

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

PHILIPPE CAILLOT  
FRANÇOISE MARTIN

## **Le modèle bayésien**

*Annales de l'I. H. P., section B*, tome 8, n° 2 (1972), p. 183-210

[http://www.numdam.org/item?id=AIHPB\\_1972\\_\\_8\\_2\\_183\\_0](http://www.numdam.org/item?id=AIHPB_1972__8_2_183_0)

© Gauthier-Villars, 1972, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## Le modèle bayésien

par

**Philippe CAILLOT**

I. U. T., Université Grenoble II, 2 rue du Général-Marchand, 38-Grenoble.

et

**Françoise MARTIN**

Laboratoire de Probabilités. Université Paris VI,  
Tour 56, 9 quai Saint-Bernard, Paris-5<sup>e</sup>.

---

**RÉSUMÉ.** — La résolution bayésienne des problèmes statistiques passe par la définition d'une transition *a posteriori*, définie comme une espérance conditionnelle. Il nous a semblé intéressant d'utiliser systématiquement des « pseudo-transitions » pour étudier les notions de réduction, d'exhaustivité, de concentration, et les propriétés asymptotiques du modèle bayésien, afin de mettre en évidence le rôle des hypothèses de régularité qui assurent l'existence de versions régulières des espérances conditionnelles.

**SUMMARY.** — The bayesian methods in statistics use the idea of a *a posteriori* transition, defined as a conditionnal expectation. It seemed to us interesting to make systematical use of « pseudo-transitions » to study the notions of reduction, sufficiency, concentration, and the asymptotical properties of the bayesian model, in order to illuminate the role of the hypothesis of regularity which ensure the existence of regular versions of conditionnal expectations.

---

## I. — PSEUDO-TRANSITIONS

## Définitions et notations.

Un espace mesurable  $\underline{A}$  est un couple  $(\underline{A}, \mathcal{A})$ , où  $\underline{A}$  est un ensemble, et  $\mathcal{A}$  une  $\sigma$ -algèbre de parties de  $\underline{A}$ ; les éléments de  $\underline{A}$  seront systématiquement repérés par la lettre minuscule  $a$ , et ceux de  $\mathcal{A}$  par la même lettre soulignée:  $a \in \underline{A}$ ;  $\underline{a} \in \mathcal{A}$ . Une mesure  $\alpha$  sur l'espace mesurable  $A$  sera notée  $\alpha_A$ . L'ensemble des classes d'équivalence de v. a. r. sur  $A$  pour la relation d'égalité  $\alpha$ -presque sûre sera noté  $L(A, \alpha)$ , et un élément  $f$  de  $L(A, \alpha)$  sera noté  $f^A$ . Nous utiliserons la notation  $(\forall_x a)$  pour une relation vraie  $\alpha$ -presque partout sur  $\underline{A}$ .

(I-1) DÉFINITIONS. — Étant donnés deux espaces mesurables  $A$  et  $B$ , et une mesure  $\alpha_A$ , nous appellerons  $\alpha$ -transition de  $A$  vers  $B$ , notée  $M_B^A$  (ou  $M$ ), une application  $\underline{b} \mapsto M^A(\underline{b})$  de  $\mathcal{B}$  dans  $L(A, \alpha)$ , telle que :

- 1)  $M^A(\Phi) = 0$ ;  $(\forall \underline{b}), M^A(\underline{b}) \geq 0$ .
- 2) Pour toute suite  $\{\underline{b}_n\}$  d'éléments disjoints de  $\mathcal{B}$ , on ait :

$$M^A(\Sigma \underline{b}_n) = \Sigma M^A(\underline{b}_n)$$

Si  $M_B^A$  est une  $\alpha$ -transition, on appelle mesure image de  $\alpha_A$  par  $M_B^A$  la mesure  $\beta_B$  définie par :

$$(\forall \underline{b}), \quad \beta(\underline{b}) = \int M^A(\underline{b}) d\alpha$$

et nous noterons:  $\beta_B = M_B^A \circ \alpha_A$ .

Nous appellerons  $\alpha$ -transition de probabilité de  $A$  vers  $B$  toute  $\alpha$ -transition  $M_B^A$  telle que  $M_A(B) = 1$ .

Un élément positif  $f^A$  de  $L(A, \alpha)$  peut s'identifier à une  $\alpha$ -transition  $M_B^A$  de  $A$  vers un espace mesurable  $B$  tel que  $\mathcal{B}$  soit la  $\sigma$ -algèbre triviale, en posant:  $M^A(B) = f^A$ .

Par abus de langage, nous utiliserons la même notation pour les éléments de  $L(\underline{A}, \alpha)$  et pour leurs représentants. En particulier si  $f^A$  est un élément de  $L(A, \alpha)$ , son intégrale sera aussi notée :

$$\int f^A d\alpha_A \quad \text{ou} \quad \int f(a) \alpha(da).$$

Par suite nous noterons encore  $M_A^B$  toute transition de  $A$  vers  $B$ . [Rappelons

qu'une transition de A vers B est une application  $(a, \underline{b}) \mapsto M(a, \underline{b})$  de  $A \times \mathcal{B}$  dans  $\overline{\mathbb{R}}$  telle que :

- 1)  $(\forall \underline{b})$ ,  $M(\cdot, \underline{b})$  est une v. a. r. sur A, notée  $M^A(\underline{b})$ .
- 2)  $(\forall a)$ ,  $M(a, \cdot)$  est une mesure sur B, notée  $M_B(a)$ .]

Notons  $\mathcal{T}(A, B)$  l'ensemble des transitions de A vers B, et  $\mathcal{T}_\alpha(A, B)$  l'ensemble des  $\alpha$ -transitions de A vers B. Alors pour toute mesure  $\alpha_A$ , on peut écrire :  $\mathcal{T}(A, B) \subset \mathcal{T}_\alpha(A, B)$ . De même si  $\alpha'_A$  est une mesure absolument continue par rapport à  $\alpha_A$ , alors on peut écrire :  $\mathcal{T}_{\alpha'}(A, B) \subset \mathcal{T}_\alpha(A, B)$ .

(I-2) DÉFINITION. — Nous dirons qu'une  $\alpha$ -transition  $M_B^A$  est régulière s'il existe une transition  $N_B^A$  telle que :

$$(\forall \underline{b}), \quad M^A(\underline{b}) = N^A(\underline{b})$$

(I-3) EXEMPLES.

1) Soient deux espaces mesurables A et B. A toute application mesurable  $f$  de A dans B, on peut associer une transition de probabilité  $f_B^A$ , définie par :

$$(\forall a), (\forall \underline{b}), \quad f(a, \underline{b}) = \begin{cases} 1 & \text{si } f(a) \in \underline{b} \\ 0 & \text{sinon} \end{cases}$$

soit encore :

- $(\forall a)$ ,  $f_B^A(a)$  est la mesure de Dirac sur B au point  $f(a)$ .
- $(\forall \underline{b})$ ,  $f^A(\underline{b})$  est la fonction indicatrice de  $f^{-1}(\underline{b})$ .

En particulier la transition de probabilité associée à l'application identique d'un espace mesurable A, notée  $I_A^A$ , sera appelée transition identité.

2) Étant donné un espace mesurable  $A = (\underline{A}, \mathcal{A})$ , une sous-tribu  $\mathcal{A}'$  de  $\mathcal{A}$ , et une probabilité  $\alpha_A$ , si  $A' = (\underline{A}, \mathcal{A}')$ , la probabilité  $\mathcal{A}'$ -conditionnelle définit une  $\alpha$ -transition  $M_A^A$  par :

$$(\forall a), \quad M^{A'}(a) = E_\alpha^{A'}(I^A(a))$$

et on écrira :  $M_A^{A'} = E_\alpha^{A'}(I_A^A)$

**Bi-mesures.**

(I-4) DÉFINITIONS. — Étant donné deux espaces mesurables A et B, on appelle bi-mesure  $\beta_{BA}$  toute fonction positive d'ensemble  $(a, \underline{b}) \mapsto \beta(a, \underline{b})$  définie sur le produit cartésien  $\mathcal{A} \times \mathcal{B}$ , vérifiant les propriétés :

- 1)  $(\forall a)$ , la section  $\beta(a, \cdot)$  est une mesure sur B, notée  $\beta_B(a)$ .
- 2)  $(\forall \underline{b})$ , la section  $\beta(\cdot, \underline{b})$  est une mesure sur A, notée  $\beta_A(\underline{b})$ .

On appelle mesure marginale sur A [respectivement sur B] d'une bi-mesure  $\beta_{BA}$  la mesure  $\beta_A(\underline{B})$  [respectivement  $\beta_B(\underline{A})$ ]. On dit qu'une bi-mesure  $\beta_{BA}$  est  $\sigma$ -finie si ses marginales  $\beta_A(\underline{B})$  et  $\beta_B(\underline{A})$  sont  $\sigma$ -finies.

(I-5) PROPOSITION. — Étant donnés deux espaces mesurables A et B, et une mesure  $\sigma$ -finie  $\alpha_A$ , la formule :

$$(\forall \underline{a}), (\forall \underline{b}), \beta(\underline{a}, \underline{b}) = \int_{\underline{a}} M^A(\underline{b}) d\alpha$$

établit une correspondance bijective entre les  $\alpha$ -transitions de probabilité de A vers B, et les bi-mesures sur AB de marginale  $\alpha_A$  sur A. Nous écrivons la formule précédente :

$$\beta_{BA} = M_B^A \circ \alpha_A$$

*Démonstration.* — Il est clair que si  $M_B^A$  est une  $\alpha$ -transition, la formule précédente définit une bi-mesure  $\beta_{BA}$ .

