وزارة التعليم العالي والبحث العلمي

Université Badji Mokhtar
Annaba

Badji Mokhtar University -
Annaba

جامعة باجي مختار
عنابة

Faculté des Sciences
Département de Mathématiques

THESE

Présentée en vue de l'obtention du diplôme de
Doctorat en Mathématiques
Option : Probabilités-Statistiques

Comparaison asymptotique des solutions bayésiennes et du maximum de vraisemblance liée au coût.

Par:
Méchakra-Tiah Hadda

Sous la direction de :
Dr Chadli Assia

Devant le jury

Présidente :	Seddik-Ameur Nacera	Prof.	UBM Annaba
Examinatrice :	Nemouchi Nahima	Prof.	Université de Constantine
Examinateur :	Mohdeb Zaher	Prof.	Université de Constantine
Examinateur :	Aissani Amar	Prof.	USTHB Alger
Examinateur :	Remita Med Riadh	MCA	UBM Annaba

Année : 2013/2014

Cette thèse est dédiée à la mémoire de mon père et de ma mère. A Samia, à Imen, à Yasmine et à Naceur.

***A* NOTRE ENCADREUR** :Le professeur Assia Chadli.

Nous la remercions de la confiance dont elle a fait preuve à notre égard. Qu'il nous soit permis de lui témoigner ici notre plus sincère reconnaissance et de lui adresser le témoignage de notre gratitude pour tous les conseils et toute l'aide qu'elle nous a prodigués pour la réalisation de ce travail.

A **NOS JUGES :**

Le professeur Seddik-Ameur Nacera :

Nous lui sommes profondément reconnaissant de l'honneur qu'elle nous a fait en acceptant de présider notre jury. Nous la remercions vivement.

Au professeur Nemouchi Nahima, au professeur Mohdeb Zaher, au professeur Aissani Amar et au professeur Rémita Riadh :

Qu' il nous soit permis de leur manifester ici notre reconnaissance pour avoir accepté de juger ce travail et siéger parmi nos juges.

Qu'ils veuillent bien trouver ici l'expression de notre gratitude et de notre respect.

Résumé

A.Wald a comparé les solutions minimax et les solutions bayésiennes . Nous essayons d'établir dans ce travail un résultat similaire à celui de A.Wald entre solutions bayésiennes et les solutions du maximum de vraisemblance.Nous prenons l'espace des paramètres un groupe compact métrisable. Nous construisons au préalable une suite de fonctions de coût bornées, invariantes par translation, continues et mesurables.

Nous prouvons alors l'existence des solutions bayésiennes et des solutions de maximum de vraisemblance sous certaines conditions de régularité.

L'idée clé est de construire une suite de fonctions de coûts à l'aide du Lemme d'Urysohn. La loi à priori est fixée ; il s'agit de la mesure Haar. Le théorème de section de K.Kuratowski et Ryll-Nardzewski permet d'établir la mesurabilité des solutions bayésiennes et des solutions du maximum de vraisemblance. Nous montrons enfin qu'elles sont asymptotiquement équivalentes.Il s'agit en fait d'un problème in verse car la loi à priori étant fixée, on construit alors la fonction de coût.

Mots-clefs

Fonction de coût, mesure de Haar, solution bayésienne, solution du maximum de vraisemblance, groupe topologique, fonction multivoque, mesurabilité, *loi à priori*

Bayesian and maximum likelihood solutions .An asymptotic comparison related to cost function

Abstract

A.Wald showed that if the rule of minimax exist, it is also a rule of bayesian decision in comparaison with the most disadvantageous law, that is to say a rule of Bayes in comparaison with the law which would maximize (among all law) the risk of Bayes. The purpose of the present work to establish a result similar, to that of A.Wald, between Bayes estimator and maximum likelihood estimator.

We compare the bayesian solution and the maximum likelihood one when the parameter space is a metrisable compact group.We first construct a sequence of bounded cost functions which are invariant by translation.

We prove the existence of bayesian solutions and maximum likelihood ones under some regularity conditions. For us, the a priori law is fixed ; it is Haar measure, and the cost function that varies.The section theorem of K.KURATOWSKI and RYLL-NARDZEWSKI allows to establish measurability of bayesian solutions as will as maximum likelihood ones. We shall show that they are asymptotically equivalent.
Key Words : cost function, Haar measure, bayesian solution, maximum likelihood solution, topological group, multivoc function, measurability, a priori law.

Keywords

Cost function, Haar measure, Bayesian solution, maximum likelihood solution, topological group, multivoc functions, measurability, *a priori law.*

Table des matières

Introduction

0.1 Position du problème

Le problème fondamental de la théorie de la décision peut être formulé comme suit : étant donné le triplet (Θ, D, C), où Θ est l'espace des paramètres, D est l'espace des décisions finales, C est la fonction de coût , et un élément aléatoire observable $\omega \in \Omega$ dont la loi de distribution P_θ dépend du paramètre $\theta \in \Theta$, quelle règle de décision $\delta(\omega) \in D$ le statisticien doit choisir ? L'espace Ω cité plus haut est l'espace des observations , nous le munissons de la tribu a.

Pour discriminer entre les différentes règles de décision , nous utilisons un critère appelé fonction de risque, qui est définie par : $R(\theta, \delta) = E(C(\theta, \delta(\omega)))$, cette quantité représente la perte moyenne subie en choisissant $\delta(\omega)$ quand le paramètre θ est inconnu , et dépend de θ . Ce qui fait que nous ne pouvons obtenir de règles de décisions optimales dans le sens où $R(\theta, \delta)$ est uniformément minimum.

Afin de contourner cette difficulté, nous ordonnons les différentes règles de décision selon d'autres principes tels que : le principe du minimax, ou le principe de Bayes, ou encore nous considérons seulement des règles de décisions basées sur les méthodes intuitives telles que le principe du maximum de vraisemblance.
Nous supposons la famille de probabilités $(P_\theta), \theta \in \Theta$ dominée par une probabilité Q sur (Ω, a).

Nous noterons $L(\theta, .)$ la densité P_θ par rapport à Q.

L'estimateur du maximum de vraisemblance $\hat{\theta}$ est la valeur de θ qui maximise $L(\theta, \omega)$ à la vue de ω.Il est facile de remarquer que cette définition est tout à fait intuitive.

Une règle de décision δ_M est dite minimax, si elle minimise (parmi tous les règles δ) le risque maximum ; c'est à dire satisfait à :

$$\underset{\theta}{sup}(R(\theta,\delta_M)) = \underset{\delta}{inf}(\underset{\theta}{sup}(R(\theta,\delta))$$

Dans le cadre bayésien le paramètre θ est considéré comme une variable aléatoire et une loi de distribution μ lui est attribuée. Cette loi est appelée loi à priori.

Une règle de décision δ_B est dite règle de Bayes, si elle minimise le risque de Bayes , c'est à dire : $r(\mu,\delta_B) = inf_\delta(r(\mu,\delta)) = inf_\delta(E_\mu(R(\theta,\delta)))$ où l'espérance mathématique$E_\mu(R(\theta,\delta))$ est calculée par rapport à la loi μ .

L'expression $r(\mu,\delta)$est appelée risque de Bayes . Pour une description plus détaillée de ces règles de décision , le lecteur peut consulter [3] ou [5].

Les trois méthodes d'estimation que nous venons de décrire ne donnent pas nécessairement les mêmes estimateurs pour le paramètre inconnu θ, même dans un cadre asymptotique [12] pour une comparaison des performances asymptotiques des méthodes de Bayes et des méthodes du maximum de vraisemblance.

A. Wald [13] a cependant montré que si la règle de décision minimax existe , elle est aussi une règle de décision bayésienne par rapport à la loi à priori la plus défavorable, c'est à dire une règle de Bayes par rapport à la loi qui maximiserait (parmi toutes les lois) le risque de Bayes .

Le but de ce travail est d'établir un résultat similaire à celui de A.Wald entre estimateur de Bayes et estimateur du maximum de vraisemblance.
Nous tenterons de trouver un cadre(Θ, D, C) pour lequel l'estimateur de Bayes et l'estimateur du maximum de vraisemblance sont égaux ou asymptotiquement égaux.

Avant de procéder à la construction de ce cadre, signalons deux choses importantes :

1. Nous réduisons le problème par invariance. Cette simplification est motivée par les considérations suivantes : en cas d'égalité des solutions de Bayes et des solutions du maximum de vraisemblance, nous pouvons dire que la loi à priori μ sur l'espace des paramètres, n'a pas d'influence.

   Dans ce cas μ peut être interprétée comme une loi à priori « non informative

» . [9] justifie l'absence d'information dans μ par des propriétés d'invariance [6].

2. Il est possible d'obtenir des solutions bayésiennes pour des fonctions de coût plus générales que celles envisagées par le modèle de décision statistique de A.Wald , où la fonction de coût s'exprimait comme la somme de deux coûts positifs l'un relatif à l'observation effectuée, et l'autre à la décision finale prise.

   Nous pouvons considérer un coût global $C(\theta, z, \omega, t)$ de la décision finale z quand le paramètre vaut θ, après observation de la réalisation ω jusqu'à l'instant [11].

   Dans notre cas, le système observé n'est pas dynamique, c'est-à-dire que l'observation est fixée en durée, et l'information disponible est exprimée par la tribu a .

   Nous avons une fonction de coût C définie ainsi :

$$C : (\theta, \delta, \omega) \in \Theta \times D \times \Omega \mapsto C(\theta, \delta, \omega) \in \mathbb{R}$$

0.2 Description des résultats

En premier dans le chapitre 1, nous rappelons toutes les notions fondamentales dont nous aurons besoin telles que : la théorie de la mesure, la théorie de la décision statistique, les fonctions multivoques, les groupes topologiques, le théorème de Radon-Nicodym, le théorème de Fubini, le lemme d'Urysohn et la mesure de Haar.

Suit alors le chapitre 2 où nous établirons la mesurabilité de la fonction de vraisemblance $L(\theta, \omega)$ par rapport au couple (θ, ω). Par le théorème 2.2 nous dégageons les conditions pour lesquelles L est mesurable dans le couple.Puis on termine par une démonstration moyennant l'approche bayesienne . Cette démonstration est plus aisée que celle établie dans le livre de P.A.Meyer [13].

Au chapitre 3 nous comparons la solution bayésienne et la solution du maximum de vraisemblance quand l'espace des paramètres est un groupe compact métrisable.

Nous construisons au préalable une suite de fonctions de coût bornée, puis nous démontrons l'existence des solutions bayésiennes et des solutions du maximum de

vraisemblance sous certaines conditions de régularité.
Nous montrons enfin que ces deux solutions sont asymptotiquement équivalentes.

La suite de fonctions de coût est construite à l'aide du Lemme d'URYSOHN, puis la mesurabilité des solutions bayésiennes et du maximum de vraisemblance est établie à l'aide du théorème de la section mesurable de K. KURATOWSKI et RYLL-NARDZEWSKI.

Chapitre 1

Notions Fondamentales

Nous rappelons ici quelques notions fondamentales telles que :la théorie de la mesure, la théorie de la décision statistique, les groupes topologiques les fonctions multivoques, la mesure de Haar, et le Lemme d'Urysohn qui seront utilisés dans les chapitres suivants.

1.1 Rappels de la théorie de la mesure

1.1.1 Produit de deux espaces mesurables

Soient (Ω, a)et $(\Theta, \tau$ deux espaces mesurables.Sur le produit cartésien$\Theta \times \Omega$, on considère la famille définie par : $\mathcal{C} = \{A \subset \Theta \times \Omega, A = A_1 \times A_2, A_1 \in a$ et $A_2 \in \tau\}$ Les éléments de $\mathcal{C}$ s'appellent pavées mesurables de $\Theta \times \Omega$.

1.1.2 Proposition

$\mathcal{C}$ est une semi-algèbre de parties de $\Theta \times \Omega$.

1.1.3 Définition

La tribu produit de a et τ est par définition la tribu sur $\Omega \times \Theta$ engendrée par $\mathcal{C}$.On la note $a \otimes \tau$.

1.1.4 Définition

Soit A une partie de $\Omega \times \Theta$:
$\forall \omega \in \Omega$, on appelle coupe de A en ω le sous ensemble de Θ définie par :

$$A_\omega = \{\theta \in \Theta, (\omega, \theta) \in A\}$$

De même on définit la coupe de A en θ par :

$$A_\theta = \{\omega \in \Omega, (\omega, \theta) \in A\}$$

.

