



HAL
open science

La décantation en estimation asymptotique

Raymond Moche

► **To cite this version:**

Raymond Moche. La décantation en estimation asymptotique. Annales de l'ISUP, 1988, XXXIII (2), pp.39-77.
<hal-03672768>

HAL Id: hal-03672768

<https://hal.science/hal-03672768v1>

Submitted on 19 May 2022

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

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



HAL Authorization

Pub. Inst. Stat. Univ.
XXXIII, fasc. 2, 1988, 39 à 77

LA DECANTATION EN ESTIMATION ASYMPTOTIQUE

Raymond MOCHÉ
Laboratoire de Statistique et Probabilités (M2)
Université des Sciences et Techniques de
Lille Flandres Artois
59655 - Villeneuve d'Ascq Cedex (France)

Summary. - First we prove a new sufficient condition for the uniform asymptotic separation of two composite hypothesis of a general statistical model $(\theta, (X_n, \mathcal{B}_n), H)$. We only require that the conditional distributions of the observed random variables satisfy decantation conditions.

In part II, we deal with the decantation problem : given two families of probability laws H_0 and H_1 on a measurable space (Ω, \mathcal{A}) , when does there exist a function $\psi : (\Omega, \mathcal{A}) \rightarrow [0, 1]$ such that :

$$\inf \left(\int \psi \cdot dP ; P \in H_1 \right) > \sup \left(\int \psi \cdot dQ ; Q \in H_0 \right) ?$$

Using convex hulls method and decanted balls method, we then prove usable corollaries of the first result.

Finally applications to the asymptotic estimation of a general parameter are given in part III.

Mots clefs : Statistique des processus, décantation, séparation asymptotique uniforme, hypothèses multiples, estimation asymptotique.

Classification AMS (MOS) : 62M09

Résumé : Nous présentons d'abord un nouveau théorème de séparation asymptotique en forme de deux hypothèses multiples H_0 et H_1 dans un modèle statistique général $(\theta, (X_n, \mathcal{B}_n), H)$ où le processus $(X_n)_{n \geq 1}$ des observations est seulement assujéti à des conditions sur ses lois conditionnelles, dites conditions de décantation, de type suivant :

$$\forall n \geq 2, \forall x \in \prod_{i=1}^{n-1} X_i, \exists \psi_{x^{(n-1)}} : (X_n, \mathcal{B}_n) \rightarrow [0, 1]$$

telle que :

$$\inf_{P \in H_1} \left(\int \psi_{x^{(n-1)}} \cdot dP_{X_n}^{(X_1, \dots, X_{n-1})=x^{(n-1)}} \right) > \sup_{P \in H_0} \left(\int \psi_{x^{(n-1)}} \cdot dP_{X_n}^{(X_1, \dots, X_{n-1})=x^{(n-1)}} \right)$$

Dans la seconde partie, nous étudions diverses conditions suffisantes pour que deux familles de lois soient décantées, portant sur les enveloppes convexes de familles, ou sur des propriétés de décantation de certaines distances, ou de certains préordres sur l'ensemble des lois concernées. Nous en déduisons des versions applicables du théorème de séparation asymptotique uniforme.

Nous les utilisons dans la troisième partie pour estimer un paramètre quelconque $f : H \rightarrow (E, d)$, à valeurs dans un espace métrique E au moins séparé. Nous obtenons des estimateurs qui convergent suivant différents modes suivant que a , ou non, certaines propriétés de compacité supplémentaires, ou que l'hypothèse générale H est munie d'une probabilité a priori.

0 - INTRODUCTION.-

L'emploi de la décantation en statistique asymptotique s'est développé à l'impulsion de J. Geffroy, à partir de 1974 ⁽⁴⁾ dans le cas où les observations sont indépendantes sans être nécessairement équidistribuées, puis, après 1976 ⁽⁵⁾ pour des observations quelconques. Rappelons que l'un des traits les plus remarquables de cette méthode est qu'elle permet d'établir l'existence de suites convergentes de tests au

estimeurs par des procédés essentiellement constructifs, et qui ne font intervenir aucun moment de lois des grands nombres. La théorie repose, outre diverses constructions, sur un théorème de séparation asymptotique uniforme de deux hypothèses multiples (⁵, th. 1) dont nous avons obtenu une version sensiblement améliorée que nous présentons ici. Nous en profitons pour faire le point sur l'ensemble de la méthode, grâce à des résultats originaux, démontrés délibérément dans le cadre le plus général possible. Ceci dit, la décantation n'est pas qu'une théorie abstraite : elle est dès à présent appliquée avec succès à de grands problèmes de la statistique asymptotique tels que l'estimation par le maximum de vraisemblance, dans le cas d'observations quelconques (¹⁹ p. 42-69), ou l'estimation d'un signal dans un bruit blanc (¹⁹ p. 70-90). Elle paraît également apte à l'évaluation des vitesses de convergence de ses estimateurs, ouvrant ainsi de très belles perspectives de développement.

I - SÉPARATION ASYMPTOTIQUE UNIFORME DE DEUX HYPOTHESES MULTIPLES.

1 - Théorème 1. - Soit (Ω, A) un espace mesurable, $(\mathcal{B}_n)_{n \geq 0}$ une suite croissante de sous-tribus de A , avec $\mathcal{B}_0 = (\emptyset, \Omega)$, $(\psi_n)_{n \geq 1}$ et $(p_n)_{n \geq 1}$ deux suites de fonctions réelles définies sur Ω , et $(a_n)_{n \geq 1}$ une suite de nombres réels vérifiant les propriétés suivantes : $\forall n \geq 1$,

- ψ_n est \mathcal{B}_n -mesurable,
- p_n est \mathcal{B}_{n-1} -mesurable,
- $0 \leq \psi_n, p_n, a_n, p_n + a_n \leq 1$.

On désigne par $(C_n)_{n \geq 1}$ la suite des événements

$$C_n = \left(\sum_{j=1}^n a_j \psi_j \geq \sum_{j=1}^n a_j \left(p_j + \frac{1}{2} a_j \right) \right) : \text{elle est adaptée à } (\mathcal{B}_n)_{n \geq 1}.$$

Soit H_0 et H_1 deux familles de lois de probabilité μ sur (Ω, A) telles que :

$\forall n \geq 1$, il existe une version régulière μ_n^B de la probabilité μ conditionnée par B_n , et telles que :

$$- \forall P \in H_0, \quad E_P^{B_{n-1}}(\psi_n) \underset{P\text{-p.s.}}{\leq} p_n$$

$$- \forall Q \in H_1, \quad E_Q^{B_{n-1}}(\psi_n) \underset{Q\text{-p.s.}}{\geq} p_n + a_n$$

Alors nous avons :

$$\cdot \text{Sup}(P(C_n) ; P \in H_0) \leq e^{-3/8 \sum_{j=1}^n a_j^2}$$

$$\cdot \text{Inf}(Q(C_n) ; Q \in H_1) \geq 1 - e^{-3/8 \sum_{j=1}^n a_j^2}$$

Nous utiliserons dans la démonstration les résultats suivants :

soit (Ω, A, μ) un espace de probabilité, B une sous-tribu de A pour laquelle il existe une version régulière μ^B de la probabilité μ conditionnée par B .

a) Inégalité de Markov conditionnelle : Si X et Y sont 2 v.a.r. sur (Ω, A, μ) , X étant de plus B -mesurable, nous avons :

$$\mu^B(Y \geq X) \underset{\mu\text{-p.s.}}{\leq} e^{-X} \cdot E_{\mu^B}(e^Y).$$

b) Une inégalité de W. Hoeffding (⁸, inégalité (4.16)) : si Z est une v.a.r. sur (Ω, A, μ) prenant ses valeurs dans l'intervalle $[a, b]$, nous avons

$$E_{\mu}(e^{Z-E(Z)}) \leq e^{1/8(b-a)^2}.$$

c) Forme conditionnelle : en appliquant cette inégalité à la loi conditionnelle μ_Z^B de Z sachant B , nous obtenons :

$$E_{\mu}^B(e^{Z-E_{\mu}^B(Z)}) \leq e^{1/8(b-a)^2} \quad (6)$$

μ -p.s.

Démonstration : $\forall n \geq 1$, posons $Y_n = \sum_{j=1}^n a_j \psi_j$.
D'après (5), et puisque $\sum_{j=1}^n a_j (p_j + \frac{1}{2} a_j)$ est B_{n-1} -mesurable, nous avons : $\forall p \in H_0$,

$$P^{B_{n-1}}(C_n) \leq e^{-\sum_{j=1}^n a_j (p_j + \frac{1}{2} a_j)} \cdot E_P^{B_{n-1}}(e^{Y_n}).$$

En intégrant, on obtient :

$$P(C_n) \leq E_P(e^{-\sum_{j=1}^n a_j (p_j + \frac{1}{2} a_j)}) \cdot E_P^{B_{n-1}}(e^{Y_n}) \quad (7)$$

$$E_P^{B_{n-1}}(e^{Y_n}) = E_P^{B_{n-1}}(e^{Y_{n-1} + a_n \psi_n})$$

$$= e^{Y_{n-1}} \cdot E_P^{B_{n-1}}(e^{a_n \psi_n})$$

$$\leq e^{Y_{n-1}} \cdot e^{a_n E_P^{B_{n-1}}(\psi_n) + \frac{1}{8} a_n^2}, \quad \text{d'après (6)}$$

$$\leq e^{Y_{n-1}} \cdot e^{a_n p_n + \frac{1}{8} a_n^2} \quad \text{d'après (1).}$$

P -p.s.