Inversement soit une bi-mesure  $\beta_{BA}$  telle que  $\beta_A(\underline{B}) = \alpha_A$ . L'inégalité :

$$(\forall \underline{a}), (\forall \underline{b}), \beta(\underline{a}, \underline{b}) \leq \beta(\underline{a}, \underline{B}) = \alpha(\underline{a})$$

entraîne l'absolue continuité des mesures  $\beta_A(\underline{b})$  par rapport à  $\alpha_A$ . Donc il existe une  $\alpha$ -transition  $M_B^A$  unique, vérifiant la relation annoncée, qui est définie par :

$$(\forall \underline{b}), M^A(\underline{b}) = \frac{d\beta_A(\underline{b})}{d\alpha_A}$$

*N. B.* — S'il n'y a aucune ambiguïté, on pourra omettre la source et le but des pseudo-transitions, et écrire par exemple :  $M \circ \alpha$  ou  $M \circ \alpha$ .

(I-6) REMARQUES.

1) La formule précédente peut s'interpréter en disant que, pour tout  $\underline{b}$ ,  $M^A(\underline{b})$  est « l'espérance  $\mathcal{A}$ -conditionnelle pour la bi-mesure  $\beta$  » de  $I^B(\underline{b})$ , et s'écrire :

$$M_B^A = E_{\beta}^A(I_B^B)$$

2) Il est clair que toute bi-mesure  $\beta_{BA}$  est additive sur la semi-algèbre  $\mathcal{A} \times \mathcal{B}$ , donc peut se prolonger de façon unique en une fonction additive d'ensembles sur l'algèbre engendrée  $s(\mathcal{A} \times \mathcal{B})$ . Mais cette fonction n'est pas nécessairement  $\sigma$ -additive, donc ne peut se prolonger en une mesure sur la  $\sigma$ -algèbre  $\mathcal{A} \vee \mathcal{B}$ . Donnons en effet un contre-exemple dont l'idée vient de [6] :

Soit X l'espace mesurable formé du segment  $[0, 1]$  muni de sa tribu

borélienne, et soit  $\lambda$  la mesure de Lebesgue sur  $X$ . Soit  $\underline{A}$  une partie de  $\underline{X}$  de mesure extérieure 1 et de mesure intérieure 0 (l'existence d'un tel ensemble est montrée dans [3]). Munissons  $\underline{A}$  et  $\underline{B} = \underline{X} - \underline{A}$  des tribus  $\mathcal{A}$  et  $\mathcal{B}$ , traces de  $\mathcal{X}$ . Enfin soit  $\alpha$  la mesure trace de  $\lambda$  sur  $A$ . On vérifie facilement que la formule :

$$(\forall \underline{x}), \quad T^A(\underline{x} \cap \underline{B}) = I^A(\underline{x} \cap \underline{A})$$

définit une  $\alpha$ -transition  $T_B^A$ . La bi-mesure  $\beta = M \circ \alpha$  est alors définie par :

$$(\forall \underline{x}), \quad (\forall \underline{x}'), \quad \beta(\underline{x} \cap \underline{A}, \underline{x}' \cap \underline{B}) = \lambda(\underline{x} \cap \underline{x}')$$

Dans l'espace mesurable produit  $Y = X \otimes X$ , la tribu  $\mathcal{C} = \mathcal{A} \vee \mathcal{B}$  sur  $\underline{C} = \underline{A} \times \underline{B}$  est évidemment la trace sur  $\underline{C}$  de la tribu  $\mathcal{Y}$ . Par suite si la bi-mesure  $\beta_{BA}$  pouvait se prolonger en une mesure  $\beta_C$ , la formule :

$$(\forall \underline{y}), \quad \beta^*(\underline{y}) = \beta(\underline{y} \cap \underline{C})$$

définirait une mesure  $\beta_C^*$ . D'après la définition de  $\beta$ , cette mesure devrait coïncider sur les pavés mesurables de  $Y$  avec la mesure  $\lambda_C^*$ , définie par :

$$(\forall \underline{x}), \quad (\forall \underline{x}'), \quad \lambda^*(\underline{x} \times \underline{x}') = \lambda(\underline{x} \cap \underline{x}')$$

Or c'est impossible car, si  $\underline{y}$  est la diagonale de  $\underline{Y}$ , on a :

$$\lambda^*(\underline{y}) = 1 \quad \text{et} \quad \beta^*(\underline{y}) = \beta(\underline{y} \cap \underline{C}) = \beta(\Phi) = 0$$

2) D'autre part, si  $T_B^A$  est une  $\alpha_A$ -transition régulière, la bi-mesure  $T_B^A \circ \alpha_A$  est  $\sigma$ -additive, mais la réciproque n'est pas vraie. En effet dans le deuxième exemple de (I-3), on peut supposer que la probabilité  $\mathcal{A}'$ -conditionnelle n'admette pas de version régulière ; cependant la bi-mesure  $\beta_{A'A} = T_{A'}^A \circ \alpha_A$ , définie par :

$$(\forall \underline{a}), \quad (\forall \underline{a}'), \quad \beta(\underline{a}, \underline{a}') = \int_{\underline{a}'} T^A(\underline{a}') d\alpha = \alpha(\underline{a} \cap \underline{a}')$$

est  $\sigma$ -additive.

Nous allons maintenant donner des conditions suffisantes, classiques, pour assurer soit la régularité d'une  $\alpha$ -transition, soit la  $\sigma$ -additivité d'une bi-mesure.

(I-7) PROPOSITION. — Étant donnés deux espaces mesurables  $A$  et  $B$ , et une mesure  $\alpha_A$ , pour qu'une  $\alpha$ -transition  $M_B^A$  soit régulière, il suffit que la tribu  $\mathcal{B}$  soit dénombrablement engendrée et contienne une classe compacte d'approximation pour la mesure  $M \circ \alpha$ . Pour qu'il en soit ainsi, il suffit que  $B$  soit un espace polonais muni de sa tribu borélienne et que la mesure  $M \circ \alpha$  soit  $\sigma$ -finie.

*Démonstration.* — Voir [8], proposition (V-4-5).

(I-8) PROPOSITION. — Étant donnés deux espaces mesurables A et B, pour qu'une bi-mesure  $\sigma$ -finie  $\beta_{BA}$  soit  $\sigma$ -additive, il suffit que, pour l'un des espaces A ou B, toute sous-tribu dénombrablement engendrée admette une classe compacte d'approximation pour sa marginale.

*Démonstration.* — Supposons que l'espace mesurable B vérifie la condition annoncée. Posant  $\alpha_A = \beta_A(\underline{B})$ , soit  $M_B^A$  la  $\alpha$ -transition définie par :  $\beta = M \circ \alpha$  (cf. (I-5)). Soient  $\underline{a} \times \underline{b} = \Sigma \underline{a}_n \times \underline{b}_n$ ,  $\mathcal{B}^*$  la sous-tribu de  $\mathcal{B}$  engendrée par la suite  $\{\underline{b}_n\}$ , et  $B^* = (\underline{B}, \mathcal{B}^*)$ . Alors, d'après la proposition précédente, la restriction de  $M_B^A$  à  $B^*$  admet une version régulière, donc la restriction  $\beta_{B^*A}$  de  $\beta_{BA}$  à  $B^*A$  est  $\sigma$ -additive. D'où le résultat.

(I-9) DÉFINITION. — Étant donnés deux espaces mesurables A et B, et une mesure  $\lambda_A$ , on dit qu'une  $\lambda$ -transition  $M_B^A$  est faiblement dominée si la bi-mesure  $\gamma = M \circ \lambda$  est  $\sigma$ -finie,  $\sigma$ -additive, et admet une densité par rapport à une mesure produit  $\mu_A \otimes \nu_B$ .

(I-10) PROPOSITION. — Avec les notations précédentes, si  $M_B^A$  est faiblement dominée, alors  $\gamma = M \circ \alpha$  admet une densité par rapport au produit de ses marginales, et par suite  $M_B^A$  est régulière.

*Démonstration.* — Posons :

$$\alpha_A = \gamma_A(\underline{B}) = M^A(\underline{B}) \circ \lambda_A \quad \text{et} \quad \beta_B = \gamma_B(\underline{A}) = M_B^A \circ \lambda_A$$

Si  $h$  est une version de la densité de  $\gamma$  par rapport à  $\mu \otimes \nu$ , et si on pose :

$$\varphi(a) = \int h(a, b) \nu(db) \quad \text{et} \quad \psi(b) = \int h(a, b) \mu(da)$$

on a :

$$\alpha_A = \varphi^A \circ \mu_A \quad \text{et} \quad \beta_B = \psi^B \circ \nu_B$$

La bi-mesure  $\gamma_{BA}$  étant  $\sigma$ -finie, on peut supposer que les fonctions  $h$ ,  $\varphi$  et  $\psi$  sont finies. Posons :

$$\underline{a}_0 = \{\varphi = 0\}, \quad \underline{b}_0 = \{\psi = 0\}, \quad \text{et} \quad C_0 = (\underline{a}_0 \times \underline{B}) \cup (\underline{A} \times \underline{b}_0)$$

On a :

$$\alpha(\underline{a}_0) = \beta(\underline{b}_0) = \alpha \otimes \beta(C_0) = \gamma(C_0) = 0$$

La fonction :

$$f(a, b) = \frac{h(a, b)}{\varphi(a)\psi(b)} 1_{C_0^c} \subset (a, b)$$

est donc partout définie, et on vérifie immédiatement que :

$$(\forall \underline{a}), (\forall \underline{b}), \int_{\underline{a} \times \underline{b}} f(a, b)\alpha(da)\beta(db) = \gamma(\underline{a}, \underline{b})$$

Enfin si  $m(a)$  est une version finie de  $M^A(\underline{B})$ , en posant :

$$(\forall \underline{a}), (\forall \underline{b}), N(a, \underline{b}) = m(a) \int_{\underline{b}} f(a, b)\beta(db)$$

on définit une transition  $N_B^A$  telle que :

$$N_B^A \circ \lambda_A = \gamma_{BA}.$$

Donc,

$$(\forall \underline{b}), N^A(\underline{b}) = M^A(\underline{b})$$

(I-11) DÉFINITION. — Si  $M_B^A$  est une  $\lambda_A$ -transition faiblement dominée, on appellera densité faible de  $M$  pour  $\lambda$  la densité de  $M \circ \lambda$  par rapport au produit de ses marginales.

