarrow \infty$ ou si L est bornée, alors $W_0^J \rightarrow V_0^\infty$ quand $J \rightarrow \infty$.

Démonstration :

Il est facile de voir, dans ce problème, que toute règle utilisée pour le problème tronqué en $J+1$ donne un risque plus petit lorsque cette même règle est

utilisée dans le problème tronqué en J avec arrêt obligatoire si J est atteint.

Donc $W_0^J \leq W_0^{J+1}$. De même $W_0^J \leq V_0^\infty$ (la première assertion du théorème est démontrée).

Pour montrer la deuxième, supposons que $r_J(\tau, (\varphi, \delta))$ représente le risque de Bayes pour le problème tronqué modifié en J.

$$r_J(\tau, (\varphi, \delta)) = \sum_{j=0}^{J-1} E_\theta \{ \psi_j(X_1, \dots, X_j) \int_{\Delta} [L(T, y) \delta_j(X_j; dy)] \} \\ + C_j(T, X_1, \dots, X_j) + E_\theta [\psi_J(X_1, \dots, X_J) C_J(T, X_1, \dots, X_J)]$$

Si (φ^J, δ^J) est une ϵ -stratégie de Bayes pour le problème tronqué modifié

$$r_J(\tau, (\varphi^J, \delta^J)) \leq W_0^J + \epsilon$$

et

$$V_0^\infty - W_0^J \leq r(\tau, (\varphi^J, \delta^J)) - r_J(\tau, (\varphi^J, \delta^J)) + \epsilon \\ = E_\theta \{ \psi_J^J(X_1, \dots, X_J) L(T, \delta_J^J(X_1, \dots, X_J)) \} + \epsilon$$

Or l'hypothèse $E_\theta [\rho_J(X_2, \dots, X_j)] \rightarrow 0$ quand $J \rightarrow \infty$ entraîne immédiatement que

$$E_\theta [\psi_J^J(X_1, \dots, X_j) \int_{\Delta} [L(T, y) \delta_J^J(X_1, \dots, X_J; dy)] \} \rightarrow 0 \quad \text{quand } J \rightarrow \infty.$$

Pour achever la démonstration supposons maintenant que L soit bornée en B, de sorte que

$$V_0^\infty - W_0^J \leq B P_{\varphi^J} \{N=J\} + \epsilon$$

Il reste à prouver que $P_{\varphi^J} \{N=J\} \rightarrow 0$ quand $J \rightarrow \infty$

soit B' un nombre arbitrairement grand et soit

alors

$$S_J = \{(T, X_1, \dots, X_J) : C_j(T, X_1, X_J) < B'\}$$

$$\begin{aligned} V_0^{\infty+\epsilon} &\geq W_0^J + \epsilon \geq r_J(T, (\varphi^{(J)}, \delta^{(J)})) \\ &\geq E_{\theta} \{ \psi_J(X_1, \dots, X_J) C_J(T, X_1, \dots, X_J) \} \\ &\geq B' E_{\theta} \{ \psi_J(X_1, \dots, X_J) \chi_{S_J}(T, X_1, \dots, X_J) \} \\ &\geq B' [P_{\varphi^{(J)}} \{N=J\} - E_{\theta} \{ \psi_J(X_1, \dots, X_J) \chi_{S_J}(T, X_1, \dots, X_J) \}] \\ &\geq B' [P_{\varphi^{(J)}} \{N=J\} + P(S_J)] \end{aligned}$$

χ_{S_J} désigne la fonction caractéristique de S_J .
Donc

$$P_{\varphi^{(J)}} \{N=J\} \leq \frac{V_0^{\infty+\epsilon}}{B'} + P(S_J)$$

comme S_J tend vers l'ensemble vide quand J tend vers l'infini, on a :
 $\lim_{J \rightarrow \infty} P(S_J) = 0$; par suite :

$$\lim_{J \rightarrow \infty} P_{\varphi^{(J)}} \{N=J\} \leq \frac{V_0^{\infty+\epsilon}}{B'}$$

Ceci achève la démonstration puisque B' est arbitraire. Nous disposons donc d'un moyen d'encadrer V_0^{∞} .

Remarque - Pour calculer W_0^J , nous pouvons définir W_J^J par

$$W_J^J = E_{\theta} \{ C_J(T, X_1, \dots, X_J) / X_1, \dots, X_J \}$$

et pour $j < J$:

$$W_{j-1}^J = \min \{ W_{j-1}^J, E[W_j^J / X_1 = x_1, \dots, X_{j-1} = x_{j-1}] \}$$

IV-1 - RECHERCHE DE STRATEGIES MINIMAX A PARTIR DE STRATEGIES DE BAYES

Théorème 1

Soit $\{\tau_j\}$ une suite de probabilités sur (Θ, \mathcal{E}) et $\{(\delta_j, \varphi_j)\}$ une suite de stratégies de Bayes pour τ_j . Si

$$R[\tau_j, (\delta_j, \varphi_j)] \xrightarrow{j \rightarrow \infty} K$$

et si $\forall \theta \in (\Theta, \mathcal{E})$, il existe (δ^0, φ^0) telle que $R[\theta, (\delta^0, \varphi^0)] = K$ alors (δ^0, φ^0) est minimax.

Démonstration

Si (δ^0, φ^0) n'est pas minimax, il existe (δ^1, φ^1) telle que

$$\sup_{\theta \in \Theta} R[\theta, (\delta^1, \varphi^1)] < K$$

il existe donc $J \in \mathbb{N}$ tel que $\forall j \geq J$, on ait

$$\sup_{\theta \in \Theta} R[\theta, (\delta^1, \varphi^1)] < R[\tau_j, (\delta_j, \varphi_j)]$$

Soit

$$\begin{aligned} R[\tau_j, (\delta^1, \varphi^1)] &\leq \int_{\Theta} \sup_{\theta} R[\theta, (\delta^1, \varphi^1)] d(\tau_j(\theta)) \\ &\leq R[\tau_j, (\delta_j, \varphi_j)] \end{aligned}$$

Il y a donc contradiction avec le fait que (δ_j, φ_j) est une stratégie de Bayes pour τ_j . Le théorème est alors démontré.

Remarque : Ce résultat reste encore valable si $R[\theta, (\delta^0, \varphi^0)] \leq K$ ou $R[\theta, (\delta^0, \varphi^0)] = K$ et (δ^0, φ^0) est ε -stratégie de Bayes pour une certaine distribution.

IV-2 - CLASSES DE STRATEGIES ESSENTIELLEMENT COMPLETES

Rappel : Inégalité de Jessen

Soit une fonction f convexe définie sur une partie convexe Δ de \mathbb{R}^j .
 P est une mesure de probabilité sur Δ de moyenne finie alors
 $\int_{\Delta} f(z) dP(z) \geq f\left(\int_{\Delta} z dP(z)\right)$ ou encore $E[f(Z)] \geq f[E(Z)]$ (Z étant une variable aléatoire associée à z).

Ici nous essayons d'étendre au cas séquentiel des résultats connus de la décision statistique classique.

Propriété de convexité

Si Δ est une partie convexe de \mathbb{R}^k et si pour chaque valeur de $\theta \in \Theta$, $L(\theta, y)$ est une fonction convexe de y , si on ne considère que la classe des stratégies (δ, φ) pour lesquelles

$$\forall j, \forall x_1, \dots, x_j \quad \int_{\Delta} \delta_j(x_j; dy) < +\infty$$

alors la classe des stratégies dont les décisions terminales ne sont pas aléatoires est essentiellement complète.

Démonstration

Soit la stratégie (δ, φ) où $\delta = \{\delta_j\}$; montrons que la stratégie (δ', φ) avec $\delta'_j(x_j) = \int_{\Delta} y \delta_j(x_j; dy)$ est aussi bonne que (δ, φ) .

En effet $\forall j, \forall x_j$, on peut appliquer l'inégalité de Jessen

$$\int_{\Delta} L(\theta, y) \delta_j(x_j; dy) \geq L(\theta, \int_{\Delta} y \delta_j(x_j; dy))$$

et en tenant compte de l'expression de $R[\theta, (\delta, \varphi)]$, on a bien

$$R[\theta, (\delta', \varphi)] \leq R[\theta, (\delta, \varphi)].$$

Ce résultat s'applique à l'exemple traité, la propriété ayant été vérifiée directement.

IV-3 - EXHAUSTIVITE ET DECISION SEQUENTIELLE

Dans l'exemple que nous avons étudié dans la section III.3, nous avons introduit la statistique ΣX_i qui est une statistique exhaustive des observations. Nous allons ici, étudier le lien entre l'exhaustivité et le problème de décision séquentielle.

Théorème 2

Soit la structure statistique

$$[\mathbb{R}^j, \mathcal{A}_j, \{p_j(\theta, d\mathbf{x}_j), \theta \in (\Theta, \mathcal{G})\}] \quad \text{et } \mathcal{A}'_j \subset \mathcal{A}_j$$

avec \mathcal{A}'_j exhaustive, alors, nous avons

- a) pour toute fonction f \mathcal{A}_j - p_j -intégrable, il existe une fonction g \mathcal{A}'_j - p_j -intégrable telle que :

$$g(\mathbf{x}_j) = E_\theta[f(\mathbf{x}_j)/\mathcal{A}'_j]$$

- b) pour toute mesure de transition $\mu(\mathbf{x}_j, D)$ sur $(\mathbb{R}^j, \mathcal{A}_j)$ et $(\mathbb{R}^k, \mathcal{A}_k)$, il existe une mesure de transition $\nu(\mathbf{x}_j, D)$ sur $\mathbb{R}^j \times \mathcal{A}_k$ et telle que :

$$\nu(\mathbf{x}_j, D) = E_\theta[\mu(\mathbf{x}_j, D)/\mathcal{A}'] \quad [\mathcal{A}_j, \theta] \quad \forall D \in \mathcal{A}_k \quad \forall \theta \in \Theta$$

Démonstration :

Pour simplifier les notations, nous notons : $(\mathbb{R}^j, \mathcal{A}_j)$

$$(\Omega, \mathcal{A}) \text{ et } \mathcal{A}'_j, \mathcal{A}'$$

et nous désignons par \mathcal{F} , l'ensemble des fonctions \mathcal{A} - p_j -intégrables telles que a) soit satisfaite pour un g \mathcal{A}' - p_j -intégrables $\mathcal{F} \neq \emptyset$ car \mathcal{F} contient les fonctions indicatrices. Donc \mathcal{F} contient aussi les fonctions étagées.

Soit une fonction positive f mesurable : elle est limite simple d'une suite croissante $\{f_j\}$ de fonctions étagées.

$$f_j \rightarrow f \Rightarrow \begin{cases} E_\theta[f_j/\mathcal{G}'] \rightarrow E_\theta[f/\mathcal{G}'] \\ [\mathcal{G}, \theta] \quad \forall \theta \in \Theta \end{cases}$$

Or

$\mathfrak{F}\{g_j\}$ \mathcal{G}' - p_j -intégrable telle que

$g_j = E_\theta[f_j/\mathcal{G}']$ donc $g_j \rightarrow g$ et a) est satisfait

Or \mathfrak{F} contenant les fonctions positives et négatives contient toutes les fonctions intégrables (f intégrable : $f = f^+ - f^-$).

Pour établir la condition b), nous allons prendre $k = 1$ et remplacer $(\mathbb{R}^k, \mathcal{G}_k)$ par $(\mathbb{R}, \mathcal{G})$.

Soit $I_r =]-\infty, r[$ posons $f(r, z) = \mu(z, I_r)$.