En reportant dans (7), on obtient :

$$\begin{aligned}
 P(C_n) &\leq e^{-3/8 a_n^2} \cdot E_P(e^{-\sum_{j=1}^{n-1} a_j (p_j + \frac{1}{2} a_j)} \cdot e^{Y_{n-1}}) \\
 &= e^{-3/8 a_n^2} \cdot E_P(e^{-\sum_{j=1}^{n-1} a_j (p_j + \frac{1}{2} a_j)} \cdot E_P^{n-2}(e^{Y_{n-1}})) \\
 &\leq e^{-3/8 (a_1^2 + \dots + a_n^2)}, \text{ en répétant } (n-1) \text{ fois le calcul}
 \end{aligned}$$

précédent, ce qui termine la démonstration de l'inégalité (3). Posons

maintenant : $\forall n \geq 1$,

$$\psi'_n = 1 - \psi_n, \quad p'_n = 1 - a_n - p_n.$$

Puisque nous avons évidemment, $\forall n \geq 1$ et $\forall Q \in H_1$:

$$\begin{aligned}
 E_Q^{n-1}(\psi'_n) &= 1 - E_Q^{n-1}(\psi_n) \\
 &\leq 1 - a_n - p_n = p'_n,
 \end{aligned}$$

on obtient, en appliquant (3), que : $\forall n \geq 1$,

$$Q\left(\sum_{j=1}^n a_j \psi'_j \geq \sum_{j=1}^n a_j (p'_j + \frac{1}{2} a_j)\right) \leq e^{-3/8 \sum_{j=1}^n a_j^2}.$$

Or nous avons :

$$\begin{aligned}
 \left(\sum_{j=1}^n a_j \psi'_j \geq \sum_{j=1}^n a_j (p'_j + \frac{1}{2} a_j)\right) &= \left(\sum_{j=1}^n a_j \psi_j \leq \sum_{j=1}^n a_j (p_j + \frac{1}{2} a_j)\right) \\
 &\supset (\Omega - C_n).
 \end{aligned}$$

On obtient donc (4) par simple passage au complémentaire.

2 - Remarques. - Dans le théorème fondamental de J. Geffroy (5, th. 1), les fonctions ψ_n sont nécessairement des fonctions indicatrices, et les p_j sont des constantes, ce qui en limite le champ d'application. L'amélioration

que nous apportons est simplement due à l'emploi de la forme conditionnelle de l'inégalité de Hoeffding, l'essentiel de la démonstration de J. Geffroy étant par ailleurs conservé.

3 - Application à la statistique des processus.-

Considérons une suite d'observations $X = (X_n)_{n \geq 1}$ à valeurs dans $(X^{(+\infty)}, \mathcal{B}^{(+\infty)}) = \bigotimes_{n=1}^{+\infty} (X_n, \mathcal{B}_n)$ dont la loi inconnue appartient à un ensemble donné H de lois sur $(X^{(+\infty)}, \mathcal{B}^{(+\infty)})$.

$\forall P \in H$ et $\forall n \geq 1$, P_n et $P^{(n)}$ désignent respectivement la loi de X_n , sur (X_n, \mathcal{B}_n) , et celle de $X^{(n)} = (X_1, \dots, X_n)$, sur $(X^{(n)}, \mathcal{B}^{(n)}) = \bigotimes_{j=1}^n (X_j, \mathcal{B}_j)$, lorsque X suit la loi P .

Nous disons que $(X^{(+\infty)}, \mathcal{B}^{(+\infty)}, H)$ est un modèle statistique si : $\forall P \in H$ et $\forall n \geq 2$, la loi conditionnelle de X_n sachant $X^{(n-1)}$, lorsque X suit la loi P , existe, et nous supposons toujours qu'une version en a été choisie. Alors, $\forall x^{(n-1)} \in X^{(n-1)}$, $P_{x^{(n-1)}}$ désignera la loi de X_n sachant que $X^{(n-1)} = x^{(n-1)}$.

L'existence de ces lois conditionnelles est assurée par exemple si $\forall n \geq 2$, $X^{(n)}$ est un espace métrique σ -compact, ou un espace polonais, dont $\mathcal{B}^{(n)}$ est la tribu borélienne (cf ⁹, cor. III).

4 - Traduisons maintenant le théorème (1) dans ce contexte statistique :

Théorème 2.- Soit H_0 et H_1 deux hypothèses d'un modèle statistique $(X^{(+\infty)}, \mathcal{B}^{(+\infty)}, H)$. On suppose que $\forall n \geq 2$, il existe

$$- \varphi_n : (X^{(n)}, \mathcal{B}^{(n)}) \longrightarrow [0, 1]$$

$$- p_n : (X^{(n-1)}, \mathcal{B}^{(n-1)}) \longrightarrow [0, 1]$$

- $a_n \in [0, 1]$ tels que :

$$- 0 \leq p_n + a_n \leq 1$$

$$- \forall P \in H_0 \text{ et } \forall x^{(n-1)} \in X^{(n-1)}, \int \psi_n(x^{(n-1)}, x_n) P_{x^{(n-1)}}(dx_n) \leq p_n(x^{(n-1)})$$

$$- \forall Q \in H_1 \text{ et } \forall x^{(n-1)} \in X^{(n-1)}, \int \psi_n(x^{(n-1)}, x_n) Q_{x^{(n-1)}}(dx_n) \geq p_n(x^{(n-1)})$$

Alors il existe une suite adaptée $(C_n)_{n \geq 1}$ de cylindres ($\forall n \geq 1, C_n \in \mathcal{C}^n$)

tels que :

$$\text{Sup}(P^{(n)}(C_n) ; P \in H_0) \leq e^{-3/8 \sum_{j=2}^n a_j^2}$$

$$\text{Inf}(Q^{(n)}(C_n) ; Q \in H_1) \geq 1 - e^{-3/8 \sum_{j=2}^n a_j^2}$$

5 - Remarques.

a) Les lois initiales ne sont pas ici utilisées (mais elles pourraient l'être) puisque l'on a choisi de poser $\psi_1 = p_1 = a_1 = 0$.

b) La distance en variations d_v sur l'ensemble P de toutes les lois sur un espace mesurable quelconque (Ω, A) est définie par :

$$\begin{aligned} \forall P, Q \in P, \quad d_v(P, Q) &= \text{Sup}(|P(A) - Q(A)| ; A \in A) \\ &= \text{Sup}(|\int \psi dP - \int \psi dQ| ; \psi \in \Phi) \end{aligned}$$

où Φ est l'ensemble des fonctions $\psi : (\Omega, A) \rightarrow [0, 1]$. D'après (10) et (11), nous avons :

$$\forall n \geq 1, d_v((P^{(n)}, P \in H_0), (Q^{(n)}, Q \in H_1)) \geq 1 - 2e^{-3/8 \sum_{j=2}^n a_j^2}$$

c) Séparation asymptotique uniforme. - On dit que deux hypothèses

H_0 et H_1 d'un modèle statistique se séparent asymptotiquement uniformément si : $\forall n \geq 1, \exists S_n \in \mathcal{B}^{(n)}$ tels que :

$$- \sup(P^{(n)}(S_n) ; P \in H_0) \xrightarrow{n \rightarrow +\infty} 0$$

$$- \inf(Q^{(n)}(S_n) ; Q \in H_1) \xrightarrow{n \rightarrow +\infty} 1 .$$

On peut alors énoncer :

Corollaire 1. - Soit H_0 et H_1 deux hypothèses d'un modèle statistique auxquelles on peut appliquer le théorème (2). Alors, elles se séparent asymptotiquement uniformément si :

$$\sum_{n=2}^{+\infty} \frac{a_n^2}{n} = +\infty .$$

En effet, d'après (10) et (11), il suffit de poser :

$$- S_1 = X_1$$

$$- \forall n \geq 2, S_n = (x^{(n)}) : \sum_{j=2}^n a_j \psi_j(x^{(j)}) \geq \sum_{j=2}^n a_j (p_j(x^{(j-1)}) + \frac{1}{2} a_j) .$$

On peut donc exhiber les cylindres suivant lesquels se fait la séparation asymptotique uniforme, et minorer l'amplitude de cette séparation (inégalité (12)). On peut aussi montrer facilement que H_0 et H_1 sont orthogonales, c'est-à-dire que : $\exists S \subset \mathcal{B}^{(+\infty)}$ tel que :

$$- \forall P \in H_0, P(S) = 1$$

$$- \forall Q \in H_1, Q(S) = 0 .$$

L'orthogonalité de H_0 et H_1 , sauf cas très particulier comme le cas de

deux hypothèses simples, n'implique pas la séparation asymptotique uniforme

II - DECANTATION.

1.- Définition.- On dit que deux familles de lois P_0 et P_1 sur un espace mesurable (Ω, A) sont décantées s'il existe une fonction $\psi \in \Phi$ telle que :

$$\text{Inf} \left(\int \psi \cdot dQ ; Q \in P_1 \right) > \text{Sup} \left(\int \psi \cdot dP ; P \in P_0 \right) .$$

C'est évidemment une notion de convexité. En effet, si $c(P_0)$ et $c(P_1)$ désignent respectivement les enveloppes convexes de P_0 et P_1 , si $\overline{c(P_0)}$ et $\overline{c(P_1)}$ désignent les adhérences de ces enveloppes convexes pour la distance d_V , nous avons $\forall \psi \in \Phi$:

$$\begin{aligned} & \text{Inf} \left(\int \psi \cdot dQ ; Q \in P_1 \right) - \text{Sup} \left(\int \psi \cdot dP ; P \in P_0 \right) \\ &= \text{Inf} \left(\int \psi \cdot dQ ; Q \in c(P_1) \right) - \text{Sup} \left(\int \psi \cdot dP ; P \in c(P_0) \right) \\ &= \text{Inf} \left(\int \psi \cdot dQ ; Q \in \overline{c(P_1)} \right) - \text{Sup} \left(\int \psi \cdot dP ; P \in \overline{c(P_0)} \right) . \end{aligned}$$

On voit que les hypothèses (8) et (9) du théorème (2) sont des hypothèses de décantation des familles de lois conditionnelles : $(P_x^{(n-1)} ; P \in H_0^x)$ et $(Q_x^{(n-1)} ; Q \in H_1^x)$, notées dans la suite $H_0^n(x^{(n-1)})$, et $H_1^n(x^{(n-1)})$. Nous présentons ci-dessous trois types de preuve de la décantation de $H_0^n(x^{(n-1)})$ et $H_1^n(x^{(n-1)})$: la méthode des convexes, la méthode des boules, et la méthode des lois ordonnées.

