ANNALES DE L'I. H. P., SECTION B

FRANÇOISE MARTIN
DANIEL VAGUELSY

Introduction à la statistique mathématique. IV. Propriétés asymptotiques du modèle statistique

Annales de l'I. H. P., section B, tome 5, n° 4 (1969), p. 357-384
<http://www.numdam.org/item?id=AIHPB_1969__5_4_357_0>

Ann. Inst. Henri Poincaré, Section B :
Vol. V, n° 4, 1969, p. 357-384. *Calcul des Probabilités et Statistique.*

Introduction à la statistique mathématique.

IV. Propriétés asymptotiques du modèle statistique

par

Mme Françoise MARTIN (1)
A. Faculté des Sciences de Paris

et

M. Daniel VAGUELSY (2)
Attaché de Recherche.

LE MODÈLE STATISTIQUE ASYMPTOTIQUE

Le modèle statistique asymptotique trouve son origine dans l'étude du comportement des stratégies (estimateurs, tests) quand le nombre des observations augmente indéfiniment.

Plusieurs types de problèmes peuvent se poser. Dans le cas de l'estimation par exemple :

a) Chercher des conditions générales d'existence d'un estimateur convergent,

b) Trouver la loi limite d'un estimateur donné pour utiliser, dès que

(1) Institut Henri Poincaré (Probabilités), 5, rue Pierre-Curie, Paris (5e).
(2) Centre universitaire expérimental de Vincennes.

la taille de l'échantillon est assez grande, cette loi limite comme approximation,

c) Chercher les procédures asymptotiquement optimales.

Ce sont les questions *b*) et *c*) qui se trouvent le plus souvent traitées dans la littérature.

Entrent en particulier dans cette catégorie les études consacrées au comportement asymptotique du maximum de vraisemblance, et celles consacrées à la recherche d'estimateurs asymptotiquement normaux et optimaux (B. A. N. estimators : best asymptoticaly normal estimators).

Les principales références bibliographiques sur ce sujet se trouvent dans [*4*], [*12*], [*1*], [*15*].

Ce qui suit est consacré à une étude générale du point *a*).

Nous donnerons tout d'abord les définitions de convergence qui généralisent celles habituellement utilisées, puis nous introduirons la notion de séparation d'une famille de mesures qui nous permettra de donner des conditions nécessaires de convergence. Après un résultat concernant les couples de mesures asymptotiquement étrangères nous étudierons quelques réciproques. Le dernier paragraphe est consacré à une application des résultats précédents à la démonstration d'une condition nécessaire et suffisante de convergence des probabilités *a posteriori*.

Nous considérons comme connus les principaux résultats concernant les martingales. On peut les trouver dans [*7*], [*11*].

IV.1. DÉFINITION DU MODÈLE. NOTATIONS

1.1. Le modèle statistique asymptotique est défini par la donnée de :

— un espace mesurable $(\Omega, \mathscr{A})$: espace des observations,

— un ensemble ordonné filtrant à droite I,

— une famille croissante $\{\mathscr{A}_i, i \in \mathrm{I}\}$ de sous-tribus de $\mathscr{A}$, telle que :

$$\mathscr{A} = \bigvee_{i \in \mathrm{I}} \mathscr{A}_i,$$

— un espace mesurable $(\Theta, \mathscr{H})$: espace des paramètres,

— une transition de probabilité P_Ω^Θ.

Nous noterons $\mathscr{P} = \{\mathrm{P}^\theta, \theta \in \Theta\}$ la famille des probabilités correspondantes sur $(\Omega, \mathscr{A})$, et, pour tout $i \in \mathrm{I}$, $\mathscr{P}_i = \{\mathrm{P}_i^\theta, \mathscr{A} \in \Theta\}$ la famille des restrictions à $(\Omega, \mathscr{A}_i)$ des probabilités P^θ.

De plus nous supposerons que l'application $\theta \rightsquigarrow P^\theta$ est injective :

— un espace mesurable $(D, \mathscr{D})$: espace des décisions,

— une application mesurable $h : (\Theta, \mathscr{H}) \to (D, \mathscr{D})$, qui identifie le problème considéré.

Nous noterons h_D^Θ la transition déterministe associée à l'application mesurable h, qui est définie par :

$$(\forall \theta \in \Theta), (\forall \underline{d} \in \mathscr{D}), h(\theta, \underline{d}) = \left| \begin{array}{l} 1 \text{ si } h(\theta) \in \underline{d}, \\ 0 \text{ sinon.} \end{array} \right.$$

1.2. Nous appellerons procédure statistique une famille

$$S = \{ S_{iD}^\Omega, i \in I \}$$

de transitions de probabilité adaptée à $\{ \mathscr{A}_i, i \in I \}$, c'est-à-dire une famille indexée par I de transitions de probabilité S_i de $(\Omega, \mathscr{A}_i)$ dans $(D, \mathscr{D})$.

Nous noterons alors, pour tout $i \in I$, Π_{iD}^Θ la transition composée $S_{iD}^\Omega \circ P_\Omega^\Theta$, qui est définie par :

$$(\forall \theta \in \Theta), (\forall \underline{d} \in \mathscr{D}), \Pi_i(\theta, \underline{d}) = \int_\Omega S_i(., \underline{d}) dP^\theta.$$

1.3. **Cas particuliers.**

a) Supposons que $D = \{ 0, 1 \}$. L'application h définit alors une partition mesurable $\{ \Theta_0, \Theta_1 \}$ de Θ. On dit dans ce cas que l'on a un problème de test. Une procédure statistique $\{ S_{iD}^\Omega, i \in I \}$ est alors entièrement déterminée par le processus $\{ \Phi_i, i \in I \}$ adapté à $\{ \mathscr{A}_i, i \in I \}$, à valeurs dans $[0, 1]$, défini par :

$$(\forall i \in I), (\forall \omega \in \Omega), \Phi_i(\omega) = S_i(\omega, \{ 1 \})$$

On dit que $\{ \Phi_i, i \in I \}$ est une fonction de test (ou plus simplement un test) adaptée de Θ_0 contre Θ_1.

b) Supposons maintenant que, l'espace des décisions étant quelconque, la procédure statistique $\{ S_{iD}^\Omega, i \in I \}$ soit déterministe, c'est-à-dire qu'il existe un processus $\{ \sigma_i, i \in I \}$ adapté à $\{ \mathscr{A}_i, i \in I \}$, à valeurs dans $(D, \mathscr{D})$, tel que :

$$(\forall i \in I), (\forall \omega \in \Omega), (\forall \underline{d} \in \mathscr{D}), S_i(\omega, \underline{d}) = \left| \begin{array}{l} 1 \text{ si } \sigma_i(\omega) \in \underline{d} \\ 0 \text{ sinon.} \end{array} \right.$$

On dit alors que le processus $\{ \sigma_i, i \in I \}$ est un estimateur adapté.

IV.2. CONVERGENCE DES PROCÉDURES

Une procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ étant une famille de transitions, nous considérons deux sortes de convergence, soit celles de familles de v. a. r. sur $(\Omega, \mathscr{A})$ de la forme $\{ S_i(., \underline{d}), i \in I \}$, soit celles des familles de mesures sur $(D, \mathscr{D})$ de la forme $\{ \Pi_i(\theta, .), i \in I \}$.

Mais de plus, quel que soit le mode de convergence considéré, nous nous imposerons la limite, définie à partir de la transition h_D^{Θ}. Nous allons donc définir des notions de procédures « consistantes ».

Nous supposerons désormais que D est un espace topologique séparé, et que $\mathscr{D}$ est sa tribu borélienne.

2.1. Convergence des V. A. R. $S_i(., \underline{d})$.

DÉFINITIONS 2.1. — $(\overline{C1})$ La procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ sera dite fortement presque sûrement consistante pour $\mathscr{P}$ si :

$$(\forall \theta \in \Theta), (\exists N \in \mathscr{A} : P^{\theta}(N) = 0), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$(\forall \omega \in N^c), \lim_{I} S_i(\omega, \underline{d}) = 1.$$

(C1). La procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ sera dite presque sûrement consistante pour $\mathscr{P}$ si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$\lim_{I} (P^{\theta}\text{-p.s.}) \; S_i(., \underline{d}) = 1$$

c'est-à-dire si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta)), (\exists N \in \mathscr{A} : P^{\theta}(N) = 0)$$
$$(\forall \omega \in N^c), \lim_{I} S_i(\omega, \underline{d}) = 1.$$

(C2). La procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ sera dite consistante en norme L^1 pour $\mathscr{P}$ si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$\lim_{I} (\text{en norme } L^1(\Omega, \mathscr{A}, P^{\theta})) \; S_i(., \underline{d}) = 1$$

c'est-à-dire, puisque $S_i(\omega, \underline{d}) \leqslant 1$ et en vertu de la définition des Π_{iD}^{Θ}, si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$\lim_{I} \Pi_i(\theta, \underline{d}) = 1.$$

(C3). La procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ sera dite consistante en probabilité pour $\mathscr{P}$ si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$\lim_{I} (\text{en proba. } P^{\theta})\, S_i(., \underline{d}) = 1$$

c'est-à-dire si :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta)), (\forall \varepsilon > 0)$$
$$\lim_{I} P^{\theta}[1 - S_i(., \underline{d}) > \varepsilon] = 0.$$

RELATIONS ENTRE LES DÉFINITIONS PRÉCÉDENTES.

Il est clair que $(\overline{C1}) \Rightarrow (C1)$. De plus, pour tout $\underline{d} \in \mathscr{D}$, les v. a. r. $\{ S_i(., \underline{d}), i \in I \}$, étant uniformément bornées par 1, sont équi-intégrables dans chacun des espaces $L^1(\Omega, \mathscr{A}, P^{\theta})$, et par suite $(C2) \rightleftarrows (C3)$.

Enfin si l'ensemble filtrant I admet une partie cofinale innombrable, alors on a :

$$(\overline{C1}) \Rightarrow (C1) \Rightarrow (C2) \rightleftarrows (C3).$$

2.2. **Convergence des mesures** $\Pi_i(\theta, .)$.