Considérons $q \in Q$, d'après a)

$$g^*(q, z) = E_\theta[f(q, z)/\mathcal{G}'] \quad [\mathcal{G}, \theta] \quad \forall \theta \in \Theta$$

Soit $q_1, q_2, \dots, q_n, \dots$ une suite d'éléments de Q .

On a :

$$f(\text{Min}(q_n, q_m), z) \leq f(\text{Max}(q_n, q_m), z)$$

d'où $\forall \theta \in \Theta$

$$g^*(\text{Min}(q_n, q_m), z) \leq g^*(\text{Max}(q_n, q_m), z) \quad [\mathcal{G}, \theta]$$

donc g^* est une fonction croissante de $q \in Q$

$$\forall l \in \mathbb{N} \text{ soit } u_l = \text{Min}(q_1, q_2, \dots, q_l)$$

On a donc

$$u_1 \geq u_2 \geq \dots \geq u_l \geq u_{l+1} \dots \text{ et } \lim_{l \rightarrow \infty} u_l = -\infty$$

Puisque $\mu(z, \mathbb{R})$ est p_j -intégrable $\forall \theta \in \Theta$, on a : $\mu(z, \mathbb{R}) < +\infty$ (sauf sur un ensemble de \mathcal{G} de mesure nulle pour chaque θ).

Or

$$I_{u_1} \supset I_{u_2} \supset \dots \supset I_{u_l} \supset I_{u_{l+1}} \supset \dots$$

et

$$\mu(z, I_{u_l}) < +\infty \quad \forall l$$

d'où $\mu(z, I_{u_1}) \geq \mu(z, I_{u_{1+1}})$ et $\mu(z, I_{u_1}) \xrightarrow{1 \rightarrow \infty} 0$

(car la mesure μ est finie).

d'où

$$\left\{ \begin{array}{l} f(u_1, z) \geq f(u_{1+1}, z) \quad \forall 1 \\ \lim_{1 \rightarrow \infty} f(u_1, z) = 0 \end{array} \right. \quad [\mathcal{A}]$$

donc

$$\lim_{1 \rightarrow \infty} g^*(u_1, z) = 0 \quad [\mathcal{A}']$$

or

$$\lim_{1 \rightarrow \infty} g^*(u_1, z) = \inf_j g^*(q_j, z) = 0 \quad [\mathcal{A}']$$

le même raisonnement permet d'en déduire que

$$\lim_{1 \rightarrow \infty} g^*(v_1, z) = \sup g^*(q_j, z) < +\infty \quad [\mathcal{A}'] \quad (\mu(z, I_{v_1}) < +\infty)$$

on posera

$$v_1 = \text{Max}(q_1, \dots, q_1).$$

Donc, sauf pour $A \in \mathcal{A}'$ de mesure nulle pour chaque θ on a :

$$\forall z \in \mathcal{C}_{\Omega}^A \quad \left\{ \begin{array}{l} g^*(q, z) \text{ fonction croissante de } q \in Q \\ \sup_{q \in Q} g^*(q, z) < +\infty, \inf_{q \in Q} g^*(q, z) = 0 \end{array} \right.$$

En définissant $g(q, z)$ par :

$$g(q, z) = \begin{cases} g^*(q, z) & \forall z \in \mathcal{C}_{\Omega}^A \\ 0 & \forall z \in A. \end{cases}$$

$g(q, z)$ a les propriétés de g^* pour toute valeur de y , et pour chaque $q \in Q$, est mesurable par rapport à \mathcal{A}' .

Enfin $g(q, z) = E_{\theta}[f(q, z)] \quad [\mathcal{A}, \theta]$

Soit $r \in \mathbb{R}$ et $\{r_j\}$ une suite croissante d'éléments de \mathbb{R} convergeant vers r (\mathbb{R} est dense dans \mathbb{R}). Définissons $h(r, z)$ par

$$h(r, z) = \lim_{j \rightarrow \infty} g(r_j; z) \quad (\text{ce qui a un sens})$$

et

$$h(r, z) \text{ est ainsi définie } \forall r \in \mathbb{R}$$

$\forall z \in \Omega$ $h(r, z)$ est croissante par rapport à r , continue à gauche par construction, et $\inf_r h(r, z) = 0$, $\sup_r h(r, z) < +\infty$.

De plus $f(r_1, z) \leq f(r_2, z) \leq \dots \leq f(r_j, z) \leq \dots \quad \forall z$

D'où d'après b) $\forall r, \forall \theta \in \Theta$

$$h(r, z) = \lim_{j \rightarrow \infty} E_{\theta}[f(r_j; z)/\mathcal{A}'] = E_{\theta}[f(r, z)/\mathcal{A}'] \quad [\mathcal{A}, \theta].$$

Nous définissons pour chaque $z \in B$, $\nu(B, z)$ la mesure unique finie sur \mathbb{R} telle que

$$\nu(I_r, z) = h(r, z)$$

h est mesurable pour chaque B , par rapport à \mathcal{A}' ; on montre (cf. H. Robbins : mixture of distributions Annales of Mathematic Statistics Vol. 19 1948 p. 360 à 369) que ν est \mathcal{A}' -mesurable $\forall B$.

Soit alors θ fixé et $A_0 \in \mathcal{A}'$ fixé : posons

$$m(B) = \int_{A_0} \mu(B, z) P_j(\theta) \quad n(B) = \int_{A_0} \nu(B, z) P_j(\theta)$$

m et n sont deux mesures finies sur \mathbb{R}

$$\begin{aligned} n(I_r) &= \int_{A_0} h(r, z) P_j(\theta) = \int_{A_0} g(r, z) P_j(\theta) = \int_{A_0} g^*(r, z) p_j(\theta) \\ &= \int_{A_0} E_{\theta}[\mu(z, I_r)/\mathcal{A}'] P_j(\theta) = \int_{A_0} \mu(z, I_r) P_j(\theta) \end{aligned}$$

d'où
$$n(I_r) = m(I_r) \quad \forall r \in \mathbb{Q} \Rightarrow n(I_r) = m(I_r) \quad \forall r \in \mathbb{R}$$

$$\Rightarrow n(B) = m(B) \quad \forall B \in \mathcal{B}$$

d'où puisque θ et A_0 sont arbitraires

$$\int_{A_0} \mu(B, z) P_j(\theta) = \int_{A_0} \nu(B, z) P_j(\theta) \quad A_0 \in \mathcal{A}', B \in \mathcal{B}, \theta \in \Theta$$

or pour chaque $B \in \mathcal{B}$ ν est \mathcal{A}' -mesurable et unique (Radon-Niky) ; on en déduit :

$$\nu(z, B) = E_\theta[\mu(z, B) / \mathcal{A}'] [\mathcal{A}, \theta] \quad \forall \theta \in \Theta \quad \forall B \in \mathcal{B}$$

et il est clair que ν est \mathcal{A}' -intégrable $\forall B$

C.Q.F.D.

Revenons au cas séquentiel

Définition

Une suite $\{T_j\}$ de statistiques ($j=1,2,\dots$) est dite exhaustive si, pour chaque valeur de j , T_j est exhaustive relativement à \mathbb{R}^j .

Ici donc, nous reconsidérons $[\mathbb{R}^\infty, \mathcal{A}_\infty, \{P_\theta, \theta \in (\Theta, \mathcal{C})\}]$ que nous notons encore $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{C})\}]$ alors que $[\Omega_j, \mathcal{A}_j, \{P_j(\theta), (\theta \in \Theta, \mathcal{C})\}]$ désignera la structure statistique $[\mathbb{R}^j, \mathcal{B}_{\mathbb{R}^j}, \{P_j(\theta), \theta \in (\Theta, \mathcal{C})\}]$.

On a évidemment :

$$\mathcal{A}_j \subset \mathcal{A}_{j+1} \subset \dots \subset \mathcal{A}_\infty$$

si T_j est exhaustive, nous notons $\mathcal{A}_j^!$ la tribu engendrée par T_j ($\mathcal{A}_j^! \subset \mathcal{A}_j$).

$\mathcal{A}_j^!$ est exhaustive par définition.

Supposons que (Δ, \mathcal{D}) soit un espace du type $(\mathbb{R}^k, \mathcal{B}_{\mathbb{R}^k})$.

Lemme 1

Si $\{T_j\}$ est exhaustive et si $\{\delta_j, \Psi_j\}$ est une stratégie, basée sur $\{T_j\}$, il existe une stratégie $\{\delta'_j, \Psi_j\}$ telle que δ'_j est une fonction \mathcal{G}'_j -mesurable et

$$\int_{\mathbb{R}^j} \Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j; D) p_j[\theta, d\mathbf{x}_j] \quad \forall \theta$$

$$= \int_{\mathbb{R}^j} \Psi_j(\mathbf{x}_j) \delta'_j(\mathbf{x}_j; A) p_j[\theta, d\mathbf{x}_j] \quad \forall D \in \mathcal{D} \quad \forall \theta \in \Theta$$

Définition

On dira que δ'_j est basée sur \mathcal{G}'_j ou T_j

Démonstration

$\Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j; D)$ est

- $\forall \mathbf{x}_j \in \mathbb{R}^j$ une mesure sur (Δ, \mathcal{D})
- $\forall D \in \mathcal{D}$ une fonction de \mathbf{x}_j \mathcal{G}_j -intégrable $\forall \theta$ puisque majorée par 1.

On peut alors appliquer le théorème 2 précédent :

il existe $\lambda_j(\mathbf{x}_j, D)$ telle que

$\forall \mathbf{x}_j$ λ_j est une mesure sur (Δ, \mathcal{D})

$\forall D \in \mathcal{D}$ λ_j est une fonction de \mathbf{x}_j \mathcal{G}'_j -intégrable $\forall \theta$

et

$$\lambda_j(z, D) = E_\theta[\Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j; D) / \mathcal{G}'_j] \quad [\mathcal{G}, \theta] \quad (1)$$

$$\forall A \in \mathcal{D} \quad \forall \theta \in \Theta$$

Posons

$$\delta_j^!(\mathbf{x}_j, D) = \begin{cases} \frac{\lambda_j(\mathbf{x}_j, D)}{\lambda_j(\mathbf{x}_j, \Delta)} & \text{si } \lambda_j(\mathbf{x}_j, \Delta) > 0 \\ \pi_j \text{ (mesure quelconque)} & \text{si } \lambda_j(\mathbf{x}_j, \Delta) = 0 \end{cases}$$

$$\int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \delta_j^!(\mathbf{x}_j, D) p_j(\theta; d\mathbf{x}_j) = \int_{\mathbb{R}^\infty} E_\theta[\Psi_j(\mathbf{x}_j) / \mathcal{G}_j] \delta_j^!(\mathbf{x}_j, D) p_j(\theta; d\mathbf{x}_j)$$

si on utilise l'égalité (1) avec $D = \Delta$

$$\lambda_j(\mathbf{x}_j; \Delta) = E_\theta[\Psi_j(\mathbf{x}_j) / \mathcal{G}_j]$$

d'où

$$\int_{\mathbb{R}^\infty} \lambda_j(\mathbf{x}_j; \Delta) \delta_j^!(\mathbf{x}_j, D) p_j(\theta; d\mathbf{x}_j) = \int_{\mathbb{R}^\infty} \lambda_j(\mathbf{x}_j, D) p_j(\theta; d\mathbf{x}_j)$$

et d'après (1) cette intégrale est égale à :

$$\int_{\mathbb{R}^\infty} E_\theta[\Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j, D) / \mathcal{G}_j] p_j(\theta; d\mathbf{x}_j)$$

$$= \int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j, D) p_j(\theta; d\mathbf{x}_j)$$

d'où

Théorème 3

Si $\{T_j\}$ est une suite de statistiques exhaustives pour θ la classe des stratégies (δ_j, φ_j) où $\delta_j, \forall j$, est basée sur T_j est essentiellement complète.