1.1.5 Remarque

Si $A \subset \Omega \times \Theta$ A est de la forme $A = A_1 \times A_2$ où $A_1 \subset \Omega$ et $A_2 \subset \Theta$.

$$A_\omega = \{ \begin{array}{l} A_2 \; si \; \omega \in A_1 \\ \emptyset \; sinon \end{array} \qquad A_\theta = \{ \begin{array}{l} A_1 \; si \; \theta \in A_2 \\ \emptyset \; sinon \end{array}$$

1.1.6 Proposition

$\forall A \in a \otimes \tau$ on a :

1. $A_\omega \in \tau, \forall \omega \in \Omega$
2. $A_\theta \in a, \forall \theta \in \Theta$

1.1.7 Définition

Soient (Ω, a) et (Θ, τ) deux espaces mesurables .
$(P_\theta) \in \Theta$ est une famille de probabilités de transition sur (Ω, a) paramétrées sur (Θ, τ) Si

1. $\forall \theta \in \Theta, A \to P_\theta(A)$ une probabilité sur $(\Omega, \mathfrak{a})$
2. $\forall A \in a, \theta \to P_\theta(A) \, est \tau$ - mesurable.

1.1.8 Proposition

Soient $(P_\theta)_{\theta\in\Theta}$ une famille de probabilité de transition sur (Ω, a) paramétrée sur (Θ, τ). Pour toute fonction mesurable f bornée ou positive de $(\Omega \times \Theta, a \otimes \tau)$ dans $\mathbb{R}$, l'application de (Θ, τ) dans $\mathbb{R}$:

$$\theta \longrightarrow \int_\Omega f(\omega, \theta) P_\theta(d\omega)$$

est définie et est mesurable.

Démonstration

1. Si f est la fonction indicatrice de $A \times B$ où $A \in a,\ B \in \tau.\ f = \chi_{A\times B}$.

$$\theta \longrightarrow \int_\Omega \chi_{A\times B}(\theta, \omega) P_\theta d(\omega) = \int_A \chi_B(\omega) P_\theta d(\omega) = \chi_B(\theta) P_\theta(A)$$

est τ-mesurable

2. Si pour $n \in \mathbb{N}, f_n = \sum_{i=1}^n \alpha_i \chi_{A_i\times B_i}, A_i \in a, B_i \in \tau, i = 1, \cdot, n.$

$$\theta \longrightarrow \int_\Omega \sum_{i=1}^n \alpha_i \chi_{A_i\times B_i}(\theta, \omega) P_{theta} d(\omega) = \sum_{i=1}^n \alpha_i \chi_{B_i}(\theta) P_\theta(A_i)$$

est τ-mesurable

3. Si f mesurable positive alors on sait que $f = \lim \uparrow f_n$ où f_n est une fonction étagée.Dans ce cas,

$$\theta \longrightarrow \int_\Omega f(\theta, \omega) P_\theta d(\omega) = \int_\Omega \lim \uparrow f_n(\theta, \omega) P_\theta d(\omega) = \lim \int_\Omega f_n(\omega, \theta) P_\theta d(\omega)$$

est τ-mesurable

4. Si f mesurable, alors elle admet la représentation $f = f^+ - f^-$.Dans ce cas, $\int_\Omega f(\theta, \omega) P_\theta d(\omega) = \int_\Omega f^+(\theta, \omega) P_\theta d(\omega) - \int_\Omega f^-(\theta, \omega) P_\theta d(\omega)$ est τ-mesurable

1.1.9 Théorème

Soient $(P_\theta)_{\theta\in\Theta}$ une famille de probabilité de transition sur (Ω, a) paramétrée sur (Θ, τ), et μ une probabilité sur (Θ, τ).
Il existe alors une unique probabilité Π sur $(\Omega \times \Theta, a \otimes \tau)$ telle que pour tout h fonction mesurable positive(ou bornée) de $(\Omega \times \Theta, a \otimes \tau)$ dans $\mathbb{R}$,

$$\int_{\Omega\times\Theta} h(\omega, \theta)\Pi(d\omega) = \int_\Theta (\int_\omega h(\omega, \theta)P_\theta(d\omega))\mu(d\theta)$$

Démonstration
D'après la proposition précédente le second membre est défini.On définit Π par :

$$\forall A \in a \otimes \tau : \Pi(A) = \int_\Theta (\int_\Omega \chi_A(\omega, \theta)P_\theta d(\omega))\mu d(\theta)$$

Π est une probabilité sur $(\Omega \times \Theta, a \otimes \tau)$ car en effet :

1. $\Pi(\emptyset) = 0$ et $\Pi(\Omega) = 1$;

2. $$\Pi(\sum A_{i=1}^n) = \int_\Theta (\int_\Omega \chi_{\sum A_i}(\omega, \theta)P_\theta d(\omega))\mu d(\theta) =$$
$$\sum_{i=1}^n \int_\Theta (\int_\Omega \chi_{A_i}(\omega, \theta)P_\theta d(\omega))\mu d(\theta) =$$
$$\sum_{i=1}^n \Pi(A_i).$$

Pour l'unicité et l'égalité citée dans le théorème, on commence par les fonctions génératrices, ensuite on généralise selon la technique classique.

1.1.10 Remarque

Ce théorème peut être prolongé au cas où μ est une mesure positive, $\sigma - finie$ sur (Θ, τ) ; dans ce cas nous obtenons une mesureΠ positive, $\sigma - finie$ sur $(\Omega \times \Theta, a \otimes \tau)$

1.1.11 Théorème Radon Nicodym

Soit une mesure positive,$\mu \quad \sigma - finie \quad sur\ \mathfrak{B}\ (tribu)$ et une mesure réelle ou complexe ν telle que $\nu \ll \mu$, alors il existe une fonction, $f \in L^1(\mu)$, telle que :

$$\nu(A) = \int_A f d\mu \ , \quad \forall\ A \in \mathfrak{B}.$$

Démonstration.

1. **Existence.**

   Nous allons envisager les cas $\mu < \infty$, $\nu \geq 0$, puis ν quelconque et $\mu\sigma - finie$ quelconque.

   (a) μ **finie** . a) On suppose d'abord que $\nu \geq 0$. Remarquons que s'il existe f répondant à la question, alors

   $\forall\varphi$ tel que $\varphi \leq f$, on aura $\nu(A) \geq \int_A \varphi d\mu$.

   On est donc amené à considérer l'ensemble Φ des fonctions φ, telles que

   $$\varphi \geq 0 \ , \nu(A) \geq \int_A f d\mu \ , \quad \forall\ A \in \mathfrak{B}.$$

   et à chercher s'il existe $f \in \Phi$, telle que

   $$\int_E \varphi d\mu \ \leq \int_E f d\mu \ .$$

   Or, $\varphi \in \Phi$ donc $(\int_A \varphi d\mu \leq \nu(E))$ et par conséquent $(\sup_{\varphi\in\Phi} \int_E \varphi d\mu \quad \leq \nu(E))$.

   Soit M définie comme suit :

   $$M = \sup_{\varphi\in\Phi} \int_E \varphi d\mu$$

   on va montrer qu'il existe $f \in \Phi$ telle que $M = \int_E f d\mu$.

   En effet :

   $$\forall n\ \exists\varphi_n \quad tel\ que\ \ \varphi_n \in \Phi,\ M - \frac{1}{n} \leq \int_E \varphi_n d\mu,$$

Soit $f_n = \sup(\varphi_1, \ldots, \varphi_n)$ et $f = \lim \quad \varphi_n$

Si $\varphi_n \in \Phi$ la propriété de Beppo-Levi entraine $f \in \Phi$ *et* $\int_E f d\mu \geq M$

Il suffit donc d'établir que $f_2 \in \Phi$ et de raisonner par récurrence.

$$\int_A f_2 d\mu = \int_A \sup(\varphi_1, \varphi_2) d\mu = \int_{A\cap B_1} \varphi_1 d\mu + \int_{A\cap B_2} d\mu$$

Avec $B_1 = \{x, \varphi_1(x) \geq \varphi_2(x)\}$ et $B_2 = \{x, \varphi_1(x) < \varphi_2(x)\}$

Donc $\int_A f d\mu \leq \nu(A \cap B_1) + \nu(A \cap B_2) = \nu(A)$

D'où l'existence de la fonction f telle que $\int_E f d\mu = M$.

La fonction f satisfait

$$\nu(A) = \int_A f d\mu \ , \quad \forall\ A \in \mathfrak{B}.$$

Comme $f \in \Phi$, il reste à établir

$$\nu(A) \geq \int_A f d\mu \ , \quad \forall\ A \in \mathfrak{B}.$$

Supposons le contraire . Alors, $\exists A_0 \ , \epsilon > 0$ tels que

$$\int_{A_0} f d\mu < \nu(A_0) - \epsilon$$

Soit $\exists A_0$ *et* $\epsilon > 0$, tels que $\int_{A_0} (f + \epsilon) d\mu < \nu(A_0)$, puisque μ est finie.

Puisque $\nu \ll \mu$ alors nécessairement $\mu(A_0) > 0$.

Appliquons le théorème de décomposition de Hahn à la fonction d'ensemble :

$$A \longrightarrow \nu(A) - \int_A (f + \epsilon) d\mu \ ,$$

Il existe alors une partition A_1 *et* A_2 de A_0, telle que :

$$\nu(A) - \int_A (f + \epsilon) d\mu \geq 0 \quad , \forall A \subset A_1$$

$$\nu(A) - \int_A (f + \epsilon) d\mu \leq 0, \ \forall A \subset A_2$$

Soit $f_\epsilon = \begin{cases} f+\epsilon \quad sur \ A_1 \\ f \ sur \ \overline{A_1} \end{cases}$

Alors $f_\epsilon \in \Phi$ et $\int_E f_\epsilon d\mu \ > \int_E f d\mu$ ce qui est contraire à la propriété de f.

b) Supposons ν réelle ou complexe. On a $\nu = \nu_1 - \nu_2$ dans le cas réel ,$\nu \ _1 et \ \ \nu_2$ sont deux mesures positives absolument continues si ν l'est (résulte du Théorème de Hahn). On peut donc appliquer le cas précédent $\nu \geq 0$ à chacune des mesures$\nu \ _1 et \ \ \nu_2$, d'où le résultat. Le cas complexe se ramène au cas réel.

(b) $\mu \ \ \sigma-$**finie**.

Soit$E = \bigcup_{k=1}^{\infty} E_k$, et $F = \bigcup_{k=1}^{m}$ où les E_k disjoints deux à deux et de mesures finies.

D'aprés le cas précédent sur chaque E_k, il existe f_k .On pose $f = f_k$ sur E_k.

Si $A \subset \bigcup_{k=1}^{m} E_k$, on a bien

$$\nu(A) = \int_A f d\mu$$

Donc $\int_F |f| \, d\mu \leq \|\nu\|$, car la variation totale de la mesure définie sur F par $A \longrightarrow \int_A f d\mu$, est égale à : $\int_F |f| \, d\mu$

L'inégalité $\int_F |f| \, d\mu \leq \|\nu\|$ valable, pour tout m, entraine que $f \in L^1(E, \mathfrak{B}, \mu)$

Pour tout $A \in \mathfrak{B}$, on a donc

$$\int_A f d\mu = \lim_{m\to\infty} \int_{F\cap A} f d\mu = \lim \nu(A \cap F) \quad = \nu(A)$$

d'après le Théorème de Lebesque (convergence dominée).

2. **UNICITE.**

   L'unicité résulte du fait que $\|\nu\| = \int_E |f| \, d\mu$.

On achève cette section par le théorème de Fubini qu'on a jugé important de rappeler car il sera utilisé dans les Chapitres suivants.

1.1.12 Théorème de Fubini

Soit f une fonction positive : $(\theta,\omega) \mapsto f(\theta,\omega)$ et $\mathfrak{B}-$ mesurable, alors :

1. $\forall\theta \in \Theta : \omega \longmapsto f(\theta,\omega)$ est une fonction $\mathfrak{B}_2-$ mesurable et $\int_\Omega f(\theta,\omega)d\mu_2$est une fonction $\mathfrak{B}_1-$mesurable.

2. $\forall\omega \in \Omega : \theta \longmapsto f(\theta,\omega)$ est une fonction $\mathfrak{B}_1-$ mesurable et $\int_\Theta f(\theta,\omega)d\mu_1$est une fonction $\mathfrak{B}_2-$mesurable.

3. $\int_{\Theta\times\Omega} f(\theta,\omega)\,d\mu \ = \int_\Theta d\mu_1 \int_\Omega f(\theta,\omega)\,d\mu_2 = \int_\Omega d\mu_2 \int_\Theta f(\theta,\omega)\,d\mu_1$

Démonstration. Cet énoncé est vraie pour les fonctions caractéristiques d'ensembles $\mathfrak{B}$ - mesurables .
Par passage à la limite à partir de fonctions étagées mesurables on voit qu'il est valable pour toute fonction $\mathfrak{B}-$ mesurable positive.