2.- Méthode des convexes. - Elle repose sur le théorème suivant, dû à L. Le Cam (¹², th. 5) : soit C_0 et C_1 deux familles convexes de lois sur un espace mesurable (Ω, A) , dominées par une mesure σ -finie sur cet espace, et $\varepsilon \geq 0$. Alors, nous avons :

$$d_V(C_0, C_1) \geq \varepsilon \iff \exists \psi \in \Phi \quad \text{telle que :}$$

$$\inf \left(\int \psi \cdot dQ ; Q \in C_1 \right) \geq \varepsilon + \sup \left(\int \psi \cdot dP ; P \in C_0 \right).$$

Cette fonction ψ peut fort bien ne pas être une indicatrice.

Exemple. - Soit $\Omega = (\omega_1, \omega_2, \omega_3)$, A étant l'ensemble de ses six parties. Considérons la loi uniforme $U = \sum_{i=1}^3 \frac{1}{3} \delta_{\omega_i}$ et l'ensemble C des lois qui chargent ω_2 deux fois plus que ω_1 .

$$(U) \text{ et } C = \left(\frac{\alpha}{4} \delta_{\omega_1} + \frac{\alpha}{2} \delta_{\omega_2} + \left(1 - \frac{3\alpha}{4}\right) \delta_{\omega_3} ; 0 \leq \alpha \leq \frac{4}{3} \right) \text{ sont}$$

deux ensembles convexes et dominés de lois, dont la distance en variations est $d_V((U), C) = \frac{1}{9}$.

Il est facile de voir qu'il existe une fonction $\psi \in \Phi$ et une seule telle que :

$\forall P \in C, \int \psi \cdot dU \geq \frac{1}{9} + \int \psi \cdot dP$: c'est la fonction définie par $\psi(\omega_1) = 1, \psi(\omega_2) = 0, \psi(\omega_3) = \frac{1}{3}$, ce qui illustre le résultat précédent. Il est tout aussi facile de vérifier que (U) et C ne sont décantées par aucune fonction indicatrice.

3.- Nous devons souvent imposer des conditions de régularité aux fonctions H_0^n et H_1^n . Nous exprimerons ces conditions à l'aide de l'écart de Hausdorff e_{HV} associé à d_V de la manière suivante : si P_0 et P_1 désignent deux familles non vides quelconques de lois sur un espace mesurable

(Ω, A) , on pose :

$$e_{HV}(P_0, P_1) = \text{Max}(\text{Sup}(d_V(P, P_1) ; P \in P_0), \text{Sup}(d_V(Q, P_0) ; Q \in P_1)) .$$

Théorème 3.- Soit H_0 et H_1 deux hypothèses non vides d'un modèle statistique $(X^{(+\infty)}, B^{(+\infty)}, H)$.

$\forall n \geq 2$, on pose :

$$\begin{aligned} - \forall x^{(n-1)} \in X^{(n-1)} , \Delta_n(x^{(n-1)}) &= d_V(c(H_0^n(x^{(n-1)})), c(H_1^n(x^{(n-1)}))) \\ - \Delta_n &= \text{Inf}(\Delta_n(x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)} . \end{aligned}$$

On suppose que les conditions suivantes sont satisfaites :

- $\forall n \geq 1$, X_n est un espace topologique à base dénombrable, B_n est sa tribu borélienne ; $X^{(n)}$ est σ -compact pour la topologie - produit,

- $\forall n \geq 2$, $c(H_0^n(x^{(n-1)}))$ et $c(H_1^n(x^{(n-1)}))$ sont des fonctions continues de $x^{(n-1)}$ pour e_{HV} ;

- $\forall n \geq 2$ et $\forall x^{(n-1)} \in X^{(n-1)}$, les deux familles de lois $H_0^n(x^{(n-1)})$ et $H_1^n(x^{(n-1)})$ sont dominées par une mesure σ -finie sur (X_n, B_n) . Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} \Delta_n^2 = +\infty .$$

Démonstration : D'après (13), il existe une suite $(\Delta'_n)_{n \geq 2}$ de réels tels que :

$$\begin{aligned} - \text{si } \Delta_n = 0 , \Delta'_n &= 0 \\ - \text{si } \Delta_n > 0 , \Delta_n > \Delta'_n > 0 . \\ - \sum_{n=2}^{+\infty} \Delta_n'^2 &= +\infty \quad (14) . \end{aligned}$$

Le problème est essentiellement de construire les fonctions

ψ_n , $n \geq 2$, décantant les deux familles de lois conditionnelles.

a) Soit $n \geq 2$ tel que $\Delta_n > 0$.

Par définition de Δ_n , $\forall x^{(n-1)} \in X^{(n-1)}$, nous avons :

$$d_V(c(H_0^n(x^{(n-1)})), c(H_1^n(x^{(n-1)}))) \geq \Delta_n.$$

Comme ces deux familles de lois sont dominées, d'après le théorème de Le Cam,

il en résulte que :

$$\exists \Psi_{x^{(n-1)}} : (X_n, \mathcal{B}_n) \longrightarrow [0, 1], \text{ fonction telle que :}$$

$$\inf \left(\int \Psi_{x^{(n-1)}} d\nu ; \nu \in c(H_1^n(x^{(n-1)})) \right) \geq \Delta_n + \sup \left(\int \Psi_{x^{(n-1)}} d\mu ; \mu \in c(H_0^n(x^{(n-1)})) \right) \quad (15)$$

$c(H_0^n(x^{(n-1)}))$ et $c(H_1^n(x^{(n-1)}))$ étant des fonctions continues de $x^{(n-1)}$,

$\forall x^{(n-1)} \in X^{(n-1)}$, on peut trouver un voisinage ouvert $V(x^{(n-1)})$ de $x^{(n-1)}$ dans $X^{(n-1)}$ tel que : $\forall y^{(n-1)} \in V(x^{(n-1)})$, on ait :

$$- e_{HV}(c(H_0^n(x^{(n-1)})), c(H_0^n(y^{(n-1)}))) < \frac{\Delta_n - \Delta'_n}{2} \quad (16)$$

$$- e_{HV}(c(H_1^n(x^{(n-1)})), c(H_1^n(y^{(n-1)}))) < \frac{\Delta_n - \Delta'_n}{2} \quad (17)$$

On déduit de (16) que :

$$\sup \left(\int \Psi_{x^{(n-1)}} \cdot d\mu ; \mu \in c(H_0^n(y^{(n-1)})), y^{(n-1)} \in V(x^{(n-1)}) \right) \leq \sup \left(\int \Psi_{x^{(n-1)}} d\mu ; \mu \in c(H_0^n(x^{(n-1)})) \right) + \frac{\Delta_n - \Delta'_n}{2} \quad (18)$$

et de même de (17) que :

$$\begin{aligned} \text{Inf} \left(\int \psi_{x^{(n-1)}} \cdot dv ; v \in c(H_1^n(y^{(n-1)})), y^{(n-1)} \in V(x^{(n-1)}) \right) \\ \geq \text{Inf} \left(\int \psi_{x^{(n-1)}} \cdot dv ; v \in c(H_1^n(x^{(n-1)})) \right) - \frac{\Delta_n - \Delta'_n}{2} \end{aligned}$$

Compte tenu de (15), (18), et (19), nous avons finalement :

$$\begin{aligned} \text{Inf} \left(\int \psi_{x^{(n-1)}} \cdot dv ; v \in c(H_1^n(y^{(n-1)})), y^{(n-1)} \in V(x^{(n-1)}) \right) \\ \geq \text{Sup} \left(\int \psi_{x^{(n-1)}} \cdot d\mu ; \mu \in c(H_0^n(y^{(n-1)})), y^{(n-1)} \in V(x^{(n-1)}) \right) + \Delta'_n \end{aligned}$$

$X^{(n-1)}$ étant σ -compact, on peut extraire de $(V(x^{(n-1)}))$, $x^{(n-1)} \in X^{(n-1)}$ un recouvrement dénombrable de $X^{(n-1)}$: $(V(x_p^{(n-1)}))$, $p \geq 1$ à partir duquel on construit une partition dénombrable de $X^{(n-1)}$ notée (V_u) , $u \geq 1$ en posant :

$$\forall u \geq 1, \quad V_u = V(x_u^{(n-1)}) - \bigcup_{p=1}^{u-1} V(x_p^{(n-1)})$$

Cette partition est $\mathcal{B}^{(n-1)}$ -mesurable car $\mathcal{B}^{(n-1)}$ est la tribu borélienne de $X^{(n-1)}$ pour la topologie - produit.