Démonstration

Si on utilise δ' définie au lemme 1 précédent, il suffit de montrer que $\forall j$, on a :

$$\int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy) = \int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \int_{\Delta} L(\theta, y) \delta_j^!(\mathbf{x}_j; dy)$$

L'égalité est évidente si $L(\theta, y)$ est, $\forall \theta$, une fonction élémentaire

$$L(\theta, y) = \sum_{i=1}^k L_i(\theta) \chi_{C_i}(\theta) \quad C_i(\theta) \in \mathcal{D}$$

l'intégrale s'écrit en effet :

$$\int_{\mathbb{R}^\infty} \psi_j(\mathbf{x}_j) \sum_{i=1}^k L_i(\theta) \delta_j(\mathbf{x}_j; C_i)$$

et il suffit d'appliquer le lemme 1.

Si $L(\theta, y)$ est une fonction quelconque, on réalisera une approximation définie par :

$$|L(\theta, y) - \sum_{i=1}^k L_i(\theta) \chi_{C_i}(\theta)| \leq \epsilon$$

ce qui est possible car $|\psi_j| \leq 1$ et δ_j est une probabilité donc une mesure finie.

Mais nous utiliserons dans la section suivante, une méthode faisant intervenir le théorème de Lebesgue pour les fonctions L positives, la propriété s'étendant ensuite aux fonctions de signe quelconque.

Nous allons maintenant étudier des stratégies entièrement basées sur $\{T_j\}$ c'est-à-dire des stratégies $\{\delta_j, \varphi_j\}$ telles que φ_j , donc ψ_j , est pour tout j , \mathcal{G}_j -mesurable. Mais avant d'en arriver là, il est nécessaire d'introduire la notion de transitivité.

IV-4 - TRANSITIVITE ET DECISION SEQUENTIELLE

Définition

Une suite $\{T_j\}$ de statistiques est dite transitive si

$$\forall j \quad E_\theta [T_{j+1}(X_1, \dots, X_{j+1}) / X_1, \dots, X_j] = E_\theta [T_{j+1}(X_1, \dots, X_{j+1}) / T_j(X_1, \dots, X_j)]$$

$$[\mathcal{G}, \theta] \quad \forall \theta \in \Theta$$

Nous noterons, pour simplifier T_j au lieu de $T_j(X_1, \dots, X_j)$.

La transitivité signifie donc, en clair que T_j est aussi bon renseignement sur T_{j+1} que $\mathbf{X}_j = (X_1, \dots, X_j)$.

La transitivité traduite en termes de tribus s'exprimera par $\{\mathcal{G}_j\}$ étant une suite de sous-tribus de \mathcal{G}_∞ ($\mathcal{G}_j \subset \mathcal{G}_{j+1}$) et $\mathcal{G}'_j \subset \mathcal{G}_j$.

On dira que $\{\mathcal{G}'_j\}$ est une suite transitive de tribus si $\forall j, \forall A \in \mathcal{G}'_{j+1}$
 $\forall \theta \in \Theta$

$$E_\theta[x_A/\mathcal{G}_j] = E_\theta[x_A/\mathcal{G}'_j] \quad [A, \theta]$$

Théorème 4

La suite $\{\mathcal{G}'_j\}$ est transitive si $\forall j, \forall A \in \mathcal{G}_j \quad \forall \theta \in \Theta$

$$E_\theta[x_A/\mathcal{G}'_{j+1}] = E_\theta[E_\theta(x_A/\mathcal{G}'_j)/\mathcal{G}'_{j+1}] \quad [A, \theta]$$

Démonstration

Soit $A \in \mathcal{G}_j$ et $B \in \mathcal{G}'_{j+1}$

On aura :

$$\int_A E_\theta[x_B/\mathcal{G}_j] p_j(\theta, d\mathbf{x}_j) = \int_A x_A p_j(\theta, d\mathbf{x}_j)$$

par définition même de l'espérance conditionnelle).

$$= \int_\Delta x_A \cdot x_B p_j(\theta, d\mathbf{x}_j) = \int_B x_A p_j(\theta, d\mathbf{x}_j) = \int_B E_\theta[x_A/\mathcal{G}'_{j+1}] p_j(\theta, d\mathbf{x}_j) \quad (1)$$

D'autre part, grâce aux mêmes techniques. :

$$\begin{aligned} & \int_A E_\theta[x_B/\mathcal{G}'_j] p_j(\theta, d\mathbf{x}_j) = \int_\Delta x_A E_\theta[x_B/\mathcal{G}'_j] p_j(\theta, d\mathbf{x}_j) \\ &= \int_\Delta E_\theta[x_A \cdot E_\theta(x_B/\mathcal{G}'_j)/\mathcal{G}'_j] p_j(\theta, d\mathbf{x}_j) = \int_\Delta E_\theta[x_B/\mathcal{G}'_j] \cdot E_\theta[x_A/\mathcal{G}'_j] p_j(\theta, d\mathbf{x}_j) \\ &= \int_B E_\theta(x_A/\mathcal{G}'_j) p_j(\theta, d\mathbf{x}_j) \quad \text{or } B \in \mathcal{G}'_{j+1} \end{aligned}$$

donc pour la même raison :

$$\int_B E_\theta(x_A/\mathcal{G}'_j) p_j(\theta, d\mathbf{x}_j) = \int_B E_\theta[E_\theta(x_A/\mathcal{G}'_j)/\mathcal{G}'_{j+1}] p_j(\theta, d\mathbf{x}_j) \quad (2)$$

Tenant compte des 2 relations (1) et (2), on a :

$$\int_A E_\theta [x_B / \mathfrak{a}_j] p_j(\theta, d\mathfrak{x}_j) = \int_A E_\theta [x_B / \mathfrak{a}_j^!] p_j(\theta, d) \quad (3)$$

$$\Leftrightarrow \int_B E_\theta [x_A / \mathfrak{a}_{j+1}^!] p_j(\theta, d) = \int_B E_\theta [E_\theta(x_A / \mathfrak{a}_j^!) / \mathfrak{a}_{j+1}^!] p_j(\theta, d) \quad (4)$$

or quand on a (3) pour tout A, on a :

$$E_\theta [x_A / \mathfrak{a}_j] = E_\theta [x_B / \mathfrak{a}_j^!] \quad [\mathfrak{a}, \theta]$$

il en est de même pour la relation (4) et on a donc bien :

$$E_\theta [x_A / \mathfrak{a}_{j+1}^!] = E_\theta [E_\theta(x_A / \mathfrak{a}_j^!) / \mathfrak{a}_{j+1}^!]$$

c.q.f.d.

Corollaire :

Pour toute fonction \mathfrak{a}_j -intégrable, on a :

$$E_\theta [f / \mathfrak{a}_{j+1}^!] = E_\theta [E_\theta(f / \mathfrak{a}_j^!) / \mathfrak{a}_{j+1}^!] \quad [\mathfrak{a}, \theta] \quad \forall \theta \in \Theta$$

La démonstration de ce corollaire se fait d'une manière classique

D'après le théorème 4, le corollaire est vrai pour les fonctions indicatrices.

Il est donc vrai pour les fonctions en escalier.

Par suite vrai pour les fonctions positives qui sont limites croissantes de fonctions en escalier.

Enfin, il est vrai pour une fonction f de signe quelconque car

$$f = f^+ - f^-.$$

Remarques

- 1) Il arrive parfois que le corollaire soit pris pour une définition.
- 2) D'après la définition, on a, pour la même raison et pour toute fonction g \mathcal{G}_{j+1} -intégrable, on a :

$$E_{\theta}[g/\mathcal{G}_j] = E_{\theta}[g/\mathcal{G}_j^!]$$

$$[\mathcal{A}, \theta] \quad \forall \theta$$

En considérant $\mathcal{G}_j^!$ comme engendrée par T_j , ces résultats se traduisent en termes de statistiques.

IV-5 - TRANSITIVITE EXHAUSTIVITE ET DECISION SEQUENTIELLE

Nous étudions maintenant le cas d'une suite $\{T_j\}$ de statistiques à la fois transitive et exhaustive.

Théorème 5

Si $\{T_j\}$ est une suite de statistiques exhaustive et transitive, à toute règle d'arrêt $\{\varphi_j\}$, on peut associer une règle d'arrêt $\{\varphi_j^!\}$ basée sur T_j et telle que :

$$\forall j \quad E_{\theta}[\varphi_j/T_j] = E_{\theta}[\varphi_j^!/T_j] \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta$$

où en termes de tribus : si $\{\mathcal{G}_j^!\}$ est une suite de tribus exhaustive et transitive, alors

$$E_{\theta}[\varphi_j/\mathcal{G}_j^!] = E_{\theta}[\varphi_j^!/\mathcal{G}_j^!] \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta \quad (5)$$

(et $\varphi_j^!$ est $\mathcal{G}_j^!$ -mesurable).

Démonstration

$$\psi_j = (1-\varphi_0) \dots (1-\varphi_{j-1}) \varphi_j$$

posons

$$\Lambda_j = (1-\varphi_0) \dots (1-\varphi_{j-1})$$

D'après le théorème 2 de la section IV, \mathcal{A}_j étant exhaustive, il existe f_j^* et f_j telles que :

$$f_j^* = E_{\theta}[\psi_j / \mathcal{A}_j] [\mathbf{a}, \theta] \quad \forall \theta \in \Theta$$

$$f_j = E_{\theta}[\Lambda_j / \mathcal{A}_j] [\mathbf{a}, \theta] \quad \forall \theta \in \Theta$$

comme $0 \leq \psi_j \leq \Lambda_j$ on a : $0 \leq f_j^* \leq f_j$

Posons

$$\varphi_j^* = \begin{cases} f_j^* / f_j & \text{si } f_j > 0 \\ 1 \text{ (par exemple)} & \text{si } f_j = 0 \end{cases}$$

alors φ_j^* est \mathcal{A}_j -mesurable. Montrons que la relation (5) est satisfaite.

Soit $\{\psi_j^*\}$ et $\{\Lambda_j^*\}$ associées à φ_j^* .

On va montrer que

$$E_{\theta}[\Lambda_j / \mathcal{A}_j] = E_{\theta}[\Lambda_j^* / \mathcal{A}_j] [\mathbf{a}, \theta] \quad \forall \theta, \forall j$$

En effet (6) \Rightarrow (5)

$$\begin{aligned} E_{\theta}[\psi_j^* / \mathcal{A}_j] &= E_{\theta}[\varphi_j^* \cdot \Lambda_j^* / \mathcal{A}_j] \\ &= \varphi_j^* E_{\theta}[\Lambda_j^* / \mathcal{A}_j] = \varphi_j^* E_{\theta}[\Lambda_j / \mathcal{A}_j] = \varphi_j^* f_j = f_j^* \\ &= E_{\theta}[\psi_j / \mathcal{A}_j]. \end{aligned}$$

Si $f_j = 0$ on a alors $f_j^* = 0$ et le résultat est évident.

Nous allons établir (6) par récurrence sur j .

Pour $j = 1$ la relation (6) est vraie car $\Lambda_1 = 1 = \Lambda_1^!$ et évidemment la relation (5) est également vraie.

Hypothèse de récurrence :

Supposons (6) vraie donc (5) vraie au rang j et montrons que (5) est vraie au rang $j+1$, pour toute valeur de θ .