DÉFINITION 2.2. — Étant donné $\mathscr{F} \subset \mathscr{L}^{\infty}(\Theta \times D, \mathscr{H} \otimes \mathscr{D})$, la procédure statistique $\{ S_{iD}^{\Omega}, i \in I \}$ sera dite $\mathscr{F}$-consistante pour $\mathscr{P}$ si :

$$(\forall \theta \in \Theta), (\forall f \in \mathscr{F}), \lim_{I} \int_{D} f(\theta, .) d\Pi(\theta, .) = f(\theta, h(\theta))$$

c'est-à-dire si :

$$(\forall \theta \in \Theta), (\forall f \in \mathscr{F}), \lim_{I} \int_{\Omega} dP^{\theta}(\omega) \int_{D} f(\theta, .) dS_i(\theta, .) = f(\theta, h(\theta)).$$

Il est clair que, si $\mathscr{F}' \subset \mathscr{F} \subset \mathscr{L}^{\infty}(\Theta \times D, \mathscr{H} \otimes \mathscr{D})$, alors si une procédure statistique est $\mathscr{F}$-consistante, elle est aussi $\mathscr{F}'$-consistante pour $\mathscr{P}$.

Remarque. — Nous n'avons pas considéré la convergence forte des probabilités $\Pi_i(\theta, .)$ vers la mesure ponctuelle $\delta_{h(\theta)}$, pour chaque $\theta \in \Theta$, définie par :

$$(\forall \theta \in \Theta), \lim_{I} \sup_{d \in \mathscr{D}} | \Pi_i(\theta, \underline{d}) - \delta_{h(\theta)}(\underline{d}) | = 0.$$

La raison en est que la condition précédente implique :

$$(\forall \theta \in \Theta), (\forall \underline{d} \in \mathscr{D} : h(\theta) \in d), \lim_{\mathrm{I}} (\text{en norme } \mathrm{L}^1(\Omega, \mathscr{A}, \mathrm{P}^\theta)) \mathrm{S}_i(., \underline{d}) = 1$$

Nous aurions donc une définition beaucoup plus exigente que (C2).

Cas particuliers.

a) Supposons que les éléments de $\mathscr{F}$ soient des fonctions constantes en θ, autrement dit prenons $\mathscr{F} \subset \mathscr{L}^\infty(\mathrm{D}, \mathscr{D})$.

— Si $\mathscr{F} = \mathscr{L}^\infty(\mathrm{D}, \mathscr{D})$, alors la $\mathscr{F}$-consistance pour $\mathscr{P}$ équivaut à la convergence faible, pour tout $\theta \in \Theta$, des mesures $\Pi_i(\theta, .)$ vers $\delta_{h(\theta)}$. Mais comme cette condition implique :

$$(\forall \theta \in \Theta), (\forall \underline{d} \in \mathscr{D} : h(\theta) \in \underline{d}), \lim_{\mathrm{I}} (\text{en norme } \mathrm{L}^1(\Omega, \mathscr{A}, \mathrm{P}^\theta)) \mathrm{S}_i(., \underline{d}) = 1$$

elle est aussi beaucoup trop forte.

— Si D est un espace métrique complet séparable, et si $\mathscr{F} = \mathscr{C}_b(\mathrm{D})$ (espace des fonctions numériques continues bornées sur D), alors la $\mathscr{F}$-consistance pour $\mathscr{P}$ est équivalente à la convergence en loi, pour tout $\theta \in \Theta$, des mesures $\Pi_i(\theta, .)$ vers $\delta_{h(\theta)}$.

— Si $\mathscr{F} = \mathscr{C}_k(\mathrm{D})$ (espace des fonctions numériques continues à support compact sur D), on retrouve la convergence étroite.

b) Supposons que l'espace topologique Θ soit métrisable, et que ρ soit une distance compatible avec sa topologie. Soit $p \geq 1$. Si nous prenons $\mathscr{F} = \{ [\rho(h(\theta), d)]^p \}$, alors la $\mathscr{F}$-consistance pour $\mathscr{P}$ d'une procédure statistique $\{ \mathrm{S}_{i\mathrm{D}}^{\Omega}, i \in \mathrm{I} \}$ s'écrit :

$$(\forall \theta \in \Theta), \lim_{\mathrm{I}} \int_{\mathrm{D}} [\rho(h(\theta), .)]^p d\Pi_i(\theta, .) = 0.$$

c) Dans le cadre décisionnel, soit $1(\theta, d)$ une fonction de coût. Si nous prenons $\mathscr{F} = \{ 1(\theta, d) \}$ alors la $\mathscr{F}$-consistance pour $\mathscr{P}$ d'une procédure statistique $\{ \mathrm{S}_{i\mathrm{D}}^{\Omega}, i \in \mathrm{I} \}$ s'écrit :

$$(\forall \theta \in \Theta), \lim_{\mathrm{I}} \mathrm{R}(\mathrm{S}_i, \theta) = 1(\theta, h(\theta))$$

où $\mathrm{R}(\mathrm{S}_i, \theta) = \int_{\mathrm{D}} 1(\theta, .) d\Pi_i(\theta, .)$ est le risque de la stratégie S_i en θ. Remarquons qu'il est en général préférable de prendre pour $\mathscr{F}$ une famille de fonctions de coût.

Remarque. — On peut construire un sous-ensemble $\mathscr{F}$ de $\mathscr{L}^\infty(\Theta \times \mathrm{D}, \mathscr{H} \otimes \mathscr{D})$ tel que, pour toute procédure statistique, la $\mathscr{F}$-consistance et la consistance en norme L^1 pour $\mathscr{P}$ soient équivalentes.

En effet, pour tout $d \in \mathrm{D}$, soit $\mathscr{V}_d$ l'ensemble des voisinages ouverts de d. A tout

$$\varphi = \{ \mathrm{V}_d,\ d \in h(\Theta) \} \in \prod_{d \in h(\Theta)} \mathscr{V}_d,$$

associons la fonction $f \in \mathscr{L}^\infty(\Theta \times \mathrm{D}, \mathscr{H} \otimes \mathscr{D})$ définie par :

$$(\forall \theta \in \Theta),\ (\forall d \in \mathrm{D}),\ f(\theta, d) = 1_{\mathrm{V}_{h(\theta)}}(d).$$

La famille $\mathscr{F}$ ainsi obtenue convient.

2.3. Cas d'un problème de test.

Dans ce cas $\mathrm{D} = \{0, 1\}$. Soit $\{\Theta_0, \Theta_1\}$ la partition mesurable de Θ induite par l'application h. Soit $\{\Phi_i,\ i \in \mathrm{I}\}$ un test adapté de Θ_0 contre Θ_1 (cf. (1.3.a)). Dans ce cas le mode de convergence le plus utilisé est celui de la définition (C2), qui s'écrit ici :

$$\begin{cases} \forall \theta \in \Theta_0,\ \lim\limits_{\mathrm{I}} \mathrm{E}^\theta[\Phi_i] = 0 \\ \forall \theta \in \Theta_1,\ \lim\limits_{\mathrm{I}} \mathrm{E}^\theta[\Phi_i] = 1 \end{cases}$$

où :

$$\mathrm{E}^\theta[\Phi_i] = \int_\Omega \Phi_i d\mathrm{P} \qquad \forall \theta \in \Theta.$$

Nous dirons alors que $\{\Phi_i,\ i \in \mathrm{I}\}$ est un test adapté consistant en moyenne de θ_0 contre Θ_1.

Dans le cas d'un problème de test, on utilise souvent un type de convergence plus faible, et essentiellement lié au critère statistique (convergence des tests de seuil α), défini de la facon suivante :

Étant donné $\alpha \in]0, 1]$, un test adapté $\{\Phi_i,\ i \in \mathrm{I}\}$ de Θ_0 contre Θ_1 sera dit consistant en moyenne au seuil α si :

$$(\forall \theta_0 \in \Theta_0),\ (\forall i \in \mathrm{I}),\ \mathrm{E}^\theta[\Phi_i] \leq \alpha.$$
$$(\forall \theta_1 \in \Theta_1),\ \lim_{\mathrm{I}} \mathrm{E}^\theta[\Phi_i] = 1.$$

2.4. Consistance des estimateurs adaptés.

Nous allons maintenant interpréter les différents modes de convergence que nous venons d'introduire dans le cas d'une procédure statistique déterministe associée à un estimateur adapté (cf. (1.3.b)).

PROPOSITION 2.4.1. — Étant donné une procédure statistique déterministe $S = \{ S_{iD}^{\Omega}, i \in I \}$ associée à un estimateur adapté $\{ \sigma_i, i \in I \}$, les propositions suivantes sont équivalentes :

a) La procédure statistique S est fortement presque sûrement consistante pour $\mathscr{P}$.

b) $(\forall \theta \in \Theta)$, $\lim_{I} (P^{\theta}\text{-p. s.}) \sigma_i = h(\theta)$.

Démonstration. — En remarquant qu'une procédure statistique déterministe ne prend que les valeurs 0 ou 1, la proposition *a)* est équivalente à :

$$(\forall \theta \in \Theta), (\exists N \in \mathscr{A} : P^{\theta}(N) = 0) : (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$(\forall \omega \in N^c), (\exists i_0 \in I), (\forall i \in I : i \geq i_0), S_i(\omega, \underline{d}) = 1.$$

Or :

$$S_i(\omega, \underline{d}) = 1 \leftrightarrows \sigma_i(\omega) \in \underline{d}.$$

Donc la proposition *a)* est encore équivalente à :

$$(\forall \theta \in \Theta), (\exists N \in \mathscr{A} : P^{\theta}(N) = 0), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta))$$
$$(\forall \omega \in N^c), (\exists i_0 \in I), (\forall i \in I : i \geq i_0), \sigma_i(\omega) \in \underline{d}.$$

D'où le résultat.

PROPOSITION 2.4.2. — Étant donné une procédure statistique déterministe $S = \{ S_{iD}^{\Omega}, i \in I \}$ associée à un estimateur adapté $\{ \sigma_i, i \in I \}$, les propositions suivantes sont équivalentes :

a) La procédure S est consistante en probabilité pour $\mathscr{P}$,

b) $(\forall \theta \in \Theta)$, $\lim_{I}$ (en proba. P^{θ}) $\sigma_i = h(\theta)$.