**Composition des pseudo-transitions.**

Étant donnés deux espaces mesurables  $A$  et  $B$ , notons  $s(\mathcal{A} \times \mathcal{B})$  l'algèbre engendrée par les pavés mesurables de  $\underline{A} \times \underline{B}$  et  $\mathcal{L}^\infty(\mathcal{A} \times \mathcal{B})$  l'espace de Banach des fonctions sur  $\underline{A} \times \underline{B}$  qui sont limites uniformes de fonctions  $s(\mathcal{A} \times \mathcal{B})$ -étagées. Une bi-mesure  $\beta$  se prolonge en un contenu positif unique sur  $s(\mathcal{A} \times \mathcal{B})$ , et par suite on peut définir l'intégrale par rapport à une bi-mesure de tout élément positif  $f$  de  $\mathcal{L}^\infty(\mathcal{A} \times \mathcal{B})$  (cf. [2]), ou de sa classe d'équivalence  $f^{AB}$  pour la relation d'égalité  $\beta$ -presque sûre, et nous écrivons :

$$f^{AB} \circ \beta_{BA} = \int f^{AB} d\beta$$

Si  $M_B^A$  et  $N_C^B$  sont deux transitions, on définit la transition composée  $P_B^A = N_C^B \circ M_B^A$  par :

$$(\forall \underline{a}), (\forall \underline{c}), P(a, \underline{c}) = \int N(b, \underline{c})M(\underline{a}, db)$$

et, si  $\alpha_A$  est une mesure, on a :

$$(\forall \underline{a}), (\forall \underline{c}), \int_{\underline{a}} P(\underline{a}, \underline{c})\alpha(da) = \int_{\underline{a}} N(b, \underline{c})M \circ \alpha(da, db)$$



Si  $M_B^A$  est une  $\alpha_A$ -transition et  $N_C^B$  une  $M \circ \alpha$ -transition, l'intégrale

$$\int I^A(\underline{a}) N^B(\underline{c}) d\beta_{BA}$$

définit une bi-mesure  $\gamma_{CA}$  de marginale  $\alpha_A$  sur  $A$ , ce qui nous conduit à poser :

(I-12) DÉFINITION. — Si  $\alpha_A$  est une mesure  $\sigma$ -finie, si  $M_B^A$  est une  $\alpha$ -transition et  $N_C^B$  une  $M \circ \alpha$ -transition, leur composée est la  $\alpha$ -transition  $P_C^A$  définie par la formule :

$$(\forall \underline{a}), (\forall \underline{c}), \int_a P^A(\underline{c}) d\alpha_A = \int I^A(\underline{a}) N^B(\underline{c}) d\beta_{BA}$$

que nous écrivons :

$$P_C^A = N_C^B \circ M_B^A$$

On peut montrer que l'opération de composition est associative, et que

$$M_B^A \circ I_A^A = I_B^B \circ M_B^A = M_B^A$$

(I-13) REMARQUES.

1) Si  $\alpha'_A$  est une mesure  $\sigma$ -finie absolument continue par rapport à  $\alpha_A$ , on peut considérer  $M_B^A$  comme une  $\alpha'$ -transition, et, la mesure  $M \circ \alpha'$  étant absolument continue par rapport à  $M \circ \alpha$ ,  $N_C^B$  comme une  $M \circ \alpha'$ -transition. Alors si  $\beta' = M \circ \alpha' \beta'$ , on a :

$$(\forall \underline{a}), (\forall \underline{c}), \int_a P^A(\underline{c}) d\alpha'_A = \int I^A(\underline{a}) N^B(\underline{c}) d\beta'_{BA}$$

En effet la formule est vraie si la densité  $f^A$  de  $\alpha'_A$  par rapport à  $\alpha_A$  est une fonction étagée, donc aussi, par passage à la limite, pour toute fonction positive  $f^A$ .

2) La définition précédente peut s'interpréter en disant que, pour tout  $\underline{c}$ ,  $P^A(\underline{c})$  est « l'espérance  $\mathcal{A}$ -conditionnelle pour la bi-mesure  $\beta$  » de  $N^B(\underline{c})$ , et s'écrire :

$$N_C^B \circ M_A^B = E_{\beta}^A(N_C^B)$$

Nous avons alors la formule (cf. I-6) :

$$N_C^B \circ E_{M \circ \alpha}^A(I_B^B) = E_{M \circ \alpha}^A(N_C^B)$$

## II. — MODÈLE BAYÉSIEN

## Introduction.

Généralement, un modèle bayésien est défini par

- une mesure *a priori*  $\mu_{\Theta}$  (qui n'est pas nécessairement une probabilité) sur l'espace mesurable des paramètres ;
- $P_x^{\Theta}$  une transition de probabilité de l'espace des paramètres  $\Theta$  dans l'espace des observations  $X$ .

La résolution bayésienne des problèmes statistiques posés (estimation, tests, ...) passe par la recherche d'une transition *a posteriori*  $Q_{\Theta}^X$  qui, à chaque observation  $x$  associe une distribution de probabilité  $Q_{\Theta}^x$  sur  $\Theta$  sensée représenter l'information que l'observation  $x$  contient sur le paramètre  $\theta$  ; cette transition est étudiée ci-dessous.

Un problème est qu'il n'existe pas nécessairement de telle transition mais seulement une P-transition, que la méthode même de construction — le conditionnement — ne peut conduire qu'à cela, et que l'existence d'une version régulière est un problème essentiellement mathématique qui n'a que peu à voir avec le fond méthodologique. Dans ces conditions il nous a semblé intéressant d'utiliser systématiquement des pseudo-transitions et de mettre en évidence le rôle des hypothèses de régularité.

(II-2) DÉFINITION. — Nous appellerons modèle bayésien un couple  $(\mu_{\Theta}, P_X^{\Theta})$  où :

- 1)  $\mu_{\Theta}$  est une mesure  $\sigma$ -finie sur un espace mesurable  $\Phi$ .
- 2)  $P_X^{\Theta}$  est une  $\mu$ -transition de probabilité de  $\Theta$  dans un espace mesurable  $X$ , telle que la mesure  $P_X^{\Theta} \circ \mu_{\Theta}$ , notée  $P_X^{\mu}$ , soit  $\sigma$ -finie.

Nous noterons  $\pi_{\Theta X}$  la bi-mesure  $P_X^{\Theta} \circ \mu_{\Theta}$ . Alors  $P_X^{\mu} = \pi_X(\Theta)$ . Il résulte de (I-5) que le modèle bayésien est défini de façon équivalente par la bi-mesure  $\sigma$ -finie  $\pi_{\Theta X}$ , par les formules :

$$\mu_{\Theta} = \pi_{\Theta}(X) \quad ; \quad P_X^{\Theta} = E_{\pi}^{\Theta}(I_X^{\Theta})$$

(II-3) DÉFINITION. — On appelle pseudo-transition *a posteriori* d'un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$  la  $P^{\mu}$ -transition  $Q_{\Theta}^X$  définie par :

$$Q_{\Theta}^X \circ P_X^{\mu} = P_X^{\Theta} \circ \mu_{\Theta}$$

soit :

$$Q_{\Theta}^X = E_{\pi}^X(I_{\Theta}^{\Theta})$$

ou encore :

$$(\forall \underline{\theta}), \quad Q^X(\underline{\theta}) = \frac{d\pi_X(\underline{\theta})}{d\pi_X(\Theta)}$$

(II-4) REMARQUES. — L'existence d'une transition *a posteriori*, c'est-à-dire la régularité de  $Q_{\Theta}^X$ , est assurée au moins dans les deux cas suivants :

1) si la tribu de  $\Theta$  est dénombrablement engendrée et possède une classe compacte d'approximation pour  $\mu_{\Theta}$  (I-7), donc en particulier si  $\Theta$  est un espace polonais muni de sa tribu borélienne ;

2) si la pseudo-transition  $P_X^{\Theta}$  est faiblement dominée, et par suite régulière (I-10). En effet, puisque  $Q \circ P^{\mu} = P \circ \mu$ ,  $Q_{\Theta}^X$  est elle-même faiblement dominée, donc régulière. Une version régulière est alors donnée par :

$$Q(x, \underline{\theta}) = \int_{\Theta} f(\theta, x) \mu(d\theta)$$

où  $f$  est une version de la densité de  $P$  pour  $\mu$ .

### Réduction par une opération bayésienne.

(II-5) DÉFINITION. — Nous appellerons opération bayésienne sur un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$  une  $P^{\mu}$ -transition de probabilité  $T_X^X$ , de  $X$  dans un espace mesurable  $X'$ , et modèle bayésien image le modèle  $(\mu_{\Theta}, P_{X'}^{\Theta})$ , où  $P_{X'}^{\Theta} = T_X^X \circ P_X^{\Theta}$ .