On a :

$$\begin{aligned} E_{\theta}[\Lambda_{j+1}/\mathfrak{a}_{j+1}^!] &= E_{\theta}[(\Lambda_j - \Psi_j)/\mathfrak{a}_{j+1}^!] \\ &= E_{\theta}[\Lambda_j/\mathfrak{a}_{j+1}^!] - E_{\theta}[\Psi_j/\mathfrak{a}_{j+1}^!] \end{aligned}$$

En appliquant le corollaire à Λ_j et Ψ_j et grâce à la transitivité

$$E_{\theta}[\Lambda_j/\mathfrak{a}_{j+1}^!] = E_{\theta}[E_{\theta}(\Lambda_j/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!]$$

$$E_{\theta}[\Psi_j/\mathfrak{a}_{j+1}^!] = E_{\theta}[E_{\theta}(\Psi_j/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!]$$

et donc

$$\begin{aligned} E_{\theta}[\Lambda_{j+1}/\mathfrak{a}_{j+1}^!] &= E_{\theta}[E_{\theta}(\Lambda_j/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!] - E_{\theta}[E_{\theta}(\Psi_j/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!] \\ &= E_{\theta}[E_{\theta}(\Lambda_j^!/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!] - E_{\theta}[E_{\theta}(\Psi_j^!/\mathfrak{a}_j^!)/\mathfrak{a}_{j+1}^!] \end{aligned}$$

en utilisant de nouveau le corollaire, il vient :

$$\begin{aligned} E_{\theta}[\Lambda_{j+1}/\mathfrak{a}_{j+1}^!] &= E_{\theta}[\Lambda_j^!/\mathfrak{a}_{j+1}^!] - E_{\theta}[\Psi_j^!/\mathfrak{a}_{j+1}^!] \\ &= E_{\theta}[\Lambda_{j+1}^!/\mathfrak{a}_{j+1}^!] \end{aligned}$$

c.q.f.d.

Théorème 6

Si $\{T_j\}$ est une suite de statistiques exhaustive et transitive, à chaque stratégie $\{\delta_j, \Psi_j\}$, on peut associer une stratégie $\{\delta_j^!, \Psi_j^!\}$ basée sur $\{T_j\}$ telle que :

$$\int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) \\ = \int_{\mathbb{R}^\infty} \Psi_j^!(\mathbf{x}_j) \delta_j^!(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) \quad \forall D \quad \forall \theta \in \Theta$$

On a donc

$$\begin{cases} \delta_j^!(\mathbf{x}_j; D), \forall D \in \mathcal{D} \text{ est } \mathcal{G}_j^! \text{-mesurable} \\ \Psi_j^!(\mathbf{x}_j) \text{ est } \mathcal{G}_j^! \text{-mesurable} \end{cases}$$

Démonstration

Comme $\mathcal{G}_j^!$ est exhaustive (d'après le lemme 1 de la section IV), il existe $\{\delta_j^!\}$ basée sur $\mathcal{G}_j^!$ telle que :

$$\int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \delta_j(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) = \int_{\mathbb{R}^\infty} \Psi_j(\mathbf{x}_j) \delta_j^!(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j)$$

D'autre part, $\mathcal{G}_j^!$ étant exhaustive et transitive, d'après le théorème 5 de la section IV il existe $\Psi_j^!$ telle que :

$$E_\theta[\Psi_j / \mathcal{G}_j^!] = E_\theta[\Psi_j^! / \mathcal{G}_j^!] \quad [\mathcal{G}_j, \theta] \quad \forall \theta \in \Theta$$

Or

$$\begin{aligned}
 \int_{\mathbb{R}^\infty} \psi_j(\mathbf{x}_j) \delta_j'(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) &= \int_{\mathbb{R}^\infty} E_\theta \{ \psi_j(\mathbf{x}_j) \delta_j'(\mathbf{x}_j; D) / \mathcal{G}_j \} p_j(\theta; d\mathbf{x}_j) \\
 &= \int_{\mathbb{R}^\infty} E_\theta [\psi_j / \mathcal{G}_j] \delta_j'(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) \\
 &= \int_{\mathbb{R}^\infty} E_\theta [\psi_j' / \mathcal{G}_j] \delta_j'(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) = \int_{\mathbb{R}^\infty} E_\theta [\psi_j'(\mathbf{x}_j) \delta_j'(\mathbf{x}_j; D) / \mathcal{G}_j] p_j(\theta; d\mathbf{x}_j) \\
 &= \int_{\mathbb{R}^\infty} \psi_j'(\mathbf{x}_j) \delta_j'(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j)
 \end{aligned}$$

c.q.f.d.

Le résultat le plus important fait l'objet du théorème suivant :

Théorème 7

Si $\{T_j\}$ est une suite exhaustive et transitive et si $C_j(\theta; \mathbf{x}_j)$ dépend, en fait, de $\mathbf{x}_j = (X_1, \dots, X_j)$ par l'intermédiaire de T_j pour tout j , alors la classe des stratégies basées sur $\{T_j\}$ est essentiellement complète.

Démonstration :

Considérons une stratégie (δ_j, φ_j) . En décomposant le coût on a :

Coût relatif aux C_j

$$\begin{aligned}
 &\sum_j \int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) C_j(\theta, \mathbf{x}_j) p_j(\theta, d\mathbf{x}_j) \\
 &= \sum_j \int_{\mathbb{R}^j} [E_\theta \psi_j(\mathbf{x}_j)] C_j(\theta, \mathbf{x}_j) / \mathcal{G}_j] p_j(\theta, d\mathbf{x}_j) \\
 &= \sum_j \int_{\mathbb{R}^j} C_j(\theta, \mathbf{x}_j) E_\theta \{ \psi_j(\mathbf{x}_j) / \mathcal{G}_j \} p_j(\theta, d\mathbf{x}_j) \\
 &= \sum_j \int_{\mathbb{R}^j} C_j(\theta, \mathbf{x}_j) E_\theta \{ \psi_j'(\mathbf{x}_j) / \mathcal{G}_j \} p_j(\theta, d\mathbf{x}_j)
 \end{aligned}$$

$$= \sum_j \int_{\mathbb{R}^j} \psi_j^!(\mathbf{x}_j) C_j(\theta, \mathbf{x}_j) p_j(\theta, d\mathbf{x}_j).$$

Coût relatif aux δ_j (dû à la décision terminale)

$$\int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j)$$

Si $L(\theta, y) = \mathcal{L}(\theta) \chi_{D(\theta)}$ $D(\theta) \in \mathcal{D}$.

alors $\int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy) = \mathcal{L}(\theta) \delta_j(\mathbf{x}_j; D)$

et d'après le théorème 5 de la section IV

$$\begin{aligned} & \int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \mathcal{L}(\theta) \delta_j(\mathbf{x}_j; D) p_j(\theta; d\mathbf{x}_j) \\ &= \int_{\mathbb{R}^j} \psi_j^!(\mathbf{x}_j) \delta_j^!(\mathbf{x}_j; D) \mathcal{L}(\theta) p_j(\theta, d\mathbf{x}_j) \end{aligned}$$

Nous venons donc de démontrer le théorème 7 pour

$L(\theta, y)$ de la forme $\mathcal{L}(\theta) \chi_{D(\theta)}$

il est facile alors de voir que le résultat est vrai pour toute fonction $L(\theta, y)$ de la forme :

$$L(\theta, y) = \sum_{i=1}^k \mathcal{L}_i(\theta) \chi_{D_i(\theta)}$$

Le théorème sera encore vrai pour une fonction $L(\theta, y)$ à valeurs positives et mesurable, car pour une telle fonction, on sait qu'il existe, pour chaque valeur de θ , une suite $\{L_j(\theta, y)\}$ du type précédent et telle que $L(\theta, y)$ soit limite simple des $L_i(\theta, y)$ (pour θ fixé).

Alors :

$$\begin{aligned} (7) \quad & \int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L_i(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} p_j(\theta; d\mathbf{x}_j) \\ &= \int_{\mathbb{R}^j} \psi_j^!(\mathbf{x}_j) \left\{ \int_{\Delta} L_i(\theta, y) \delta_j^!(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j) \end{aligned}$$

et d'après le théorème de Lebesgue relatif à la convergence monotone :

$$\int_{\Delta} L_i(\theta, y) \delta_j(\mathbf{x}_j; dy) \xrightarrow{i \rightarrow \infty} \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dx) \quad \forall \mathbf{x}_j \in \mathbb{R}^j$$

on en déduit, puisque $\psi_j \geq 0$

$$\psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L_i(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} \xrightarrow{i \rightarrow \infty} \psi_j(\mathbf{x}_j) \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy)$$

appliquant à nouveau le théorème de Lebesgue, pour θ fixé, il vient :

$$\int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L_i(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} p_j(\theta; d\mathbf{x}_j) \xrightarrow{i \rightarrow \infty}$$

$$\int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j)$$

De même relativement à ψ_j' et δ_j' , on aura :

$$\int_{\mathbb{R}^j} \psi_j'(\mathbf{x}_j) \left\{ \int_{\Delta} L_i(\theta, y) \delta_j'(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j) \xrightarrow{i \rightarrow \infty}$$

$$\int_{\mathbb{R}^j} \psi_j'(\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, y) \delta_j'(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j)$$

et par conséquent, en utilisant l'égalité (7)

$$\int_{\mathbb{R}^j} \psi_j(\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, y) \delta_j(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j)$$

$$= \int_{\mathbb{R}^j} \psi_j'(\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, y) \delta_j'(\mathbf{x}_j; dy) \right\} p_j(\theta, d\mathbf{x}_j) \quad \forall \theta \in \Theta \quad \forall j$$

Nous avons ainsi établi le théorème pour les fonctions positives. Si $L(\theta, y)$ est intégrale et de signe quelconque, pour chaque $\theta \in \Theta$, nous écrirons

$$L(\theta, y) = L^+(\theta, y) - L^-(\theta, y)$$

ce qui nous permettra d'appliquer le résultat précédent.

IV-9 - TRIBUS EXHAUSTIVES MINIMALES

Considérons la structure $[\Omega, \mathcal{A}, \{P_\theta, \theta \in (\Theta, \mathcal{E})\}]$.

Définition

Une tribu \mathcal{A}_1 de \mathcal{A} est une tribu exhaustive minimale si toute tribu exhaustive \mathcal{A}' de \mathcal{A} est telle que :

$$\mathcal{A}_1 \subset \mathcal{A}' \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta.$$

Une notion identique pour les statistiques peut également être définie :

Lemme 2

$$\text{Si } \mathcal{A}_1 \subset \mathcal{A}_2 \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta.$$

- à toute fonction f \mathcal{A}_1 -mesurable et telle que $c \leq f \leq d$ ($(c,d) \in \mathbb{R} \times \mathbb{R}$), on peut associer une fonction g \mathcal{A}_2 -mesurable, telle que $c \leq g \leq d$ et $g = f$ $[\mathcal{A}, \theta] \quad \forall \theta \in \Theta$.
- à toute probabilité ou mesure de transition μ \mathcal{A}_1 -mesurable, on peut associer une probabilité de transition ν \mathcal{A}_2 -mesurable, telle que :

$$\mu(z, A) = \nu(z, A) \quad \forall A \quad [\mathcal{A}, \theta] \quad \forall \theta \in \Theta.$$

La démonstration de ce théorème se fait comme celle du théorème de IV.7.

Définition

La structure statistique $[\Omega, \mathcal{A}, \{p_\theta, \theta \in (\Theta, \mathcal{E})\}]$ est dite régulière s'il existe une suite $\{\mathcal{A}_j^0\}$ de sous-tribus de \mathcal{A} vérifiant $\mathcal{A}_j^0 \subset \mathcal{A}_j$ et telle que \mathcal{A}_j^0 , pour tout j , soit exhaustive, minimale et transitive.