Démonstration. — Comme les définitions (C2) et (C3) sont équivalentes, la proposition *a)* s'écrit (cf. (2.1)) :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta)), \lim_{I} \Pi_i(\theta, \underline{d}) = 1.$$

Or, comme :

$$S_i(\omega, \underline{d}) = 1 \rightleftarrows \sigma_i(\omega) \in \underline{d}$$

on a :

$$(\forall i \in I), (\forall \theta \in \Theta), (\forall \underline{d} \in \mathscr{D}), \Pi_i(\omega, \underline{d}) = \int_{\Omega} S_i(., \underline{d}) dP^{\theta} = P^{\theta}[\sigma_i \in \underline{d}].$$

Par suite la proposition *a)* est encore équivalente à :

$$(\forall \theta \in \Theta), (\forall \underline{d} \text{ ouvert de D contenant } h(\theta)), \lim P^{\theta}[\sigma_i \in \underline{d}] = 1$$

d'où le résultat.

On voit facilement que, étant donné $\mathscr{F} \subset \mathscr{L}^\infty(\Theta \times D, \mathscr{H} \otimes \mathscr{D})$, la $\mathscr{F}$-consistance pour $\mathscr{P}$ d'une procédure statistique déterministe se traduit, pour l'estimateur adapté correspondant $\{\sigma_i, i \in I\}$, par :

$$(\forall \theta \in \Theta), (\forall f \in \mathscr{F}), \lim_{I} \int_\Omega f(\theta, \sigma_i(.))dP^\theta = f(\theta, h(\theta)).$$

En particulier, si D est métrisable, ρ étant une distance compatible avec sa topologie, et si, étant donné $p \geq 1$, $\mathscr{F} = \{[\rho(h(\theta), d)]^p\}$, alors la $\mathscr{F}$-consistance pour $\mathscr{P}$ d'une procédure statistique déterministe est équivalente à la convergence, pour tout $\theta \in \Theta$, en norme dans $L^p_D(\Omega, \mathscr{A}, P^\theta)$ de l'estimateur adapté correspondant vers $h(\theta)$.

IV.3. SÉPARATION DES MESURES

Nous allons dans ce paragraphe, définir des propriétés de la famille qui nous permettront de donner des conditions d'existence de procédures consistantes.

Nous considérons maintenant un espace mesurable $(\Omega, \mathscr{A})$ et une famille $\mathscr{P}$ de probabilités sur $(\Omega, \mathscr{A})$.

3.1

DÉFINITION 3.1.1. — Deux parties $\mathscr{P}_1$ et $\mathscr{P}_2$ de $\mathscr{P}$ seront dites étrangères, ce que nous noterons $\mathscr{P}_1 \perp \mathscr{P}_2$, si :

$$(\forall P_1 \in \mathscr{P}_1), (\forall P_2 \in \mathscr{P}_2), (\exists N \in \mathscr{A}) : P_1(N) = P_2(N^c) = 0.$$

DÉFINITION 3.1.2. — Deux parties $\mathscr{P}_1$ et $\mathscr{P}_2$ de $\mathscr{P}$ seront dites fortement étrangères, ce que nous noterons $\mathscr{P}_1 \perp\!\!\!\perp \mathscr{P}_2$, si :

$$(\exists N \in \hat{\mathscr{A}}) : (\forall P_1 \in \mathscr{P}_1), (\forall P_2 \in \mathscr{P}_2), P_1(N) = P_2(N^c) = 0$$

où $\hat{\mathscr{A}}$ est la complétée universelle de $\mathscr{A}$ pour la famille $\mathscr{P}$ (cf. ch. II.I.4.*a*).

PROPRIÉTÉS ÉLÉMENTAIRES.

a) Il est clair que :

$$\mathscr{P}_1 \perp\!\!\!\perp \mathscr{P}_2 \Rightarrow \mathscr{P}_1 \perp \mathscr{P}_2 \Rightarrow \mathscr{P}_1 \cap \mathscr{P}_2 = \emptyset.$$

De plus on voit facilement que, si $\mathscr{P}_1$ et $\mathscr{P}_2$ sont des parties dénombrables de $\mathscr{P}$, alors :

$$\mathscr{P}_1 \perp \mathscr{P}_2 \Rightarrow \mathscr{P}_1 \perp\!\!\!\perp \mathscr{P}_2.$$

b) Soient $\{\mathscr{P}_\alpha, \alpha \in A\}$ et $\{\mathscr{P}_\beta, \beta \in B\}$ deux familles de parties de $\mathscr{P}$. Il est clair que :

$$[(\forall \alpha \in A), (\forall \beta \in B), \mathscr{P}_\alpha \perp \mathscr{P}_\beta] \Rightarrow \left[\bigcup_{\alpha \in A} \mathscr{P}_\alpha \perp \bigcup_{\beta \in B} \mathscr{P}_\beta\right].$$

Si A et B sont dénombrables, alors on a aussi :

$$[(\forall \alpha \in A), (\forall \beta \in B), \mathscr{P}_\alpha \underline{\perp} \mathscr{P}_\beta] \Rightarrow \left[\bigcup_{\alpha \in A} \mathscr{P}_\alpha \underline{\perp} \bigcup_{\beta \in B} \mathscr{P}_\beta\right].$$

En effet l'hypothèse nous dit que, à tout couple $(\alpha, \beta) \in A \times B$, nous pouvons associer un élément $N_{\alpha\beta}$ de $\hat{\mathscr{A}}$ tel que :

$$(\forall P \in \mathscr{P}_\alpha), (\forall Q \in \mathscr{P}_\beta), P(N_{\alpha\beta}) = Q(N_{\alpha\beta}^c) = 0.$$

Soit

$$N = \bigcap_{\alpha \in A} \bigcup_{\beta \in B} N_{\alpha\beta}.$$

Comme A et B sont dénombrables, $N \in \hat{\mathscr{A}}$. Si

$$P \in \bigcup_{\alpha \in A} \mathscr{P}_\alpha,$$

alors il existe α_0 dans A tel que : $P \in \mathscr{P}\alpha_0$, donc

$$P(N) \leq \sum_{\beta \in B} P(N_{\alpha_0\beta}) = 0.$$

De même si

$$Q \in \bigcup_{\beta \in B} \mathscr{P}_\beta,$$

alors il existe β_0 dans B tel que $Q \in \mathscr{P}_{\beta_0}$, donc :

$$P(N^c) \leq \sum_{\alpha \in A} P(N_{\alpha\beta_0}^c) = 0.$$

3.2

DÉFINITION 3.2.1. — Soit $\mathscr{C}$ une classe de parties de $\mathscr{P}$. Nous dirons que la famille $\mathscr{P}$ est séparée (respectivement fortement séparée) pour $\mathscr{C}$ si :

$$(\forall (\mathscr{P}_1, \mathscr{P}_2) \in \mathscr{C} \times \mathscr{C} : \mathscr{P}_1 \cap \mathscr{P}_2 = \varnothing), \qquad \mathscr{P}_1 \perp \mathscr{P}_2$$

(respectivement, $\mathscr{P}_1 \underline{\perp} \mathscr{P}_2$).

DÉFINITION 3.2.2. — Soit $\mathscr{C}$ une classe de parties de $\mathscr{P}$. Nous dirons que la famille $\mathscr{P}$ est bien séparée (respectivement fortement bien séparée) pour $\mathscr{C}$ si :

$$(\forall \mathscr{P}_1 \in \mathscr{C}),\ \mathscr{P}_1 \perp (\mathscr{P} \setminus \mathscr{P}_1)$$

(respectivement $\mathscr{P}_1 \perp\!\!\!\perp (\mathscr{P} \setminus \mathscr{P}_1)$).

Remarque. — Soit $\mathscr{C}$ une classe de parties de $\mathscr{P}$. Notons $\overline{\mathscr{C}}$ la fermeture de $\mathscr{C}$ pour la complémentation. Alors il est clair que les propriétés pour la famille $\mathscr{P}$ d'être bien séparée (respectivement fortement bien séparée) pour $\mathscr{C}$ et séparée (respectivement fortement séparée) pouf $\overline{\mathscr{C}}$, sont équivalentes.

PROPOSITION 3.2. — Supposons la famille $\mathscr{P}$ fortement bien séparée pour une classe $\mathscr{C}$ de parties de $\mathscr{P}$. Soit $\mathscr{M}$ la plus grande classe de parties de $\mathscr{P}$ telle que $\mathscr{P}$ soit fortement bien séparée pour $\mathscr{M}$. Alors $\mathscr{M}$ est une tribu de parties de $\mathscr{P}$ contenant la tribu engendrée par $\mathscr{C}$.

Démonstration. — Il est clair que $\mathscr{M}$ contient $\mathscr{C}$ et que $\mathscr{M}$ est fermée pour la complémentation. Soit $\{\mathscr{P}_n, n \geq 1\}$ une suite d'éléments de $\mathscr{M}$. Posons

$$\mathscr{P}_0 = \bigcap_{n \geq 1} \mathscr{P}_n.$$

Pour tout $n \geq 1$ $\mathscr{P}_0 \subset \mathscr{P}_n$, donc $\mathscr{P}_0 \perp\!\!\!\perp \mathscr{P} \setminus \mathscr{P}_n$ et par suite (cf. 3.2.*b*)

$$\mathscr{P}_0 \perp\!\!\!\perp \bigcup_n (\mathscr{P} \setminus \mathscr{P}_n) = \mathscr{P} \setminus \mathscr{P}_0.$$

CAS PARTICULIERS.

a) Soit $\mathscr{C}$ la classe de parties de $\mathscr{P}$ réduites à un élément. Alors les propriétés pour la famille $\mathscr{P}$ d'être séparée, bien séparée ou fortement séparée pour $\mathscr{C}$ sont équivalentes à la propriété que les probabilités de la famille sont deux à deux étrangères. Mais la propriété d'être fortement bien séparée pour $\mathscr{C}$ est plus forte, elle s'écrit :

$$(\forall P \in \mathscr{P}),\ (\exists N \in \mathscr{A}) : (\forall Q \in \mathscr{P} : Q \neq P),\ P(N) = Q(N^c) = 0.$$

b) Soit $\mathscr{C}$ la classe de toutes les parties de $\mathscr{P}$. Alors les propriétés pour la famille $\mathscr{P}$ d'être séparée pour $\mathscr{C}$ sont équivalentes à la propriété que les probabilités de la famille sont deux à deux étrangères. Les propriétés pour $\mathscr{P}$ d'être fortement séparée ou fortement bien séparée pour $\mathscr{C}$ sont équivalentes et plus fortes que les précédentes.

3.3

DÉFINITION 3.3. — Soit $\{\mathscr{P}_t, t\in T\}$ une partition de $\mathscr{P}$.