1.1.13 Définition

Soient (Ω, a) et (Θ, τ) deux espaces mesurables et $(P_\theta) \in \Theta$ une famille de probabilités de transition sur (Ω, a) paramétrées sur (Θ, τ) . Nous appelons structure statistique (ou modèle) le triplet (le système) $\{\Omega, a, (P_\theta) \in \Theta\}$. Nous disons que la structure statistique $\{\Omega, a, (P_\theta) \in \Theta\}$ est dominée si :

il existe une mesure positive , σ -finie Q sur (Ω, a) telleque :

$$\forall\theta \in \Theta, P_\theta << Q.$$

D'après le théorème de RADON-NIKODYM, il existe une application , f_θ de (Ω, a) sur $\left(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+}\right)$ mesurable telle que :

$$\forall A \in a, P_\theta(A) = \int_A f_\theta(\omega)\, Q(d(\omega))\,.$$

f_θ est appelée densité de probabilité de P_θ par rapport à Q ; Nous appelons fonction de vraisemblance l'application :

$$L : (\omega,\theta) \in \Omega\times\Theta \mapsto L(\omega,\theta) = f_\theta(\omega) \in \mathbb{R}^+.$$

1.1.14 Remarques

1. La mesure dominante n'est pas unique .Si :

$$\forall \theta \in \Theta, P_\theta << Q.$$

où Q est une mesure positive σ -finie sur $(\Omega, \mathfrak{a})$ et si $Q << \nu$ où ν est une mesure positive σ-finie sur (Ω, a), alors

$$\forall \theta \in \Theta, P_\theta << \nu.$$

Nous pouvons toujours et ce sera commode pour les calculs théoriques de choisir comme mesure dominante une loi de probabilité équivalente à la famille $(P_\theta), \theta \in \Theta$ [1].

2. La fonction de vraisemblance n'est mesurable que par rapport à ω. Pouvons nous affirmer qu'elle est mesurable dans le couple.C'est l'objet du chapitre 2.

1.2 Théorie de la décision statistique

Cette section rappelle certaines notions de la théorie statistique . Soient :
(Ω, a)un espace mesurable (espace des observations).
(Θ, τ) un espace mesurable (espace des paramètres).
$(D, \mathcal{D})$ un espace mesurable (espace des décisions possibles).
$(P_\theta), \theta \in \Theta$ une famille de probabilités de transition sur(Ω, a) paramétrée sur (Θ, τ) .
$P_\theta(\cdot)$ régit l'observation ω dans Ω quand θ est la valeur (inconnue) du paramètre .
C une fonction mesurable :

$$C : (\theta, \delta, \omega) \in \ \Theta \times D \times \Omega \mapsto C(\theta, \delta, \omega) \in \mathbb{R}^+$$

où $C(\theta, \delta, \omega)$ représente le coût de la décision finale δ quand le paramètre vaut θ et quand on observe ω.

Une règle de décision (ou stratégie) δ est une application de (Ω, a) dans (D, D) mesurable ; elle consiste à décider $\delta(\omega)$ après avoir observé ω.

Nous noterons par $C(\theta, \omega, \delta(\omega))$ le coût de la décision $\delta(\omega)$ quand nous avons observé ω et quand le paramètre vaut θ et par Δ l'ensemble des règles de décision.

Résoudre un problème de décision statistique consiste à choisir une règle de décision dans Δ (selon des critères reconnus comme raisonnables).

Nous disposons de la fonction de risque :

$$R(\theta, \omega) = \int_\Omega C(\theta, \omega, \delta(\omega)) P_\theta d\omega.$$

Comme critère de choix, nous définissons une relation de préférence dans Δ. Soient δ et $\delta' \in \Delta$:

δ' est préféré à δ, on note ($\delta \ \leq \delta'$), si et seulement si

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

δ' est équivalente à δ , on note ($\delta \ \sim \delta'$), si et seulement si

$$[\ R(\theta, \delta') = R(\theta, \delta) \ \ , \ \forall \theta \in \Theta \ ]$$

δ' est strictement préférée , on note ($\delta < \delta'$), si et seulement si

$$[\ R(\theta, \delta') \leq \ R(\theta, \delta) \ , \ \forall \theta \in \Theta \ ] \quad \text{et} \ [\ \exists \theta \in \Theta, \ R(\theta, \delta') < \ R(\theta, \delta) \ ].$$

1.2.1 Définition

$\delta \in \Delta$ est appelée règle de décision admissible, s'il n'existe pas $\delta' \in \Delta$ telle que $\delta < \delta'$, autrement dit il n'existe pas de $\delta' \in \Delta$ telle que nous ayons :

$$\{ \ [\ R(\theta, \delta') \leq R(\theta, \delta), \quad \forall \theta \in \Theta \ ] \quad et \quad [\ \exists \theta \in \Theta \ \ , \ R(\theta, \delta') < \ R(\theta, \delta) \ ] \ \}.$$

1.2.2 Définition

Une règle de décision $\delta_0 \in \Delta$ est uniformément optimale si :

$$\forall \delta \in \Delta \quad , \delta \ \leq \ \delta_0.$$

C'est à dire :

$$R(\theta, \delta_0) \leq R(\theta, \delta), \forall \delta \in \Delta, \forall \theta \in \Theta.$$

Il n'existe pas en général de règle de décision optimale, puisque la fonction de risque $R(\theta, \delta)$ dépend d'un paramètre inconnu θ ; mais des règles de décision admissibles. Pour choisir entre celles-ci nous sommes amenés à :

1. Soit à raisonner de manière conservatrice, c'est à dire en fournissant une solution minimax .

   Nous posons :
   $$\forall \ \ \delta \in \Delta \ \ , \quad \rho(\delta) = \sup_{\theta \in \Theta} R(\theta, \delta).$$

   δ_M est appelé solution minimax si

   $$\rho(\delta_M) \quad \leq \quad \rho(\delta) \quad , \quad \forall \delta \in \Delta$$

   Autrement dit :

   $$\sup_{\theta \in \Theta} R(\theta, \delta_M) \quad = \quad \inf_{\delta} (\sup_{\theta \in \Theta} R(\theta, \delta) \ ).$$

2. Soit à donner une idée préconçue de la solution par exemple en estimation ponctuelle du paramètre θ.Nous choisissons la valeur de θ qui donne la plus grande probabilité du système étudié. Cette procédure permet d'obtenir un estimateur du maximum de vraisemblance.

3. Soit à partir du principe que sur l'espace des paramètres Θ, nous avons une probabilité μ qui pondère entre les valeurs possibles du paramètre. C'est la théorie de la décision bayésienne .

Dans ce cas nous considérons le risque de Bayes défini ainsi :

$$r(\mu, \delta) = \int_{\Theta} R(\theta, \delta)\, \mu(d\theta) = \int_{\Theta} \left[\int_{\Omega} C(\theta, \omega, \delta(\omega))\, P_{\theta}(d\omega) \right] \mu(d\theta).$$

δ_B est appelée solution bayésienne si :

$$r(\mu, \delta_B) = \inf_{\delta}(\quad r(\mu, \delta)).$$

De telles situations sont très fréquentes en statistique pour plus de détails [7] et [13].

1.3 Groupes Topologiques

Comme celà a été indiqué en introduction, nous sommes amenés à étudier des lois à priori satisfaisant à des propriétés d'invariance ; d'où l'intérêt de rappeler quelques notions sur les groupes topologiques et mesure de Haar. C'est l'objet des deux sections suivantes.
Soit G un groupe noté multiplicativement (pas nécessairement commutatif) et e son élément neutre.

1.3.1 Définition

Nous disons qu'une topologie sur G est compatible avec la structure de groupe si les deux applications suivantes : a) L'application :

$$(x, y) \in G \times G \mapsto xy \in G.$$

b) L'application :

$$(x, y) \in G \times G \mapsto x^{-1} \in G.$$

sont continues.

Un groupe muni d'une topologie compatible avec la structure de groupe est appelé groupe topologique.

1.3.2 Remarques

1. Pour tout $a \in G$, la translation à gauche : $x \mapsto ax$ (respectivement la translation à droite $x \mapsto xa$) est continue. C'est un homéomorphisme de G sur lui même .
2. Pour tout $a, b \in G$, l'application : $x \mapsto axb$ (en particulier $x \mapsto axa^{-1}$) est un homéomorphisme de G sur lui-même.
3. L'application :

$$x \mapsto x^{-1}$$

est bijective, et est égale à l'application réciproque.C'est un isomorphisme de G sur lui-même .

1.3.3 Proposition

Soit G un groupe topologique .

1. Pour toute partie ouverte (respectivement fermée) A de G et pour tout $x \in G$: les ensembles xA, , Ax et A^{-1} sont des ouverts (respectivement des fermés).
2. Pour toute partie ouverte A de G et pour toute partie B de G, les ensembles AB et BA sont des ouverts de G.
3. Si V est un voisinage de e dans G, et A une partie non vide quelconque de G, alors VA et AV sont des voisinages de A. En effet si W est un voisinage ouvert de e contenu dans V, on a WA et AW sont des ouverts et contiennent A, $WA \subset VA$ et $AW \subset AV$.

1.3.4 Remarques

1. Soit $a \in G$ et V un voisinage de e.Puisque $x \mapsto ax$ et $x \mapsto xa$ sont des homéomorphismes alors aV et Va sont des voisinages de a.
2. Si on exprime que $x \mapsto xy^{-1}$ $x \mapsto x$ sont continues pour $x = y = e$ on obtient :

i) $\forall U$ *(voisinage de e)* $\exists V$ *(voisinage de e) telque* $V^{-1} \subset U$

ii) $\forall U$ *(voisinage de e) on a* U^{-1} *(voisinage de e)*

1.3.5 Définition

On dit qu'un voisinage V de e est symétrique si $V = V^{-1}$.

1.3.6 Définition

Une famille $\mathcal{F}$ de voisinage de x forme un système fondamental de voisinage de x si :
$\forall U$ voisinage de x $\exists V \in \mathcal{F}$ tel que $V \subset U$.

1.3.7 Proposition

Les voisinages symétriques de e forment un système fondamental de voisinages de e.

En effet puisque pour tout voisinage U de e, U^{-1} est aussi voisinage de e donc $W = U \cap U^{-1}$ est un voisinage de e contenu dans U et est symétrique.

1.3.8 Groupes métrisables

Pour que la topologie d'un groupe topologique soit métrisable il faut et il suffit qu'il existe un système fondamental dénombrable de voisinages de l'élément neutre e dont l'intersection est réduite à e. Lorsque qu'il en est ainsi, la topologie du groupe G peut être définie par une distance invariante à gauche et à droite.

1.3.9 Définition

Soient G un groupe et E un ensemble. Une opération (ou opération à gauche ou action) de G dans E est une application : $(s, x) \in G \times E \mapsto sx \in E$ telle que :
Si e est l'élément neutre de G, on a : $ex = x, \forall x \in E$.
$\forall s, t \in G$ on a : $s(tx) = (st)x, \forall x \in E$.
Nous dirons que G opère sur E.

1.3.10 Définition

. Soit un G groupe topologique et E un espace topologique. Nous dirons que G opère continument dans E si l'application :

$$(s,x) \in G \times E \mapsto sx \in E$$

est continue.

1.3.11 Lemme

Si G est un groupe topologique opérant continument dans un espace topologique E alors pour tout $s \in G$, l'application

$$x \in E \mapsto sx \in E$$

est un homéomorphisme.
En effet c'est une bijection continue, ainsi que sa bijection réciproque.

1.3.12 Remarque

Une opération à droite d'un groupe G dans un ensemble E est une application définie comme suit :

$$(s,x) \in G \times E \mapsto xs \in E$$

telle que :

1. $xe = x, \forall x \in E.$;
2. $x(st) = (xs)t, \forall s,t \in G, \forall x \in E$.

On transpose aussitôt à ces opérations tout ce qui a été dit ci dessus, c'est à dire que toutes les définitions et propriétés énoncées plus haut restent valables dans le cas d'une opération à droite.

1.4 Mesure de Haar

La notion de « mesure de Haar » a été introduite par Alfréd Haar en 1933, ainsi que la démonstration de son existence pour tout groupe localement compact, sous l'hypothèse restrictive que le groupe est métrisable et séparable (ce qui équivaut en fait à le supposer seulement à base dénombrable).

La démonstration de son « unicité » (à la proportionnalité près) a été faite par John von Neumann en 1935. En 1940, André Weil a démontré le théorème en toute généralité et Henri Cartan en a produit une preuve sans l'axiome du choix et démontrant simultanément l'existence et l'unicité.

Une mesure de Haar μ sur un groupe localement compact est une mesure de Borel quasi-régulière non nulle, invariante par translation à gauche.
Autrement dit, pour toute partie borélienne B de G, et pour tout g dans G, on a :

$$\mu(gB) = \mu(B)$$

L'existence d'une mesure de Haar μ est assurée dans tout groupe localement compact. Elle est finie sur les parties compactes de G. De plus, toute mesure borélienne complexe invariante par translation à gauche s'écrit $\alpha\mu$ où α est un nombre complexe.
Bien qu'elle ne soit définie qu'à un coefficient multiplicateur près, de nombreux ouvrages parlent de la mesure de Haar. Cet usage est justifié pour un groupe compact ou pour un groupe discret, où des normalisations peuvent être effectuées.