Théorème 8

Si la structure $[\Omega, \mathcal{A}, \{p_\theta, \theta \in (\Theta, \mathcal{E})\}]$ est régulière et si $\{\mathcal{A}_j\}$ est une suite exhaustive, la classe des stratégies basées sur $\{\mathcal{A}_j\}$ est essentiellement complète.

Démonstration

Soit $\{\mathcal{A}_j^0\}$ la suite exhaustive minimale définissant la régularité de la structure $[\Omega, \mathcal{A}, \{p_\theta, \theta \in (\Theta, \mathcal{E})\}]$.

Nous supposons, de plus, que les $C_j(\theta, \mathcal{X}_j)$ sont \mathcal{A}_j^0 -mesurables. Soit donc une stratégie $\{\delta_j, \varphi_j\}$, d'après le théorème 4 de III il existe une stratégie $\{\delta_j^!, \varphi_j^!\}$ basée sur \mathcal{A}_j^0 et équivalente à $\{\delta_j, \varphi_j\}$.

Or puisque \mathcal{A}_j^0 est exhaustive minimale, on a :

$$\mathcal{A}_j^0 \subset \mathcal{A}_j \quad [\mathcal{A}_j, \theta] \quad \forall \theta \in \Theta$$

d'après le lemme 2 ci-dessus, il existe $\{\varphi_j^!\}$, telle que $\varphi_j^!$ soit $\mathcal{A}_j^!$ -mesurable et il existe $\{\delta_j^!\}$ ($\delta_j^!$ basée sur $\mathcal{A}_j^!$), telles que :

$$\varphi_j^! = \varphi_j^! \quad [\mathcal{A}_j, \theta] \quad \forall \theta \in \Theta$$

$$\delta_j^!(\mathcal{X}_j, D) = \delta_j^!(\mathcal{X}_j, D) \quad \forall D \in \mathcal{D} \quad [\mathcal{A}_j, \theta] \quad \forall \theta \in \Theta$$

On a d'une part :

$$\forall j \quad \int_{\mathbb{R}^j} \varphi_j^!(\mathcal{X}_j) C_j(\theta, \mathcal{X}_j) p_j(\theta, d\mathcal{X}_j) =$$

$$\int_{\mathbb{R}^j} \varphi_j^!(\mathcal{X}_j) C_j(\theta, \mathcal{X}_j) p_j(\theta, d\mathcal{X}_j)$$

D'autre part :

$$\begin{aligned} \forall j \quad & \int_{\mathbb{R}^j} \psi_j^! (\mathbf{x}_j) \left\{ \int_{\Delta} L(\theta, dy) \delta_j^! (\mathbf{x}_j, dy) \right\} p_j(\theta; d\mathbf{x}_j) \\ & = \int_{\mathbb{R}^j} \psi_j^{!!} \left\{ \int_{\Delta} L(\theta, y) \delta_j^{!!} (\mathbf{x}_j, dy) \right\} p_j(\theta; d\mathbf{x}_j) \end{aligned}$$

En effet les mesures $\delta_j^!$ et $\delta_j^{!!}$ sont les mêmes sauf sur un ensemble de \mathbb{R}^j de mesure nulle, pour tout θ .

Les deux stratégies sont donc équivalentes.

BIBLIOGRAPHIE

- T.S. FERGUSON Mathematic Statistics Decision Theoric.
- BAHADUR Sufficiency and Statistical Decision Functions
- J.R. BARRA Notions fondamentales de statistique mathématiques
Dunod Paris 1971
- J.R. BARRA et A. BAILLE - Problèmes de statistiques mathématiques
Dunod Paris 1969
- J.R. BARRA Contributions à l'étude des lois de probabilité empiriques
Gap Imp. Louis Jean 1961 (Thèses Sciences Appliquées
Paris 1961 n° 3768)
- C. FOURGEAUD et A. FUCHS - Statistique
Dunod Paris 1967
- P.L. HENNEQUIN et A. TORTRAT - Théorie des probabilités et quelques applications
Masson, Paris 1965 Inc 8 VIII.
- H. ROBBIN Mixture of distributions
Annals of mathematic statistic Volume 19 - 1948

DEFAUTS DE PAGE EN PROGRAMMATION

ET CHAINES DE MARKOV

D. NIANGO

INTRODUCTION

-!-!-!-!-!-!-

L'étude statistique de la succession des caractères utilisés pour composer un texte dans un langage particulier a suggéré à Markov, un schéma d'enchaînement aléatoire avec liaison (par exemple en anglais, la fréquence de la lettre h après un t est très supérieure à ce qu'elle est après un d).

De nombreux autres problèmes statistiques peuvent également être traités par la méthode des chaînes de Markov.

Nous essaierons, ici, d'appliquer cette méthode à l'étude des défauts de page en programmation.

Rappelons rapidement qu'un défaut de page se définit de la façon suivante :

Un programme d'une taille importante est décomposé en un certain nombre de morceaux appelés pages ; si une instruction à exécuter n'est pas présente en mémoire réelle (par opposition à la mémoire virtuelle de la machine) et si aucune page réelle n'est libre, le problème est de sacrifier une page du programme qui est dans la mémoire réelle. Le choix le plus judicieux est celui qui consiste à remplacer la page la plus anciennement utilisée par une nouvelle. Tel est le mécanisme du défaut de page dû en général à plusieurs paramètres : la machine elle-même, le rapport taille mémoire réelle à la taille mémoire virtuelle, taille d'une page, le langage utilisé, le programmeur, etc...).

Madame Martins dans un rapport publié en 1977 a étudié le problème de la simulation des instants de défaut de page par les instants de retour à l'état 0 d'une chaîne de Markov dont les états sont notés 0, 1, ..., r. Ces travaux n'ont pas été concluants.

Nous nous proposons, dans cet exposé, d'effectuer la même démarche mais en simulant par les instants de retour en 0 ou 1.

Nous diviserons l'exposé en trois paragraphes :

- 1°) Etude de l'intervalle entre deux réalisations (passage à l'état 0 ou 1).
- 2°) Corrélation entre deux intervalles successifs.
- 3°) Etude d'un exemple simple.

I - ETUDE DE L'INTERVALLE ENTRE DEUX REALISATIONS

I.1 - Définition d'une chaîne de Markov

Processus de Markov discret

Soit $X_0, X_1, \dots, X_n, \dots$ une suite de variables aléatoires prenant les valeurs $0, 1, \dots, n, \dots$ X_n représentant l'état d'un système à l'instant n . Cette suite constitue un processus de Markov discret si la probabilité conditionnelle de X_n , connaissant X_0, X_1, \dots, X_{n-1} est égale à la probabilité conditionnelle de X_n connaissant X_{n-1} , c'est-à-dire :

$$P\{X_n / X_0, X_1, \dots, X_{n-1}\} = P\{X_n / X_{n-1}\}.$$

Chaîne de Markov

Une chaîne de Markov est un processus de Markov discret ne prenant qu'un nombre fini de valeurs.

Nous étudions ici une chaîne à $r+1$ états notés $0, 1, \dots, r$.

Pour simplifier, nous supposons que la chaîne de Markov est homogène, c'est-à-dire que la probabilité d'aller de l'état i à l'état j en n étapes ne dépend que de i, j et n , soit pour tout k ,

$$P_{i,j}^n = P\{X_{n+k} = j / X_k = i\},$$

où p_{ij} est la probabilité d'aller de l'état i à l'état j en une étape, p_{ij}^n étant alors la probabilité d'aller de l'état i à l'état j en n étapes. Nous rappelons au paragraphe suivant une méthode classique pour le calcul des p_{ij}^n .

I.2 - Calcul des p_{ij}^n

Nous allons calculer les p_{ij}^n par la méthode des fonctions génératrices.
Pour cela posons :

$$P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^n z^n \quad \text{avec} \quad p_{ij}^0 = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Multiplions par $z p_{ki}$, soit

$$z p_{ki} P_{ij}(z) = \sum_{n=0}^{\infty} p_{ki} p_{ij}^n z^{n+1}$$

et sommions par rapport à i ($i = 1, 2, \dots, r$)

$$\begin{aligned} \sum_{i=0}^r z p_{ki} P_{ij}(z) &= \sum_{i=0}^r \sum_{n \geq 0} p_{ki} p_{ij}^n z^{n+1} \\ &= \sum_{n \geq 0} \left(\sum_{i=0}^r p_{ki} p_{ij}^n \right) z^{n+1} = P_{kj}(z) - p_{kj}^0 \end{aligned}$$

Si on fixe j et z , les $P_{ij}(z) = z_i$ sont solutions du système linéaire suivant :

$$(1.1) \quad z_k - \sum_{i=0}^r z p_{ki} z_i = p_{kj}^0 \quad (k = 0, 1, \dots, r)$$

La matrice du système (1.1) est la matrice $D(z)$ suivante :

$$D(z) = \begin{pmatrix} 1-zp_{00} & -zp_{01} & \dots & -zp_{0r} \\ -zp_{10} & 1-zp_{11} & -zp_{12} & \dots & -zp_{1r} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -zp_{r0} & \dots & \dots & \dots & 1-zp_{rr} \end{pmatrix}$$

qu'on peut mettre sous la forme

$$D(z) = I - zA$$

avec

$$A = \begin{pmatrix} P_{00} & P_{01} & \dots & P_{0r} \\ P_{10} & P_{11} & \dots & P_{1r} \\ \dots & \dots & \dots & \dots \\ P_{r0} & P_{r1} & \dots & P_{rr} \end{pmatrix}$$

On remarque que A est la matrice stochastique d'ordre r+1 (associée à la chaîne (matrice dont tous les termes sont positifs et tels que la somme des éléments d'une ligne quelconque est égale à 1)).

On a donc :

$$D(z) = z \left(\frac{1}{z} I - A \right).$$

Calculons le déterminant $\Delta(z)$ de D(z).

$$\Delta(z) = z^{r+1} \det \left(\frac{1}{z} I - A \right).$$

Introduisons pour cela le polynôme caractéristique

$$P(\lambda) = \det (A - \lambda I).$$

Nous supposons pour simplifier que les racines de P(λ) sont toutes simples et non nulles ; de sorte que :

$$P(\lambda) = \prod_{i=0}^r (\lambda_i - \lambda)$$

ce qui donne :

$$\Delta(z) = (-1)^{r+1} z^{r+1} P\left(\frac{1}{z}\right) = (-1)^{r+1} z^{r+1} \prod_{i=0}^r \left(\lambda_i - \frac{1}{z}\right) = \prod_{i=0}^r (1 - \lambda_i z)$$

Sous les hypothèses faites, $\Delta(z)$ est un polynôme de degré $r+1$.

Revenons au système linéaire (1.1) dont le second membre est, en notation matricielle

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

matrice colonne dont tous les termes sont nuls sauf celui situé à la j ème ligne, qui est égal à 1.

Ce système est de Cramer, si $|z| < \inf_i \frac{1}{|\lambda_i|}$ et les formules de Cramer permettent d'écrire z_k sous la forme du quotient d'un polynôme de degré au plus égal à r , par $\Delta(z)$.

Soit

$$z_k = P_{kj}(z) = \sum_{i=0}^r \frac{b_{kj}^i}{1 - \lambda_i z} = \sum_{i=0}^r b_{kj}^i (1 + \lambda_i z + \dots + \lambda_i^n z^n + \dots)$$

si

$$|z| < \frac{1}{\max_i |\lambda_i|} = 1$$

En effet les valeurs propres de A sont toutes de module ≤ 1 et la valeur propre 1 est évidente).