On pose alors, $\forall u \geq 1$, $\forall y^{(n-1)} \in V_u$, $\forall y_n \in X_n$:

$$- \varphi_n(y^{(n-1)}, y_n) = \psi_{x_u^{(n-1)}}(y_n)$$

$$- p_n(y^{(n-1)}) = \text{Sup} \left(\int \psi_{x_u^{(n-1)}} \cdot dP_y^{(n-1)} ; y^{(n-1)} \in V_u, P \in H_0 \right)$$

Par construction, les fonctions φ_n et p_n sont respectivement $\mathcal{B}^{(n)}$ et $\mathcal{B}^{(n-1)}$ -mesurables et vérifient :

$\forall u \geq 1, \forall y^{(n-1)} \in V_u, \forall P \in H_0, \forall Q \in H_1 :$

$$- \int \psi_n(y^{(n-1)}, y_n) \cdot P_{y^{(n-1)}}(dy_n) \leq p_n(y^{(n-1)}) \quad , \quad \text{d'après (21)}$$

$$- \int \psi_n(y^{(n-1)}, y_n) \cdot Q_{y^{(n-1)}}(dy_n) \geq \inf \left(\int \psi_{x_u}^{(n-1)} \cdot dQ_{y^{(n-1)}} ; Q \in H_1, y^{(n-1)} \in V_u \right) \\ \geq \Delta'_n + p_n(y^{(n-1)}) \quad , \quad \text{d'après (20).}$$

Autrement dit, nous avons :

$\forall x^{(n-1)} \in X^{(n-1)}, \forall P \in H_0, \forall Q \in H_1,$

$$\int \psi_n(x^{(n-1)}, x_n) P_{x^{(n-1)}}(dx_n) \leq p_n(x^{(n-1)}) \quad (22)$$

$$\int \psi_n(x^{(n-1)}, x_n) Q_{x^{(n-1)}}(dx_n) \geq \Delta'_n + p_n(x^{(n-1)}) \quad (23)$$

b) Soit $n \geq 2$ tel que $\Delta_n = 0$: alors (22) et (23) sont trivialement vérifiées si l'on pose $\psi_n = p_n = 0$.

c) D'après le théorème (2), il existe une suite adaptée de cylindres

$(C_n)_{n \geq 1}$ tels que :

$$\forall P \in H_0, \quad P^{(n)}(C_n) \leq e^{-3/8 \sum_{j=2}^n \Delta_j^2}$$

$$\forall Q \in H_1, \quad Q^{(n)}(C_n) \geq 1 - e^{-3/8 \sum_{j=2}^n \Delta_j^2}$$

Compte-tenu de (14), il y a donc séparation asymptotique uniforme de H_0

et H_1 .

4.- La distance de Hellinger d_H entre deux lois P et Q sur un espace mesurable quelconque (Ω, A) est définie (cf. ¹⁰) par :

$$d_H(P, Q) = \sqrt{1 - \rho_H(P, Q)} \quad \text{où}$$

$$\rho_H(P, Q) = \int \sqrt{\frac{dP}{d\mu} \cdot \frac{dQ}{d\mu}} d\mu, \quad \mu \text{ désignant une mesure } \sigma\text{-finie quelconque}$$

sur (Ω, A) dominant à la fois P et Q .

Sur l'ensemble \mathcal{P} de toutes les lois sur (Ω, A) , d_V et d_H sont des distances topologiquement équivalentes puisqu'elles vérifient :

$$\forall P, Q \in \mathcal{P}, \quad d_H^2(P, Q) \leq d_V(P, Q) \leq \sqrt{2} d_H(P, Q) \quad (24) \quad (\text{cf } ^{12}, \text{ p. } 127).$$

Les écarts de Hausdorff e_{HV} et e_{HH} (calculé à partir de d_H) sur l'ensemble des parties non vides de \mathcal{P} sont eux aussi topologiquement équivalents puisque pour tout couple de familles non vides \mathcal{P}_0 et \mathcal{P}_1 de lois sur (Ω, A) , on a :

$$e_{HH}^2(\mathcal{P}_0, \mathcal{P}_1) \leq e_{HV}(\mathcal{P}_0, \mathcal{P}_1) \leq \sqrt{2} e_{HH}(\mathcal{P}_0, \mathcal{P}_1).$$

Il est alors évident que dans le théorème (3), on peut substituer d_H à d_V , ce qui est intéressant car d_H se calcule presque toujours plus facilement que d_V (voir par exemple la distance de Hellinger de deux lois normales non dégénérées sur \mathbb{R}^n : ¹², p. 129 ou ¹³, p. 191). On obtient alors :

Corollaire 2.- Soit H_0 et H_1 deux hypothèses non vides d'un modèle statistique $(X^{(+\infty)}, \mathcal{B}^{(+\infty)}, H)$.

$\forall n \geq 2$, on pose :

$$- \forall x^{(n-1)} \in X^{(n-1)}, \quad \delta_n(x^{(n-1)}) = d_H(c(H_0^n(x^{(n-1)})), c(H_1^n(x^{(n-1)})))$$

$$- \delta_n = \text{Inf}(\delta_n(x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)}.$$

On suppose que les conditions suivantes sont satisfaites :

- $\forall n \geq 1$, X_n est un espace topologique à base dénombrable, B_n est sa tribu borélienne ; $X^{(n)}$ est σ -compact pour la topologie - produit ;
- $\forall n \geq 2$, $c(H_0^n(x^{(n-1)}))$ et $c(H_1^n(x^{(n-1)}))$ sont des fonctions continues de $x^{(n-1)}$ pour e_{HH} ;
- $\forall n \geq 2$, $\forall x^{(n-1)} \in X^{(n-1)}$, les deux familles de lois $H_0^n(x^{(n-1)})$ et $H_1^n(x^{(n-1)})$ sont dominées par une mesure σ -finie sur (X_n, B_n) .

Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} \delta_n^4 = +\infty \quad (25)$$

Remarque. - Dans les applications, on pourra souvent remplacer (25) par $\sum_{n=2}^{+\infty} \delta_n^2 = +\infty$. C'est le cas par exemple dans le problème de l'estimation d'un signal dans un bruit gaussien, cas où les lois conditionnelles seront des lois gaussiennes de même variance sur \mathbb{R} .

5.- Distances décantantes. - Les hypothèses du théorème (3) comprennent des hypothèses de continuité qui sont dans une large mesure étrangères au problème de la séparation asymptotique uniforme de H_0 et H_1 : elles ne sont utilisées que pour assurer la mesurabilité des fonctions ψ_n . Par conséquent, elles deviennent inutiles si les espaces X_n , $n \geq 1$ sont dénombrables. On obtient ainsi un résultat général que l'on pourrait par exemple appliquer au cas où la suite $(X_n)_{n \geq 1}$ des observations est une chaîne de Markov dont l'espace des états est dénombrable.

Corollaire 3.- Soit H_0 et H_1 deux familles de lois sur $\prod_{n=1}^{+\infty} (X_n, \mathcal{B}_n)$, chacun des espaces X_n , $n \geq 1$, étant dénombrable et muni de la tribu \mathcal{B}_n de toutes ses parties.

Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} \Delta_n^2 = +\infty \quad (13)$$

De notre point de vue, ce genre de modèle, où n'interviennent que des conditions de décantation, est très représentatif de la théorie.

En outre, on peut remplacer la distance d_V , à partir de laquelle sont calculés les Δ_n , par de nombreuses autres distances.

Définition.- Nous dirons qu'une distance d sur un ensemble P de lois sur un espace mesurable (Ω, \mathcal{A}) est décantante s'il existe une fonction $f :]0, +\infty[\rightarrow]0, +\infty[$ telle que :

$$\forall P, Q \in P \text{ et } \forall \alpha > 0, \quad d(P, Q) \geq \alpha \implies d_V(P, Q) \geq f(\alpha).$$

Si d est décantante, on choisira la meilleure fonction f possible, à savoir $f_d :]0, +\infty[\rightarrow]0, +\infty[$ définie par

$$f_d(\alpha) = \begin{cases} 1 & \text{si } ((P, Q) ; P, Q \in P \text{ et } d(P, Q) \geq \alpha) = \emptyset \\ \text{Inf}(d_V(P, Q) ; P, Q \in P \text{ et } d(P, Q) \geq \alpha) & \text{, sinon.} \end{cases}$$

f_d est croissante et vérifie évidemment :

$$P \neq Q \implies d_V(P, Q) \geq f_d(d(P, Q)).$$

D'après (24), la distance de Hellinger d_H est décantante, et il est facile de voir que f_{d_H} est la fonction $x \mapsto \text{Min}(x^2, 1)$, si $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ (cf. ¹, distances décantantes).

Les distances usuelles suivantes sont décantantes :

- la distance de Prohorov d_P sur l'ensemble de toutes les lois boréliennes sur tout espace métrique (X, δ) :

$$\forall P, Q, \quad d_P(P, Q) = \text{Inf}(\epsilon > 0 ; \forall F \in \mathcal{F}, P(F) \leq Q(F^\epsilon) + \epsilon) \\ = \text{Inf}(\epsilon > 0 ; \forall F \in \mathcal{F}, Q(F) \leq P(F^\epsilon) + \epsilon),$$

où \mathcal{F} est l'ensemble des fermés de (X, δ) et où

$$\forall F \in \mathcal{F} \quad \text{et} \quad \forall \epsilon > 0, \quad F^\epsilon = \{x ; x \in X \text{ et } \delta(x, F) < \epsilon\}.$$

On sait que si (X, δ) est séparable, d_P définit la topologie de la convergence faible (cf ², App. III).

- les distances de Lévy d_L et de Kolmogorov d_K sur l'ensemble de toutes les lois boréliennes sur \mathbb{R}^n , ainsi définies : si F_P désigne la fonction de répartition de toute loi P , on pose, $\forall P, Q$:

$$- d_L(P, Q) = \text{Inf}(h ; h > 0 \text{ et } \forall x \in \mathbb{R}^n, F_Q(x - \vec{h}) - h \leq F_P(x) \leq F_Q(x + \vec{h}) + h,$$

où \vec{h} désigne le n -uple (h, \dots, h) ,

$$- d_K(P, Q) = \text{Sup}(|F_P(x) - F_Q(x)| ; x \in \mathbb{R}^n).$$

d_L définit comme d_P la topologie de la convergence faible (cf ⁷, chap. 2, § 9) ; d_K est la distance du théorème de Glivenko - Cantelli (cf ¹¹).