L'existence d'une mesure de Haar sur un groupe compact peut être déduite du théorème du point fixe de Kakutani.
Comme le groupe G est compact, $\mu(G)$ est finie, et quitte à effectuer une normalisation, il est possible de supposer μ de probabilité.

Sur tout groupe topologique localement compact et séparable G, il existe une mesure borélienne invariante par les translations à gauche, appelée mesure de Haar, unique à coefficient multiplicatif près. Elle est finie sur les parties compactes de G. Si G est lui-même compact, toute mesure de Haar est finie, et il est possible de la normaliser. On dispose ainsi sur tout groupe compact G d'une unique mesure de probabilité qui soit invariante par translations à gauche, qu'on nomme la mesure de Haar de G, notée μ dans la suite de l'article.

En pratique, la mesure de Haar permet de moyenner les objets sur G pour obtenir des objets invariants.

1.4.1 Notations

Soit G un groupe topologique opérant continument à gauche dans un espace localement compact, nous noterons par $\gamma(s)$ l' homéomorphisme de E dans E, définie par :

$$\gamma(s)x = sx, \forall s \in G, \forall x \in E.$$

Si μ est une mesure défine sur E, nous noterons par $\gamma(s)\mu$ la mesure image de μ par $\gamma(s)$ .On a :

$$(\gamma(s)\mu(A) = \mu(s^{-1}A), \forall A.$$

où A un ensemble mesurable de E.

1.4.2 Définition

Soit μ une mesure sur E Nous dirons que μ est invariante par G si

$$\gamma(s)\mu = \mu \forall s \in G.$$

1.4.3 Remarque

Soit G est un groupe topologique opérant continument à droite dans E.Nous noterons par $\delta(s)$ l'homéomorphisme de E dans E, défini par :

$$\delta(s)x = xs^{-1}.$$

Soit μ est une mesure sur E, nous noterons par $\delta(s)\mu$ la mesure image de μ par $\delta(s)$

1.4.4 Définition

Si μ est une mesure sur E.Nous dirons que μ est invariante à droite par G si

$$(\delta(s)\mu) = \mu, \forall s \in G$$

Lorsque G est un groupe localement compact opérant sur lui-même par translation à gauche et à droite c'est à dire :

$$\gamma(s)x = sx \; et \; \delta(s)x = xs^{-1}.$$

Alors nous pouvons définir sur G les notions de mesures invariantes à gauche et à droite.

1.4.5 Définition

Soit G un groupe localement compact. Nous appelons mesure de Haar à gauche (respectivement à droite) sur G une mesure positive non nulle, σ-finie sur G invariante à gauche (respectivement à droite).

L'existence et l'unicité d'une pareille mesure sont données par le théorème suivant. La démonstration de ce résultat est donnée dans [8].

1.4.6 Théorème

Sur tout groupe localement compact, il existe une mesure de Haar à gauche (respectivement à droite) et à un facteur constant prés, il n'en existe qu'une.

1.4.7 Remarque Importante

Si G est compact , il existe une mesure de Haar et une seule μ sur G telle que $\mu(G) = 1$.On l'appelle mesure de Haar normalisée.

1.5 Applications multivoques

Nous nous servons des applications multivoques pour établir certains résultats importants du Chapitre 3.Il nous parait nécessaire d'en faire un petit paragraphe. Pour plus de détails le lecteur peut consulter [2].

1.5.1 Définition

Soient Xet Y deux ensembles quelconques . Pour tout $x \in X$ nous faisons correspondre un sous ensemble $\Phi(x)$ de Y .Nous dirons que la correspondance

$$x \in X \to \Phi(x) \subset Y$$

est une application multivoque de X dans Y. Ainsi on définit les ensembles :

$$X^\star = \{x \in X/\Phi(x) \neq \emptyset\},$$

appelé ensemble de définition de Φ.

$$Y^\star = \underset{x \in X}{\cup} \Phi\left(x\right),$$

appelé ensemble des valeurs de Φ.
Si $\Phi(x)$ est composé d'un seul élément, nous dirons que Φ est une application univoque de X dans Y.
L'inverse supérieur Φ^+ est par définition

$$\Phi^+(B) = \{x \in X/\Phi(x) \subset B\}, \forall B \subset Y.$$

L'inverse inférieur Φ^- est par définition

$$\forall B \subset Y : \Phi^-(B) = \{x \in X/\Phi(x) \cap B \neq \emptyset\}.$$

1.5.2 Théorème

$$\Phi^{-}\left(\underset{i}{\cup}A_i\right)=\underset{i}{\cup}\Phi^{-}\left(A_i\right).$$

La démonstration est immédiate.
Pour terminer ce chapitre, nous citons un théorème de KURATOWSKI et RYLL-NARDZEWSKI, qui sera utilisé dans le chapitre 3. Avant de citer le théorème, donnons une définition.

1.5.3 Définition

Un espace polonais est un espace mesurable, séparable, sur lequel il existe un métrique compatible avec la topologie pour laquelle l'espace est complet.

Soit X un espace polonais et $(U,\mathcal{U})$ un espace mesurable .
Nous considérons l'application multivoque de U dans X telle que :
$\forall u\in U,\Phi(u)$ est fermé et non vide dans X .

1.5.4 Théorème

Si $\Phi^{-}(G)=\{u/\Phi(u)\cap G=\emptyset\}\in\mathcal{U}$ pour tout G ouvert dans X, alors Il existe application ϕ mesurable de (U,U) dans $(X,\mathcal{B}_X)$ telle que :

$$\forall u\in U,\phi(u)\in\Phi(u).$$

Pour la démonstration, le lecteur peut consulter [10] .On achève cette section en rappellant le Lemme d'Urysohn.

1.6 Lemme d'Uryshon

Ce lemme établit que pour deux fermés disjoints F et G d'un espace normal X il existe une application continue de X dans $[0,1]$ qui vaut 0 sur F et 1 sur G.On le considère parfois comme le premier résultat non trivial sur les espaces topologiques.

Ce lemme permet d'étendre aux espaces normaux le théorème de Tietze initialement démontré en 1914 par Heinrich Tietze pour les espaces métriques .
Pavel Urysohn trouve une nouvelle démonstration et énonce son lemme un peu plus tard,dans un texte mathématique dont l'objectif est la démonstration des théorèmes sur l'invariance de la dimension d'un espace topologique localement homéomorphe à un espace euclidien.
Il existe un premier énoncé spécifique aux espaces T_4 Théorème

1.6.1 Lemme d'URysohn

Soient X un espace T_4 et F , G deux fermés disjoints de X.Alors il existe une fonction continue$f$ de X dans $[0,1]$, définie comme suit :

$$f : x \in X \mapsto \text{f(x)} = \begin{cases} 1 & si \quad x \ \in \text{G} \\ 0 & si \quad x \ \in F \end{cases}$$

Démonstration
Soit D une partie dénombrable dense de $]0,1[$. On définit une famille$(U\left(r\right))_{r \in D}$ d'ouverts telle que :

1. $\forall r \in D, F \subset U\left(r\right)$ et $\overline{\left(U\left(r\right)\right)} \subset X \backslash G$;
2. $\forall r, s \in D, \ r < s \quad \Rightarrow \ \overline{\left(U\left(r\right)\right)} \ \subset \ U\left(s\right)$.

Comme D est dénombrable , on choisit de noter par r une bijection de $\mathbb{N}$ dans D, puis on raisonne par récurrence.
Soient n$\in$ $\mathbb{N}$, (r,s) $\in D^2$ et supposons que les $U(r_k)$ pour k< n vérifient les propriétés 1. et 2. ci-dessus .
Ainsi $F_n = F \cup (\bigcup_{k<n, r_{k<} r_n} \overline{U(r_k)})$ est fermé et est inclus dans l'ouvert suivant :

$$O_n = (\bigcap_{k<n, r_{k>} r_n} U(r_{_k}) n (X \backslash G)$$

Puisque X est un T_4 , il existe donc un ouvert $U(r_n)$tel que :
$F_n \subset U\left(r_n\right)$ et $\overline{U(r_n)} \subset O_n$.
S'ensuit alors la définition de la fonction$f$de Xdans $[0,1]$ par :

$f(x) = \inf(\{s \in D \ , x \in U(s) \ \cup \{1\}) = sup(\left\{s \in D, x \notin \overline{U(r)} \ \cup \{\ 0\right\})$.

Par construction f vaut 0 sur F et 1 sur G. Elle est continue car elle est semi-continue inférieurement et supérieurement.

Un autre corollaire est le suivant .

1.6.2 Corollaire

Soient X un espace localement et K une partie compact de X, alors il existe une application continue de X dans $[0, 1]$ à support compact et qui vaut 1 sur K.

Chapitre 2

Mesurabilité de la fonction de vraisemblance

2.1 Introduction

Comme cela a été indiqué en chapitre 1, le problème de la mesurabilité de la fonction de vraisemblance L par rapport au couple (θ, ω) n'est pas évident .Elle a été abordée et démontrée par P.A.MEYER [13].Sa démonstration est basée sur les théorèmes de convergence des martingales.Notre travail ici consiste a élaborér une démonstration plus directe de ce même résultat moyennant l'approche bayésienne.

2.2 Définition

Soit (Ω, a) un espace mesurable. Nous dirons qu'une sous tribu $\mathcal{F}$ de a est séparable si $\mathcal{F}$ est engendrée par une suite $(A_n)_n$ d'éléments de a.

2.3 Théorème

Soient(Ω, a) *et* (Θ, τ) *deux espaces mesurables. Nous supposerons que la tribu* a *est séparable.*
Si $(P_\theta)\theta \in \Theta$ *est une famille de probabilités de transition sur* (Ω, a) *paramétrée sur* (Θ, τ) *telle que :* $\forall \theta \in \Theta, P_\theta \ll Q$ *où* Q *est une probabilité sur* (Ω, a) *, alors :*

pour toute probabilité(ou mesure positive, $\sigma - finie$) μ sur (Θ, τ) il existe une application L mesurable, positive sur $(\Theta \times \Omega, \tau \otimes a)$ dans $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ telle que :
μ presque tout $\theta \in \Theta$, $L(\theta, .)$ est la densité de P_θ par rapport à Q

2.3.1 preuve

Nous traitons le cas où μ est une probabilité sur (Θ, τ) .Le cas où μ est une mesure positive $\sigma - finie$ se traite de la même manière. D'après le théorème 1.1.9, il existe une probabilité Π sur $(\Theta \times \Omega, \tau \otimes a)$ telle que pour toute fonction f mesurable positive(ou bornée) de $\Omega \times \Theta$ dans $\mathbb{R}$ nous avons

$$\int_{\Omega \times \Theta} f(\omega, \theta)\Pi(d\omega\theta) = \int_{\Theta}[\int_{\Omega} f(\omega, \theta)P_\theta(d\omega)]\mu(d\theta)$$

Nous pouvons affirmer que Π est absolument continue par rapport $\mu \otimes Q$. En effet, soit $E \in a \otimes \tau$ tel que $\mu \otimes Q(E) = 0$.
Considérons la coupe E_θ de E en θ c'est à dire : $E_\theta = \{\omega : (\omega, \theta) \in E\} \ \in a, \ \forall \theta \in \Theta$.
pour θ fixé, $\chi_E(\omega, \theta) = \chi_{E_{\omega,\theta}}(\omega)$

$$\mu \otimes Q(E) = \int_{\Theta}[\int_{\Omega} \chi_E(\omega, \theta)Q(d\omega)]\mu(d\theta) = \int_{\Theta}[\int_{\Omega} \chi_{E_\theta}(\omega)Q(d\omega)]\mu(d\theta) = \int_{\Theta} Q(E_\theta)\mu(d\theta) = 0$$

d'où $Q(E_\theta) = 0 \mu$ presque partout
Par conséquent $P_\theta(E_\theta) = 0\mu$ p.p.(car $P_\theta \ll Q$) et donc $\int_\Theta P_\theta(E_\theta)\mu(d\theta) = 0$.
Or :

$$\int_{\Theta} P_\theta(E_\theta)\mu(d\theta) = \int_{\Theta}[\int_{\Omega} \chi_{E_\theta}(\omega)P_\theta(d\omega)]\mu(d\theta) = \int_{\Theta}[\int_{\Omega} \chi_E(\omega, \theta)P_\theta(d\omega)]\mu(d\theta) = \Pi(E)$$

Puisque $\mu \otimes Q(E) = 0$, alors $\Pi(E) = 0$ .Ceci prouve bien l'absolue continuité de Π par rapport à $\mu \otimes Q$.
Puisque $\Pi \ll \mu \otimes Q$, il existe une application L de $\Theta \times \Omega$ dans $\mathbf{R}^+$ mesurable telle que

$$\forall E \in a \otimes \tau, \Pi(E) = \int_E L(\theta, \omega)\mu \otimes Q(d\theta\omega).$$

Donc $\forall A \in \tau, \forall B \in a$, nous avons d'une part :

$$\Pi(A \times B) = \int_{A\times B} L(\theta,\omega)\mu \otimes Q(d\theta\omega) = \int_A [\int_B L(\theta,\omega)Q(d\omega)]\mu(d\theta).$$

D'autre part :

$$\Pi(A \times B) = \int_\Theta [\int_\Omega \chi_{A\times B}(\omega,\theta)P_\theta(d\omega)]\mu(d\theta) = \int_A [\int_B P_\theta(d\omega)]\mu(d\theta) = \int_A P_\theta(B)\mu(d\theta).$$

Par conséquent on obtient :$P_\theta(B) = \int_B L(\theta,\omega)Q(d\omega)\mu$ p.p.