Par identification des développements, on a :

$$P_{kj}^n = \sum_{i=0}^r b_{kj}^i (\lambda_i)^n$$

On voit donc, que du point de vue pratique, il est nécessaire de calculer les valeurs propres λ_i de A et les coefficients b_{kj}^i .

I.3 - Etude de l'instant du premier retour en 0

On rappelle ici la méthode utilisée par Madame Martins.
Le système étant en 0, à l'instant 0, nous allons étudier l'instant du premier passage en 0.

Désignons par f_{ij}^n la probabilité du premier passage au temps n en j partant de i à l'instant 0.

On doit déterminer la suite $\{f_{00}^n\}_{n=1,2,\dots}$.

Ici encore nous introduisons la fonction génératrice :

$$V(z) = \sum_{n=1}^{\infty} f_{00}^n z^n$$

On a :

$$p_{00}^n = f_{00}^n + f_{00}^{n-1} p_{00}^1 + \dots + f_{00}^1 p_{00}^{n-1} \quad n \geq 1$$

D'où

$$\begin{aligned} P_{00}(z) = \sum_{n=0}^{\infty} p_{00}^n z^n &= 1 + f_{00}^1 z + (f_{00}^2 + f_{00}^1 p_{00}^1) z^2 \\ &\quad + (f_{00}^3 + f_{00}^2 p_{00}^1 + f_{00}^1 p_{00}^2) z^3 + \dots \end{aligned}$$

$$p_{00}(z) = 1 + V(z) + p_{00}^1 z V(z) + p_{00}^2 z^2 V(z) + \dots$$

$$= V(z) (1 + z p_{00}^1 + z^2 p_{00}^2 + \dots) + 1 = 1 + V(z) P_{00}(z)$$

D'où :

$$(1.2) \quad V(z) = \frac{P_{00}(z) - 1}{P_{00}(z)}$$

Si $P_{\infty}(z)$ est connu on peut donc théoriquement déterminer $V(z)$ qui est une fraction rationnelle, et en déduire la suite $\{f_{\infty}^n\}_n$.

Madame Martins à l'aide des formules déduites de (1.2) a tenté de déterminer des matrices A permettant d'obtenir des valeurs données par l'expérience pour la suite $\{f_{\infty}^n\}_n$. C'est sur ce travail de détermination des chaînes qu'elle a achoppé.

I.4 - Etude du processus ponctuel des instants de passage en 0 et 1.

On désigne par T_n l'instant du même passage en 0 ou 1 du système régi par A lorsque ce système Σ est en 0 au temps 0.

On note g_{ij}^n la probabilité d'aller de i en j en n étapes sans passer par i ou j , et par G_{ij} les fonctions génératrices correspondantes.

On a :

$$\begin{aligned}
 p_{\infty}^n &= g_{\infty}^n + g_{\infty}^{n-1} p_{\infty}^1 + g_{\infty}^{n-2} p_{\infty}^2 + \dots + g_{\infty}^1 p_{\infty}^{n-1} \\
 &+ g_{01}^{n-1} p_{10}^1 + g_{01}^{n-2} p_{10}^2 + \dots + g_{01}^1 p_{10}^{n-1} \quad (n \geq 1) ;
 \end{aligned}$$

ce qui donne :

$$\begin{aligned}
 P_{\infty}(z) &= 1 + g_{\infty}^1 z + (g_{\infty}^2 + g_{\infty}^1 p_{\infty}^1 + g_{01}^1 p_{10}^1) z^2 + \\
 &+ (g_{\infty}^3 + g_{\infty}^2 p_{\infty}^1 + g_{\infty}^1 p_{\infty}^2 + g_{01}^2 p_{10}^1 + g_{01}^1 p_{10}^2) z^3 \\
 &+ \dots
 \end{aligned}$$

Soit :

$$\begin{aligned}
 P_{\infty}(z) &= 1 + G_{\infty}(z) + p_{\infty}^1 z G_{\infty}(z) + p_{\infty}^2 z^2 G_{\infty}(z) + \dots \\
 &+ p_{10}^1 z G_{01}(z) + p_{10}^2 z^2 G_{01}(z) + \dots
 \end{aligned}$$

$$P_{\infty}(z) = 1 + G_{\infty}(z) P_{\infty}(z) + G_{01}(z) P_{10}(z) \quad (2.1)$$

De même

$$P_{11}(z) = 1 + G_{11}(z) P_{11}(z) + G_{10}(z) P_{01}(z) \quad (2.2)$$

$$P_{01}(z) = G_{01}(z) P_{11}(z) + G_{\infty}(z) P_{01}(z) \quad (2.3)$$

$$P_{10}(z) = G_{11}(z) P_{10}(z) + G_{10}(z) P_{\infty}(z) \quad (2.4)$$

Les relations (2.1) et (2.3) donnent G_{∞} et G_{01} , tandis que les relations (2.2) et (2.4) donnent G_{11} et G_{10} .

Dans les deux cas, le déterminant est le même :

$$\delta(z) = P_{\infty}(z) P_{11}(z) - P_{01}(z) P_{10}(z).$$

Or on a :

$$\left\{ \begin{array}{l} P_{\infty}(0) = P_{11}(0) = 1 \\ P_{10}(0) = P_{01}(0) = 0 \end{array} \right.$$

Ces relations montrent que $\delta(0) \neq 0$, donc $\delta(z)$ est non nul dans un voisinage de l'origine.

On peut donc calculer les $G_{ij}(z)$ dans un voisinage de $z = 0$; en fait, bien que les $P_{kl}(z)$ ne soient pas définis pour $z = 1$

($P_{kl}(z) = z_k = \sum_{i=0}^r \frac{b_{kl}^i}{1-\lambda_i z}$ et 1 est valeur propre de A), on doit obtenir les $G_{ij}(z)$ pour $|z| \leq 1$.

On note $\Sigma(T_n)$ l'état de Σ au temps n .

Soit

$$\psi_n(i) = P\{\Sigma(T_n) = i\} \quad (i = 0, 1)$$

On a

$$P\{T_{n+1} = T_n + k, \Sigma(T_{n+1}) = j / \Sigma(T_n) = i\} = g_{ij}^k$$

d'où

$$P\{T_{n+1} = T_n + k, \Sigma(T_{n+1}) = j\} = g_{0j}^k \psi_n(0) + g_{1j}^k \psi_n(1) = \phi_n^k(j)$$

soit

$$\left\{ \begin{array}{l} P\{T_{n+1} = T_n + k\} = \phi_n^k(0) + \phi_n^k(1) = h_n(k) \\ \psi_{n+1}(j) = \sum_{k=1}^{+\infty} \phi_n^k(j) \end{array} \right.$$

On a donc un procédé récurrent pour calculer la suite $\{\psi_n\}$ et les lois des $T_{n+1} - T_n$.

On a :

$$\psi_{n+1}(0) = \sum_{k=1}^{+\infty} \phi_n^k(0) = \sum_{k \geq 1} g_{00}^k \psi_n(0) + \sum_{k \geq 1} g_{10}^k \psi_n(1)$$

$$\psi_{n+1}(0) = \psi_n(0) G_{00}(1) + \psi_n(1) G_{10}(1)$$

$$(S) \quad \left\{ \begin{array}{l} \psi_{n+1}(0) = \psi_n(0) [G_{00}(1) - G_{10}(1)] + G_{10}(1) \\ \psi_1(0) = G_{00}(1) \end{array} \right.$$

$$\text{On a} \quad 0 \leq G_{00}(1) \leq 1,$$

$$0 \leq G_{10}(1) \leq 1,$$

et donc $-1 \leq G_{\infty}(1) - G_{10}(1) \leq 1$.

(si $\forall i, j \quad p_{ij} > 0$, on a $G_{\infty}(1) < 1$
 $G_{10}(1) < 1$).

Si $-1 < G_{\infty}(1) - G_{10}(1) < 1$, ψ_n tend vers une limite finie ψ qui correspond à un état stationnaire.

Si $G_{\infty}(1) - G_{10}(1) = \alpha \neq 1$, la résolution de (S) donne :

$$\psi_{n+1}(0) - \psi_n(0) = \alpha(\psi_n(0) - \psi_{n-1}(0))$$

soit

$$\psi_{n+1}(0) - \psi_n(0) = \alpha^{n-1}(\psi_2(0) - \psi_1(0)),$$

ou encore

$$\psi_{n+1}(0) - \psi_1(0) = (\psi_2(0) - \psi_1(0)) \frac{1 - \alpha^n}{1 - \alpha};$$

ce qui donne :

$$\psi_{n+1}(0) = (\alpha G_{\infty}(1) + G_{10}(1)) \frac{1 - \alpha^n}{1 - \alpha} + G_{\infty}(1) \frac{\alpha^n - \alpha}{1 - \alpha}$$

$$\psi_{n+1}(0) = G_{\infty}(1) \alpha^n + G_{10}(1) \frac{1 - \alpha^n}{1 - \alpha}$$

si

$$\alpha < 1 \quad \psi_n(0) = G_{10}(1) \frac{1}{1 - \alpha} = \psi(0)$$

et

$$h_n(k) = g_{\infty}^k \psi_n(0) + g_{10}^k \psi_n(1) + g_{01}^k \psi_n(0) + g_{11}^k \psi_n(1)$$

On peut donc calculer les $h_n(k)$, ainsi la loi de $T_{n+1}^{-1} T_n$, si on connaît les g_{ik}^k et donc les G_{ij} . Il est donc essentiel d'obtenir les $G_{ij}(z)$.

Ici, encore, nous utiliserons la méthode des fonctions génératrices ; la fonction génératrice des $h_n(k)$ est la suivante :

$$H_n(z) = \psi_n(0) \sum_{k=1}^{\infty} g_{00}^k z^k + \dots = \sum_{k \geq 1} h_n(k) z^k.$$

D'où

$$H_n(z) = \psi_n(0) [G_{00}(z) + G_{01}(z) - G_{10}(z) - G_{11}(z)] + G_{10}(z) + G_{11}(z)$$

La résolution des équations (2.1), (2.2), (2.3) et (2.4) donne :

$$G_{00}(z) = \frac{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z) - P_{11}(z)}{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z)} = \frac{\delta(z) - P_{11}(z)}{\delta(z)}$$

$$G_{01}(z) = \frac{P_{01}(z)}{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z)} = \frac{P_{01}(z)}{\delta(z)}$$

$$G_{11}(z) = \frac{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z) - P_{00}(z)}{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z)} = \frac{\delta(z) - P_{00}(z)}{\delta(z)}$$

$$G_{10}(z) = \frac{P_{10}(z)}{P_{00}(z) P_{11}(z) - P_{01}(z) P_{10}(z)} = \frac{P_{10}(z)}{\delta(z)}$$

Rappelons que

$$P_{\infty\infty}(z) = \frac{1}{\Delta(z)} \begin{vmatrix} 1 & -z_{p_{o1}} & \dots & -z_{p_{or}} \\ 0 & 1-z_{p_{11}} & & \\ \vdots & & & \\ \vdots & & & \\ 0 & -z_{p_{r1}} & - & - & - & 1-z_{p_{rr}} \end{vmatrix}$$

$$P_{10}(z) = \frac{1}{\Delta(z)} \begin{vmatrix} 1-z_{p_{\infty o}} & 1 & -z_{p_{o2}} & \dots & -z_{p_{or}} \\ -z_{p_{1o}} & 0 & -z_{p_{12}} & \dots & -z_{p_{1r}} \\ \vdots & & & & \\ \vdots & & & & \\ -z_{p_{ro}} & 0 & - & - & - & 1-z_{p_{rr}} \end{vmatrix}$$

$$P_{11}(z) = \frac{1}{\Delta(z)} \begin{vmatrix} 1-z_{p_{\infty o}} & 0 & -z_{p_{o2}} & \dots & -z_{p_{or}} \\ -z_{p_{1o}} & 1 & -z_{p_{12}} & \dots & -z_{p_{1r}} \\ \vdots & & & & \\ \vdots & & & & \\ -z_{p_{ro}} & 0 & -z_{p_{r2}} & \dots & 1-z_{p_{rr}} \end{vmatrix}$$

$$P_{o1}(z) = \frac{1}{\Delta(z)} \begin{vmatrix} 0 & -z_{p_{o1}} & \dots & z_{p_{or}} \\ 1 & 1-z_{p_{11}} & & -z_{p_{1r}} \\ \vdots & & & \\ \vdots & & & \\ 0 & -z_{p_{r1}} & \dots & 1-z_{p_{rr}} \end{vmatrix}$$

Ces formules permettent de calculer $G_{00}(z)$, $G_{01}(z)$, $G_{10}(z)$ et $G_{11}(z)$ et donc, par l'intermédiaire de (S), d'obtenir $H_n(z)$. On dispose ainsi d'un moyen de déterminer la loi pour tout n de $T_{n+1} - T_n$.