Autrement dit (I-12), la  $\mu$ -transition  $P_{X'}^{\Theta}$  est définie par :

$$(\forall \underline{\theta}), \quad (\forall \underline{x}'), \quad \int_{\Theta} P_{X'}^{\Theta}(\underline{x}') d\mu_{\Theta} = \int I^{\Theta}(\underline{\theta}) T_X^X(\underline{x}') d\pi_{\Theta X}$$

ce qui s'écrit encore, en posant  $\pi' = P' \circ \mu$  :

$$(\forall \underline{\theta}), \quad \pi_{X'}(\underline{\theta}) = T_X^X \circ \pi_X(\underline{\theta})$$

(II-6) PROPOSITION. — Soit  $T_X^X$ , une opération bayésienne sur un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$ . La pseudo-transition *a posteriori* du modèle image est donnée par la formule :

$$(\forall \underline{\theta}), \quad Q^{X'}(\underline{\theta}) = E_{T \circ P^{\mu}}^{X'}(Q^X(\underline{\theta}))$$

*Démonstration.* — On doit montrer que :

$$(\forall \underline{\theta}), \quad (\forall \underline{x}'), \quad \int_{X'} Q^{X'}(\underline{\theta}) d(T \circ P^{\mu}) = \int_{X'} Q^X(\underline{\theta}) d(T \circ P^{\mu})$$

La marginale sur  $X'$  de la bi-mesure  $T \circ P^\mu$  est la mesure  $T \circ P^\mu = P'^\mu$ ; donc, par définition de  $Q_\Theta^{X'}$ , le premier membre de l'égalité précédente est  $\pi'(\theta, \underline{x}')$ . D'autre part le deuxième membre s'écrit :

$$\int Q^X(\theta) T^X(\underline{x}') dP_X^\mu = \int T^X(\underline{x}') d\pi_X(\theta) = \pi'(\theta, \underline{x}')$$

(II-7) REMARQUE. — Notons  $S_X^{X'}$  la  $P'^\mu$ -transition définie par :

$$T_{X'}^X \circ P_X^\mu = S_X^{X'} \circ P_X'^\mu$$

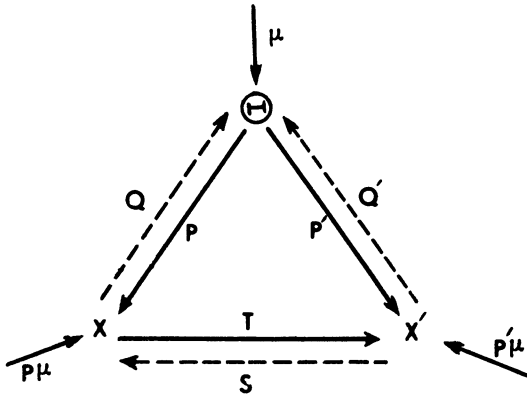
soit :

$$S_X^{X'} = E_{T \circ P^\mu}^{X'}(I_X^X)$$

Alors la relation de la proposition (II-6) s'écrit encore (cf. I-13) :

$$Q_\Theta^{X'} = Q_\Theta^X \circ S_X^{X'}$$

Représentons les éléments que nous venons d'introduire par le schéma suivant :



Alors les deux diagrammes, direct et inverse, sont commutatifs. De plus les pseudo-transitions  $P$  et  $Q$  sont réciproques par rapport au couple de mesures  $(\mu_\Theta, P_X^\mu)$ , au sens suivant :

$$P \circ \mu = Q \circ P^\mu$$

De même  $P'$  et  $Q'$  sont réciproques par rapport au couple  $(\mu_\Theta, P_X'^\mu)$  et  $T$  et  $S$  par rapport au couple  $(P_X^\mu, P_X'^\mu)$ .

Il est intéressant de chercher à quelles conditions les pseudo-transitions  $T$  et  $S$  sont réciproques l'une de l'autre non seulement pour le couple  $(P_X^\mu, P_X'^\mu)$ ,

mais aussi pour tous les couples  $(\pi_{\mathbf{x}}(\theta), \pi'_{\mathbf{x}'}(\theta))$ . Si  $P_{\mathbf{x}}^{\theta}$  est une transition, alors  $P'_{\mathbf{x}'}^{\theta}$  est aussi une transition, et dire que les pseudo-transitions  $T$  et  $S$  sont réciproques pour tous les couples  $(\pi_{\mathbf{x}}(\theta), \pi'_{\mathbf{x}'}(\theta))$  équivaut à dire qu'elles sont réciproques pour « presque » tous les couples  $(P_{\mathbf{x}}(\theta), P'_{\mathbf{x}'}(\theta))$  au sens suivant :

$$(\forall \underline{x}), (\forall \underline{x}'), (\forall_{\mu} \theta), \quad T \circ P(\theta)(\underline{x}, \underline{x}') = S \circ P'(\theta)(\underline{x}, \underline{x}')$$

Si  $X' = (\underline{X}, \mathcal{X}')$  est un sous-espace mesurable de  $X$ , et si  $T_{X'}^X = I_{X'}^X$  est la transition associée à l'application identique de  $X$  dans  $X'$ , alors la relation précédente s'écrit :

$$(\forall \underline{x}), (\forall \underline{x}'), (\forall_{\mu} \theta), \quad P^{\theta}(\underline{x} \cap \underline{x}') = \int_{\underline{x}'} S^{X'}(\underline{x}) dP_{X'}^{\theta}$$

ce qui correspond à l'idée d'exhaustivité  $\mu$ -presque sûre de la sous-tribu  $\mathcal{X}'$  de  $\mathcal{X}$ .

(II-8) DÉFINITION. — Etant un modèle bayésien  $(\mu_{\Theta}, P_{\mathbf{x}}^{\Theta})$ , nous dirons qu'une opération bayésienne  $T_{X'}^X$  est  $\mu$ -exhaustive s'il existe une  $T \circ P^{\mu}$ -transition  $S_{X'}^X$  telle que :

$$(\forall \theta), \quad T_{X'}^X \circ \pi_{\mathbf{x}}(\theta) = S_{X'}^X \circ \pi'_{\mathbf{x}'}(\theta)$$

ce qui s'écrit encore :

$$(\forall \theta), \quad S_{X'}^X = E_{T \circ \pi(\theta)}^{X'}(I_{X'}^X)$$

(II-9) PROPOSITION. — Pour qu'une opération bayésienne  $T_{X'}^X$  sur un modèle bayésien  $(\mu_{\Theta}, P_{\mathbf{x}}^{\Theta})$  soit  $\mu$ -exhaustive, il faut et il suffit qu'il existe une  $P^{\mu}$ -transition de probabilité  $\Sigma_{X'}^X$  telle que

$$(\forall \theta), \quad \Sigma_{X'}^X \circ \pi'_{\mathbf{x}'}(\theta) = \pi_{\mathbf{x}}(\theta)$$

*Démonstration.* — La mesure  $\mu_{\Theta}$  étant  $\sigma$ -finie, il existe une suite  $\{\theta_n\}$  d'ensemble mesurables disjoints telle que :

$$(\forall n), \quad \mu(\theta_n) < +\infty.$$

On définit alors une probabilité équivalente  $\hat{\mu}_{\Theta}$  à  $\mu_{\Theta}$ , en posant (avec la convention  $\frac{0}{0} = 0$ ) :

$$(\forall \theta), \quad \hat{\mu}(\theta) = \sum_n \frac{\mu(\theta \cap \theta_n)}{\mu(\theta_n) 2^n}$$

Si  $(\mu, P')$  est le modèle bayésien image de  $(\mu, P)$  par  $T$ , alors  $(\hat{\mu}, P')$  est le modèle bayésien image de  $(\hat{\mu}, P)$  par  $T$  (cf. I-13). On vérifie facilement que pour une  $P''$ -transition de probabilité  $S_X^{X'}$ , les deux relations suivantes sont équivalentes :

$$\begin{aligned} (\forall \theta), \quad T \circ \pi(\theta) &= S \circ \pi'(\theta) \\ (\forall \theta), \quad T \circ \hat{\pi}(\theta) &= S \circ \hat{\pi}'(\theta) \end{aligned}$$

ainsi que les relations :

$$\begin{aligned} (\forall \theta), \quad T \circ \pi(\theta) &= S \circ \pi'(\theta) \\ (\forall \theta), \quad T \circ \hat{\pi}(\theta) &= S \circ \hat{\pi}'(\theta) \end{aligned}$$

On peut donc supposer que  $\mu_\theta$  est une probabilité. Posons alors, pour tout  $\theta$  :

$$R_X^\theta = \begin{cases} \frac{1}{\mu(\theta)} \pi_X(\theta) & \text{si } \mu(\theta) > 0 \\ P_X^\mu & \text{si } \mu(\theta) = 0 \end{cases}$$

On a :

$$T_X^{X'} \circ R_X^\theta = \begin{cases} \frac{1}{\mu(\theta)} (T_X^{X'} \circ \pi_X(\theta)) & \text{si } \mu(\theta) > 0 \\ T_X^{X'} \circ P_X^\mu & \text{si } \mu(\theta) = 0 \end{cases}$$

Donc la  $\mu$ -exhaustivité de  $T$  est équivalente à l'existence d'une  $P''$ -transition de probabilité  $S_X^{X'}$  telle que :

$$(\forall \theta), \quad T \circ R^\theta = S \circ (T \circ R^\theta)$$

Comme  $\{R_X^\theta; \theta \in \mathcal{H}\}$  est une famille de probabilités sur  $X$  dominée par  $P_X^\mu$ , la  $\mu$ -exhaustivité de  $T$  est encore équivalente à l'existence d'une  $P''$ -transition  $\Sigma_X^{X'}$  telle que [7] :

$$(\forall \theta), \quad T_X^{X'} \circ R_X^\theta = \Sigma_X^{X'} \circ (T_X^{X'} \circ R_X^\theta)$$

D'où le résultat.