Soit A une algèbre dénombrable engendrant $\mathbf{a}$, $A = (A_n, n \in \mathbf{N})$.

Pour A_n , nous avons N_n négligeable appartenant à τ tel que :

$$\forall\, \theta \in \Theta - N_n,\; P_\theta(A_n) = \int_{A_n} L(\theta,\omega)Q(d\omega)$$

Posons $N = \bigcup N_n$ (négligeable)

$$\forall\, \theta \in \Theta - N,\, \forall n \in \mathbf{N}, P_\theta(A_n) = \int_{A_n} L(\theta,\omega)Q(d\omega).$$

Une probabilité sur une algèbre s'étend d'une façon unique sur la σ-algèbre engendrée. Donc :

$$\forall \theta \in \Theta - N, P_\theta(B) = \int_B L(\theta,\omega)Q(d\omega), \forall B \in \mathbf{a}$$

Alors : $\forall \theta \in \Theta - N, L(\theta,\omega)$ est la densité de P_θ par rapport à Q.

2.3.2 Remarque

Notons que les conditions du théorème 2.3 sont les mêmes que celles citées dans le livre de P.A.Meyer[13]. Nos résultats sont valables μ presque partout.

La présente démonstration est plus simple.Cette simplicité est due au fait que nous prenions une mesure μ sur (Θ, τ), ceci n'est autre que l'approche bayesienne en décision statistique.

Chapitre 3

Solutions bayesiennes et solutions du maximum de vraisemblance quand Θ est compact

Dans ce chapitre, nous donnons un cadre mathématique qui nous permet d'affirmer que les solutions bayésiennes et les solutions du maximum de vraisemblance sont asymptotiquement égales .

Nous construisons une suite de fonctions de coût $C_n(\theta, z, \omega)$ bornées, invariantes par translation, continues en θ et z et mesurables en ω(voir Lemme 3.3.2).

Pour ce type de fonction de coût, nous obtenons une suite de solutions bayésiennes $(\tilde{\theta}_n)_n$ qui converge simplement vers la solution du maximum de vraisemblance $\hat{\theta}$ (voir Théorème 3.2).

Dans la section 3.5, nous généralisons le Théorème 3.2, en supprimant l'hypothèse restrictive d'unicité de la solution du maximum de vraisemblance.

Nous nous assurons au préalable dans les Lemmes 3.3.1 et 3.3.3 de l'existence et de la mesurabilité des solutions du maximum de vraisemblance et des solutions bayésiennes.

3.1 Notations

Ω l'espace des observations, on le muni de la tribu a.

Θ l'espace des paramètres, munissons le de la tribu borélien τ et on le suppose un groupe (noté multiplicativement) compact et métrisable.

e l'élément neutre de Θ et par d la distance invariante par translation.

$(\varepsilon_n)_n$ une suite de nombres réels positifs telle que :

i) $(\varepsilon_n)_n$ décroit strictement vers 0 ;

ii) $2\varepsilon_{n+1}<\varepsilon_n$;

iii) $\varepsilon_0<$ (diamètre $\Theta \wedge 1$) .

$(V_n=B(e,(\varepsilon_n))_{n\in N}$ est une base de voisinage de e telle que

$$B^*(e,\varepsilon_n)\cdot B^*(e,\varepsilon_n)\subset B(e,\varepsilon_{n-1})$$

Où $B^*(e,\varepsilon_n)$ est la boule fermée de centre e et de rayon ε_n.

$(P_\theta)_{\theta\in\Theta}$ est une famille de probabilités de transition sur (Ω,a) paramétrée (Θ,τ).

$L(\theta,\cdot\,)$ la densité de probabilité P_θ par rapport à Q.

Quand θ est la valeur inconnue du paramètre, elle est estimée par d appelée décision finale et appartient à D.

3.2 Hypothèses

H1 : Nous supposons qu'il existe une probabilité Q sur (Ω,a) telle que :

$$\forall\theta\in\Theta, P_\theta<<Q$$

H2 : Nous supposons que :

$$\forall\omega\in\Omega, \theta\mapsto L(\theta,\omega)$$

est continue sur Θ.

H3 : Nous supposons que $D=\Theta$, qui est une condition courante de régularité.

H4 : Nous supposons l'unicité de la solution du maximum de vraisemblance c'est à dire :

$$\Phi(\omega) = \{\theta : L(\theta, \omega) = \sup(L(x, \omega), x \in \Theta\} = \{\hat{\theta}(\omega)\}$$

H5 : Nous supposons que la mesure à priori sur (Θ, τ) est la mesure de Haar.

3.3 Théorème

Quelque soit la fonction de vraisemblance L vérifiant H2, et sous , les hypothèses, H1,H3,H4,H5, il existe une suite de fonctions de coût$(C_n)_{n\in N}$ *définies sur* $\Theta\times\Theta\times\Omega$ *à valeurs dans* $\mathbb{R}$*, vérifiant les conditions suivantes :*
i) bornées et invariantes par translation ;
ii)Continues sur $\Theta\times\Theta$;
iii)Mesurables surΩ
et telles que : $\lim_{n\to\infty} \tilde{\theta}_n = \hat{\theta}$ *simplement*, où $(\tilde{\theta}_n)_n$ est une suite de solutions bayésiennes.

3.4 Lemmes

Avant de démontrer ce résultat, nous devons donner une série de lemmes, c'est l'objet de la section suivante. Lemme 3.3.1 donne l'existence et la mesurabilité de la solution du maximum de vraisemblance.

3.4.1 lemme

Sous l'hypothèse H2, il existe une application $\hat{\theta}$ de (Ω, a) *dans*(Θ, τ) *mesurable telle que*

$$\forall \omega \in \Omega,\ L(\hat{\theta}(\omega), \omega) \geq L(\theta, \omega),, \forall \theta \in \Theta.$$

Démonstration

$$\forall \omega \in \Omega\ , \exists\ \hat{z}_\omega \in \Theta \quad telque\ L(\hat{z}_\omega, \omega) \geq L(\theta, \omega), \forall \theta \in \Theta$$

car $\forall \omega \in \Omega$,$\theta$: $\mapsto L(\theta,\omega)$ est continue sur le compact Θ .

Il reste à prouver que nous pouvons choisir une solution $\hat{z}_\omega$ mesurable en ω. Considérons l'application multivoque Φ suivante :

$$\begin{array}{ccl} \Omega & \rightarrow & \Theta \\ \omega & \mapsto & \Phi(\omega) = \{\theta : L(\theta,\omega) = M(\omega)\} \end{array}$$

où $M(\omega) = \sup\{(L(x,\omega),x \in \Theta\}$.

Pour tout $\omega \in \Omega, \Phi(\omega)$ est un fermé non vide de Θ .

Soit G un ouvert dans Θ .Dans un espace métrique, tout ouvert est réunion dénombrable de fermés $G = \cup_k F_k$ où F_k est fermé dans Θ .

Pour prouver que

$$\Phi^-(G) = \{\omega/\Phi(\omega) \cap G \neq \emptyset\} \in a$$

il suffit de prouver que :

$$\Phi^-(F) = \{\omega/\Phi(\omega) \cap F \neq \emptyset\} \in a \ \ pour\ tout\ fermé\ dans\ \Theta$$

car $\Phi^-(G) = \cup_k \Phi^-(F_k)$ (voir théorème.5.2).

F étant un fermé dans Θ compact, il est lui même compact .

$$\Phi^-(F) = \{\omega/\Phi(\omega) \cap F \neq \emptyset\} = \{\omega/\sup_{\theta \in F} L(\theta,\omega) = M(\omega)\} = \{\omega/M_F(\omega) = M(\omega)\}$$

où $\quad M_F(\omega) = \sup_{\theta \in F} L(\theta,\omega)$.

Soit $D = \{\theta_i \ / \ i \in \mathbb{N}\}$ un sous ensemble dénombrable et dense dans Θ .

$$\sup_{\theta \in F} L(\theta,\omega) = \sup_{i \in \mathbb{N}} L(\theta_i,\omega) .$$

Donc :

$$\omega \rightarrow M(\omega) = \sup_{\theta \in \Theta} L(\theta,\omega) = \sup_{i \in \mathbb{N}} L(\theta_i,\omega)$$

est mesurable (c'est un suprémum dénombrable de fonctions mesurables).

Il en est de même pour

$$M_F(\omega) = \sup_{\theta \in F} L(\theta,\omega) = \sup\{L(\theta,\omega), \theta \in D_F\}$$

où D_F est un sous ensemble dénombrable dense dans F. Par conséquent

$$\Phi^{-}(F) = \{\omega / M_F(\omega) = M(\omega)\} \in a$$

car M est M_F sont deux variables aléatoires définies sur (Ω, a) à valeurs dans un espace métrique séparable $\mathbb{R}^{+}$.

Nous avons donc $\Phi^{-}(G) \in a$ pour chaque ouvert G dans Θ. D'après le théorème 1.5.4, il existe une application $\hat{\theta}$ de (Ω, a) dans (Θ, τ) mesurable telle que :

$$\forall \omega \in \Omega, \overset{\wedge}{\theta}(\omega) \in \Phi(\omega)$$

c'est à dire :

$$\forall \omega \in \Omega, L(\overset{\wedge}{\theta}(\omega), \omega) \geq L(\theta, \omega), \forall \theta \in \Theta.$$

Le Lemme 3.3.2 est crucial pour la démonstration du Théorème 3.2. Il donne la construction de la fonction de coût .

3.4.2 Lemme

Sous les hypothèses H1, H2, H3, H4, H5, il existe une suite de fonction de coût $(C_n)_{n \in N}$ *définies* dans $\Theta \times \Theta \times \Omega$ *à valeurs dans* $\mathbb{R}$, *tel que* $\forall n$ *:*

i) C_n *est bornée et invariante par translation ;*

ii) C_n *est continue dans* $\Theta \times \Theta$ *;*

iii) C_n *est mesurable sur* Ω.

Démonstration D'aprés l'hypothèse H5 nous avons :

$$\sup\{L(x, \omega), x \notin \overset{\wedge}{\theta}(\omega) \cdot B(e, \varepsilon_n)\} < L(\overset{\wedge}{\theta}(\omega), \omega)$$

ce suprémum est atteint car $\Theta - \hat{\theta}(\omega) \cdot B(e, \varepsilon_n)$ est compact non vide .

Soit $\forall \omega \in \Omega$

$$e_n(\omega) = L(\overset{\wedge}{\theta}(\omega), \omega) - \sup\{L(x, \omega), x \notin \overset{\wedge}{\theta}(\omega) \cdot B(e, \varepsilon_n)\}.$$

$$\forall\omega \in \Omega,\ e_n(\omega) > 0$$

Il existe un voisinage $U_{n,\omega}$ de e, contenu dans $B\left(e, \varepsilon_n\right)$ tel que :

$\forall\theta \in \hat{\theta}(\omega) \cdot U_{n,\omega} \quad L(\theta,\omega) > [L(\hat{\theta}(\omega),\omega) - e_n(\omega)/2]$.

$\hat{\theta}(\omega) \cdot U_{n,\omega}$ peut être choisi comme la boule ouverte de centre $\hat{\theta}(\omega)$ et de rayon $\rho_n(\omega)$ où :

$$\rho_n\left(\omega\right) = \sup\{r{>}0 : d((\theta, \hat{\theta}(\omega)) < r \Rightarrow \quad L(\theta,\omega) > [L(\hat{\theta}(\omega),\omega) - e_n(\omega)/2]\}$$

$$= \sup\ \{r{>}0 : \inf_{\theta \in B\left(\hat{\theta}(\omega), r\right)} L\left(\theta,\omega\right) > \left[L\left(\overset{\wedge}{\theta}\left(\omega\right), \omega\right) - \frac{e_n\left(\omega\right)}{2}\right]\}.$$

D'aprés le lemme D'Urysohn, il existe une application $\gamma_{n,\omega}$ de Θ dans $[0,1]$ continue, à support dans $U_{n,\omega}$ et égal à 1 en e. Autrement dit nous avons :

$$\gamma_{n,\omega}\colon \Theta \to [0,1]$$

$$\theta \mapsto \gamma_{n,\omega}(\theta) = \begin{cases} 1 & si\ \theta{=}e \\ 0 & si\ \theta \notin U_{n,\omega} \end{cases}$$

où $U_{n,\omega}{=}B(e, \rho_n(\omega))$.