Nous dirons que la famille $\mathscr{P}$ est totalement séparée pour la partition $\{\mathscr{P}_t, t\in T\}$ s'il existe une partition $\hat{\mathscr{A}}$-mesurable $\{N_t, t\in T\}$ de Ω telle que :

$$(\forall t\in T), (\forall P\in\mathscr{P}_t), P(N_t) = 1.$$

Si les éléments de la partition sont les parties de $\mathscr{P}$ réduites à un point, nous dirons que la famille $\mathscr{P}$ est totalement séparée. On voit facilement que si $\mathscr{P}$ est totalement séparée pour une partition $\{\mathscr{P}_t, t\in T\}$, alors $\mathscr{P}$ est fortement bien séparée pour la tribu engendrée par cette partition.

3.4

DÉFINITION 3.4. — Soit $(\Theta, \mathscr{H}, \mu)$ un espace de probabilité, et soit g une application de Θ dans $\mathscr{P}$. Nous dirons que la famille $\mathscr{P}$ est μ-fortement séparée si :

$$(\forall T\in\mathscr{H} : \mu(T)\in]0, 1[), (\exists N\in\mathscr{H} : \mu(N) = 0) : g(T\setminus N) \perp g(T^c\setminus N).$$

3.5. Exemples.

a) Soient $(\Omega_0, \mathscr{A}_0)$ un espace mesurable et $\mathscr{P}_0$ une famille de probabilités sur $(\Omega_0, \mathscr{A}_0)$. Supposons que $(\Omega, \mathscr{A})$ soit l'espace produit

$$\left(\prod_1^\infty \Omega_0, \bigotimes_1^\infty \mathscr{A}_0\right)$$

et que la famille $\mathscr{P}$ soit la famille des probabilités produit $\prod_1^\infty P_0$. Alors la famille $\mathscr{P}$ est séparée pour la classe de toutes ses parties (on peut consulter à ce sujet [*7*], [*2*], [*14*]).

b) Supposons que Ω soit le carré du plan construit sur [0, 1], et $\mathscr{A}$ sa tribu borélienne. Soit Θ l'ensemble des couples $(\theta_1, \theta_2)\in\mathbb{R}^2$ tels que l'intersection de la droite $y = \theta_1 x + \theta_2$ et de Ω soit non vide. A tout point θ de Θ, associons la probabilité P^θ sur $(\Omega, \mathscr{A})$ qui est uniforme sur l'intersection avec Ω de la droite $y = \theta_1 x + \theta_2$. Alors la famille

$$\mathscr{P}_0 = \{P^\theta, \theta\in\Theta\cap\{\theta_2 = 0\}\}$$

est totalement séparée. Mais ce n'est pas vrai pour la famille $\mathscr{P} = \{P^\theta, \theta\in\Theta\}$. La famille $\mathscr{P}$ est fortement bien séparée pour la classe de ses parties réduites à un point, et seulement séparée pour la classe de toutes ses parties.

IV.4. CONDITIONS NÉCESSAIRES DE CONSISTANCE

Nous allons maintenant donner quelques conditions nécessaires d'existence de procédures consistantes au moyen de propriétés de séparation de la famille $\mathscr{P}$.

PROPOSITION 4.1. — Soit un modèle statistique asymptotique (cf. 1.1). Soit δ_h la partition mesurable de Θ induite par l'application h. S'il existe un estimateur adapté $\{\sigma_i, i \in \mathrm{I}\}$ tel que la procédure statistique déterministe associée soit fortement presque sûrement consistante pour $\mathscr{P}$, alors la famille $\mathscr{P}$ est totalement séparée pour la partition δ_h.

Démonstration. — En vertu de la proposition (2.4.1), si la procédure déterministe associée à $\{\sigma_i, i \in \mathrm{I}\}$ est fortement presque sûrement consistante pour $\mathscr{P}$, alors :

$$(\forall \theta \in \Theta), \lim_{\mathrm{I}} (\mathrm{P}^\theta\text{-p. s.})\, \sigma_i = h(\theta).$$

A tout point $d \in \mathrm{D}$, associons l'ensemble

$$\mathrm{N}_d = \{\omega \in \Omega : \lim_{\mathrm{I}} \sigma_i(\omega) = d\}.$$

La famille $\{\mathrm{N}_d, d \in \mathrm{D}\}$ forme une partition de Ω.

Étant donné un élément $d \in \mathrm{D}$ et un élément $\theta \in \Theta$, on a ou bien $h(\theta) = d$, alors $\mathrm{N}_d \in \widehat{\mathscr{A}}_\theta$ (tribu complétée de $\mathscr{A}$ pour P^θ) et $\mathrm{P}^\theta(\mathrm{N}_d) = 1$ ou bien

$$h(\theta) = d' \neq d,$$

alors, comme $\mathrm{N}_d \subset \mathrm{N}_{d'}^c$, $\mathrm{N}_d \in \widehat{\mathscr{A}}_\theta$ et $\mathrm{P}^\theta(\mathrm{N}_d) = 0$. D'où le résultat.

PROPOSITION 4.2. — Soit un modèle statistique asymptotique. S'il existe une procédure statistique presque sûrement consistante pour $\mathscr{P}$, alors la famille est fortement séparée pour la classe $\mathscr{C} = \{h^{-1}(\underline{d}), \underline{d}$ ouvert de $\mathrm{D}\}$.

Démonstration. — Soit $\{\mathrm{S}_{i\mathrm{D}}^{\Omega}, i \in \mathrm{I}\}$ une procédure statistique presque sûrement consistante pour $\mathscr{P}$. La définition (C1) (cf. 2.1) nous permet alors d'associer à tout ouvert $\underline{d}$ de D et à tout point $\theta \in h^{-1}(\underline{d})$ un ensemble $\mathrm{N}_{\underline{d},\theta} \in \mathscr{A}$, P^θ-négligeable, tel que sur $\mathrm{N}_{\underline{d},\theta}^c$,

$$\lim_{\mathrm{I}} \mathrm{S}_i(., \underline{d}) = 1.$$

Posons

$$N_{\underline{d}} = \bigcap_{\theta \in h^{-1}(\underline{d})} N_{\underline{d},\theta}.$$

Sur $N_{\underline{d}}^c$, $\lim_I S_i(.,\underline{d}) = 1$. Donc si $\underline{d}$ et $\underline{d}'$ sont deux ouverts disjoints de D, alors, puisque $S_i(\omega, \underline{d}) + S_i(\omega, \underline{d}') \leq 1$, les ensembles $N_{\underline{d}}^c$ et $N_{\underline{d}}^c$ sont disjoints.

Soit $\theta \in \Theta$. Si $h(\theta) \in \underline{d}$, alors, d'après la définition de $N_{\underline{d}}$, il est clair que $N_{\underline{d}} \in \widehat{\mathscr{A}}_\theta$ et $P^\theta(N_{\underline{d}}) = 0$. Si $h(\theta) \notin \underline{d}$, alors, $\underline{d}'$ étant un voisinage ouvert de $h(\theta)$ ne rencontrant pas $\underline{d}$, on a : $N_{\underline{d}}^c \subset N_{\underline{d}'}$, donc :

$$N_{\underline{d}} \in \widehat{\mathscr{A}}_\theta \qquad \text{et} \qquad P^\theta(N_{\underline{d}}^c) = 0.$$

La proposition est donc démontrée.

PROPOSITION 4.3. — *Soit un modèle statistique asymptotique. S'il existe une procédure statistique consistante en probabilité pour $\mathscr{P}$, alors la famille $\mathscr{P}$ est séparée pour la classe $\mathscr{C} = \{ h^{-1}(d) : d \in D \}$.*

Démonstration. — Soit $\{ S_{iD}^\Omega, i \in I \}$ une procédure statistique consistante en probabilité pour $\mathscr{P}$. Soient θ_1 et θ_2 deux points de Θ tels que $h(\theta_1) \neq h(\theta_2)$. Soient $\underline{d}_1$ et $\underline{d}_2$ des voisinages ouverts disjoints de $h(\theta_1)$ et $h(\theta_2)$ respectivement.

Alors, en vertu de la définition (C3) (2.1), pour tout $\varepsilon \in]0, 1[$, on a :

$$\lim_I P^{\theta_1}[S_i(.,\underline{d}_1) \leq 1 - \varepsilon] = 0,$$

et

$$\lim_I P^{\theta_2}[1 - S_i(.,\underline{d}_2) \leq \varepsilon] = 1.$$

Par suite, puisque $\underline{d}_2^c \supset \underline{d}_1$, en prenant $\varepsilon \in \left]0, \frac{1}{2}\right[$ et en posant, pour tout $i \in I$, $A_i = \{ \omega \in \Omega : S_i(\omega, \underline{d}_1) \leq \varepsilon \}$, on a :

$$A_i \in \mathscr{A}, \qquad \lim_I P^{\theta_1}(A_i) = 0 \qquad \text{et} \qquad \lim_I P^{\theta_2}(A_i) = 1.$$

Dans ces conditions, en utilisant la décomposition de Jordan de la mesure $(P^{\theta_1} - P^{\theta_2})$, on a :

$$(P^{\theta_1} - P^{\theta_2})^-(\Omega) = \sup_{A \in \mathscr{A}} (P^{\theta_2} - P^{\theta_1})(A) = 1,$$

d'où

$$(P^{\theta_1} \wedge P^{\theta_2})(\Omega) = 1 - (P^{\theta_1} - P^{\theta_2})^-(\Omega) = 0,$$

c'est-à-dire $P^{\theta_1} \perp P^{\theta_2}$.

IV.5. COUPLE DE MESURES ASYMPTOTIQUEMENT ÉTRANGÈRES

Soit un modèle asymptotique statistique pour lequel $\mathscr{P} = \{ P, Q \}$. Le but du présent paragraphe est de caractériser le fait que P et Q sont étrangères au moyen de propriétés limite des couples $\{ P_i, Q_i \}$ de probabilités sur $(\Omega, \mathscr{A}_i)$, quand i parcourt I.

DÉFINITION 5.1. — Soit $(\Omega, \mathscr{A})$ un espace mesurable, et soient P et Q deux probabilités sur $(\Omega, \mathscr{A})$. Nous appellerons pseudo-densité de P par rapport à Q toute v. a. r., notée $\Pi^{P/Q}$, sur $(\Omega, \mathscr{A})$, Q-presque sûrement positive, et telle que :

$$(\forall A \in \mathscr{A}), \qquad P(A) = \int_A \Pi^{P/Q} dQ + P(A \cap N^{P/Q})$$

où $N^{P/Q} \in \mathscr{A}$ et $Q(N^{P/Q}) = 0$.