*Remarque.* — Si  $X' = (\underline{X}, \mathcal{X}')$  est un sous-espace mesurable de  $X$ , et si  $T_X^{X'} = I_X^{X'}$  est la transition associée à l'application identique de  $X$  dans  $X'$ , alors la  $\mu$ -exhaustivité de  $T$  équivaut à l'exhaustivité au sens classique [4] de la sous-tribu  $\mathcal{X}'$  de  $\mathcal{X}$  dans le modèle statistique « lissé par  $\mu$  »  $(X, \{R_X^\theta; \theta \in \mathcal{H}\})$ . La proposition suivante exprime que  $\mathcal{X}'$  est une sous-tribu  $\mu$ -exhaustive de  $\mathcal{X}$  si et seulement si la pseudo-transition *a posteriori* possède une version  $\mathcal{X}'$ -mesurable.

(II-10) PROPOSITION. — Pour qu'une opération bayésienne  $T_X^{X'}$  sur

un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$  soit  $\mu$ -exhaustive, il faut et il suffit que la pseudo-transition *a posteriori* du modèle initial et du modèle image vérifient :

$$Q_{\Theta}^X = Q_{\Theta}^{X'} \circ T_X^X.$$

*Démonstration.* — On doit montrer que l'opération bayésienne  $T_X^X$  est  $\mu$ -exhaustive si et seulement si on a la relation :

$$(Q_{\Theta}^{X'} \circ T_X^X) \circ P_X^{\Theta} = (P_X^{\Theta} \circ \mu_{\Theta})$$

ce qui s'écrit encore, en utilisant (I-12) :

$$(\forall \theta), (\forall x), \int I^X(x) Q^{X'}(\theta) d(T \circ P^{\mu})_{XX'} = \pi(\theta, x)$$

Soit  $S_X^{X'} = E_{T \circ P^{\mu}}^{X'}(I_X^X)$ . Le premier membre de l'égalité précédente s'écrit encore, en utilisant (II-3) :

$$\begin{aligned} \int I^X(x) Q^{X'}(\theta) d(S \circ P^{\mu})_{XX'} &= \int Q^{X'}(\theta) S^{X'}(x) dP_X^{\mu} \\ &= \int S^{X'}(x) d\pi_{X'}(\theta) = (S \circ \pi'(\theta))(x) \end{aligned}$$

La relation annoncée est donc équivalente à la suivante

$$(\forall \theta), S_X^{X'} \circ \pi_{X'}'(\theta) = \pi_X(\theta)$$

Si T est  $\mu$ -exhaustive, alors cette relation est vérifiée. Inversement cette relation entraîne la  $\mu$ -exhaustivité de T, en vertu de la proposition (II-9).

### Opérations bayésiennes et décision.

(II-11) DÉFINITIONS. — Étant donné un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$  et un espace mesurable A (espace des décisions), nous appellerons  $\mu$ -stratégie une  $P^{\mu}$ -transition de probabilité  $U_A^X$ ; étant donnée de plus une fonction  $L \in \mathcal{L}^{\infty}(\mathcal{H} \times \mathcal{A})$  (fonction de coût), nous appellerons risque bayésien de la  $\mu$ -stratégie  $U_A^X$  l'expression :

$$R(U_A^X) = \int L^{\Theta A} du_{\Theta A}$$

où  $u_{\Theta A}$  est la bi-mesure  $(U_A^X \circ P_X^{\Theta}) \circ \mu_{\Theta}$ .

Remarquons que, si les pseudo-transitions  $U_A^X$  et  $P_X^{\Theta}$  sont régulières, alors on retrouve l'expression classique du risque bayésien :

$$R(U_A^X) = \int L(\theta, a) U(x, da) P(\theta, dx) \mu(d\theta)$$

Étant donnés un modèle bayésien  $(\mu_{\Theta}, P_{\mathbf{X}}^{\Theta})$ , un espace de décision  $A$ , et une opération bayésienne  $T$  sur  $(\mu_{\Theta}, P_{\mathbf{X}}^{\Theta})$ , on peut faire correspondre à toute  $\mu$ -stratégie  $V_{\mathbf{A}}^{\mathbf{X}'}$  une  $\mu$ -stratégie  $U_{\mathbf{A}}^{\mathbf{X}}$  ayant, pour toute fonction de coût,  $L \in \mathcal{L}^{\infty}(\mathcal{H} \times \mathcal{A})$ , même risque bayésien. En effet il suffit de prendre  $U_{\mathbf{A}}^{\mathbf{X}} = V_{\mathbf{A}}^{\mathbf{X}'} \circ T_{\mathbf{X}}^{\mathbf{X}'}$ . On s'intéresse habituellement à l'inverse de cette propriété.

(II-12) PROPOSITION. — Soient un modèle bayésien  $(\mu_{\Theta}, P_{\mathbf{X}}^{\Theta})$  et une opération bayésienne  $T_{\mathbf{X}}^{\mathbf{X}'}$  sur lui. Pour que, à tout espace de décision  $A$  et à toute  $\mu$ -stratégie  $U_{\mathbf{A}}^{\mathbf{X}}$ , on puisse faire correspondre une  $\mu$ -stratégie  $V_{\mathbf{A}}^{\mathbf{X}'}$  ayant même risque bayésien que  $U_{\mathbf{A}}^{\mathbf{X}}$  pour toute fonction de coût  $L \in \mathcal{L}^{\infty}(\mathcal{H} \times \mathcal{A})$ , il faut et il suffit que l'opération bayésienne  $T_{\mathbf{X}}^{\mathbf{X}'}$  soit  $\mu$ -exhaustive.

*Démonstration.*

1) Supposons  $T_{\mathbf{X}}^{\mathbf{X}'}$   $\mu$ -exhaustive. Soit  $S_{\mathbf{X}}^{\mathbf{X}'}$  la  $P^{\mu}$ -transition réciproque. Posons  $V_{\mathbf{A}}^{\mathbf{X}'} = U_{\mathbf{A}}^{\mathbf{X}} \circ S_{\mathbf{X}}^{\mathbf{X}'}$ . Il suffit de montrer que :

$$v_{\Theta A} = (V_{\mathbf{A}}^{\mathbf{X}'} \circ P_{\mathbf{X}'}^{\Theta}) \circ \mu_{\Theta} = (U_{\mathbf{A}}^{\mathbf{X}} \circ P_{\mathbf{X}}^{\Theta}) \circ \mu_{\Theta} = u_{\Theta A}$$

Or pour tout couple  $(\underline{\theta}, \underline{a})$ , on a :

$$\begin{aligned} v(\underline{\theta}, \underline{a}) &= \int_{\underline{\theta}} (V \circ P')^{\Theta}(\underline{a}) d\mu_{\Theta} = \int I^{\Theta}(\underline{\theta}) V^{\mathbf{X}'}(\underline{a}) d(P' \circ \mu)_{\Theta \mathbf{X}'} \\ &= \int V^{\mathbf{X}'}(\underline{a}) d\pi'_{\mathbf{X}'}(\underline{\theta}) = \int U^{\mathbf{X}}(\underline{a}) d(S \circ \pi'(\underline{\theta}))_{\mathbf{X}} \\ &= \int U^{\mathbf{X}}(\underline{a}) d\pi_{\mathbf{X}}(\underline{\theta}) = u(\underline{\theta}, \underline{a}) \end{aligned}$$

2) Inversement, prenons  $A = X$  et  $U_{\mathbf{A}}^{\mathbf{X}} = I_{\mathbf{X}}^{\mathbf{X}}$ . Par hypothèse il existe alors une  $P^{\mu}$ -transition  $V_{\mathbf{X}}^{\mathbf{X}'}$  ayant, pour toute fonction de coût

$$L \in \mathcal{L}^{\infty}(\mathcal{H} \times \mathcal{A}),$$

même risque bayésien que  $I_{\mathbf{X}}^{\mathbf{X}}$ . Donc, pour tout couple  $(\underline{\theta}, \underline{a})$ , on a :

$$\int I^{\mathbf{X}}(\underline{x}) d\pi_{\mathbf{X}}(\underline{\theta}) = \int V^{\mathbf{X}'}(\underline{x}) d\pi'_{\mathbf{X}'}(\underline{\theta})$$

ce qui s'écrit encore :

$$(\forall \underline{\theta}), \quad \pi_{\mathbf{X}}(\underline{\theta}) = V_{\mathbf{X}}^{\mathbf{X}'} \circ \pi'_{\mathbf{X}'}(\underline{\theta})$$

La  $\mu$ -exhaustivité de  $T$  résulte alors de la proposition (II-9).