On verra que d_P, d_L, d_K sont décantantes parce qu'elles vérifient une propriété plus forte, à savoir qu'elles décantent leurs boules. Lorsqu'il

s'agit des distances sur l'ensemble de toutes les lois sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, nous avons :

$$\forall x > 0, \quad f_{d_P}(x) = f_{d_L}(x) = f_{d_K}(x) = \text{Min}(x, 1) \quad (\text{cf } ^1, \text{ déjà cité}).$$

Distances non décantantes. - Parmi les distances qui ne sont pas décantantes mais qui sont utiles en statistique asymptotique, citons la distance de Fortet - Mourier (cf ³). Elle est définie sur l'ensemble \mathcal{P}^* des lois boréliennes P sur un espace métrique quelconque (X, δ) telles que :

$$\int \delta(x, x_0) P(dx) < +\infty$$

où x_0 est un point arbitrairement choisi dans X .

L'ensemble E_{x_0} des fonctions réelles lipschitziennes définies sur (X, δ) , et qui s'annulent en x_0 , est un espace de Banach pour la norme :

$$\forall f \in E_{x_0}, \quad \|f\|_{x_0} = \text{Sup} \left(\frac{f(x) - f(y)}{\delta(x, y)} ; x, y \in X \text{ et } x \neq y \right).$$

Toute loi P de \mathcal{P} peut être identifiée à un élément du dual topologique $(E_{x_0}^*, \| \cdot \|_{x_0}^*)$ de cet espace de Banach, à savoir la forme linéaire continue $f \rightsquigarrow \int f \cdot dP$.

Enfin, il est clair que :

$$\forall P, Q \in \mathcal{P}, \quad \|P-Q\|_{x_0}^* = \text{Sup} \left(\left| \int f \cdot dP - \int f \cdot dQ \right| ; f \in E_{x_0} \text{ et } \|f\|_{x_0} = 1 \right)$$

ne dépend pas de x_0 . Par définition, on pose $d_{FM}(P, Q) = \|P-Q\|_{x_0}^*$.

On démontre facilement que d_{FM} est décantante si et seulement si (X, δ) est borné.

Par exemple, si $(X, \delta) = (\mathbb{R}, | \cdot |)$, \mathcal{P} est l'ensemble des lois P telles que : $\int |x| P(dx) < +\infty$, et d_{FM} est la distance non décantante $d_{FM}(P, Q) = \int |F_P(x) - F_Q(x)| dx$ (3, 4.7).

D'après (26) et le corollaire précédent, il est clair que nous avons :

Corollaire 4. - Soit H_0 et H_1 deux familles de lois sur $\Theta(X_n, \mathcal{B}_n)$, chacun des espaces X_n , $n \geq 1$, étant dénombrable et muni de la tribu \mathcal{B}_n de toutes ses parties.

$\forall n \geq 2$, on suppose que $c(P_{x^{(n-1)}}; P \in H_0 \cup H_1, x^{(n-1)} \in X^{(n-1)})$ est munie d'une distance décantante d_n , et on pose :

$$\alpha_n = \inf(d_n(c(H_0^n(x^{(n-1)})), c(H_1^n(x^{(n-1)}))) ; x^{(n-1)} \in X^{(n-1)}).$$

Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} f_{d_n}^2(\alpha_n) = +\infty.$$

6.- Cas des lois - produits. - Une loi P sur $(X^{(+\infty)}, \mathcal{B}^{(+\infty)})$ est dite loi - produit si $P = \prod_{n=1}^{+\infty} P_n$. Si P est la loi d'une suite d'observations $X = (X_n)_{n \geq 1}$, P est une loi - produit si et seulement si les v.a. X_n , $n \geq 1$, sont indépendantes. Il est clair que dans ce cas le théorème (3) se simplifie considérablement, et que seules des conditions de décantation seront imposées. On obtient ainsi le corollaire (5), où l'on utilise seulement la distance d_V , et le corollaire (6), où l'on utilise des distances décantantes quelconques.

Corollaire 5.- Soit H_0 et H_1 deux familles de lois - produits sur $(X^{(+\infty)}, B^{(+\infty)})$.

$\forall n \geq 2$, on pose $\Delta_n = d_V(c(P_n; P \in H_0), c(Q_n; Q \in H_1))$,

et on suppose que ces familles de lois sont dominées par une mesure σ -finie sur (X_n, B_n) .

Alors H_0 et H_1 se séparent asymptotiquement uniformément

si $\sum_{n=2}^{+\infty} \Delta_n^2 = +\infty$.

Corollaire 6.- Soit H_0 et H_1 deux familles de lois - produits sur $(X^{(+\infty)}, B^{(+\infty)})$.

$\forall n \geq 2$, on suppose que $c(P_n; P \in H_0 \cup H_1)$ est munie d'une distance décantante d_n et est dominée par une mesure σ -finie sur (X_n, B_n) , et on pose :

$$\alpha_n = d_n(c(P_n; P \in H_0), c(Q_n; Q \in H_1)) .$$

Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} f_d^2(\alpha_n) = +\infty .$$

7.- Méthode des boules.- Dans ce qui suit, les familles de lois $H_0^n(x^{(n-1)})$ et $H_1^n(x^{(n-1)})$ seront décantées parce qu'elles seront contenues dans des boules définies à partir de distances ad hoc.

Définition.- Soit P un ensemble de lois sur un espace mesurable quelconque (Ω, A) . On dit qu'une distance d sur P décante ses boules

il existe une fonction $F_d :]0, +\infty[\rightarrow]0, +\infty[$ telle que :

quelles que soient les boules fermées $B(P_0, \alpha_0)$ et $B(P_1, \alpha_1)$ de \mathcal{P} et quel que soit $\alpha > 0$ tel que $d(P_0, P_1) > \alpha + \alpha_0 + \alpha_1$, on peut trouver

une fonction $\psi \in \Phi$ telle que :

$$\forall P \in B(P_0, \alpha_0) \text{ et } \forall Q \in B(P_1, \alpha_1), \int \psi \cdot dQ \geq \int \psi \cdot dP + F_d(\alpha) \quad (28)$$

Remarques.-

- Toute distance vérifiant (28) est évidemment décantante.

- Nous ne faisons plus d'hypothèses de domination sur les boules de lois considérées qui, de plus, ne sont pas nécessairement convexes. Nous ne sommes donc plus dans le cadre d'application du théorème de Le Cam.

- La généralité de la définition précédente n'est pas restreinte si l'on impose à F_d , ce que nous ferons dans la suite, d'être continue à gauche, et de s'annuler à l'origine.

Exemples de distances qui décantent leurs boules.- Il est facile de démontrer (cf ¹⁴, p. 66-69) que d_V , sur l'ensemble de toutes les lois sur un espace mesurable quelconque, d_P , sur l'ensemble de toutes les lois boréliennes sur un espace métrique quelconque, d_L et d_K sur l'ensemble de toutes les lois boréliennes sur \mathbb{R}^n décantent leurs boules, l'une quelconque d de ces distances vérifiant la propriété plus forte suivante :

$$\forall P_0, P_1 \in \mathcal{P}, \forall \alpha_0, \alpha_1 \geq 0, \forall \alpha > 0 \text{ tels que } d(P_0, P_1) > \alpha_0 + \alpha_1 + \alpha,$$

$\exists A \in \mathcal{A}$ tel que :

$$\forall P \in B(P_0, \alpha_0) \text{ et } \forall Q \in B(P_1, \alpha_1), \quad Q(A) \geq P(A) + \alpha \quad (29)$$

Par contre la distance d_{FM} ne décant pas ses boules sur \mathcal{P}^* (voir 27)

puisque'elle n'y est pas décantante, si (X, δ) n'est pas borné.

8.- Théorème 4.- Soit H_0 et H_1 deux hypothèses non vides d'un modèle statistique $(X^{(+\infty)}, \mathcal{B}^{(+\infty)}, H)$.

On suppose que les hypothèses suivantes sont satisfaites :

- $\forall n \geq 1$, X_n est un espace topologique à base dénombrable, \mathcal{B}_n est une tribu borélienne, \mathcal{P}_n est un ensemble de lois sur (X_n, \mathcal{B}_n) muni d'une distance d_n qui décante ses boules ;

- $\forall n \geq 1$, $X^{(n)}$ est σ -compact pour la topologie - produit ;

- $\forall n \geq 2$ et $\forall x \in X^{(n-1)}$, il existe deux boules fermées de l'espace métrique (\mathcal{P}_n, d_n) notées $B_n(x^{(n-1)})$ et $B'_n(x^{(n-1)})$ centrées respectivement en $\mu_n(x^{(n-1)})$ et en $\mu'_n(x^{(n-1)})$ et de rayons respectifs $r_n(x^{(n-1)})$ et $r'_n(x^{(n-1)})$ et telles que :

$$- H_0^n(x^{(n-1)}) \subset B_n(x^{(n-1)})$$

$$- H_1^n(x^{(n-1)}) \subset B'_n(x^{(n-1)})$$

- les centres μ_n et μ'_n et les rayons r_n et r'_n de ces boules sont des fonctions continues de $x^{(n-1)}$.