Prenons $\gamma_{n,\omega}(\theta) = \frac{\rho_n(\omega) - d(\theta,e)}{\rho_n(\omega)} \vee 0$. Cette application suggère la fonction de coût suivante :

$$C_n\colon \Theta{\times}\Theta{\times}\Omega \to [0,1]\ \ telle\ \ que\ (\theta,z,\omega) \mapsto C_n(\theta,z,\omega) = 1 - \gamma_{n,\omega}(z^{-1}\theta)$$

c'est à dire :

$$C_n(\theta,z,\omega) = 1 - \frac{\rho_n(\omega) - d(\theta,z)}{\rho_n(\omega)} \vee 0. \qquad (1.1)$$

Il est clair que C_n est continue en θ et z, bornée et invariante par translation c'est à dire :

$$C_n(x\theta, xz, \omega) = C_n(\theta,z,\omega), \forall x \in \Theta$$

Puisque nous sommes amenés à évaluer les quantités telles que :

$$\int C_n(\theta,z,\omega) \quad L(\theta,\omega)\xi(\theta)d\omega$$

Il est important que $C_n(.,.,\omega)$ soit mesurable en ω .

Ainsi prouver que $C_n(.,.,\omega)$ est mesurable en ω, revient à prouver la mesurabilité de $\rho_n(\omega)$ par rapport ω, puisque C_n (θ, z, ω) est écrite sous la forme (1.1) .

Prouvons donc que l'application ρ_n de (Ω, a) dans $\left(\mathbb{R}_+^*, \mathcal{B}_{\mathbb{R}_+^*}\right)$ définie comme suit :

$$\rho_n(\omega) = \sup\{r > 0 : \inf_{\theta \in B\left(\hat{\theta}(\omega), r\right)} (L(\theta, \omega)) > [L(\overset{\wedge}{\theta}(\omega),\omega) - e_n(\omega)/2]\}$$

est mesurable.

Posons

$X(s, \omega) = \inf_{\theta \in B(\hat{\theta}(\omega), s)} (L(\theta, \omega))$.

Pour montrer que $\omega \to \rho_n(\omega)$ est mesurable, nous devons prouver que :

$$\forall s \in \mathbb{R}_+^*, \quad \{\omega{:}\rho_n(\omega) > s\} \quad \in a.$$

Nous remarquons que :

$$\{\omega{:}\rho_n(\omega) > s\} = \bigcap_{s_i < s} \left\{\omega{:}X(s_i, \omega) > [L(\overset{\wedge}{\theta}(\omega),\omega) - e_n(\omega)/2], s_i \in \mathbb{Q}\right\}$$

Le problème revient ainsi à montrer que les deux applications suivantes :

$$\omega \to e_n(\omega) = L(\overset{\wedge}{\theta}(\omega),\omega) - \sup\{L(x, \omega), x \notin \theta(\omega) \cdot B(e, \varepsilon_n)\}$$

$$\omega \to X(s, \omega) = \inf_{\theta \in B(\hat{\theta}(\omega), s)} (L(\theta, \omega))$$

sont mesurables. La preuve de la mesurabilité de la première application est établie dans la section 3.5.

Montrons que pour s fixé $\omega \to X(s, \omega)$ est mesurable.

Soit Θ^o un sous ensemble dénombrable dense dans $B(e, s)$. L'ensemble :

$\{\hat{\theta}(\omega)\theta_i/\theta_i \in \Theta^o\}$ est dense dans $\hat{\theta}(\omega)B(e, s) = B(\hat{\theta}(\omega), s)$. L' application suivante :

$$\Phi_i(\omega) : \omega \to (\overset{\wedge}{\theta}(\omega)\theta_i, \omega) \to L(\overset{\wedge}{\theta}(\omega)\theta_i, \omega) \textit{ est mesurable.}$$

Par conséquent

$$\omega \to \inf_i \Phi_i(\omega) = \inf_{\theta \in B(\hat{\theta}(\omega),s)} (L(\theta, \omega))$$

est mesurable.

Rappelons que $\tilde{\theta}_n$ est un estimateur de Bayes (par rapport à μ) si :

$$r\left(\mu, \tilde{\theta_n}\right) = \inf_{\delta} r\left(\mu, \delta\right) = \inf_{\delta} \int_{\Theta} [\int_{\Omega} C_n(\theta, \delta(\omega), \omega) L(\theta, \omega) Q(d\omega)] \mu(d\theta)$$

$$= \inf_{\delta} \int_{\Omega} [\int_{\Theta} C_n(\theta, \delta(\omega), \omega) L(\theta, \omega) \mu(d\theta)] Q(d\omega).$$

D'après le théorème de Fubini puisque

$$\inf_{\delta} \int_{\Theta} [\int_{\Omega} C_n(\theta, \delta(\omega), \omega) L(\theta, \omega) Q(d\omega)] \mu(d\theta) \ < \ \infty$$

Minimiser le risque de Bayes $r(\mu, \delta)$ revient à minimiser :

$$\int_{\Theta} C_n(\theta, d, \omega) L(\theta, \omega) \mu(d\theta)$$

pour ω fixé $\delta(\omega) = d \in \Theta$.

Le lemme suivant donne l'existence et la mesurabilité de la solution bayésienne $\tilde{\theta}_n$.

3.4.3 Lemme

Sous les hypothèses H, H2, ..., H5, il existe une application mesurable $\tilde{\theta}_n$ de(Ω, a) dans (Θ, τ) mesurable telle que :

$$\forall \omega \in \Omega, \exists\, d \in \Theta, \int_{\Theta} C_n(\theta, \widetilde{\theta}_n\left(\omega\right), \omega) L(\theta, \omega) \mu(d\theta) \leq \int_{\Theta} C_n(\theta, d, \omega) L(\theta, \omega) \mu(d\theta)$$

Démonstration

Il est clair que minimiser :

$$\int_{\Theta} C_n(\theta,\ d,\ \omega)\ \ L(\theta,\omega)\mu(d\theta)$$

revient , par définition de C_n , à maximiser :

$$\int_{\Theta} \gamma_{n,\omega}(d^{-1}\theta)\ \ L(\theta,\omega)\mu(d\theta).$$

Le maximum de cette dernière intégrale est atteint car :

$$\forall\omega \in \Omega, d \to \int_{\Theta} \gamma_{n,\omega}(d^{-1}\theta)\ \ L(\theta,\omega)\mu(d\theta)$$

est continue dans le compact Θ.

Par conséquent :

$$\forall\omega \in \Omega, \exists \widetilde{d_{\omega}}^{n} \in \Theta \ \ \text{tel que :}$$

$$\int_{\Theta} \gamma_{n,\omega}(\widetilde{d_{\omega}}^{n-1}\theta)\ \ L(\theta,\omega)\mu(d\theta) \geq \int_{\Theta} \gamma_{n,\omega}(x^{-1}\theta)\ \ L(\theta,\omega)\mu(d\theta), \forall x \in \Theta$$

Il reste à prouver que nous pouvons choisir un $\tilde{d}_{\omega}^{n}$ mesurable en ω.

Pour cela considérons l'application multivoque suivante Φ_n :

$$\begin{array}{lcl} \Omega & \to & \Theta \\ \omega & \mapsto & \Phi_n(\omega) = \{d : \int_{\Theta} \gamma_{n,\omega}(d^{-1}\theta) L(\theta,\omega)\mu(d\theta) = g_n(\omega)\} \end{array}$$

où

$$g_n(\omega) = \sup_{x\in\Theta} \int_{\Theta} \gamma_{n,\omega}(x^{-1}\theta)\ \ L(\theta,\omega)\mu(d\theta).$$

Nous vérifions les conditions du Théorème 1.5.4 de la même manière que pour l'application Φ (Lemme 3.3.1). Nous pouvons affirmer qu'il existe une application mesurable $\tilde{\theta}_n$ de (Ω, a) dans (Θ, τ) telle que :

$$\forall\omega \in \Omega\ \ , \widetilde{\theta}_n(\omega)\ \in \Phi_n(\omega).$$

3.5 Démonstration du Théorème 3.2

L'existence de la suite de fonctions de coût $(C_n)_{n\in\mathbb{N}}$ est établie dans le Lemme 3.3.2. L'existence et la mesurabilité des solutions bayésiennes $(\tilde{\theta}_n)_{n\in\mathbb{N}}$ et la solution du maximum de vraisemblance $\hat{\theta}$ sont établies dans le Lemme 3.3.3 et le Lemme 3.3.1.

Il reste à montrer que $\lim_{n\to\infty} \tilde{\theta}_n = \hat{\theta}$ simplement. Rappelons la définition de $\Phi_n(\omega)$ et $\Phi(\omega)$

$$\Phi_n(\omega) =\{d\colon \int_\Theta \gamma_{n,\omega}(d^{-1}\theta)L(\theta,\omega)\mu(d\theta) = \sup_{x\in\Theta}\int_\Theta \gamma_{n,\omega}(x^{-1}\theta)L(\theta,\omega)\mu(d\theta)\}$$

$$\Phi\,(\omega) =\{\theta\colon L(\theta,\omega) = \sup\{L(x,\omega), x\in\Theta\} =\{\overset{\wedge}{\theta}\,(\omega)\}$$

et supposons qu'il existe :

$$g_n \in \Phi_n\,(\omega) \ \ tel\ que\ g_n B(e,\rho_n(\omega)) \cap \overset{\wedge}{\theta}\,(\omega)\, B(e,\varepsilon_n) = \emptyset.$$

Pour ces considérations nous avons :

a)D'une part :

$$\int_\Theta \gamma_{n,\omega}(g_n^{-1}\theta) \ \ L(\theta,\omega)\mu(d\theta) = \int_{g_n B(e,\rho_n(\omega))} \gamma_{n,\omega}(g_n^{-1}\theta) \ \ L(\theta,\omega)\mu(d\theta)$$

car la fonction $\gamma_{n,\omega}$ est à support dans $B(e,\rho_n(\omega))$, et nous avons

$$\int_{g_n B(e,\rho_n(\omega))} \gamma_{n,\omega}(g_n^{-1}\theta)L(\theta,\omega)\mu(d\theta) \leq L(\overset{\wedge}{\theta}\,(\omega)\,,\omega) - e_n\,(\omega)\,] \int_{g_n B(e,\rho_n(\omega))} \gamma_{n,\omega}(g_n^{-1}\theta)\mu(d\theta)$$

car

$$\sup\{L(\theta,\omega), \theta \notin (\overset{\wedge}{\theta}\,(\omega) \cdot B\,(e,\varepsilon_n)\} = [L(\overset{\wedge}{\theta}\,(\omega)\,,\omega) - e_n\,(\omega)\,]$$

b) D'autre part :

$$\int_{(\,\hat{\theta}(\omega)B(e,\rho_n(\omega))} \gamma_{n,\omega}((\overset{\wedge}{\theta}\,(\omega)\,)^{-1}\theta) \ \ L(\theta,\omega)\mu(d\theta) \dot{\geq}$$

$$L(\overset{\wedge}{\theta}(\omega),\omega){-}e_n(\omega)/2]\int_{(\hat{\theta}(\omega)B(e,\rho_n(\omega))}\gamma_{n,\omega}((\overset{\wedge}{\theta}(\omega))^{-1}\theta)\mu(d\theta)$$

car

$$\forall\theta\in(\overset{\wedge}{\theta}(\omega)\,B(e,\rho_n(\omega)),\Rightarrow L(\theta,\omega)>L(\overset{\wedge}{\theta}(\omega),\omega){-}e_n(\omega)/2$$

$$>[L(\overset{\wedge}{\theta}(\omega),\omega){-}e_n(\omega)]\int_{(\hat{\theta}(\omega)B(e,\rho_n(\omega))}\gamma_{n,\omega}((\overset{\wedge}{\theta}(\omega))^{-1}\theta)\mu(d\theta)$$

Notons que :

$$\int_{(\hat{\theta}(\omega)B(e,\rho_n(\omega))}\gamma_{n,\omega}((\overset{\wedge}{\theta}(\omega))^{-1}\theta)\mu(d\theta))=\int_{\Theta}\gamma_{n,\omega}((\hat{\theta}(\omega))^{-1}\theta)\mu(d\theta)=\int_{\Theta}\gamma_{n,\omega}(\theta)\mu(d\theta)$$

et aussi :

$$\int_{g_nB(e,\rho_n(\omega))}\gamma_{n,\omega}(g_n^{-1}\theta)\mu(d\theta)=\int_{\Theta}\gamma_{n,\omega}(g_n^{-1}\theta)\mu(d\theta{=}\int_{\Theta}\gamma_{n,\omega}(\theta)\mu(d\theta)$$