### Fonction de concentration.

(II-13) DÉFINITION. — Soit  $(\mu_{\Theta}, P_X^{\Theta})$  un modèle bayésien dont  $Q_{\Theta}^X$  est la pseudo-transition *a posteriori*. On appelle  $\mu$ -fonction de concentration de seuil  $\alpha$  ( $\alpha \geq 0$ ) la fonction  $C_{\alpha}^X \in L^{\infty}(X, P^{\mu})$  définie par :

$$C_{\alpha}^X = P^{\mu}\text{-ess. sup. } \{ Q^X(\theta) : \mu(\theta) \leq \alpha \}$$

(II-14) PROPOSITION. — Soit  $T_X^X$  une opération bayésienne sur un modèle bayésien  $(\mu_{\Theta}, P_X^{\Theta})$ . Pour tout  $\alpha > 0$  soient  $C_{\alpha}^X$  et  $C_{\alpha}^{X'}$  les fonctions de concentration du modèle initial et du modèle image. On a alors, en posant  $S_X^{X'} = E_{T \cdot P^{\mu}}^{X'}(I_X^X)$  :

$$E_{T \cdot P^{\mu}}^{X'}(C_{\alpha}^X) = C_{\alpha}^X \circ S_X^{X'} \geq C_{\alpha}^{X'}$$

De plus si  $T_X^X$  est  $\mu$ -exhaustive, on a :

$$C_{\alpha}^X \circ S_X^{X'} = C_{\alpha}^{X'} \quad \text{et} \quad C_{\alpha}^{X'} \circ T_X^X = C_{\alpha}^X$$

*Démonstration.* — Soit  $\alpha \geq 0$ ; pour tout  $\theta$  tel que  $\mu(\theta) \leq \alpha$ , on a par définition :  $C_{\alpha}^X \geq Q^X(\theta)$ , et par suite :

$$(1) \quad C_{\alpha}^X \circ S_X^{X'} \geq T \circ P^{\mu}\text{-ess. sup. } \{ Q^{X'}(\theta) ; \mu(\theta) \leq \alpha \} = C_{\alpha}^{X'}$$

Si on suppose  $T$   $\mu$ -exhaustive, on a :  $Q_{\Theta}^X = Q_{\Theta}^{X'} \circ T_X^X$  (II-10), et le même raisonnement que précédemment conduit à :

$$(2) \quad C_{\alpha}^{X'} \circ T_X^X \geq C_{\alpha}^X$$

En composant les deux membres de ces inégalités avec  $P^{\mu}$ , il vient

$$C_{\alpha}^X \circ P_X^{\mu} \leq C_{\alpha}^{X'} \circ T_X^X \circ P_X^{\mu} \leq C_{\alpha}^X \circ S_X^{X'} \circ T_X^X \circ P_X^{\mu}$$

ce qui, en utilisant l'identité  $S \circ T \circ P^{\mu} = P^{\mu}$ , et en tenant compte de (1) et (2), conduit au résultat.

### Ensembles de concentration.

(II-15) DÉFINITION. — Si  $(\mu_{\Theta}, P_X^{\Theta})$  est un modèle bayésien dont la  $\mu$ -transition  $P_X^{\Theta}$  est faiblement dominée (cf. I-9), pour tout  $\alpha \geq 0$  et tout point  $x$  de  $X$ , on appelle ensemble de concentration de seuil  $\alpha$  en  $x$  l'ensemble mesurable  $\theta(x)$  défini à partir d'une version  $f$  de la densité faible de  $P_X^{\Theta}$  par rapport à  $\mu_{\Theta}$  de la façon suivante : si  $k_{\alpha}(x)$ ,  $a$  et  $b$  sont les nombres réels positifs définis par les inégalités :

$$a = \mu \{ f(\theta, x) > k_{\alpha}(x) \} \leq \alpha \leq \mu \{ f(\theta, x) \geq k_{\alpha}(x) \} = b$$

on pose :

$$\underline{\theta}_a(x) = \begin{cases} \{ f(\theta, x) > k_a(x) \} & \text{si } a = \alpha \\ \{ f(\theta, x) \geq k_a(x) \} & \text{sinon} \end{cases}$$

Il faut regarder comment les  $\underline{\theta}_a(x)$  dépendent de la version choisie. Soient  $f_1$  et  $f_2$  deux versions de la densité faible de  $P_x^\Theta$  par rapport à  $\mu_\Theta$ . Soit  $\underline{\omega} = \{ f_1 \neq f_2 \} \in \mathcal{H} \vee \mathcal{X}$ . Comme  $(\mu \otimes P^\mu)(\underline{\omega}) = 0$ , on a :

$$(\exists \underline{x}^0 : P^\mu(\underline{x}^0) = 0), \quad (\forall x \notin \underline{x}^0), \quad \mu(\underline{\omega}(x)) = 0$$

La relation  $\mu(\underline{\omega}(x)) = 0$  exprime que les sections de  $f_1$  et  $f_2$  en  $x$  sont  $\mu$ -presque partout égales.

Soient  $k_{1\alpha}(x)$  et  $k_{2\alpha}(x)$  correspondant à  $f_1$  et  $f_2$  respectivement. Soit  $x \notin \underline{x}^0$  : on peut supposer  $k_{1\alpha}(x) < k_{2\alpha}(x)$ . Alors d'après ce qui précède, on a les inclusions suivantes  $\mu$ -presque partout :

$$\begin{aligned} \{ f_2(\theta, x) > k_{2\alpha}(x) \} &\subset \{ f_2(\theta, x) \geq k_{2\alpha}(x) \} \\ &\subset \{ f_1(\theta, x) > k_{1\alpha}(x) \} \subset \{ f_1(\theta, x) \geq k_{1\alpha}(x) \} \end{aligned}$$

Donc :

$$\mu \{ f_2(\theta, x) \geq k_{2\alpha}(x) \} = \mu \{ f_1(\theta, x) > k_{1\alpha}(x) \} = \alpha$$

et :

$$\{ f_2(\theta, x) \geq k_{2\alpha}(x) \} = \{ f_1(\theta, x) < k_{1\alpha}(x) \} = \underline{\theta}_\alpha(x)$$

(II-16) REMARQUE. — Si  $f$  est une version de la densité faible de  $P_x^\Theta$  par rapport à  $\mu_\Theta$ , alors une version régulière de la  $P^\mu$ -transition *a posteriori* est donnée par :

$$Q(x, \underline{\theta}) = \int_{\underline{\theta}} f(x, \theta) d\mu(\theta)$$

Le lemme de Neyman et Pearson permet donc d'interpréter la fonction indicatrice  $1_{\underline{\theta}_\alpha(x)}$  comme un test optimum de seuil  $\alpha$  (ou  $b$  si  $\mu \{ f(\theta, x) > k_\alpha(x) \} < \alpha$ ) de la mesure  $Q_\Theta(x)$  contre  $\mu_\Theta$ . Par suite on a :

$$(\forall \underline{\theta} : \mu(\underline{\theta}) \leq \alpha), \quad (\forall P^\mu x), \quad Q(x, \underline{\theta}) \leq Q(x, \underline{\theta}_\alpha(x))$$

donc pour toute version de  $C_\alpha^X$ , on a :

$$(\forall_{P^\mu x}, \quad C_\alpha(x) \leq Q(x, \underline{\theta}_\alpha(x))$$

avec égalité si  $\mu(\underline{\theta}(x)) = \alpha$ .

(II-17) PROPOSITION. — Si  $(\mu_\Theta, P_x^\Theta)$  est un modèle bayésien faiblement dominé, et si  $f$  est une version de la densité faible de  $P_x^\Theta$  par rapport à

$\mu_{\Theta}$ , alors, pour tout point  $x$  de  $\underline{X}$ , l'ensemble de concentration de seuil  $\alpha$  associé à  $f$  décroît, quand  $\alpha$  décroît vers 0, vers

$$\underline{\theta}_\alpha(x) = \{ \underline{\theta} : f(\underline{\theta}, x) = \mu\text{-ess. sup. } f(\underline{\theta}, x) \}$$

*Démonstration.* — Soit  $x \in \underline{X}$ . Il est clair que si  $\alpha \leq \alpha'$ , alors  $k_\alpha(x) \leq k_{\alpha'}(x)$ . Posons :  $k_0(x) = \lim_{\alpha \searrow 0} k_\alpha(x)$ . Pour  $k > k_0(x)$ , on a :

$$(\forall \alpha > 0), \quad \{ f(\underline{\theta}, x) > k \} \subset \{ f(\underline{\theta}, x) > k_\alpha(x) \} \subset \underline{\theta}_\alpha(x)$$

donc :

$$(\forall \alpha > 0), \quad \mu \{ f(\underline{\theta}, x) > k \} \leq \alpha$$

et par suite :

$$\mu \{ f(\underline{\theta}, x) > k \} = 0$$

Pour  $k < k_0(x)$ , on peut trouver  $\alpha > 0$  tel que  $k < k_\alpha(x)$ , et par suite :

$$\mu \{ f(\underline{\theta}, x) > k \} \geq \{ f(\underline{\theta}, x) \geq k_\alpha(x) \} \geq \alpha$$

D'où le résultat.