$\forall n \geq 2$, on pose :

$$\rho_n = \inf(\max(0, d_n(\mu_n(x^{(n-1)}), \mu'_n(x^{(n-1)})) - r_n(x^{(n-1)}) - r'_n(x^{(n-1)}))) ; x^{(n-1)} \in X^{(n-1)}$$

Alors, pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} F_d^2(\rho_n) = +\infty .$$

Démonstration : D'après (32) et puisque les fonctions F_{d_n} sont continues à gauche, il existe une suite $(\rho'_n)_{n \geq 2}$ de réels tels que :

- si $\rho_n = 0$, $\rho'_n = 0$
 - si $\rho_n > 0$, $\rho_n > \rho'_n > 0$
 - $\sum_{n=2}^{+\infty} F_{d_n}^2(\rho'_n) = +\infty$.
- (33)

a) Soit $n \geq 2$ tel que $\rho_n > 0$. Par définition de ρ_n , nous avons :

$$\forall x^{(n-1)} \in X^{(n-1)} , d_n(\mu_n(x^{(n-1)}), \mu'_n(x^{(n-1)})) \geq \rho_n + r_n(x^{(n-1)}) + r'_n(x^{(n-1)}) \quad (34)$$

Comme μ_n, μ'_n, r_n, r'_n sont des fonctions continues, on peut trouver un voisinage ouvert $V(x^{(n-1)})$ de tout point $x^{(n-1)}$ dans $X^{(n-1)}$ tel que, d'après (30) et (31), on ait :

$$\forall y^{(n-1)} \in V(x^{(n-1)}) ,$$

$$- H_0^n(y^{(n-1)}) \subset B_n(y^{(n-1)}) \subset B(\mu_n(x^{(n-1)}), r_n(x^{(n-1)}) + \frac{\rho_n - \rho'_n}{2}) \quad (35)$$

$$- H_1^n(y^{(n-1)}) \subset B'_n(y^{(n-1)}) \subset B(\mu'_n(x^{(n-1)}), r'_n(x^{(n-1)}) + \frac{\rho_n - \rho'_n}{2}) \quad (36)$$

Comme d_n décante ses boules (28) et compte tenu de (34), (35), (36), il existe une fonction $\Psi_{x^{(n-1)}} : (X_n, B_n) \rightarrow [0, 1]$ telle que :

$$\forall y^{(n-1)} , z^{(n-1)} \in V(x^{(n-1)}) , \forall P \in H_0 , \forall Q \in H_1$$

$$\int \Psi_{x^{(n-1)}} \cdot dQ_{z^{(n-1)}} \geq F_{d_n}(\rho'_n) + \int \Psi_{x^{(n-1)}} \cdot dP_{y^{(n-1)}} \quad (37)$$

En utilisant la σ -compacité de $X^{(n-1)}$, on peut extraire du recouvrement ouvert $(V(x^{(n-1)}), x^{(n-1)} \in X^{(n-1)})$ de $X^{(n-1)}$ un recouvrement dénombrable

noté $(V(x_p^{(n-1)}), p \geq 1)$ à partir duquel on construit comme précédemment une partition dénombrable $B^{(n-1)}$ -mesurable de $X^{(n-1)}$, notée $(V_u, u \geq 1)$ en posant :

$$\forall u \geq 1, \quad V_u = V(x_u^{(n-1)}) - \bigcup_{p=1}^{u-1} V(x_p^{(n-1)})$$

On définit alors deux fonctions $\psi_n : (X^{(n)}, B^{(n)}) \rightarrow [0, 1]$ et $p_n : (X^{(n-1)}, B^{(n-1)}) \rightarrow [0, 1]$ par :

$$\forall u \geq 1, \quad \forall y^{(n-1)} \in V_u \text{ et } \forall y_n \in X_n :$$

$$- \psi_n(y^{(n-1)}, y_n) = \psi_{x_u^{(n-1)}}(y_n)$$

$$- p_n(y^{(n-1)}) = \text{Sup} \left(\int \psi_n(y^{(n-1)}, y_n) P_{y^{(n-1)}}(dy_n) ; P \in H_0, y^{(n-1)} \in V_u \right)$$

Ces fonctions vérifient : $\forall x^{(n-1)} \in X^{(n-1)}$,

$$- \forall P \in H_0, \int \psi_n(x^{(n-1)}, x_n) P_{x^{(n-1)}}(dx_n) \leq p_n(x^{(n-1)})$$

$$- \forall Q \in H_1, \int \psi_n(x^{(n-1)}, x_n) Q_{x^{(n-1)}}(dx_n) \geq F_{d_n}(\rho'_n) + p_n(x^{(n-1)})$$

d'après (37).

b) Si $\rho_n = 0$, on pose $\psi_n = p_n = 0$. Les inégalités (38) et (39) sont alors trivialement vérifiées.

D'après le théorème (2), on conclut qu'il existe une suite adaptée $(C_n)_{n \geq 1}$ de cylindres tels que $\forall n \geq 1$:

$$\text{Sup}(P^{(n)}(C_n) ; P \in H_0) \leq e^{-3/8 \sum_{j=2}^n F_{d_j}^2(\rho'_j)}$$

$$\text{Inf}(Q^{(n)}(C_n) ; Q \in H_1) \geq 1 - e^{-3/8 \sum_{j=2}^n F_{d_j}^2(\rho'_j)}$$

D'après (33), H_0 et H_1 se séparent donc asymptotiquement uniformément.

9.- Dans le cas de lois - produits, ou lorsque les X_n , $n \geq 1$ sont dénombrables, le théorème (4) se simplifie considérablement. A titre d'exemple, si l'on utilise seulement la distance d_V pour des lois - produits, on obtient :

Corollaire 7.- Soit H_0 et H_1 deux familles de lois - produits sur $(X^{(+\infty)}, B^{(+\infty)})$.

On suppose que $\forall n \geq 2$, il existe deux lois μ_n et μ'_n sur (X_n, B_n) et deux réels r_n et r'_n tels que :

$$-\forall P \in H_0, \quad d_V(\mu_n, P_n) \leq r_n$$

$$-\forall Q \in H_1, \quad d_V(\mu'_n, Q_n) \leq r'_n,$$

et on pose $\rho_n = \text{Max}(0, d_V(\mu_n, \mu'_n) - r_n - r'_n)$.

Alors pour que H_0 et H_1 se séparent asymptotiquement uniformément, il suffit que :

$$\sum_{n=2}^{+\infty} \rho_n^2 = +\infty.$$

10.- Méthode des lois ordonnées.- Soit P un ensemble de lois sur un espace mesurable quelconque (Ω, A) , muni d'un préordre \leq . Ce préordre peut avoir ou non certaines propriétés de décantation. On pourrait par exemple imaginer que deux familles quelconques de lois de P , convexes pour \leq (cf ¹⁸, p. 19 - 21) sont décantées dès qu'elles sont disjointes.

Choisissons une définition moins exigeante :

Définition.- Nous dirons que \leq décante ses sections si $\forall P_0, P_1 \in P$ telles que $P_0 < P_1$, la section finissant en P_0 soit $s(P_0) = (P, P \in P, P \leq P_0)$ et la section commençant en P_1 soit $S(P_1) = (P, P \in P, P \geq P_1)$ sont décantées.

Exemple.- Soit \leq l'ordre défini sur l'ensemble P de toutes les lois sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ par :

$$P \leq Q \iff F_P \leq F_Q.$$

Il est facile de vérifier que cet ordre décante ses sections et vérifie la propriété plus forte suivante :

$\forall P_0, P_1 \in P$ telles que $P_0 < P_1$, $\exists a \in \mathbb{R}$ tel que :

$$P \leq P_0 \text{ et } Q \geq P_1 \implies Q(-\infty, a|) \geq d_K(P_0, P_1) + P(-\infty, a|),$$

où $]-\infty, a|$ est soit $]-\infty, a]$, soit $]-\infty, a[$.

De plus cet ordre devient total si on le restreint à l'ensemble des translatées d'une même loi, ce qui est le cas dans le problème généralisé de l'estimation d'un paramètre de translation (cf. ⁶) ou dans l'estimation d'un signal dans un bruit gaussien, où on appliquera tout naturellement cette technique. Nous l'avons aussi utilisée autrefois (¹⁴, p. 101) pour estimer un paramètre réel dans un modèle statistique associé à des observations indépendantes, les lois de chacune des observations étant des fonctions monotones du paramètre. On pourrait plus généralement obtenir de même des conditions suffisantes de séparation asymptotique uniforme de deux hypothèses multiples H_0 et H_1

lorsque l'hypothèse générale H est ordonnée, les familles de lois conditionnelles $(P_x^{(n-1)}; P \in H)$, $\forall n \geq 2$, $\forall x^{(n-1)} \in X^{(n-1)}$, sont ordonnées, les fonctions $P \rightsquigarrow P_x^{(n-1)}$ étant de plus monotones.

Nous ne développerons pas cette idée ici.

III - APPLICATION A L'ESTIMATION ASYMPTOTIQUE GENERALE.-

1.- Rappels.- Soit $(X^{(+\infty)}, B^{(+\infty)}, H)$ un modèle statistique, (E, d) un espace métrique, et $f : H \rightarrow E$ un paramètre que l'on désire estimer.

Supposons que ce modèle vérifie la condition suivante :

$$(40) \quad \left\{ \begin{array}{l} \forall P_0, P_1 \in H, \forall \alpha_0, \alpha_1 > 0 \text{ tels que } d(f(P_0), f(P_1)) > \alpha_0 + \alpha_1, \\ \text{les hypothèses } f^{-1}(B(f(P_0), \alpha_0)) \text{ et } f^{-1}(B(f(P_1), \alpha_1)) \text{ se séparent} \\ \text{asymptotiquement uniformément.} \end{array} \right.$$

Nous avons alors les résultats généraux suivants :

(41) si l'espace métrique (E, d) est précompact, il existe une suite $(\hat{f}_n)_{n \geq 1}$ d'estimateurs de f qui converge uniformément p.s., c'est-à-dire qui vérifie :

$$\forall \varepsilon > 0, \sup_{k \geq n} (P(\bigcup_{k \geq n} ((d(\hat{f}_k, f(P)) > \varepsilon) \times \prod_{j=k+1}^{+\infty} X_j)); P \in H) \xrightarrow[n \rightarrow +\infty]{} 0,$$

(42) si (E, d) est seulement σ -précompact, il existe une suite p.s. convergente d'estimateurs de f ,

(43) plus généralement, si (E, d) est séparable, il existe une suite convergente d'estimateurs de f . De plus, si l'hypothèse générale H est munie de la tribu A engendrée par les fonctions $P \rightsquigarrow P^{(n)}(B)$, $n \geq 1$, $B \in \mathcal{B}^{(n)}$, et d'une probabilité a priori μ sur A , il existe une suite d'estimateurs de f qui converge P - p.s. pour μ -presque toute probabilité P .