L'égalité précédente est la décomposition de Lebesgue de P par rapport à Q. Nous savons qu'une v. a. r. $\Pi^{P/Q}$ ainsi définie est Q-presque sûrement unique.

Si $P \ll Q$, alors $\Pi^{P/Q} \in \dfrac{dP}{dQ}$. Par extension nous noterons encore $\dfrac{dP}{dQ}$ la classe des pseudo-densités de P par rapport à Q.

Soit λ une mesure σ-finie sur $(\Omega, \mathscr{A})$ dominant P et Q (par exemple $\lambda = P + Q$).

Si $X \in \dfrac{dP}{d\lambda}$ et $Y \in \dfrac{dQ}{d\lambda}$, alors on a :

$$\Pi^{P/Q} = \frac{X}{Y} [Q].$$

LEMME 5.2. — Soit un modèle statistique asymptotique pour lequel $\mathscr{P} = \{ P, Q \}$. Alors :

a) $\{ \Pi^{P_i/Q_i}, \mathscr{A}_i, i \in I \}$ est une surmartingale positive dans $(\Omega, \mathscr{A}, Q)$.

b) Pour tout $\alpha \in]0, 1[$, $\{ (\Pi^{P_i/Q_i})^\alpha, \mathscr{A}_i, i \in I \}$ est une surmartingale positive équi-intégrable dans $(\Omega, \mathscr{A}, Q)$.

Démonstration.

a) Soient i et j deux éléments de I tels que $i \leq j$. Soit $A_i \in \mathscr{A}_i$. On a :

$$\int_{A_i} E_Q^{\mathscr{A}_i}[\Pi^{P_j/Q_j}] dQ = \int_{A_i} \Pi^{P_j/Q_j} dQ \leq P(A_i).$$

Or on sait que la pseudo-densité Π^{P_i/Q_i} est Q_i-presque sûrement la plus grande des v. a. r. Z_i sur $(\Omega, \mathscr{A}_i)$ telles que :

$$(\forall A_i \in \mathscr{A}_i), \qquad \int_{A_i} Z_i dQ \leq P(A_i).$$

Par conséquent :

$$(\forall (i, j) \in I \times I : i \leq j), \qquad E_Q^{\mathscr{A}_i}[\Pi^{P_j/Q_j}] \leq \Pi^{P_i/Q_i}[Q].$$

b) On déduit de *a*), par l'inégalité de Jensen, que pour tout $\alpha \in]0, 1[$, $\{ (\Pi^{P_i/Q_i})^\alpha, \mathscr{A}_i, i \in I \}$ est encore une surmartingale positive dans $(\Omega, \mathscr{A}, Q)$. Soit λ une mesure σ-finie sur $(\Omega, \mathscr{A})$ dominant P et Q. Soient, pour tout $i \in I$, λ_i sa restriction à $(\Omega, \mathscr{A}_i)$,

$$X_i \in \frac{dP_i}{d\lambda_i} \quad \text{et} \quad Y_i \in \frac{dQ_i}{d\lambda_i}.$$

Alors, pour tout $i \in I$ et tout $A_i \in \mathscr{A}_i$, on a :

$$\int_{A_i} (\Pi^{P_i/Q_i})^\alpha dQ = \int_{A_i} \left(\frac{X_i}{Y_i}\right)^\alpha Y_i d\lambda_i = \int_{A_i} (X_i)^\alpha . (Y_i)^{1-\alpha} d\lambda$$

d'où, en appliquant l'inégalité de Hölder :

$$\int_{A_i} (\Pi^{P_i/Q_i})^\alpha dQ \leq \left(\int_{A_i} X_i d\lambda\right)^\alpha . \left(\int_{A_i} Y_i d\lambda\right)^{1-\alpha} = (P(A_i))^\alpha . (Q(A_i))^{1-\alpha}$$

d'où enfin :

$$\int_{A_i} (\Pi^{P_i/Q_i})^\alpha dQ \leq (Q(A_i))^{1-\alpha}.$$

On en déduit que la surmartingale est équi-intégrable.

THÉORÈME 5.3. — Soit un modèle asymptotique statistique pour lequel $\mathscr{P} = \{ P, Q \}$. Pour que les probabilités P et Q soient étrangères, il faut et il suffit que :

$$(\exists \alpha \in]0, 1[) : \lim_I E_Q[(\Pi^{P_i/Q_i})^\alpha] = 0.$$

Si I est dénombrable et totalement ordonné, cette condition est équivalente à la suivante :

$$\lim_I \,(Q\text{-p. s.})\; \Pi^{P_i/Q_i} = 0.$$

Démonstration. — Puisque :

$$P \perp Q \rightleftharpoons \Pi^{P/Q} = 0 \;[Q],$$

il suffit de montrer que, pour $\alpha \in]0, 1[$, on a :

$$\lim_{I} E_Q[|(\Pi^{P_i/Q_i})^\alpha - (\Pi^{P/Q})^\alpha|] = 0.$$

Or le lemme précédent nous permet d'affirmer qu'il existe une v. a. r. notée $\Pi^{(\alpha)}$, sur $(\Omega, \mathscr{A})$ telle que :

$$\lim_{I} E_Q[|(\Pi^{P_i/Q_i})^\alpha - \Pi^{(\alpha)}|] = 0.$$

ou, ce qui est équivalent, telle que :

$$\lim_{I} \text{(en proba. Q)}\ (\Pi^{P_i/Q_i})^\alpha = \Pi^{(\alpha)}.$$

Par suite on peut trouver dans I une suite croissante $\{ i_n, n \geq 1 \}$ telle que

$$\lim_{n} \text{(Q-p. s.)}(\Pi^{P_{i_n}/Q_{i_n}})^\alpha = \Pi^{(\alpha)}.$$

Or, en considérant de nouveau une mesure σ-finie λ dominant P et Q, $\{ X_{i_n}, \mathscr{A}_{i_n}, n \geq 1 \}$ et $\{ Y_{i_n}, \mathscr{A}_{i_n}, n \geq 1 \}$ sont deux martingales équi-intégrables dans $(\Omega, \mathscr{A}, \lambda)$, telles que :

$$\lim_{n} (\lambda\text{-p. s.})X_{i_n} = X \qquad \text{et} \qquad \lim_{n} (\lambda\text{-p. s.})Y_{i_n} = Y.$$

On en déduit que

$$\Pi^{(\alpha)} = \left(\frac{X}{Y}\right)^\alpha [Q],$$

c'est-à-dire :

$$\Pi^{(\alpha)} = (\Pi^{P/Q})^\alpha [Q].$$

Si I est dénombrable totalement ordonné, la convergence Q-presque sûre de $(\Pi^{P_i/Q_i})^\alpha$ vers $(\Pi^{P/Q})^\alpha$ est bien équivalente à la dernière assertion du théorème.

IV.6. EXISTENCE DE TESTS ADAPTÉS CONSISTANTS

LEMME 6.1. — Soit un modèle statistique asymptotique tel que $\mathscr{P} = \{ P, Q \}$. Les probabilités P et Q sont étrangères si et seulement si il existe un test adapté consistant en moyenne de P contre Q (cf. 2.3.1).

Démonstration. — La condition suffisante résulte de la proposition 4.3.

Supposons $P \perp Q$. Soit λ une mesure σ-finie dominant P et Q, et soient, pour tout $i \in I$,

$$X_i \in \frac{dP_i}{d\lambda_i} \quad \text{et} \quad Y_i \in \frac{dQ_i}{d\lambda_i}.$$

Soit $r > 0$, soit $\Phi_i = 1_{\left\{\frac{X_i}{Y_i} \leq r\right\}}$ pour tout $i \in I$. Le processus $\{\Phi_i, i \in I\}$ est un test adapté de P contre Q, consistant en moyenne car :

$$E_Q[\Phi_i] = 1 - Q\left[\sqrt{\Pi^{P_i/Q_i}} > \sqrt{r}\right] \geq 1 - \frac{1}{\sqrt{r}} E_Q\left[\sqrt{\Pi^{P_i/Q_i}}\right]$$

$$E_P[\Phi_i] = P\left[\sqrt{\Pi^{Q_i/P_i}} \geq \frac{1}{\sqrt{r}}\right] \leq \sqrt{r} E_P\left[\sqrt{\Pi^{Q_i/P_i}}\right]$$

et le théorème 5.3 entraîne la conclusion.

Proposition 6.2 [9]. — Soit un modèle statistique asymptotique dans lequel $\mathscr{P}$ est dénombrable. Soit $\{\mathscr{P}_0, \mathscr{P}_1\}$ une partition de $\mathscr{P}$. Pour qu'il existe un test adapté consistant en moyenne de $\mathscr{P}_0$ contre $\mathscr{P}_1$, il faut et il suffit que les familles $\mathscr{P}_0$ et $\mathscr{P}_1$ soient étrangères.

Démonstration. — La condition nécessaire résulte de la proposition 4.3.

Inversement supposons $\mathscr{P}_0 \perp \mathscr{P}_1$, ce qui est équivalent à $\mathscr{P}_0 \perp\!\!\!\!\!\underline{\;\;\;} \mathscr{P}_1$ (déf. 3.1.2.*a*). Soit $\{C_P, P \in \mathscr{P}\}$ une suite de nombres réels strictement positifs, telle que

$$\sum_{P \in \mathscr{P}_0} C_P = \sum_{P \in \mathscr{P}_1} C_P = 1.$$

Soient R_0 et R_1 les probabilités sur $(\Omega, \mathscr{A})$ définies par :

$$R_0 = \sum_{P \in \mathscr{P}_0} C_P . P, \qquad R_1 = \sum_{P \in \mathscr{P}_1} C_P . P.$$

Il est clair que $R_0 \perp R_1$. Le lemme précédent nous assure alors l'existence d'un test adapté convergent en moyenne de R_0 contre R_1, soit $\{\Phi_i, i \in I\}$. On a donc :

$$\lim_I E_{R_0}[\Phi_i] = \lim_I E_{R_1}[1 - \Phi_i] = 0.$$

Or

$$E_{R_0}[\Phi_i] = \sum_{P \in \mathscr{P}_0} C_P . E_P[\Phi_i],$$

donc :

$$(\forall P \in \mathscr{P}_0), \qquad E_P[\Phi_i] \leq \frac{1}{C_P} E_{R_0}[\Phi_i].$$

De même :

$$(\forall P \in \mathscr{P}_1), \qquad E_P[1 - \Phi_i] \leq \frac{1}{C_P} E_{R_1}[1 - \Phi_i].$$

La proposition en résulte.