II - CORRELATION ENTRE DEUX INTERVALLES SUCCESSIFS

Dans le cas de l'étude des retours en 0 les intervalles successifs sont indépendants. Il n'en est plus de même pour les retours en 0 ou 1. On étudie ici la corrélation entre deux intervalles successifs dans ce cas.

On a

$$\begin{aligned} P\{T_{n+1} = T_n + p, \Sigma(T_{n+1}) = \gamma, T_{n+2} = T_{n+1} + q, \Sigma(T_{n+2}) = k / \Sigma(T_n) = i\} \\ = g_{ij}^p \cdot g_{jk}^q \end{aligned}$$

et

$$\begin{aligned} P\{T_{n+1} = T_n + p, T_{n+2} = T_{n+1} + q, \Sigma(T_{n+1}) = j, \Sigma(T_{n+2}) = k\} = \\ g_{0j}^p \cdot g_{jk}^q \cdot \psi_n(0) + g_{1j}^p \cdot g_{jk}^q \cdot \psi_n(1) = L_n^{p,q}(j,k) \end{aligned}$$

D'où

$$\begin{aligned} P\{T_{n+1} = T_n + p, T_{n+2} = T_{n+1} + q\} = L_n^{p,q}(0,0) + L_n^{p,q}(0,1) + L_n^{p,q}(1,0) \\ + L_n^{p,q}(1,1) = h_n(p,q) \end{aligned}$$

On a donc :

$$h_n(p,q) = \psi_n(0) [g_{00}^p g_{00}^q + g_{00}^p g_{01}^q + g_{01}^p g_{01}^q + g_{01}^p g_{11}^q] \\ + \psi_n(1) [g_{10}^p g_{00}^q + g_{10}^p g_{01}^q + g_{11}^p g_{10}^q + g_{11}^p g_{11}^q]$$

Introduisant la fonction génératrice de $h_n(p,q)$

$$H_n(z_1, z_2) = \sum_{p,q} h_n(p,q) z_1^p z_2^q,$$

on a :

$$H_n(z_1, z_2) = \psi_n(0) [G_{00}(z_1)G_{00}(z_2) + G_{00}(z_1)G_{01}(z_2) + G_{01}(z_1)G_{10}(z_2) \\ + G_{01}(z_1)G_{11}(z_2)] + \psi_n(1) [G_{10}(z_1)G_{00}(z_2) + G_{10}(z_1)G_{01}(z_2) + G_{11}(z_1)G_{10}(z_2) \\ + G_{11}(z_1)G_{11}(z_2)]$$

$$\text{Posons } \Delta_{n+1} = T_{n+1}^{-1} T_n, \quad \Delta_{n+2} = T_{n+2}^{-1} T_{n+1}.$$

Le coefficient de corrélation de Δ_{n+1} et Δ_{n+2} est défini par :

$$S_n = \frac{E(\Delta_{n+1} \Delta_{n+2}) - E(\Delta_{n+1})E(\Delta_{n+2})}{\sigma(\Delta_{n+1}) \sigma(\Delta_{n+2})} = \frac{\tau_n}{\sigma_{n+1} \sigma_{n+2}}$$

On a donc

$$E(\Delta_{n+1} \Delta_{n+2}) = \sum_{p,q} pq \cdot h_n(p,q) = \frac{\delta^2 H_n(1,1)}{\delta z_1 \delta z_2}$$

$$E(\Delta_{n+1} \Delta_{n+2}) = \psi_n(0) [G'_{00}(1)G'_{00}(1) + G'_{00}(1)G'_{01}(1) + G'_{01}(1)G'_{10}(1) + G'_{01}(1)G'_{11}(1)] \\ + \psi_n(1) [G'_{10}(1)G'_{00}(1) + G'_{10}(1)G'_{01}(1) + G'_{11}(1)G'_{10}(1) + G'_{11}(1)G'_{11}(1)]$$

III - EXEMPLE

L'exemple simple suivant illustre l'application des méthodes décrites en 2 et 3.

Soit

$$A = \frac{1}{4} \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 3 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

Un calcul élémentaire montre que A a pour valeurs propres $1, \frac{1}{4}, \frac{\lambda_1}{4}, \frac{\lambda_2}{4}$ où $\lambda_1 + \lambda_2 = -1$ et $\lambda_1 \lambda_2 = 3$.

D'où

$$\Delta(z) = (1-z) \left(1 - \frac{z}{4}\right) \left(1 - \frac{\lambda_1}{4} z\right) \left(1 - \frac{\lambda_2}{4} z\right)$$

$$\Delta(z) = (1-z) \left(1 - \frac{z}{4}\right) \left(\frac{3}{16} z^2 + \frac{z}{4} + 1\right).$$

On a

$$\Delta(z)P_{\infty}(z) = \begin{vmatrix} 1 & -\frac{z}{4} & -\frac{z}{2} & 0 \\ 0 & 1 - \frac{z}{2} & -\frac{z}{4} & 0 \\ 0 & 0 & 1 & -\frac{3}{4}z \\ 0 & -\frac{z}{4} & 0 & 1 - \frac{z}{4} \end{vmatrix} = \begin{vmatrix} 1 - \frac{z}{2} & -\frac{z}{4} & 0 \\ 0 & 1 & -\frac{3}{4}z \\ -\frac{z}{4} & 0 & 1 - \frac{z}{4} \end{vmatrix}$$

$$= \left(1 - \frac{z}{2}\right) \left(2 - \frac{z}{4}\right) - 3 \frac{z^3}{64}$$

$$\Delta(z)P_{11}(z) = \begin{vmatrix} 1 - \frac{z}{4} & 0 & -\frac{z}{2} & 0 \\ -\frac{z}{4} & 1 & -\frac{z}{4} & 0 \\ -\frac{z}{4} & 0 & 1 & -\frac{3}{4}z \\ -\frac{z}{2} & 0 & 0 & 1 - \frac{z}{4} \end{vmatrix} = \begin{vmatrix} 1 - \frac{z}{4} & -\frac{z}{2} & 0 \\ -\frac{z}{4} & 1 & -\frac{3}{4}z \\ -\frac{z}{2} & 0 & 1 - \frac{z}{4} \end{vmatrix}$$

$$= \left(1 - \frac{z}{4}\right)^2 - \frac{z^2}{4} \left(1 + \frac{5}{4}z\right)$$

$$\Delta(z)P_{10}(z) = \begin{vmatrix} 1 - \frac{z}{4} & 1 & -\frac{z}{2} & 0 \\ -\frac{z}{4} & 0 & -\frac{z}{4} & 0 \\ -\frac{z}{4} & 0 & 1 - \frac{3}{4}z & \\ -\frac{z}{2} & 0 & 0 & 1 - \frac{z}{4} \end{vmatrix} = \frac{z}{4} \begin{vmatrix} 1 & -\frac{z}{4} & 0 \\ 1 & 1 & -\frac{3}{4}z \\ 2 & 0 & 1 - \frac{z}{4} \end{vmatrix}$$

$$= \frac{z}{4} \left[1 - \frac{z}{4} + \frac{z}{4} \left(1 + \frac{5}{4}z \right) \right]$$

$$= \frac{z}{4} \left[1 + \frac{5}{16}z^2 \right]$$

$$\Delta(z)P_{01}(z) = \begin{vmatrix} 0 & -\frac{z}{4} & -\frac{z}{2} & 0 \\ 1 & 1 - \frac{1}{2}z & -\frac{3}{4} & 0 \\ 0 & 0 & 1 & -\frac{3}{4}z \\ 0 & -\frac{z}{4} & 0 & 1 - \frac{z}{4} \end{vmatrix} = \frac{z}{4} \begin{vmatrix} 1 & -\frac{z}{2} & 0 \\ 0 & 1 & -\frac{3}{4}z \\ 1 & 0 & 1 - \frac{z}{4} \end{vmatrix}$$

$$= \frac{z}{4} \begin{vmatrix} 1 & -\frac{z}{2} & 0 \\ 0 & 1 & -\frac{3}{4}z \\ 0 & \frac{z}{2} & 1 - \frac{z}{4} \end{vmatrix} = \frac{z}{4} \left[1 - \frac{z}{4} + \frac{3}{8}z^2 \right]$$

D'où

$$\delta(z) = \left[1 - \frac{3}{4}z + \frac{z^2}{8} - \frac{3}{64}z^3 \right] \left[1 - \frac{z}{2} - \frac{1}{16}z^2 - \frac{5}{32}z^3 \right] - \frac{z^2}{16} \left(1 + \frac{5}{16}z^2 \right) \left(1 - \frac{z}{4} + \frac{3}{8}z^2 \right)$$

$$= 1 - \frac{5}{4}z + \frac{3}{8}z^2 - \frac{13}{64}z^3 + \frac{15}{(16)^2}z^4 - \frac{3}{8 \times 32}z^5$$

$$= \left(1 - \frac{z}{4} \right) (1-z) \left(1 + \frac{1}{8}z^2 - \frac{3}{64}z^3 \right)$$

$$= \left(1 - \frac{z}{4} \right)^2 (1-z) \left(1 + \frac{1}{4}z + \frac{3}{16}z^2 \right)$$

$$= \left(1 - \frac{z}{4} \right) \Delta(z).$$

On en déduit

$$G_{\infty}(z) = 1 - \frac{(1 - \frac{z}{4})^2 - \frac{z^2}{8} (1 + \frac{5}{4} z)}{1 - \frac{z}{4}} = \frac{z}{4} + \frac{\frac{z^2}{8} (1 + \frac{5}{4} z)}{1 - \frac{z}{4}}$$

$$G_{11}(z) = 1 - \frac{(1 - \frac{z}{2}) (1 - \frac{z}{4}) - \frac{3z^3}{64}}{1 - \frac{z}{4}} = \frac{z}{2} + \frac{3}{64} \frac{z^3}{1 - \frac{z}{4}}$$

$$G_{01}(z) = \frac{z}{4} + \frac{3}{32} \frac{z^3}{1 - \frac{z}{4}}$$

$$G_{10}(z) = \frac{z}{4} + \frac{1}{16} \frac{z^2 (1 + \frac{5}{4} z)}{1 - \frac{z}{4}} .$$