2.- Remarques.-

(i) Ces suites convergentes d'estimateurs sont entièrement construites à partir des cylindres qui séparent asymptotiquement uniformément les couples d'hypothèses auxquels (40) s'applique. Ces constructions sont dues à J. Geffroy et nous-même en ce qui concerne (41) et (42) (cf ¹⁴, p. 29 - 41) et s'appuient directement sur deux résultats de H. Schmerkotte (²¹, th. 4.8 sur l'algèbre des hypothèses que l'on peut tester, et lemme 2.5) en ce qui concerne (43).

(ii) Concrètement ces constructions sont inutilisables, d'autant plus que pour estimer f , il faudra leur ajouter celles qui figurent dans la méthode des convexes ou la méthode des boules. On retiendra de l'hypothèse (40) que les seules hypothèses qui importent sont les images réciproques par f des boules fermées $B(f(P), \alpha)$ de (E, d) .

(iii) Enfin, rappelons que pour certains modèles statistiques, l'existence d'une suite convergente d'estimateurs de f implique l'existence d'une suite p.s. convergente d'estimateurs de f (cf. ¹⁵).

3.- Application de la méthode des convexes.-

$\forall P_0 \in H$, $\forall \alpha_0 > 0$ notons, $\forall n \geq 2$ et $\forall x^{(n-1)} \in X^{(n-1)}$, $c(P_0, \alpha_0, x^{(n-1)})$ l'enveloppe convexe de $(P_x^{(n-1)} ; P \in f^{-1}(B(f(P_0), \alpha_0)))$. Le résultat suivant est une conséquence immédiate du théorème (3) qui implique (40), et des rappels ci-dessus.

Théorème 5.- Soit $(X^{(+\infty)}, B^{(+\infty)}, H)$ un modèle statistique, (E, d) un espace métrique séparable, $f : H \rightarrow E$ un paramètre à estimer. On suppose que :

- $\forall n \geq 1$, X_n est un espace topologique à base dénombrable, B_n est sa tribu borélienne, $X^{(n)}$ est σ -compact pour la topologie - produit ;

- $\forall n \geq 2$, $\forall P \in H$, $\forall \alpha > 0$, $x^{(n-1)} \rightsquigarrow c(P, \alpha, x^{(n-1)})$ est une fonction continue pour e_{HV} ;

- $\forall n \geq 2$, $\forall x^{(n-1)} \in X^{(n-1)}$, $(P_{x^{(n-1)}} ; P \in H)$ est dominée par une mesure σ -finie sur (X_n, B_n) ;

$$-\forall P_0, P_1 \in H, \forall \alpha_0, \alpha_1 > 0 \text{ tels que } d(f(P_0), f(P_1)) > \alpha_0 + \alpha_1, \\ \sum_{n=2}^{+\infty} (\text{Inf}(d_V(c(P_0, \alpha_0, x^{(n-1)}), c(P_1, \alpha_1, x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)}))^2 = +\infty \quad (44)$$

Alors il existe une suite convergente d'estimateurs de f (et même une suite p.s. convergente, si (E, d) est σ -précompact). De plus, si H est muni d'une probabilité a priori μ (cf. rappels ci-dessus), il existe une suite d'estimateurs de f qui converge P -p.s. pour μ -presque toute loi possible P .

Bien entendu, d'après le corollaire (2), on peut dans cet énoncé remplacer e_{HV} par e_{HH} et (44) par :

$$\sum_{n=2}^{+\infty} (\text{Inf}(d_H(c(P_0, \alpha_0, x^{(n-1)}), c(P_1, \alpha_1, x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)}))^4 = +\infty$$

4.- Le théorème (5) se simplifie dans le cas où les espaces des observations sont dénombrables, ou dans le cas où ces observations sont indépendantes. On en obtient alors des versions qui ne contiennent que des conditions de décantation ((45) et (46) ci-dessous).

Corollaire 8.- Soit $(X^{(+\infty)}, B^{(+\infty)}, H)$ un modèle statistique dans lequel chacun des espaces X_n , $n \geq 1$ est dénombrable et muni de la

tribu de toutes ses parties, (E, d) un espace métrique séparable,
 $f : H \rightarrow (E, d)$ un paramètre à estimer. On suppose que :

- $\forall n \geq 2$, $c(P_{x^{(n-1)}}; P \in H, x^{(n-1)} \in X^{(n-1)})$ est munie d'une distance
 décantante d_n ;

- $\forall P_0, P_1 \in H, \forall \alpha_0, \alpha_1 > 0$ tel que $d(f(P_0), f(P_1)) > \alpha_0 + \alpha_1$,
 $\sum_{n=2}^{+\infty} \int d_n^2(\text{Inf}(d_n(c(P_0, \alpha_0, x^{(n-1)}), c(P_1, \alpha_1, x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)})) = +\infty$

Alors il existe une suite convergente d'estimateurs de f , et même, si (E, d)
 est σ -précompact, une suite p.s. convergente d'estimateurs de f .

Corollaire 9.- Soit $(X^{(+\infty)}, B^{(+\infty)}, H)$ un modèle statistique associé
 à une suite d'observations indépendantes, (E, d) un espace métrique séparable
 et $f : H \rightarrow (E, d)$ un paramètre à estimer.

On suppose que :

- $\forall n \geq 2$, $c(P_n; P \in H)$ est munie d'une distance décantante d_n et
 dominée par une mesure σ -finie sur (X_n, B_n) ;

- $\forall P_0, P_1 \in H, \forall \alpha_0, \alpha_1 > 0$ tels que $d(f(P_0), f(P_1)) > \alpha_0 + \alpha_1$,
 $\sum_{n=2}^{+\infty} \int d_n^2(d_n(c(P_n; d(f(P), f(P_0)) \leq \alpha_0), c(Q_n; d(f(Q), f(P_1)) \leq \alpha_1))) = +\infty$

Alors il existe une suite convergente d'estimateurs de f , et même, si (E, d)
 est σ -précompact, une suite p.s. convergente d'estimateurs de f .

5. - Application de la méthode des boules.- Au lieu de donner une application
 directe du théorème (4) et de ses diverses variantes, nous allons démontrer un
 résultat moins général mais plus maniable, en supposant que les lois conditionnelles

nelles sont des lois à densité. Nous utiliserons le fait bien connu suivant :
 si P et Q sont deux lois de densités f et g par rapport à une même
 mesure σ -finie, alors $d_V(P, Q) = Q(g \geq f) - P(g \geq f)$.

Théorème 6. - Soit $(X^{(+\infty)}, B^{(+\infty)}, H)$ un modèle statistique, (E, d)
 un espace métrique séparable, $f : H \rightarrow E$ un paramètre à estimer. On pose :

$$- \forall P \in H, \forall \alpha > 0, \forall n \geq 2, \forall x^{(n-1)} \in X^{(n-1)},$$

$$r_n(P, \alpha, x^{(n-1)}) = \text{Sup}(d_V(P_{x^{(n-1)}}^{(n-1)}, Q_{x^{(n-1)}}^{(n-1)} ; Q \in f^{-1}(B(f(P), \alpha))) ;$$

$$- \forall P, Q \in H, \forall \alpha, \beta > 0, \forall n \geq 2, \rho_n(P, \alpha; Q, \beta) =$$

$$\text{Inf}(\text{Max}(0, d_V(P_{x^{(n-1)}}^{(n-1)}, Q_{x^{(n-1)}}^{(n-1)}) - r_n(P, \alpha, x^{(n-1)}) - r_n(Q, \beta, x^{(n-1)})) ; x^{(n-1)} \in X^{(n-1)}).$$

On suppose que :

$\forall n \geq 2$, l'ensemble $(P_{x^{(n-1)}}^{(n-1)} ; P \in H, x^{(n-1)} \in X^{(n-1)})$ de toutes
 les lois conditionnelles à l'instant n est dominé par une mesure σ -finie
 μ_n sur (X_n, B_n) et que $\forall P \in H$, on peut choisir une version $f_{n,P}(x^{(n-1)}, \cdot)$

de $\frac{dP_{x^{(n-1)}}}{d\mu_n}$ telle que la fonction $f_{n,P}$ soit $B^{(n)}$ -mesurable ;

$$- \forall n \geq 2, \forall P \in H, \forall \alpha > 0, r_n(P, \alpha, \cdot) \text{ est } B^{(n-1)}\text{-mesurable ;}$$

$$- \forall P, Q \in H, \forall \alpha, \beta > 0 \text{ tels que } d(f(P), f(Q)) > \alpha + \beta,$$

$$\sum_{n=2}^{+\infty} \rho_n^2(P, \alpha; Q, \beta) = +\infty \quad (47)$$

Alors il existe une suite convergente d'estimateurs du paramètre f . De plus,

a) si H est munie d'une probabilité a priori μ , il existe une suite
 d'estimateurs de f convergeant P -p.s. pour μ -presque toute loi possible
 P ;

b) si (E, d) est σ -précompact, il existe une suite p.s. convergente d'estimateurs de f .

Démonstration : Il suffit de montrer que ce modèle vérifie (40). Pour cela, considérons $P, Q \in H$ et $\alpha, \beta > 0$ tels que $d(f(P), f(Q)) > \alpha + \beta$.

a) Soit $n \geq 2$ tel que $\rho_n(P, \alpha; Q, \beta) > 0$.