Dans la démonstration de la proposition précédente, nous avons utilisé l'orthogonalité des barycentres sur $\mathscr{P}_0$ et $\mathscr{P}_1$. Nous allons étendre cette méthode au cas d'une famille $\mathscr{P}$ quelconque.

Considérons un modèle statistique asymptotique. Soit μ une probabilité sur l'espace de paramètres $(\Theta, \mathscr{H})$. Pour tout $T \in \mathscr{H}$, définissons une mesure R_T sur $(\Omega, \mathscr{A})$ de la façon suivante :

$$\begin{array}{lll} \text{si } \mu(T) > 0, & R_T(A) = \dfrac{1}{\mu(T)} \displaystyle\int_T P^\theta(A) d\mu(\theta), & (\forall A \in \mathscr{A}), \\ \text{si } \mu(T) = 0, & R_T = 0. & \end{array}$$

Il est facile de voir que : $(\forall T \in \mathscr{H})$, $R_T \ll R_\Theta$.

Soit donc, pour chaque $T \in \mathscr{H}$, une v. a. r. positive Q_T sur $(\Omega, \mathscr{A})$ telle que :

$$Q_T \in \mu(T) \frac{dR_T}{dR_\Theta}$$

Pour tout $i \in I$ et tout $T \in \mathscr{H}$, notons R_{iT} la restriction de R_T à $(\Omega, \mathscr{A}_i)$ et Q_{iT} une densité correspondante.

LEMME 6.3. — Pour que la famille $\mathscr{P}$ soit μ-fortement séparée (déf. 3.4), il faut et il suffit que :

$$(\forall T \in \mathscr{H} : \mu(T) \in]0, 1[) \lim_I \text{ (en norme } L^1(\Omega, \mathscr{A}, R_T)) \, Q_{iT} = 1.$$

Si de plus l'ensemble I est dénombrable totalement ordonné, la condition précédente est équivalente à :

$$\begin{array}{l} (\forall T \in \mathscr{H} : \mu(T) \in]0, 1[), (\exists H \in \mathscr{H} : \mu(H) = 0), \\ (\forall \theta \in T \setminus H), \lim_I (p^\theta\text{-p. s.}) Q_{iT} = 1. \end{array}$$

Démonstration. — Pour tout $T \in \mathscr{H}$ tel que $\mu(T) \in]0, 1[$, $\{ Q_{iT}, \mathscr{A}_i, i \in I \}$ est une martingale positive équi-intégrable dans $(\Omega, \mathscr{A}, R_\Theta)$, et par suite $\lim_I$ (norme $L^1(\Omega, \mathscr{A}, R_\Theta)) Q_{iT} = Q_T$.

Comme $\mu(\mathrm{T})\mathrm{R}_\mathrm{T} \leq \mathrm{R}_\Theta$, on a aussi :

$$\lim_{\mathrm{I}} \left(\text{norme } \mathrm{L}^1(\Omega, \mathscr{A}, \mathrm{R}_\mathrm{T})\right) \mathrm{Q}_{i\mathrm{T}} = \mathrm{Q}_\mathrm{T}.$$

Or les propositions suivantes sont clairement équivalentes :

a) $\mathscr{P}$ est μ-fortement séparée.

b) ($\forall \mathrm{T} \in \mathscr{H} : \mu(\mathrm{T}) \in]0, 1[$), $\mathrm{R}_\mathrm{T} \perp \mathrm{R}_{\mathrm{T}^c}$.

c) ($\forall \mathrm{T} \in \mathscr{H} : \mu(\mathrm{T}) \in]0, 1[$), $\mathrm{Q}_\mathrm{T} = 1[\mathrm{R}_\mathrm{T}]$.

La 1re partie du lemme en résulte.

Dans le cas ou l'ensemble I est dénombrable totalement ordonné, la martingale $\{\mathrm{Q}_{i\mathrm{T}}, \mathscr{A}_i, i \in \mathrm{I}\}$ converge aussi R_Θ-presque sûrement, donc R_T-presque sûrement, vers Q_T. Alors pour que la famille $\mathscr{P}$ soit μ-fortement séparée, il faut et il suffit que :

$$(\forall \mathrm{T} \in \mathscr{H} : \mu(\mathrm{T}) \in]0, 1[), \lim_{\mathrm{I}} (\mathrm{R}_\mathrm{T}\text{-p. s.})\, \mathrm{Q}_{i\mathrm{T}} = 1.$$

La 2e partie du lemme en résulte.

PROPOSITION 6.4. — Soit un modèle statistique asymptotique tel que l'ensemble d'indices I soit dénombrable totalement ordonné. Les propositions suivantes sont équivalentes :

a) $\mathscr{P}$ est μ-fortement séparée.

b) ($\forall \mathrm{T} \in \mathscr{H} : \mu(\mathrm{T}) \in]0, 1[$), ($\exists \mathrm{N} \in \mathscr{H} : \mu(\mathrm{N}) = 0$) (il existe un test adapté consistant en moyenne de $\mathrm{T} \setminus \mathrm{N}$ contre $\mathrm{T}^c \setminus \mathrm{N}$).

Démonstration.

a) $\Rightarrow$ *b)*. Soit $\mathrm{T} \in \mathscr{H}$ tel que $\mu(\mathrm{T}) \in]0, 1[$, et soit $r \in]0, 1[$. Posons

$$\Phi_i = 1_{\{\mathrm{Q}_{i\mathrm{T}} \leq r\}} \qquad (i \in \mathrm{I}).$$

Alors, le lemme précédent nous affirme l'existence d'une partie μ-négligeable $\mathrm{H} \in \mathscr{H}$ telle que :

$$(\forall \theta \in \mathrm{T} \setminus \mathrm{N}), \lim_{\mathrm{I}} \mathrm{E}^\theta[\Phi_i] = \lim_{\mathrm{I}} \mathrm{P}^\theta[1 - \mathrm{Q}_{i\mathrm{T}} \geq 1 - r] = 0$$

et :

$$(\forall \theta \in \mathrm{T}^c \setminus \mathrm{N}), \lim_{\mathrm{I}} \mathrm{E}^\theta[\Phi_i] = 1 - \lim_{\mathrm{I}} \mathrm{P}^\theta[\mathrm{Q}_{i\mathrm{T}} > r] = 1.$$

b) $\Rightarrow$ *a)*. Soit $\mathrm{T} \in \mathscr{H}$ tel que $\mu(\mathrm{T}) \in]0, 1[$, et soit $\{\Phi_i, i \in \mathrm{I}\}$ un test adapté convergent de $\mathrm{T} \setminus \mathrm{N}$ contre $\mathrm{T}^c \setminus \mathrm{N}$, ou N est un ensemble μ-négligeable de $\mathscr{H}$.

On a :

$$\mathrm{E}_{\mathrm{R}_\mathrm{T}}[\Phi_i] = \frac{1}{\mu(\mathrm{T})} \int_\mathrm{T} \mu(d\theta) \int_\Omega \Phi_i(\omega) \mathrm{P}^\theta(d\omega) = \frac{1}{\mu(\mathrm{T})} \int_\mathrm{T} \mu(d\theta) \mathrm{E}^\theta[\Phi_i].$$

Par suite

$$\lim_{I} E_{R_T}[\Phi_i] = 0.$$

De même

$$\lim_{I} E_{R_{T^c}}[\Phi_i] = 1.$$

On en déduit que $R_T \perp R_{T^c}$ (lemme 6.1), donc que la famille $\mathscr{P}$ est μ-fortement séparée.

IV.7. CONVERGENCE DES TRANSITIONS *A POSTERIORI*

Les variables aléatoires $\{Q_{iT}, i \in I, T \in \mathscr{H}\}$ définies au paragraphe précédent peuvent s'interpréter de la facon suivante.

Considérons, sur l'espace mesurable $(\Theta \times \Omega, \mathscr{H} \otimes \mathscr{A})$ la probabilité P^μ définie par ses valeurs sur les pavés mesurables :

$$P^\mu(T \times A) = \int_T d\mu(\theta) P^\theta(A) \qquad (T \in \mathscr{H}, A \in \mathscr{A}).$$

La restriction de P^μ à la tribu des cylindres de base dans $\mathscr{A}$ (que nous noterons encore $\mathscr{A}$) donne R_Θ, et on a, compte tenu de la définition de Q_T :

$$P^\mu(T \times A) = \mu(T) R_T(A) = \int_A Q_T dR_\Theta \qquad (T \in \mathscr{H}, A \in \mathscr{A})$$

d'où

$$Q_T \in (P^\mu)^{\mathscr{A}}(T) \qquad (T \in \mathscr{H}).$$

Autrement dit, pour tout $T \in \mathscr{H}$, la variable aléatoire Q_T est une version de la P^μ-probabilité $\mathscr{A}$-conditionnelle de l'ensemble $T \in \mathscr{H}$.

DÉFINITION 7.1. — Avec les notations précédentes, nous dirons que le modèle statistique asymptotique vérifie la condition (Δ) pour μ, si, pour tout $i \in I$, la restriction P_i^μ de la probabilité P^μ à $(\Theta \times \Omega, \mathscr{H} \otimes \mathscr{A}_i)$ admet une désintégration $Q_{i\Theta}^\Omega$ par rapport à la projection de $\Theta \times \Omega$ sur Ω.

Cette désintégration est alors appelée transition *a posteriori* (à l'instant i).

PROPOSITION 7.2. — Pour que le modèle statistique asymptotique vérifie la condition (Δ) il suffit que l'une des conditions suivantes soit réalisée :

a) Θ est un espace localement compact, et $\mathscr{H}$ sa tribu borélienne. Pour tout $i \in I$, la tribu $\mathscr{A}_i$ est complète pour la probabilité $R_{i\Theta}$.

b) Pour tout $i \in \mathrm{I}$, il existe une mesure σ-finie λ_i sur $(\Omega, \mathscr{A}_i)$ telle que, pour tout $\theta \in \Theta$, la probabilité P_i^θ est absolument continue par rapport à λ_i, et la densité $\dfrac{d\mathrm{P}_i^\theta}{d\lambda_i}$ possède une version qui est $\mathscr{H} \otimes \mathscr{A}_i$-mesurable.