D'où

$$G_{10}(1) = \frac{1}{4} + \frac{1}{16} \frac{1 + \frac{5}{4}}{1 - \frac{1}{4}} = \frac{7}{16}$$

$$G_{\infty}(1) = \frac{1}{4} + \frac{\frac{1}{8} (1 + \frac{5}{4})}{1 - \frac{1}{4}} = \frac{5}{8}$$

$$G_{11}(1) = \frac{1}{2} + \frac{3}{64} \frac{1}{\frac{3}{4}} = \frac{9}{16}$$

$$G_{01}(1) = \frac{1}{4} + \frac{3}{32} \cdot \frac{1}{\frac{3}{4}} = \frac{3}{8}$$

et

$$\alpha = G_{\infty}(1) - G_{10}(1) = \frac{3}{16} .$$

D'autre part :

$$H_n(z) = \psi_n(0) [G_{\infty}(z) + G_{01}(z) - G_{10}(z) - G_{11}(z)] + G_{10}(z) + G_{11}(z)$$

Or
$$\Psi_n(0) \xrightarrow{n} \frac{1}{1 - \frac{3}{16}} G_{10}(1) = \frac{7}{13}.$$

Par conséquent

$$\begin{aligned} H_n(z) \xrightarrow{n} \frac{7}{13} \left[\frac{z}{4} + \frac{z^2}{8} \frac{1 + \frac{5}{4}z}{1 - \frac{z}{4}} + \frac{z}{4} + \frac{3}{32} \frac{z^3}{1 - \frac{z}{4}} - \frac{z}{4} - \frac{z^2}{16} \frac{1 + \frac{5}{4}z}{1 - \frac{z}{4}} - \frac{z}{2} - \frac{3}{16} \frac{z^3}{1 - \frac{z}{4}} \right. \\ \left. + \frac{z}{4} + \frac{z^2}{16} \frac{1 + \frac{5}{4}z}{1 - \frac{z}{4}} + \frac{z}{2} + \frac{3}{64} \frac{z^3}{1 - \frac{z}{4}} \right] \end{aligned}$$

$$H_n(z) \xrightarrow{n} \frac{7}{13} \left[-\frac{z}{4} + \frac{z^2}{16} \frac{1 + \frac{5}{4}z}{1 - \frac{z}{4}} + \frac{3}{64} \frac{z^3}{1 - \frac{z}{4}} \right] + \frac{3}{4}z + \frac{z^2}{16} \frac{1 + \frac{5}{4}z}{1 - \frac{z}{4}} + \frac{3}{64} \frac{z^3}{1 - \frac{z}{4}}$$

$$H_n(z) \xrightarrow{n} \frac{7}{13} \left[-\frac{z}{4} + \frac{z^2}{16} (1+2z) \frac{1}{1 - \frac{z}{4}} \right] + \frac{3}{4}z + \frac{z^2}{16} (1+2z) \frac{1}{1 - \frac{z}{4}}$$

$$H_n(z) \xrightarrow{n} \frac{z}{4} \left(3 - \frac{7}{13} \right) + \left[\left(\frac{z^2}{16} + \frac{z^3}{8} \right) \frac{7}{13} + \frac{z^3}{8} \right] \left[1 + \frac{z}{4} + \frac{z^2}{16} + \dots \right]$$

D'où

$$H_n(z) \xrightarrow{n} \left(3 - \frac{7}{13} \right) \frac{z}{4} + \frac{z^2}{16} \left(1 + \frac{7}{13} \right) + z^3 \left(\frac{1}{16} \cdot \frac{7}{13} + \frac{1}{4} + \frac{1}{8} \cdot \frac{7}{13} + \frac{1}{16} \cdot \frac{1}{4} + \frac{1}{8} \right)$$

$$+ z^4 \left(\frac{7}{13} \left(\frac{1}{16} \cdot \frac{1}{16} + \frac{1}{32} \right) + \frac{1}{16^2} + \frac{1}{32} \right) + z^5 \left(\frac{7}{13} \left(\frac{1}{16} \times \frac{1}{64} + \frac{1}{128} \right) + \frac{1}{16} \cdot \frac{1}{64} \right)$$

$$+ \frac{1}{8} \cdot \frac{1}{16}) + \dots$$

On en déduit :

$$h_n(1) = \frac{1}{4} \left(3 - \frac{7}{13}\right) = \frac{1}{4} \cdot \frac{32}{13} = \frac{8}{13} = 0,61538$$

$$h_n(2) = \frac{1}{16} \left(2 + \frac{7}{13}\right) = \frac{20}{16 \times 13} = 0,09615$$

$$h_n(3) = \frac{7}{13} \cdot \frac{1}{8} \left(\frac{1}{8} + 1\right) + \frac{1}{8} \left(\frac{1}{8} + 1\right) = \frac{1}{8} \times \frac{9}{8} \cdot \frac{20}{13} = 0,2163$$

$$h_n(4) = \frac{7}{13} \frac{1}{32} \left(\frac{1}{8} + 1\right) + \frac{1}{32} \cdot \frac{9}{8} = 0,0540$$

$$h_n(5) = \frac{7}{13} \frac{1}{128} \left(\frac{1}{8} + 1\right) + \frac{1}{128} \left(\frac{1}{8} + 1\right) = \left(1 + \frac{1}{8}\right) \times \frac{1}{128} \left(1 + \frac{7}{13}\right) = 0,01352$$

Le comportement des premiers termes de la suite $\{h_n(i)\}_i$ est de même nature que celui observé pour l'intervalle de temps entre deux défauts de page successifs dans de nombreux exemples.

On a d'autre part :

$$G'_{\infty}(z) = \frac{1}{4} + \frac{\frac{z}{4} + \frac{15}{32} z^2}{1 - \frac{z}{4}} + \frac{\frac{1}{4} \frac{z^2}{8} (1 + \frac{5}{4} z)}{(1 - \frac{z}{4})^2}$$

et

$$G'_{\infty}(1) = \frac{4}{3}.$$

$$G'_{11}(z) = \frac{1}{2} + \frac{3 \times 3}{64} \cdot \frac{z^2}{1 - \frac{z}{4}} + \frac{1}{4} \times \frac{3}{64} \frac{z^3}{(1 - \frac{z}{4})^2}$$

$$G'_{11}(1) = \frac{17}{24}.$$

$$G'_{01}(z) = \frac{1}{4} + \frac{9}{32} \frac{z^2}{1 - \frac{z}{4}} + \frac{1}{4} \times \frac{3}{32} \frac{z^3}{(1 - \frac{z}{4})^2}$$

$$G'_{01}(1) = \frac{2}{3}.$$

$$G'_{10}(z) = \frac{1}{4} + \frac{1}{16} \frac{2z + \frac{15}{4} z^2}{1 - \frac{z}{4}} + \frac{1}{4} \times \frac{1}{16} \frac{z^2 (1 + \frac{5}{4} z)}{(1 - \frac{z}{4})^2}$$

et $G'_{10}(1) = \frac{19}{24}$.

On peut alors calculer S_n .

On a :

$$E(\Delta_{n+1} \Delta_{n+2}) = \frac{7}{13} = 13 \times 38$$

$$= 3,1455127$$

$$E(\Delta_{n+1}) = H'_n(1) = \psi_n(0) [G'_{00}(1) + G'_{01}(1) - G'_{10}(1) - G'_{11}(1)] + G'_{10}(1) + G'_{11}(1)$$

$$= 1,7692.$$

$$(H'_n(1))^2 = 3,1301753.$$

$$H''_n(1) = \psi_n(0) [G''_{00}(1) + G''_{01}(1) - G''_{10}(1) - G''_{11}(1)] + G''_{10}(1) + G''_{11}(1)$$

$$G''_{00}(z) = \frac{\frac{1}{4} + \frac{30}{32} z}{1 - \frac{z}{4}} + \frac{1}{4} \frac{\frac{z}{4} + \frac{15}{32} z^2}{(1 - \frac{z}{4})^2} = \frac{1}{32} \frac{2z + \frac{15}{4} z^2}{(1 - \frac{z}{4})^2} + \frac{1}{64} \frac{z^2 (1 + \frac{5}{4} z)}{(1 - \frac{z}{4})^3}$$

et

$$G''_{00}(1) = 2,305$$

$$G''_{11}(z) = \frac{9}{32} \frac{z}{1 - \frac{z}{4}} + \frac{9}{4 \times 64} \frac{z^2}{(1 - \frac{z}{4})^2} + \frac{9}{4 \times 64} \frac{z^2}{(1 - \frac{z}{4})^2} + \frac{3}{2 \times 4 \times 64} \frac{z^3}{(1 - \frac{z}{4})^3}$$

$$G''_{11}(1) = 0,513.$$

$$G''_{01}(z) = \frac{9}{16} \frac{z}{1 - \frac{z}{4}} + \frac{9}{4 \times 32} \frac{z^2}{(1 - \frac{z}{4})^2} + \frac{9}{4 \times 32} \frac{z^2}{(1 - \frac{z}{4})^2} + \frac{3}{2 \times 4 \times 32} \frac{z^3}{(1 - \frac{z}{4})^3}$$

$$G''_{01}(1) = 1,027.$$

$$G''_{10}(z) = \frac{1}{8} \cdot \frac{1 + \frac{15}{4} z}{1 - \frac{z}{4}} + \frac{1}{4 \times 16} \frac{2z + \frac{15}{4} z^2}{(1 - \frac{z}{4})^2} + \frac{1}{4 \times 16} \frac{2z + \frac{15}{4} z^2}{(1 - \frac{z}{4})^2} + \frac{1}{8 \times 16} \frac{z^2 (1 + \frac{5}{4} z)}{(1 - \frac{z}{4})^3}$$

$$G''_{10}(1) = 1,319.$$

D'où

$$\begin{aligned} H''_n(1) &= \frac{7}{13} [2,305 + 1,027 - 1,319 - 0,513] + 1,319 + 0,513 \\ &= \frac{7}{13} \times 1,500 + 1,832 = 2,6397. \end{aligned}$$

et

$$\begin{aligned} \lim_{n \rightarrow \infty} \sigma_{n+1}^2 &= H''_n(1) + H'_n(1) - (H'_n(1))^2 \\ &= 2,6397 + 1,7692 - 3,1302 = 1,2787 \end{aligned}$$

$$\lim_{n \rightarrow \infty} \sigma_{n+2}^2 = \lim_{n \rightarrow \infty} \sigma_{n+1}^2 = 1,2787$$

D'où

$$\lim_{n \rightarrow \infty} \tau_n = \tau = 3,1955 - 1,1801 = 0,0654$$

et

$$S_n \rightarrow \frac{0,0654}{1,2787} = 0,0512$$

ce qui est très faible.

En conclusion, nous voyons par l'exemple que nous venons de traiter, que la méthode exposée laisse espérer une bonne adéquation du modèle à la simulation des instants de défaut de pages. Il reste bien évidemment ici aussi à traiter le difficile problème de la détermination d'une matrice A permettant de simuler un exemple concret.

BIBLIOGRAPHIE

W. FELLER

Introduction to probability theory and its applications.
Vol. 1, second Edition J. Wiley

D. ISAACSON and R. W. MADSEN

Markov chains. Theory and applications
Probability and Mathematical statistics set - 1976

MARTINS

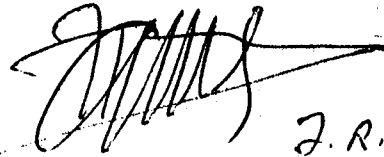
Rapport D.E.A. Statistiques
Juin 1977 - Grenoble

Dernière page d'une thèse

VU

Grenoble, le 20/6/78

Le Président de la thèse



J. R. BARRA

Vu, et permis d'imprimer,

Grenoble, le 27 juin 1978

Le Président de l'Université
Scientifique et Médicale



D^r G. CAU