Par définition de ρ_n , $\forall x^{(n-1)} \in X^{(n-1)}$, nous avons :

$$d_V^{(P)}(x^{(n-1)}, Q_{x^{(n-1)}}) \geq \rho_n(P, \alpha; Q, \beta) + r_n(P, \alpha, x^{(n-1)}) + r_n(Q, \beta, x^{(n-1)})$$

On définit une fonction $\beta^{(n)}$ -mesurable ψ_n en posant :

$$\forall x^{(n-1)} \in X^{(n-1)} \text{ et } \forall x_n \in X_n, \psi_n(x^{(n-1)}, x_n) = 1_{(f_{n,Q} \geq f_{n,P})}(x^{(n)})$$

L'inégalité (48) s'écrit alors :

$$\int \psi_n(x^{(n-1)}, \cdot) dQ_{x^{(n-1)}} - r_n(Q, \beta, x^{(n-1)}) \geq \rho_n(P, \alpha; Q, \beta) + \int \psi_n(x^{(n-1)}, \cdot) dP_{x^{(n-1)}} + r_n(P, \alpha, x^{(n-1)})$$

D'autre part, par définition de r_n , nous avons, $\forall P', Q' \in H$:

$$d(f(P), f(P')) \leq \alpha \implies d_V^{(P)}(x^{(n-1)}, P'_{x^{(n-1)}}) \leq r_n(P, \alpha, x^{(n-1)})$$

$$\text{donc } \int \psi_n(x^{(n-1)}, \cdot) dP'_{x^{(n-1)}} \leq \int \psi_n(x^{(n-1)}, \cdot) dP_{x^{(n-1)}} + r_n(P, \alpha, x^{(n-1)})$$

et de même $d(f(Q), f(Q')) \leq \beta \implies$

$$\int \psi_n(x^{(n-1)}, \cdot) dQ'_{x^{(n-1)}} \geq \int \psi_n(x^{(n-1)}, \cdot) dQ_{x^{(n-1)}} - r_n(Q, \beta, x^{(n-1)})$$

en posant, $\forall x^{(n-1)} \in X^{(n-1)}$,

$$p_n(x^{(n-1)}) = \int \psi_n(x^{(n-1)}, \cdot) dP_{x^{(n-1)}} + r_n(P, \alpha, x^{(n-1)})$$

$$= \int \psi_n(x^{(n-1)}, x_n) f_{n,P}(x^{(n-1)}, x_n) \mu_n(dx_n) + r_n(P, \alpha, x^{(n-1)}) ,$$

on définit une fonction $\mathcal{B}^{(n-1)}$ -mesurable p_n qui, d'après (49), (50) et

(51) vérifie :

$$-\forall P' \in \mathcal{F}^{-1}(B(f(P), \alpha)), \int \psi_n(x^{(n-1)}, \cdot) dP'_{x^{(n-1)}} \leq p_n(x^{(n-1)}) \quad (52)$$

$$-\forall Q' \in \mathcal{F}^{-1}(B(f(Q), \beta)), \int \psi_n(x^{(n-1)}, \cdot) dQ'_{x^{(n-1)}} \geq p_n(x^{(n-1)}) + \rho_n(P, \alpha; Q, \beta) \quad (53)$$

b) Si pour $n \geq 2$, on a $\rho_n(P, \alpha; Q, \beta) = 0$, on pose $\forall n = p_n = 0$, ce qui fait que (52) et (53) sont encore trivialement vérifiées.

c) Finalement, d'après le théorème (2) et compte tenu de (47), il existe une suite adaptée $(C_n)_{n \geq 1}$ suivant laquelle se séparent asymptotiquement uniformément les hypothèses $f^{-1}(B(f(P), \alpha))$ et $f^{-1}(B(f(Q), \beta))$, ce qui prouve bien que (40) est vérifiée.

IV - CONCLUSION.-

Nous venons de décrire le cadre général de l'utilisation de la décantation en estimation asymptotique, et, de manière sous-jacente, dans la construction de suites convergentes de tests. Il faudrait compléter cette étude en ré-actualisant notre travail sur les modèles de Geffroy, modèles statistiques eux aussi très généraux où l'on suppose (cf ^{16, 17}) que les

observations X_n , $n \geq 1$ sont toutes à valeurs dans \mathbb{R}^s , et que toutes les lois conditionnelles du modèle sont les images d'une loi connue sur \mathbb{R}^s par des transformations dont les propriétés géométriques permettront, ou non, d'estimer la loi de $(X_n)_{n \geq 1}$.

A moins d'une nouvelle amélioration du théorème de base (le théorème (2)), les développements de la théorie, attendus ou en cours, concernent les applications que nous avons signalées précédemment et l'étude dans certains cas des vitesses de convergence des estimateurs ainsi construits.

REFERENCES.

1. R. de Almeida - Décantation asymptotique dans les chaînes de Markov ;
thèse de 3ème cycle en cours de rédaction à l'Université
des Sciences et Techniques de Lille Flandres Artois.
2. P. Billingsley - Convergence of probability measures ; Wiley, New-York
(1968).
3. R. Fortet et E. Mourier - Convergence de la fonction de répartition empirique
vers la fonction de répartition théorique ; Ann. Sc.
Ecole Normale Sup. 70 (1953) p. 267-285.
4. J. Geffroy et R. Moché - Sur la séparation asymptotique uniforme des produits
de lois de probabilité ; C.R. Acad. Sc. Paris, t. 278
(1974) p. 969-971.
5. J. Geffroy - Inégalités pour le niveau de signification et la puis-
sance de certains tests reposant sur des données
quelconques ; C.R. Acad. Sc. Paris, t. 282 (1976)
p. 1299-1301.
6. " - Asymptotic separation of distributions and convergence
properties of tests and estimators ; in : Asymptotic
theory of statistical tests and estimation, I.M.
Chakravarti ed. ; Academic Press (1980) p. 159-177.
7. B.V. Gnedenko et A.N. Kolmogorov - Limit distributions for sums of independent
random variables ; Addison-Wesley, Reading (Mass., U.S.A.)
(1954).
8. W. Hoeffding - Probability inequalities for sums of bounded random
variables ; Amer. Stat. Ass. J. (1963) p. 13-30.

9. M. Jirina - Conditional probabilities on σ -algebras with counting basis ; Selected Transl. Math. Stat. and Prob., Vol. (1962) p. 79-85.
10. S. Kakutani - On equivalence of infinite product measures ; Ann. Math., t. 49 (1948) p. 214-224.
11. J. Kiefer et J. Wolfowitz - On the deviations of the empiric distribution function of vector chance variables ; Trans. Amer. Math. Soc. (1958) p. 173-186.
12. C. Kraft - Some conditions for consistency and uniform consistency of statistical procedures ; Univ. California Pub. in Stat., Univ. Calif. Press, Berkeley, Vol. 2 (1955) p. 125-141.
13. K. Matusita - A distance and related statistics in multivariate analysis ; Proc. Int. Symp. on Multivariate Analysis, New-York (1966) p. 187-200.
14. R. Moché - Décantation et séparation asymptotiques uniformes, ... thèse de Doctorat d'Etat (1ère partie) présentée à l'Université des Sciences et Techniques de Lille (1977)
15. " - (s)-modèles statistiques ; Pub. Inst. Stat. Univ. Paris vol. XXVIII , fasc. 1,2 (1983), p. 71-80.
16. " - Asymptotic Estimation in some affine Geffroy models ; Proc. Third Prague Symp. Asympt. Stat. (1983) ; M. Huskova Ed., North-Holland (Amsterdam) (1984) p. 363-369.
17. " - Estimation asymptotique dans certains modèles de Geffroy linéaires ; Statistique et Analyse des Données, Vol. 10 n° 2 (1985) p. 45-62.

18. L. Nachbin - Topology and order ; Van Nostrand, Princeton (New-Jersey, U.S.A.) (1965).
19. W. Pieczynski - Sur diverses applications de la décantation des lois de probabilité dans la théorie générale de l'estimation statistique ; thèse de doctorat d'Etat présentée à l'Université Pierre et Marie Curie (Paris VI). (1986).
20. " - On the maximum likelihood estimation for dependent observations ; Pub. Inst. Stat. Univ. Paris, à paraître.
21. H. Schmerkotte - Interpedependences between consistent tests and estimates ; Z. Wahrscheinlichkeitstheorie verw. Geb. 16 (1970) p. 293-306.

Reçu en Mars 1988

I - Introduction

Several authors have shown interest in the problem of the behaviour of the m.l.e. in the case of dependent observation. Silvey (1961) gave some conditions under which the m.l.e. is asymptotically normal and weakly consistent, but they do not seem to be readily applicable. Barr-Shalom (1972) and Ibragimov (1974) pointed out sufficient conditions for consistency and asymptotic normality. A form of the law of large numbers for martingales is the main tool of their approach. Asymptotic normality is studied with a form of the limit central theorem. Crowder (1975) uses a two term Taylor expansion of the likelihood function.

The approach of this paper is quite different : it is based on Geoffroy's denaturation theorem. This theorem, recently improved by Moche (7), specifies, in the case of dependent random variables, the construction of a consistent estimator whose exponential rate of convergence is ensured by some mild conditions expressed in terms of conditional probabilities.

II - Theoretical investigations.

We consider a statistical model (X, B, P_0) on Θ , with $(X, B) = \bigotimes_{n=1}^{\infty} (X_n, B_n)$.

For each $n \geq 1$ let μ_n be a positive σ -finite measure on (X_n, B_n) , $P_n^{(n)}$ the restriction of P_0 on $(X^{(n)}, B^{(n)}) = \bigotimes_{i=1}^n (X_i, B_i)$. $P_0^{(n)}$ is assumed to be