*Remarque.* — L'intérêt de ce résultat est le lien qu'il met en évidence entre la méthode d'estimation dite du maximum de vraisemblance et le modèle bayésien. Dans les conditions classiques d'application de la méthode du maximum de vraisemblance ( $\Theta = \mathbb{R}^n$ , modèle dominé, existence d'une famille de densités  $p(x, \theta)$  « fortement » régulières, admettant pour tout  $x$  un maximum unique en  $\theta, \dots$ ), on va voir que cette méthode est extrêmement proche d'une technique « bayésienne » utilisant la mesure de Lebesgue  $1_{\Theta}$  comme mesure *a priori*. En effet, pour tout  $x$ , la résolution de

l'équation  $\frac{\partial p}{\partial \theta}(x, \theta) = 0$  est équivalente à celle de l'équation  $\frac{\partial f}{\partial \theta}(x, \theta) = 0$ ,

où  $f(x, \theta) = \frac{p(x, \theta)}{\int p(x, \theta) d\theta}$  est une version de la densité faible par rapport

à  $1_{\Theta}$ . Or, pour 1-presque tout  $\theta$ , on a :

$$\frac{dQ_{\Theta}(x)}{d1_{\Theta}}(\theta) = \lim_{\substack{I(\theta) \rightarrow 0 \\ I \in \mathcal{I}(\theta)}} \frac{Q(x, I)}{1(I)}$$

où  $\mathcal{I}(\theta)$  est la famille des intervalles fermés de  $\mathbb{R}$  contenant  $\theta$  [2]. Alors trouver  $\hat{\theta}(x)$ , supposé unique, vérifiant  $\frac{\partial f}{\partial \theta}(x, \theta) = 0$  est équivalent, puis-

que  $f$  est supposé dérivable, à chercher une famille d'intervalles  $I_i$  telle que  $1(I_i) \rightarrow 0$  et  $\hat{\theta}(x) = \bigcap_i I_i$ . Une telle famille se déduit de façon évidente des ensembles de concentration.

### III. — MODÈLE BAYÉSIEN ASYMPTOTIQUE

(III-1) DÉFINITION. — Nous appellerons modèle bayésien asymptotique un couple  $((\mu_\theta, P_X^\theta), \{X_i; i \in I\})$ , où :

1)  $I$  est un ensemble ordonné filtrant à droite.

2)  $\{X_i; i \in I\}$  est une famille de sous-espaces mesurables de  $X$  croissant vers  $X$  (i. e. pour tout  $i \in I, X_i = (\underline{X}, \mathcal{X}_i)$ , où  $\mathcal{X}_i$  est une sous-tribu de  $\mathcal{X}$ ,

et  $\mathcal{X} = \bigvee_{i \in I} \mathcal{X}_i$ .

3)  $(\mu_\theta, P_X^\theta)$  est un modèle bayésien tel que, si  $P_{X_i}^\theta$  est la restriction de  $P_X^\theta$  à  $X_i$ ,  $(\mu_\theta, P_{X_i}^\theta)$  soit un modèle bayésien.

S'il n'y a aucune ambiguïté, on écrira  $P$  et  $P_i$  au lieu de  $P_X^\theta$  et  $P_{X_i}^\theta$ . La dernière hypothèse implique que  $P_{X_i}^\mu = P_{X_i}^\theta \circ \mu_\theta$  notée aussi  $P_i^\mu$  est une mesure  $\sigma$ -finie. Si  $\pi = P \circ \mu$ , sa restriction  $\pi_{\theta X_i}$  sera aussi notée  $\pi_i$ . Si  $Q_\theta^X$  est la  $P^\mu$ -transition *a posteriori* du modèle  $(\mu, P)$ , la  $P_i^\mu$ -transition *a posteriori* du modèle  $(\mu, P_i)$  est, d'après (II-6) :  $Q_{i\theta}^{X_i} = E_{P_i^\mu}^{X_i}(Q_\theta^X)$ . On écrira donc :

$$Q \circ P^\mu = P \circ \mu \quad \text{et} \quad (\forall i \in I), \quad Q_i \circ P_i^\mu = P_i \circ \mu$$

L'objet de ce chapitre est l'étude de  $\{Q_{i\theta}^{X_i}, i \in I\}$ .

#### (III-2) REMARQUES.

1) Si  $I$  a un plus petit élément, noté 1, l'hypothèse de  $\sigma$ -finitude des mesures  $P_i^\mu$  est satisfaite si et seulement si  $P_1^\mu$  est  $\sigma$ -finie.

2) Le modèle bayésien asymptotique est en général défini de la façon constructive suivante pour  $I = \mathbb{N}$  : on suppose qu'il existe une suite  $\{Y_i; i \in I\}$

d'espaces mesurables telle que, pour tout entier  $n, X_n = \bigotimes_1^n Y_i$  (ce qui

entraîne  $X = \bigotimes_{i \in \mathbb{N}} Y_i$ ) et que, pour tout entier  $n, P_{X_n}^\theta$  est une transition,

et enfin qu'il existe une transition de probabilité  $R_{nY_{n+1}}^{\theta X_{n+1}}$  telle que

$$(\forall \theta), \quad P_{X_{n+1}}(\theta) = R_{nY_{n+1}}^{\theta X_{n+1}}(\theta) \circ P_{X_n}(\theta)$$

ce que l'on écrira :

$$P_{X_{n+1}}^\theta = R_{nY_n}^{\theta X_n} \circ P_{X_n}^\theta$$

En posant :  $R_{1Y_1}^{\theta X_0} = P_{X_1}^\theta$ , cela revient à écrire (Inoescu Tulcea) :

$$P_{X_n}^\theta = \bigcirc_{i=1}^n R_{iY_i}^{\theta X_{i-1}} \quad \text{et} \quad P_X^\theta = \bigcirc_{i \in \mathbb{N}} R_{iY_i}^{\theta X_{i-1}}$$

Dans ces conditions il est facile de montrer que les pseudo-transitions *a posteriori* vérifient la relation de récurrence suivante :

$$Q_{n+1\theta}^{X_n} \circ (R_{nY_{n+1}}^{\theta X_n} \circ Q_{n\theta}^{X_n}) = R_{nY_{n+1}}^{\theta X_n} \circ Q_{n\theta}^{X_n}$$

Dans le cas où, pour tout  $n$ ,  $Q_n$  est régulière, cette relation peut s'interpréter ainsi : quelle que soit « l'observation  $x_n = (y_1, y_2, \dots, y_n)$  à l'instant  $n$  »,  $Q_{n\theta}^{X_n}(y_1, y_2, \dots, y_n)$  est la pseudo-transition *a posteriori* du modèle bayésien à l'étape  $n$  ( $Q_{n\theta}(y_1, \dots, y_n)$ ,  $R_{nY_{n+1}}^\theta(y_1, \dots, y_n)$ ) dont la mesure *a priori* est la mesure *a posteriori* « sachant qu'on a observé  $(y_1, y_2, \dots, y_n)$  à  $n^{\text{ième}}$  étape ».

Le cas particulier important de cette construction est celui du « modèle bayésien produit » où tous les espaces  $Y_i$  sont identiques à un même espace  $Y$ , et où les transitions  $R_{nY_{n+1}}^{\theta X_n}$  sont constantes donc égales à une même transition  $P_Y^\theta$ .

On a alors :

$$P_{X_n}^\theta = \bigotimes_{k=1}^n P_{Y_k}^\theta$$

et :

$$Q_{n+1\theta}^{X_n} \circ (P_{Y_{n+1}}^\theta \circ Q_{n\theta}^{X_n}) = P_{Y_{n+1}}^\theta \circ Q_{n\theta}^{X_n}$$

### Comportement asymptotique des pseudo-transitions *a posteriori*.

(III-3) PROPOSITION. — Soit un modèle bayésien asymptotique  $((\mu_\theta, P_X^\theta), \{X_i, i \in I\})$ . Pour toute partie mesurable  $\underline{\theta}$  de  $\Theta$  de mesure  $\mu(\underline{\theta})$  finie, la famille  $\{Q_i^{X_i}(\underline{\theta}), i \in I\}$  forme une martingale positive équi-intégrable dans  $(X, P^\mu)$ ; donc elle converge dans  $L^1(X, P^\mu)$  et en mesure vers  $Q^X(\underline{\theta})$ . Si  $I = \mathbb{N}$ , la convergence précédente a lieu aussi  $P^\mu$ -presque partout.

*Démonstration.* — Soit  $\underline{\theta}$  de mesure  $\mu(\underline{\theta})$  finie. Comme

$$E_{P^\mu}^{X_i}(Q^X(\underline{\theta})) = Q_i^{X_i}(\underline{\theta}),$$

$\{Q_i^{X_i}(\underline{\theta}), i \in I\}$  est bien une martingale positive. D'autre part  $Q^X(\underline{\theta})$  est  $P^\mu$ -intégrable puisque

$$\int Q^X(\underline{\theta}) dP_X^\mu = \mu(\underline{\theta}) < +\infty.$$

La proposition se déduit alors des théorèmes (6-3), (6-2) et du lemme (6-8) de [5].

(III-4) REMARQUES.

1) Une partie mesurable  $\underline{x}$  de  $X$  est  $P$ -négligeable si et seulement si  $P^\mu(\underline{x}) = 0$ , soit, si  $P_X$  est une transition, si :  $(\forall \mu \underline{\theta}), P(\underline{\theta}, \underline{x}) = 0$ .