Démonstration. — C'est une conséquence du théorème de désintégration (cf. III.2.1).

Soit f_i^θ une version de la densité $\dfrac{d\mathrm{P}_i^\theta}{d\lambda_i}$ telle que l'application

$$(\theta, \omega) \rightsquigarrow f_i^\theta(\omega)$$

soit $\mathscr{H} \otimes \mathscr{A}_i$-mesurable.

Soit $\mathrm{T} \in \mathscr{H}$ tel que $\mu(\mathrm{T}) > 0$. En appliquant le théorème de Fubini, on a, $\forall i \in \mathrm{I}$ et $\forall \mathrm{A} \in \mathscr{A}$

$$\mathrm{R}_{i\mathrm{T}}(\mathrm{A}) = \frac{1}{\mu(\mathrm{T})} \int_{\mathrm{T}} \mathrm{P}_i^\theta(\mathrm{A}) d\mu(\theta) = \frac{1}{\mu(\mathrm{T})} \int_{\mathrm{A}} d\lambda_i(\omega) \int_{\mathrm{T}} f_i^\theta(\omega) d\mu(\theta).$$

par suite $\mathrm{R}_{i\mathrm{T}} \ll \mathrm{R}_{i\Theta} \ll \lambda_i$

$$\frac{1}{\mu(\mathrm{T})} \int_{\mathrm{T}} f_i^\theta d\mu(\theta) \in \frac{d\mathrm{R}_{i\mathrm{T}}}{d\lambda_i}.$$

donc

$$\mathrm{Q}_{i\mathrm{T}} = \frac{\displaystyle\int_{\mathrm{T}} f_i^\theta d\mu(\theta)}{\displaystyle\int_{\Theta} f_i^\theta d\mu(\theta)} [\mathrm{R}_{i\Theta}].$$

Cette formule restant vraie même si $\mu(\mathrm{T}) = 0$, nous avons exhibé, pour chaque $\mathrm{T} \in \mathscr{H}$, une version de la variable aléatoire $\mathrm{Q}_{i\mathrm{T}}$ qui définit une transition de probabilité $\mathrm{Q}_{i\Theta}^\Omega$.

La proposition est démontrée.

Remarque. — Sous la condition *b*) de la proposition précédente, la transition de probabilité que nous venons d'exhiber est habituellement appelée transition *a posteriori*, et donné par le théorème de Bayes. Ceci justifie notre définition dans le cas général.

Si le modèle statistique asymptotique vérifie la condition (Δ), et si l'on suppose que l'espace des décisions $(\mathrm{D}, \mathscr{D})$ est l'espace $(\Theta, \mathscr{H})$, l'application étant l'application identique, alors la famille des mesures *a posteriori* $\{\mathrm{Q}_{i\Theta}^\Omega, i \in \mathrm{I}\}$ est une procédure statistique (cf. 1.2). Nous allons donner une condition nécessaire et suffisante pour que cette procédure soit consistante, ce qui implique en particulier que la limite est indépendante de la mesure μ (dite *a priori*).

La démonstration de ce résultat dans des cas particuliers est connue depuis longtemps ([*3*], [*5*], [*6*] par exemple).

LEMME 7.3. — Soit un modèle statistique asymptotique vérifiant la condition (Δ), tel que l'ensemble I soit dénombrable et totalement ordonné, que la tribu $\mathscr{H}$ sur Θ admette une base dénombrable $\mathscr{B}$, et que l'espace des décisions (D, $\mathscr{D}$) soit isomorphe à (Θ, $\mathscr{H}$). Alors la famille $\mathscr{P}$ est μ-fortement séparée si et seulement si la famille $\{ Q_{i\Theta}^{\Omega}, i \in I \}$ des transitions *a posteriori* vérifie la condition suivante :

$$(\exists H \in \mathscr{H} : \mu(H) = 0), (\forall T \in s(\mathscr{B}) : \mu(T) > 0), (\forall \theta \in T \setminus H)$$
$$\lim_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 1,$$

où $s(\mathscr{B})$ est l'algèbre de parties de Θ engendrée par $\mathscr{B}$.

Démonstration.

1) *Condition nécessaire.* — En vertu du lemme (6.3), la propriété « $\mathscr{P}$ est μ-fortement séparée » est équivalente à :

$$(\forall T \in \mathscr{H} : \mu(T) > 0), (\exists H \in \mathscr{H} : \mu(H) = 0), (\forall \theta \in T \setminus H)$$
$$\lim_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 1.$$

Comme le modèle vérifie la condition (Δ), la propriété précédente est équivalente à :

$$(\forall T \in \mathscr{H}), (\exists H \in \mathscr{H} : \mu(H) = 0) :$$
$$\begin{cases} (\forall \theta \in T \setminus H, \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 1, \\ (\forall \theta \in T^c \setminus H, \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 0. \end{cases}$$

Cette dernière condition permet d'associer à tout élément B de $\mathscr{B}$ un élément H_B de $\mathscr{H}$, μ-négligeable, tel que :

$$\begin{cases} (\forall \theta \in B \setminus H_B), \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., B) = 1, \\ (\forall \theta \in B^c \setminus H_B), \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., B) = 0. \end{cases}$$

Soit $H_0 = \bigcup\limits_{B \in \mathscr{B}} H_B$. Évidemment $H_0 \in \mathscr{H}$ et $\mu(H_0) = 0$.

Soit

$$\mathscr{C} = \left\{ T \in \mathscr{H} \,\middle|\, \begin{array}{l} (\forall \theta \in T \setminus H_0), \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 1 \\ (\forall \theta \in T^c \setminus H_0), \lim\limits_{I} (P^{\theta}\text{-p. s.}) Q_i(., T) = 0 \end{array} \right\}$$

Alors $\mathscr{C}$ est une algèbre de parties de Θ. En effet $\mathscr{C}$ est évidemment fermé pour la complémentation. Soient T et T' deux éléments de $\mathscr{C}$. Si $\theta \in T \setminus H_0$, alors $\lim_I (P^\theta\text{-p. s.}) Q_i(., T) = 1$, donc *a fortiori* $\lim_I (P^\theta\text{-p. s.}) Q_i(., T \cup T') = 1$.
De même si $\theta \in T' \setminus H_0$.
Si $\theta \in (T \cup T')^c \setminus H_0$, alors

$$\lim_I (P^\theta\text{-p. s.}) Q_i(., T) = 0 \qquad \text{et} \qquad \lim_I (P^\theta\text{-p. s.}) Q_i(., T') = 0,$$

donc

$$\lim_I (P^\theta\text{-p. s.}) Q_i(., T \cup T') = 0.$$

Par suite $T \cup T' \in \mathscr{C}$. D'autre part il est clair que $\mathscr{C}$ contient $\mathscr{B}$. Par suite $\mathscr{C}$ contient $s(\mathscr{B})$ et la condition nécessaire en résulte.

2) *Condition suffisante.* — Soit $T \in \mathscr{H}$ tel que : $0 < \mu(T) < 1$. En vertu du lemme (6.3), nous devons lui associer un ensemble $H_T \in \mathscr{H}$, μ-négligeable, tel que

$$(\forall \theta \in T \setminus H_T), \lim_I (P^\theta\text{-p. s.}) Q_i(., T) = 1.$$

Or à tout $\varepsilon \in]0, \mu(T)[$, nous pouvons associer un élément $T_\varepsilon \in s(\mathscr{B})$ tel que :

$$\mu(T \triangle T_\varepsilon) \leq \varepsilon,$$

donc tel que :

$$(\forall i \in I), Q_i(., T_\varepsilon \triangle T) \leq \varepsilon[R_\Theta].$$

Par suite on a :

$$(\forall \varepsilon \in]0, \mu(T)[), (\forall i \in I), Q_i(., T) \geq Q_i(., T_\varepsilon) - \varepsilon[R_\Theta].$$

Soit $\{ \varepsilon_n, n \in \mathbb{N} \}$ une suite de nombres réels positifs telle que :

$$\sum_n \varepsilon_n \in]0, \mu(T)[.$$

D'après ce qui précède, nous pouvons associer à tout couple $(n, i) \in \mathbb{N} \times I$ un élément $H_{n,i} \in \mathscr{H}$, μ-négligeable, tel que :

$$(\forall \theta \in H_{n,i}^c), Q_i(., T) \geq Q_i(., T_{\varepsilon_n}) - \varepsilon_n[P^\theta].$$

Posons $H' = \bigcup_{(n,i)} H_{n,i}$. Il est clair que H' est un élément de $\mathscr{H}$, μ-négligeable, et que l'on a :

$$(\forall \theta \in H'^c), (\forall (n, i) \in \mathbb{N} \times I), Q_i(., T) \geq Q_i(., T_{\varepsilon_n}) - \varepsilon_n[P^\theta].$$

Or par hypothèse, pour tout $n \in \mathbb{N}$, on a :

$$(\forall \theta \in \mathrm{T}_{\varepsilon_n} \setminus \mathrm{H}), \lim_{\mathrm{I}} (\mathrm{P}^\theta\text{-p. s.})\, \mathrm{Q}_i(., \mathrm{T}_{\varepsilon_n}) = 1.$$

où H est un élément μ-négligeable de $\mathscr{H}$ donné par la condition du lemme. Par suite, en posant

$$\mathrm{T}' = \liminf_n \mathrm{T}_{\varepsilon_n}, \qquad \text{si} \qquad \theta \in \mathrm{T}' \setminus (\mathrm{H} \cup \mathrm{H}'),$$

à partir d'un certain rang n, on a :

$$\liminf_{\mathrm{I}} \mathrm{Q}_i(., \mathrm{T}) \geq 1 - \varepsilon_m[\mathrm{P}^\theta],$$

d'où :

$$\lim_{\mathrm{I}} (\mathrm{P}^\theta\text{-p. s.})\, \mathrm{Q}_i(., \mathrm{T}) = 1.$$