Donc nous avons :

$$\int_{\hat{\theta}(\omega)B(e,\rho_n(\omega))}\gamma_{n,\omega}(((\hat{\theta}(\omega))^{-1}\theta))L(\theta,\omega)\mu(d\theta)>[L(\hat{\theta}(\omega),\omega){-}e_n(\omega)]\int_{\Theta}\gamma_{n,\omega}(\theta)\mu(d\theta).$$

et :

$$\int_{g_nB(e,\rho_n(\omega))}\gamma_{n,\omega}(g_n^{-1}\theta)L(\theta,\omega)\mu(d\theta)\leq L(\overset{\wedge}{\theta}(\omega),\omega){-}e_n(\omega)]\int_{\Theta}\gamma_{n,\omega}(\theta)\mu(d\theta)$$

Donc

$$\int_{\theta^{\wedge}(\omega)B(e,\rho_n(\omega))}\gamma_{n,\omega}((\overset{\wedge}{\theta}(\omega))^{-1}\theta))L(\theta,\omega)\mu(d\theta)>\int_{g_nB(e,\rho_n(\omega))}\gamma_{n,\omega}(g_n^{-1}\theta)L(\theta,\omega)\mu(d\theta)$$

Ce qui contredit la définition de g_n. Nous avons nécessairement :

$$\forall g_n\in\Phi_n(\omega):g_nB(e,\rho_n(\omega))\cap\overset{\wedge}{\theta}(\omega)\,B(e,\varepsilon_n)\neq\emptyset\;\Rightarrow$$

$$\forall g_n\in\Phi_n(\omega):(\overset{\wedge}{\theta}(\omega))^{-1}g_n\in B(e,\varepsilon_n)\cdot B(e,\rho_n)^{-1}\subset B(e,\varepsilon_n)\cdot B(e,\varepsilon_n)\subset B(e,\varepsilon_{n-1})$$

$$\Rightarrow\;\forall g_n\in\Phi_n(\omega):(\overset{\wedge}{\theta}(\omega))^{-1}g_n\in B(e,\varepsilon_{n-1})$$

$$\Rightarrow \ \forall g_n \in \Phi_n(\omega) : \lim_{n\to\infty} g_n = \overset{\wedge}{\theta}(\omega)$$

Comme $\theta^{\sim}{}_n(\omega) \in \Phi_n(\omega)$. Nous avons $\lim_{n\to\infty} \theta^{\sim}{}_n(\omega) = \theta^{\wedge}(\omega)$.
Par conséquent

$$\lim_{n\to\infty} \overset{\sim}{\theta}_n = \overset{\wedge}{\theta} \ \textit{simplement}$$

3.5.1 Remarques

1. Le résultat reste valable pour une suite quelconque de fonctions $\ddot{\theta}_n$ de Ω dans Θ non nécessairement mesurables telle que :

$$(\ddot{\theta}_n)_n(\omega) \in \Phi_n(\omega)$$

2. Nous démontrons que $\hat{\theta}$ est mesurable, car elle est la limite simple de fonctions mesurables.

3.6 Comparaison asymptotique de $\tilde{\theta}_n$ et $\hat{\theta}$ quand $\hat{\theta}$ n'est pas unique

Dans cette section , nous généralisons le Théorème 3.2., dans le sens où nous n'envisagerons plus l' hypothèse H4.

3.6.1 Théorème

Quelque soit la fonction de L vérifiant H2, et sous les hypothèses H1, H3, H4, H5, il existe : une suite de fonctions de coût $(C_n)_{n\in N}$ *définies dans* $\Theta\times\Omega$ *à valeurs dans* $\mathbb{R}$*, tel que* $\forall n$ *:*

i) C_n *est bornée et invariante par translation.*;

ii) C_n *est continue sur* $\Theta\times\ \Theta$;

iii) C_n *est mesurable sur* Ω*.*

et telle que : $\lim_{n\to\infty} \hat{\theta}_n$ *(si elle existe) est un estimateur du maximum de vraisemblance.*

Démonstration

La construction de la fonction de coût est établie en utilisant la même technique du **(Lemme 3.3.2)**. Nous posons :

$$\Phi(\omega) = \{\theta : L(\theta,\omega) = M(\omega)\} \; ou \; M(\omega) = \sup\{L(x,\omega), x \in \Theta\}$$

$$e_n(\omega) = M(\omega) - \sup\{L(\theta,\omega), \theta \notin \Phi(\omega) B(e,\varepsilon_n)\}$$

$$\forall \omega \in \Omega \; e_n(\omega) > 0$$

$$Z_n(\omega) = \inf_{m>1} \sup\{L(\theta,\omega), \theta \notin \Phi(\omega) B^\star(e, \varepsilon_n(1-1/m))\}.$$

$$\rho_n(\omega) = \sup\left\{ r>0 : \inf_{\theta \in \Phi(\omega) B(e,r)} L(\theta,\omega) > M(\omega) - e_n(\omega)/2 \right\} \wedge \varepsilon_n$$

Il existe une application $\gamma_{n,\omega}$ de Θ dans $[0,1]$ continue, à support dans $B(e,\rho_n(\omega))$, et est égal à 1 en e. Autrement dit nous avons :

$$\gamma_{n,\omega} : \Theta \to [0,1]$$

$$\theta \mapsto \gamma_{n,\omega}(\theta) = \begin{cases} 1 & si \theta = e \\ 0 & si \theta \notin B(e,\rho_n(\omega)) \end{cases}$$

Considérons

$$\gamma_{n,\omega}(\theta) = \frac{\rho_n(\omega) - d(\theta,e)}{\rho_n(\omega)} \vee 0$$

Cette application suggère la fonction de coût suivante :

$$C_n : \Theta \times \Theta \times \Omega \mapsto [0,1]$$

$$(\theta,z,\omega) \mapsto C_n(\theta,z,\omega) = 1 - \gamma_{n,\omega}(z^{-1}\theta)$$

c'est à dire

$$C_n(\theta, \quad z, \; \omega) = 1 - \frac{\rho_n(\omega) - d(\theta,e)}{\rho_n(\omega)} \vee 0.$$

Il est clair $\forall n \in \mathbb{N}$, C_n est continue sur θ et z. C_n bornée, invariante par translation, c'est à dire :

$$C_n(x\theta, xz, \omega) = C_n(\theta, z, \omega), \forall x \in \Theta$$

La mesurabilité de C_n sur ω est déduite des deux applications suivantes :

$$\omega \mapsto e_n(\omega) = M(\omega) - \sup\{L(\theta,\omega), \theta \notin \Phi(\omega)B(e,\varepsilon_n)\}$$

$$\omega \mapsto \inf\{L(\theta,\omega), \theta \in \Phi(\omega)B(e,r)\}$$

Clarifions ceci séparément dans les problème 1 et 2. En effet :

Problème1.

Nous devons prouver que $\omega \mapsto e_n(\omega)$ est mesurable. Nous commençons par :

$$\omega \mapsto M(\omega) = \sup\{L(\theta,\omega), \theta \in \Theta\}$$

est mesurable. Soit $D = \{\theta_i, i \in \mathbb{N}\}$ un sous ensemble dénombrable dense dans Θ

$$\omega \mapsto M(\omega) = \sup_i L(\theta_i, \omega)$$

est donc mesurable.

Montrons maintenant que : $\omega \mapsto \sup\{L(\theta,\omega), \theta \notin \Phi(\omega)B(e,\varepsilon_n)\}$ est mesurable.

a)Nous avons d'une part : pour ω fixé et n fixé :

$$\sup\{L(\theta,\omega), \theta \notin \Phi(\omega)B(e,\varepsilon_n)\} = Z_n(\omega) \quad (3.5.a)$$

La preuve de(3.5.a) se trouve dans l'appendice .

b)Deuxièmement : Soit Ψ :

$$\begin{array}{ccc} \Theta \times \Omega & \to & \mathbb{R} \\ (\theta,\omega) & \mapsto & d(\theta), \Phi(\omega)) \end{array}$$

est mesurable. La preuve de la de Ψ est traitée dans l'appendice. Donc :

$$\forall m > 1, B_m = \Psi^{-1}[0, \varepsilon_n(1-\frac{1}{m})] = \{(\theta,\omega) : d(\theta,\Phi(\omega))\} \leq \varepsilon_n(1-\frac{1}{m})\} \in \tau \times a.$$

Modifions L sur B_m

$$\widetilde{L}_m(\theta,\omega) = \begin{cases} L(\theta,\omega) \text{ si } (\theta,\omega) \in B_m. \\ -\infty \; sinon \end{cases}$$

$\forall m>1$, $\tilde{L}_m$ est mesurable sur $\Theta\times\Omega$. Soit $A_m^\omega = \Theta - \Phi(\omega)B^*(e, \varepsilon_n(1-1/m))$, A_m^ω est ouvert dans Θ .

$A_m^\omega \cap \; D$ est dense dans A_m^ω. Par conséquent, $\omega \to Z_n(\omega)$

$$= \inf_{m>1} \sup_{\theta \in A_m^\omega} L(\theta,\omega)) = \inf_{m>1} \sup_{\theta \in A_m^\omega \cap D} L(\theta,\omega))$$

$$= \inf_{m>1} (\underset{\theta\in D}{s} up \tilde{L}_m(\theta,\omega))$$

est mesurable sur Ω.

Ainsi :

$$\omega \to e_n(\omega) = M(\omega) - Z_n(\omega)$$

est mesurable sur Ω. Ce qui achève la démonstration du Problème1 .

Problème2 .

Montrons que, pour r fixé :

$$\omega \mapsto \inf\{L(\theta,\omega), \theta \in \Phi(\omega)B(e,r)\}$$

est mesurable. Soit :

$$B = \Psi^{-1}(0,r[) = \{ (\theta,\omega) : d(\theta, \Phi(\omega)) < r\} \in \tau \times a$$

Modifions L sur B :

$$\hat{L}_m(\theta,\omega) = \begin{cases} L(\theta,\omega) \text{ si } (\theta,\omega) \in B \\ +\infty \; sinon \end{cases}$$

$\hat{L}$ est mesurable sur $\Theta{\times}\Omega$. Et

$$\omega \mapsto \inf\{L(\theta,\omega),\theta \in \Phi(\omega)B\,(e,r)\}= \inf\{L(\theta,\omega),\theta \in \Phi(\omega)B\,(e,r) \cap D\}$$

$$= \inf\{\hat{L}\,(\theta,\omega),\theta \in \mathcal{D}\}\ \textit{est mesurable.}$$

Nous avons pu construire une suite de fonctions de coût, qui vérifie les conditions du théorème. Montrons enfin que pour ce type de fonctions de coût, $\lim_{n\to\infty}\tilde{\theta}_n$ si elle existe est un estimateur du maximum de vraisemblance.

Supposons qu'il existe $g_n \in \Phi_n\,(\omega)$ tel que

$$g_n B\,(e,\rho_n\,(\omega)) \cap \Phi\,(\omega)\,B\,(e,\varepsilon_n) = \emptyset$$

Dans ces considérations nous avons :

$$\int_{g_n B(e,\rho_n(\omega))} \gamma_{n,\omega}(g_n^{-1}\theta)\mathrm{L}(\theta,\omega)\mu(d\theta) \leq (M\,(\omega) - e_n\,(\omega)\,)\int_{\Theta}\gamma_{n,\omega}\,(\theta)\quad \mu(d\theta)$$

et $\forall f \in \Phi\,(\omega)$:

$$\int_{fB(e,\rho_n(\omega))} \gamma_{n,\omega}(f^{-1}\theta)\quad L(\theta,\omega)\mu(d\theta) \geq [M\,(\omega) - e_n\,(\omega)\,/2]\int_{\Theta}\gamma_{n,\omega}(\theta)\quad \mu(d\theta)$$

$$>[M\,(\omega) - e_n\,(\omega)\,]\int_{\Theta}\gamma_{n,\omega}(\theta)\quad \mu(d\theta)$$

Ceci donne :

$$\forall f \in \Phi\,(\omega)\,, \int_{fB(e,\rho_n(\omega))} \gamma_{n,\omega}(f^{-1}\theta)\quad L(\theta,\omega)\mu(d\theta) > \int_{g_n B(e,\rho_n(\omega))} \gamma_{n,\omega}(g_n^{-1}\theta)(\theta,\omega)\mu(d\theta)$$

Ceci contredit la définition de g_n $(g_n \in \Phi_n\,(\omega))$.

Nous avons forcément :

$$\forall g_n \in \Phi_n\,(\omega)\,, g_n B\,(e,\rho_n\,(\omega)) \cap \Phi\,(\omega)\,B\,(e,\varepsilon_n) \neq \emptyset \Rightarrow$$

$$\forall g_n \in \Phi_n\,(\omega) : \lim_{n\to\infty} g_n \quad \text{(si elle existe)} \in \Phi\,(\omega)$$

Nous savons qu'il existe $\tilde{\theta}_n$ de (Ω, a) dans (Θ, τ) mesurable tel que $\tilde{\theta}_n(\omega) \in \Phi_n(\omega)$. D'après ce qui précède la limite de$\lim_{n\to\infty} \tilde{\theta}_n(\omega)$ si elle existe, appartient à $\Phi(\omega)$.Comme $\lim_{n\to\infty} \tilde{\theta}_n(\omega)$ est encore une mesurable application de (Ω, a) dans (Θ, τ), alors la limite $\lim_{n\to\infty} \tilde{\theta}_n$ si elle existe, est un estimateur du maximum de vraisemblance.

Annexe A

Appendice

A.1 Lemme

(**Preuve de (3.2.a)**) *Soit* Θ *un espace métrique compact*, A est *un sous ensemble compact de* Θ, f *une application continue* de Θ dans $\mathbb{R}$.
Posons :

$$F = \{\theta : d(\theta, A) \geq \varepsilon_n\}$$

$$G_m = \{\theta : d(\theta, A) > \varepsilon_n(1 - \frac{1}{m})\}$$

$$\tilde{G}_m = \{\theta : d(\theta, A) \geq \varepsilon_n(1 - \frac{1}{m})\}$$

Alors :

$$\sup_{x \in F} f(x) = \inf_{m>1} (\sup_{x \in G_m} f(x)) = \inf_{m>1} (\sup_{\tilde{G}_m} f(x))$$

Démonstration

$$\inf_{m>1} (\sup_{x \in G_m} f(x)) = \inf_{m>1} (\sup_{\tilde{G}_m} f(x)) \;\; car \;\; G_{m+1} \subset \tilde{G}_{m+1} \subset G_m.$$

Puis selon F considérons les deux cas $F \neq \emptyset$ et $F = \emptyset$. En effet :

1. Si $F \neq \emptyset$, donc

$$F = \bigcap_m G_m = \bigcap_m \widetilde{G}_m$$

Nous avons d'une part :

$$\forall m>1, \sup_{x\in \tilde{G}_m} f(x) \geq \sup_{x\in F} f\,(x) \quad \Rightarrow$$

$$\forall m>1, \inf_{m>1} \left(\sup_{\tilde{G}_m} f(x) \right) \geq \sup_{F} f\,(x) \tag{1}$$

b)D'autre part :

$$\sup_{\tilde{G}_m} f(x) = f(x_m), x_m \in \widetilde{G}_m.$$

Nous pouvons supposer que la suite $(x_m)_m$ converge vers x.

$$d(x_m, A) \geq \varepsilon_n\,(1 - 1/m)$$

$$d(x_m, A) \to d\,(x, A)\,, \quad m \to \infty.$$

Alors

$$d\,(x, A) \geq \varepsilon_n \Rightarrow x \in F.$$

$$\sup_{\tilde{G}_m} f(x) = f(x_m) \to f\,(x) \quad ,m \to \infty$$

puisque f est continue.

$$\Rightarrow \inf_{m>1} (\, \sup_{\tilde{G}_m} f(x) = \inf_{m>1} (f\,(x_m)\,) = f\,(x)$$

car $(f\,(x_m)\,)_m$ est décroissante. Mais

$$f\,(x) \leq \sup_{F} f\,(x) \;\; (x \in F) \tag{2}$$

$$(1)\,\text{et} \quad (2) \Rightarrow \sup_{x\in F} f(x) = \inf_{m>1} (\sup_{x\in G_m} f(x)) = \inf_{m>1} (\sup_{\tilde{G}_m} f(x))$$

2. Si $F = \emptyset$, il existe nécessairement un m tel que $G_m = \emptyset$ .En effet : $(\tilde{G}_m)_m$ est une suite décroissante de compacts. Si $\tilde{G}_m$est non vide, donc $\cap\tilde{G}_m \neq \emptyset$ (propriété de la compacité).

Mais pour :

$$x_m \in \tilde{G}_m,\ \forall m > 1$$

nous avons

$$d\left(x, A\right) \geq \varepsilon_n \left(1 - 1/m\right), \forall m \succ 1$$

par conséquent

$$d\left(x, A\right) \geq \varepsilon_n \Rightarrow x \in F \textit{ ceci contredit l'hypothèse.}$$

Par conséquent

$$F = \bigcap_m G_m = \bigcap_m \widetilde{G}_m = \emptyset.$$

et

$$\sup_{x \in F} f(x) = \inf_{m>1} \left(\sup_{x \in G_m} f(x) \right) = \inf_{m>1} \left(\sup_{\tilde{G}_m} f(x) \right) = -\infty$$

A.2 Preuve de la mesurabilité Ψ

Soit Ψ :

$$\begin{array}{ccc} \Theta \times \Omega & \to & \mathbb{R} \\ (\theta, \omega) & \mapsto & d(\theta, \Phi(\omega)) \end{array}$$

Considérons $\mathcal{K}\left(\Theta\right)$une famille de compact de Θ, munie de la topologie de Hausdorff , dont une sous-base est $g \cup g'$, où g et $g^{'}$ sont définies par :

$$g \text{ ouvert} \overset{def}{=} \left\{K \in \mathcal{K}\left(\Theta\right) \ \ / \ \ K \cap G \neq \emptyset, G \text{ouvert dans} \, \Theta\right\}$$

$$g^{'} \text{ ouvert} \overset{def}{=} \left\{K \in \mathcal{K}\left(\Theta\right) \ \ / \ \ K \subset G^{'}, G^{'} \text{ ouvert dans} \Theta\right\}$$

Soient A et B deux fermés non vides de Θ

$$\delta\left(A, B\right) = \max\left\{\sup_{x \in A} d\left(x, B\right), \sup_{y \in B} d\left(y, A\right)\right\}$$

définit une distance sur $\mathcal{K}\left(\Theta\right)$, appelée la distance de Hausdorff.

La topologie induite par δ est celle définie plus haut (on rappelle Θ est un espace métrique) voir [4].

Montrons alors que l'application définie ci-dessous est mesurable : On définit Φ :

$$\begin{array}{ccl} \Omega & \to & \mathcal{K}(\Theta) \\ \omega & \mapsto & \Phi(\omega) = \{\theta : L(\theta,\omega) = M(\omega)\} \in \mathcal{K}(\Theta)\} \end{array}$$

$$\Phi^{-1}\left\{K/K \cap F\right\} = \left\{\omega{:}\Phi\left(\omega\right) \cap F)\right\} \in a$$

(Voir la démonstration du **Lemme 3.3.1.**

$$A^c = \left\{\omega{:}\Phi(\omega) \cap F^c \neq \emptyset\right\} = \bigcup_n \left\{\omega{:}\Phi\left(\omega\right) \cap F_n \neq \emptyset)\right\}$$

Car $F^c = \cup_n F_n$ où F_n est fermé dans Θ.

$$\Rightarrow A^c \in a, \text{car} \left\{\omega{:}\Phi\left(\omega\right) \cap F_n \neq \emptyset)\right\} \in a, \forall n$$

$$\Rightarrow \quad A = \Phi^{-1}\left\{K/K \subset F\right\} \in a.$$

Par conséquent Φ est mesurable.

Montrons maintenant que l'application suivante d :

$$\begin{array}{lcl} \mathcal{K}(\times) \times \Theta & \to & \mathcal{R} \\ (K,\theta) & \mapsto & d(K,\theta) \end{array}$$

est continue (uniformément).

Pour cela définissons sur $K(\Theta) \times \Theta$ la distance suivante :

$$d^{'}\left((K,\theta), \left(K^{'},\theta^{'}\right)\right) = d\left(\theta,\theta^{'}\right) + \delta\left(K,K^{'}\right).$$

$$\left|d\left(\theta,K\right) - d\left(K^{'},\theta^{'}\right)\right| \leq \left|d\left(\theta,K\right) - d\left(\theta^{'},K\right)\right| + \left|d\left(\theta^{'},K\right) - d\left(K^{'},\theta^{'}\right)\right|$$

$$\left|d(\theta,K) - d\left(\theta^{'},K\right)\right| \leq d\left(\theta,\theta^{'}\right).$$

Soient $x_1 \in K, x_2 \in K^{'}$ tels que $d\left(\theta^{'},x_1\right) = d\left(\theta^{'},K\right)$ et $d\left(\theta^{'},x_2\right) = d\left(\theta^{'},K^{'}\right)$

Soient $x_1^{'} \in K$ tels que $d\left(K, x_2\right) = d\left(x_2, x_1^{'}\right)$. Nous avons :

$$
\begin{aligned}
d\left(\theta^{'}, K\right) - d\left(\theta^{'}, K^{'}\right) &= d\left(\theta^{'}, x_1\right) - d\left(\theta^{'}, x_2\right) \\
&\leq d\left(\theta^{'}, x_1^{'}\right) - d\left(\theta^{'}, x_2\right) \\
&\leq d(x_1^{'}, x_2) = d\left(x_2, K\right) \\
&\leq \sup_{x \in K^{'}} d\left(x, K\right) \leq \delta\left(K, K^{'}\right)
\end{aligned}
$$

Idem pour :

$$
\begin{aligned}
d\left(\theta^{'}, K^{'}\right) - d\left(\theta^{'}, K\right) &\leq \delta\left(K, K^{'}\right) \\
\Rightarrow \quad \left| d\left(\theta^{'}, K\right) - d\left(\theta^{'}, K'\right) \right| &\leq \delta\left(K, K^{'}\right)
\end{aligned}
$$

Par conséquent :

$$
\left| d\left(\theta, K\right) - d\left(K^{'}, \theta^{'}\right) \right| \leq d\left(\theta, \theta^{'}\right) + \delta\left(K, K^{'}\right) = d^{'}\left(\left(K, \theta\right), \left(K^{'}, \theta^{'}\right)\right)
$$

Ceci prouve l'uniforme continuité de :

$$
(K, \theta) \mapsto d(K, \theta)
$$

Donc finalement l'application suivante :

$$
\Psi: \qquad \left(\theta, \omega\right) \in \Theta \times \Omega \mapsto \left(\theta, \Phi\left(\omega\right)\right) \in \Theta \times \mathcal{K}\left(\Theta\right) \mapsto d\left(\theta, \Phi\left(\omega\right)\right) \in \mathbb{R}
$$

est mesurable, comme composée de deux applications mesurables.

A.3 Conclusion

Dans le problème de décision bayésienne classique, la fonction de coût est une donnée intrinsèque du problème et l'attention se focalise sur la mesure à priori.

Nous procédons dans notre travail autrement, on fixe la mesure à priori c'est la mesure de Haar, et c'est la fonction de coût qui est choisie et construite à partir du Lemme d'Urysohn.

On peut généraliser ce résultat à un espace des paramètres Θ localement compact. On peut le compactifier par un point à l'infini et on se ramène à un espace compact.

1. **References**

[1] J.R,Barra, Notions Fondamentales de Statistique Mathématique, *Dunod,Paris*, 1971.

[2] C.Berge, Espaces topologiques.Fonctions Multivoques, *Dunod,Paris*, 1959.

[3] J.O.Berger , Statistical Decision Theory, *Spring-verlag Dunod,New-york*, 1980.

[4] CHRISTENSEN, Topology and Borel structure, *North Holland, Amesterdam*, 1974.

[5] T.S.Fergusson, Mathematical Statistics : A decision -Theoric-Approch, *Academic Press,New-york*, 1967.

[6] J.P.Florens, Mesures à priori et invariance dans une expérience bayesienne , *Pub.I.S.U.P Spring-Verlag, New-york*, 13, fasc 1-2, p :29-55, 1978.

[7] C.Fougereaud, A.Fuchs, Statistique, *Dunod,Paris*, 1967.

[8] P.R.Halmos, Measure Theory , *Spring-Verlag, New-york*, 1994.

[9] H.Jeffreys, Theory of Probability, *Oxford University Press, London*, 3d.ed, 1961.

[10] K.Kuratowski, C.Ryll-Nardewski, A general theorem on selectors, *Bull.Acad.Sci.Polon.*, t 13, p :397-403, 1965.

[11] E.Lanery, Solutions bayesiennes sans désintégration, *Université Paris 6 Dauphine, Paris*, 1984.

[12] L.Lecam, On some asymptotic properties of maximum likelihood estimates and related estimates, *Univ of California in stat.*vol 1. p :277-330, 1953.

[13] P.A.Meyer, Probabilités et Potentiels, *Hermann, Paris.*, 1966.

[14] J.Ulmo, J.Bernier, Eléments de décision, *Puf,Paris.*,1973.

[15] A.Wald, Statistical Decision Function, *J.Wiley and Sons, New-York.*, 1950.