La condition suffisante est alors démontrée, puisque :

$$\mathrm{T}' \setminus (\mathrm{H} \cup \mathrm{H}') \supset \mathrm{T} \setminus [(\mathrm{T} \setminus \mathrm{T}') \cup \mathrm{H} \cup \mathrm{H}']$$

et que, pour tout $n \in \mathbb{N}$, on a :

$$\mu[(\mathrm{T} \setminus \mathrm{T}') \cup \mathrm{H} \cup \mathrm{H}'] = \mu(\mathrm{T} \setminus \mathrm{T}') \leq \sum_{m \geq n} \mu(\mathrm{T} \setminus \mathrm{T}_{\varepsilon_m}) \leq \sum_{m \geq n} \varepsilon_m,$$

ce qui entraîne : $\mu[(\mathrm{T} \setminus \mathrm{T}') \cup \mathrm{H} \cup \mathrm{H}'] = 0.$

DÉFINITION 7.4. — Étant donné un modèle statistique asymptotique et une probabilité μ sur l'espace de paramètres $(\Theta, \mathscr{H})$, nous dirons qu'une procédure statistique $\{ \mathrm{S}_{i\mathrm{D}}^{\Omega}, i \in \mathrm{I} \}$ est μ-presque sûrement consistante pour $\mathscr{P}$ si :

$$(\exists \mathrm{H} \in \mathscr{H} : \mu(\mathrm{H}) = 0) : (\forall \underline{d} \text{ ouvert de D} : \mu[h^{-1}(\underline{d})] > 0),$$
$$(\forall \theta \in h^{-1}(\underline{d}) \setminus \mathrm{H}), \lim_{\mathrm{I}} (\mathrm{P}^\theta\text{-p. s.})\, \mathrm{S}_i(., \underline{d}) = 1.$$

THÉORÈME 7.5 [*10*]. — Soit un modèle statistique asymptotique vérifiant la condition (Δ), tel que l'ensemble I soit dénombrable et totalement ordonné, que la tribu $\mathscr{H}$ admette une base dénombrable $\mathscr{B}$, et que l'espace des décisions $(\mathrm{D}, \mathscr{D})$ soit isomorphe à $(\Theta, \mathscr{H})$, et vérifiant la condition suivante :

$$(\mathrm{A}) \qquad (\forall \mathrm{T} \text{ ouvert de } \Theta : \mu(\mathrm{T}) > 0), (\forall \theta \in \mathrm{T}),$$
$$(\exists \mathrm{T}_0 \in s(\mathscr{B}) : \mu(\mathrm{T}_0) > 0) : \theta \in \mathrm{T}_0 \subset \mathrm{T}.$$

Alors si la famille $\mathscr{P}$ est μ-fortement séparée, la famille $\{ \mathrm{Q}_{i\Theta}^{\Omega}, i \in \mathrm{I} \}$ des transitions *a posteriori* est μ-presque sûrement consistante pour $\mathscr{P}$. Si de plus la mesure μ est régulière, alors la réciproque est vraie.

Remarque. — Pour que la condition (A) soit vérifiée, et que la tribu $\mathscr{H}$ des boréliens de Θ soit à base dénombrable, il suffit que l'espace topologique Θ soit métrisable et séparable.

En effet si ρ est une distance sur Θ compatible avec sa topologie et si Θ_0 est une suite dense dans Θ, alors la famille $\mathscr{B}$ des boules ouvertes de centres $\theta \in \Theta_0$ et de rayons rationnels est bien une base dénombrable de la tribu borélienne $\mathscr{H}$ de Θ. D'autre part si T est un ouvert de Θ de μ-mesure positive, comme l'ouvert T est la réunion des boules de $\mathscr{B}$ qu'il contient, on pourra trouver un nombre fini de boules de $\mathscr{B}$ dont la réunion T_0 satisfait à la condition (A).

Démonstration du théorème.

1) Supposons la famille $\mathscr{P}$ μ-fortement séparée. Alors le lemme précédent nous affirme l'existence d'un élément $H \in \mathscr{H}$, μ-négligeable, tel que :

$$(\forall T \in s(\mathscr{B}) : \mu(T) > 0), (\forall \theta \in T \setminus H), \lim_{I} (P^\theta\text{-p. s.}) \, Q_i(., T) = 1.$$

Soit T un ouvert de Θ tel que $\mu(T) > 0$, et soit $\theta \in T \setminus H$.

En vertu de la condition (A), on peut trouver un élément $T_0 \in s(\mathscr{B})$ tel que :

$$\mu(T_0) > 0 \qquad \text{et} \qquad \theta \in T_0 \subset T.$$

Alors, en vertu du lemme (7.3), on a :

$$\lim_{I} (P^\theta\text{-p. s.}) \, Q_i(., T_0) = 1,$$

donc *a fortiori* :

$$\lim_{I} (P^\theta\text{-p. s.}) \, Q_i(., T) = 1.$$

2) Inversement, la mesure μ étant régulière, supposons que la famille des transitions *a posteriori* soit μ-presque sûrement consistante pour $\mathscr{P}$. Soit $T \in \mathscr{H}$ tel que $\mu(T) \in]0, 1[$. Soit $\{ \varepsilon_n, n \in \mathbb{N} \}$ une suite de nombres réels positifs décroissant vers 0. Comme la mesure μ est régulière, nous pouvons trouver une suite décroissante $\{ T_n, n \in \mathbb{N} \}$ d'ouverts contenant T telle que :

$$(\forall n \in \mathbb{N}), \mu(T_n \setminus T) \leq \varepsilon_n,$$

donc telle que :

$$(\forall n \in \mathbb{N}), (\forall i \in I), Q_i(., T_n \setminus T) \leq \varepsilon_n[R_\Theta].$$

Par suite, à tout couple $(n, i) \in \mathbb{N} \times I$, nous pouvons associer un élément $H_{n,i} \in \mathscr{H}$, μ-négligeable, tel que :

$$(\forall \theta \in H^c_{n,i}), Q_i(., T_n \setminus T) \leq \varepsilon_n[P^\theta].$$

Posons $H' = \bigcup_{(n,i)} H_{n,i}$. Évidemment $H' \in \mathscr{H}$ et $\mu(H') = 0$, et on a :

$$(\forall \theta \in H'^c), (\forall (n, i) \in \mathbb{N} \times I), Q_i(., T) \geq Q_i(., T_n) - \varepsilon_n[P^\theta].$$

Nous en déduisons que, puisque la famille des transitions *a posteriori* est μ-presque sûrement consistante pour $\mathscr{P}$, nous pouvons trouver un élément $H \in \mathscr{H}$, μ-négligeable, tel que :

$$(\forall n \in \mathbb{N}), (\forall \theta \in T \setminus (H \cup H')), \liminf_{I} Q_i(., T) \geq 1 - \varepsilon_n[P^\theta],$$

donc tel que :

$$(\forall \theta \in T \setminus (H \cup H')), \lim_{I} (P^\theta\text{-p. s.}) Q_i(., T) = 1,$$

ce qui, en vertu du lemme (6.3), démontre la deuxième partie du théorème.

Remarque. — Si $\mu(T) = 0$, on a :

$$(\forall i \in I), Q_i(., T) = 0[R_\Theta].$$

Donc si la possibilité *a priori* μ ne charge pas un ensemble $T \in \mathscr{H}$, on ne peut espérer la convergence des mesures *a posteriori* de T vers 1. Par suite il est naturel de choisir, si cela est possible, une probabilité *a priori* μ chargeant tout ouvert non vide de Θ. C'est par exemple le cas si Θ est un groupe topologique localement compact et si μ est équivalente à sa mesure de Haar. Alors, avec les notations précédentes, si $\mathscr{P}$ est μ-fortement séparée, et si on considère la restriction du modèle statistique asymptotique à $\Theta' = \Theta \setminus H$, que l'on munit de la topologie induite, alors la procédure statistique $\{ Q_{i\Theta}^{\Omega}, i \in I \}$ est, sous les conditions du théorème (7.5), μ-presque sûrement consistante pour $\mathscr{P}' = \{ P^\theta, \theta \in \Theta' \}$. Cela explique le rôle joué par la mesure de Lebesgue dans les techniques dites « bayésiennes ».

BIBLIOGRAPHIE

[1] E. W. BARANKIN et J. GURLAND, *On asymptotically normal, efficient estimators, I.* Univ. Calif. Publ. Stat., vol. **1**, 1951, p. 89-130.

[2] A. BERGER, On uniformly consistent tests. *Ann. Math. Stat.*, vol. **22**, 1951, n° 2, p. 289-293.

[3] L. LE CAM, *On some asymptotic properties of maximum likelikood and related Bayes estimates.* Univ. Calif. Publ. Stat., vol. **1**, 1953, p. 277-300.

[4] L. LE CAM, *Asymptotic theory of estimation and testing hypotheses. Proceedings 3e Berkeley Symposium.* Univ. Calif. Press, 1956, p. 129-156.

[5] L. LE CAM, *Les propriétés asymptotiques des solutions de Bayes.* Public. de l'I. S. U. P., vol. **7**, 1958, fasc. 3-4, p. 17-35.

[*6*] J. L. Doob, Applications of the theory of martingales. Le calcul des probabilités et ses applications. Coll. Int. du C. N. R. S., Paris, 1949, p. 23-27.

[*7*] G. A. Hunt, *Martingales et processus de Markov*. Dunod, Paris, 1966.

[*8*] S. Kakutani, On equivalence of infinite product measures. *Ann. Math. Stat.*, vol. **22**, 1948, p. 214-224.

[*9*] C. Kraft, *Some conditions for consistency and uniform consistency of statistical procedures*. Univ. Calif. Publ. Stat., vol. **2**, 1955, p. 125-142.

[*10*] F. Martin et D. Vaguelsy, Sur la convergence des mesures *a posteriori*. *C. R. Acad. Sci. Paris*, 1968.

[*11*] P. A. Meyer, *Probabilité et potentiel*. Hermann, Paris, 1966.

[*12*] J. Neyman, Contributions to the theory of χ^2 test. *Proceedings Berkeley Symposium, Univ. Calif. Press*, 1949, p. 239-273.

[*14*] D. Vaguelsy, Exhaustivité asymptotique et sous-tribus convergentes. *C. R. Acad. Sci. Paris*, t. **264**, 1967, p. 764-766.

[*15*] Y. Ramakrishna Sarma, Sur les tests et sur l'estimation de paramètres pour certains processus stochastiques stationnaires. *Thèse de Doctorat*, Paris, 1968.

Manuscrit reçu le 11 juin 1969.