Par suite si  $I = \mathbb{N}$  et si  $P_X^\theta$  est une transition, la dernière partie de la proposition précédente entraîne :

$$(\forall \underline{\theta} : \mu(\underline{\theta}) < +\infty), (\forall \mu \underline{\theta}), Q_n^{X_n}(\underline{\theta}) \rightarrow Q^X(\underline{\theta}) \quad P_X(\underline{\theta})\text{-p. p.}$$

2) On pourrait souhaiter, en prenant des hypothèses convenables, inverser les quantificateurs dans l'expression précédente, c'est-à-dire isoler un ensemble  $\underline{\theta}^0$   $\mu$ -négligeable tel que, pour toute partie mesurable  $\underline{\theta}$  et tout  $\theta \notin \underline{\theta}^0$ , la convergence précédente soit  $P_X(\theta)$ -presque sûre :

$$(\forall \mu \underline{\theta}), (\forall \underline{\theta} : \mu(\underline{\theta}) < +\infty), Q_n^{X_n}(\underline{\theta}) \rightarrow Q^X(\underline{\theta}) \quad P_X(\underline{\theta})\text{-p. s.}$$

Ceci est illusoire comme le montre le contre-exemple suivant :

Soit un modèle bayésien produit (cf. III-2) où  $I = \mathbb{N}$ ,  $\Theta = \mathbb{R}$ ,  $\mu_\Theta$  est la mesure de Lebesgue,  $X_1 = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , et  $P_X^\Theta$  est une transition dominée. Alors, pour tout entier  $n$ ,  $Q_n$  est une transition dominée par la mesure de Lebesgue. On démontre dans ce cas que (cf. III-11)

$$(\forall \underline{\theta}), (\forall \mu \underline{\theta}), Q^X(\underline{\theta}) = I(\underline{\theta}, \underline{\theta}) \quad P_X(\underline{\theta})\text{-p. s.}$$

Alors, en prenant  $\underline{\theta} = \{\theta\}$ , on a :

$$I(\theta, \theta) = 1 \quad \text{et} \quad (\forall n), Q_n^X(\theta) = 0$$

On va maintenant établir que, dans le cas où  $\Theta$  est L. C. D. les transitions *a posteriori*  $Q_n$  convergent en loi vers  $Q$ .

(III-5) PROPOSITION. — Si  $((\mu_\Theta, P_X^\Theta), \{X_n, n \in \mathbb{N}\})$  est un modèle bayésien asymptotique où  $\Theta$  est un espace L. C. D. muni de sa tribu borélienne et  $\mu_\Theta$  une mesure de Radon, alors les pseudo-transitions sont régulières et, pour toute version, on a :

$$(\forall \underline{x}^0 : P^\Theta(\underline{x}^0) = 0 \quad \mu\text{-p. s.}), (\forall x \notin \underline{x}^0), Q_{n\Theta}(x) \xrightarrow{\text{étroitement}} Q_\Theta(x)$$

*Démonstration.* — Tout espace L. C. D. étant polonais, on est assuré de la régularité des pseudo-transitions *a posteriori* (I-7). Pour toute version régulière des  $Q_n$  et de  $Q$ , on a, d'après (III-3):

$$(\forall \underline{\theta}: \mu(\underline{\theta}) < +\infty), (\forall \underline{x}: P^\mu(\underline{x}) = 0), (\forall x \notin \underline{x}), Q_n(x, \underline{\theta}) \mapsto Q(x, \underline{\theta})$$

Soient  $d$  une distance compatible avec la topologie de  $\Theta$ ,  $\Theta^0$  un sous-ensemble dénombrable dense de  $\Theta$ , et  $\mathcal{H}^0$  la sous-algèbre dénombrable de  $\mathcal{H}$  engendrée par les boules ouvertes centrées sur  $\Theta^0$  et de rayons rationnels. Soit  $\underline{x}^0$  la réunion des ensembles  $P^\mu$ -négligeables  $\underline{x}$  associés par la relation précédente aux éléments  $\underline{\theta}$  de  $\mathcal{H}^0$ , qui est  $P^\mu$ -négligeable. Soit enfin  $x \notin \underline{x}^0$ . On a :

$$(\forall \underline{\theta} \in \mathcal{H}^0: \mu(\underline{\theta}) < +\infty), Q_n(x, \underline{\theta}) \rightarrow Q(x, \underline{\theta})$$

Ceci entraîne que, pour toute fonction  $h$  sur  $\Theta$ , étagée,  $\mathcal{H}^0$ -mesurable et à support compact, on a :

$$\int h dQ_n(x) \rightarrow \int h dQ(x)$$

Toute fonction  $g$  sur  $\Theta$ , continue et à support compact, étant uniformément continue, est limite uniforme de fonctions  $\mathcal{H}^0$ -mesurables à supports compacts. On en déduit que la suite de probabilités  $\{Q_n(x)\}$  converge vaguement vers la probabilité  $Q_\Theta(x)$ , donc aussi étroitement.

### Comportement asymptotique des fonctions de concentration.

(III-6) PROPOSITION. — Soit un modèle bayésien asymptotique  $((\mu_\Theta, P_X^\Theta), \{X_i, i \in I\})$ . Pour  $\alpha > 0$  notons  $C_{i\alpha}^{X_i}$  et  $C_\alpha^X$  les fonctions de concentration de seuil  $\alpha$ . Alors  $\{C_{i\alpha}^{X_i}, i \in I\}$  est une sous-martingale positive dans  $(X, P^\mu)$ . Si  $C_\alpha^X$  est  $P^\mu$ -intégrable, elle est équi-intégrable, donc converge dans  $L^1(X, P^\mu)$ . Si  $I = \mathbb{N}$ , elle converge  $P^\mu$ -presque partout vers  $C_\alpha^X$ .

*Démonstration.* — Pour tout  $i \leq j$ , dans  $I$ , on a :

$$Q_{i\Theta}^{X_i} = E_{P^\mu}^{X_i}(Q_{j\Theta}^{X_j}) = E_{P^\mu}^{X_i}(Q_\Theta^X)$$

La proposition (II-14) appliquée à  $I_{X_i}^X$  donne :

$$C_{i\alpha}^{X_i} \leq E_{P^\mu}^{X_i}(C_{i\alpha}^{X_i}) \leq E_{P^\mu}^{X_i}(C_\alpha^X)$$

Par suite  $\{C_{i\alpha}^{X_i}, i \in I\}$  est une sous-martingale positive dans  $(X, P^\mu)$ . La première assertion résulte alors immédiatement de [5]. Si  $I = \mathbb{N}$ , comme

cette sous-martingale est bornée par 1 et que la mesure  $P_{X_1}^\mu$  est  $\sigma$ -finie, elle converge  $P^\mu$ -presque partout vers une limite  $L_\alpha^X \leq C_\alpha^X$ . D'autre part, pour tout  $\underline{\theta}$  tel que  $\mu(\underline{\theta}) \leq \alpha$ , on a :

$$Q_n^{X_n}(\underline{\theta}) \leq C_{n\alpha}^{X_n} \quad P^\mu\text{-p. p.}$$

donc :

$$E_{P^{X_n}}(Q^X(\underline{\theta})) \leq C_{n\alpha}^{X_n} \quad P^\mu\text{-p. p.}$$

d'où à la limite :

$$Q^X(\underline{\theta}) \leq L_\alpha^X \quad P^\mu\text{-p. p.}$$

Cette inégalité étant vraie pour tout  $\underline{\theta}$  de mesure  $\mu(\underline{\theta}) \leq \alpha$ , on en déduit :

$$C_\alpha^X \leq L_\alpha^X \quad P^\mu\text{-p. p.}$$

D'où le résultat.

### Modèle bayésien séparé.

Il est intéressant de savoir sous quelles conditions les pseudo-transitions *a posteriori* permettent de « séparer » les diverses valeurs possibles du paramètre  $\theta$ . On aura une telle séparation si, les hypothèses de régularité de la proposition (III-5) étant satisfaites, pour  $\mu$ -presque tout  $\theta$  les transitions *a posteriori* convergent pour  $P(\theta)$ -presque tout  $x$  vers une mesure de Dirac en  $\theta$ . Mais cela impose que  $Q_\theta^X$  soit très particulière. Nous allons mettre ce problème en forme.

(III-7) DÉFINITION. — On dit qu'un modèle bayésien  $(\mu_\theta, P_X^\theta)$  est séparé si, pour tout  $\underline{\theta}$ , les mesures  $\pi_{X^c}(\underline{\theta})$  et  $\pi_{X^c}(\underline{\theta}^c)$  sont étrangères.

(III-8) PROPOSITION. — Soient  $(\mu_\theta, P_X^\theta)$  un modèle bayésien et  $Q_\theta^X$  sa pseudo-transition *a posteriori*. Les propositions suivantes sont équivalentes :

a) Le modèle est séparé.

b)  $Q_\theta^X \circ P_X^\theta = I_\theta^\theta$ .

De plus si  $P_X^\theta$  est une transition, elles sont encore équivalentes à :

c)  $(\forall \underline{\theta}), (\forall \mu(\underline{\theta}), Q^X(\underline{\theta}) = I(\underline{\theta}, \underline{\theta}) \quad P_X(\underline{\theta})\text{-p. s.}$

*Démonstration.*

a)  $\Rightarrow$  b) : la condition a) s'écrit :

$$(\forall \underline{\theta}), (\exists \underline{x}) \quad \pi(\underline{\theta}, \underline{x}^c) = \pi(\underline{\theta}^c, \underline{x}) = 0$$

ce qui, puisque  $\pi = Q \circ P^\mu$ , est équivalent à :

$$(\forall \underline{\theta}), (\exists \underline{x}), \quad Q^X(\underline{\theta}) = I^X(\underline{x}) \quad P^\mu\text{-p. s.}$$



On doit montrer que  $Q_{\Theta}^X \circ P_X^{\Theta} = I_{\Theta}^{\Theta}$ , c'est-à-dire que :

$$(\forall \underline{\theta}), (\forall \underline{\theta}'), \int_{\underline{\theta}'} I^{\Theta}(\underline{\theta}) d\mu_{\Theta} = \int Q^X(\underline{\theta}) I^{\Theta}(\underline{\theta}') d\pi_{\Theta X}$$

Soit  $x$  associé à  $\underline{\theta}$  par l'hypothèse  $a$ ). Le deuxième membre de la relation précédente s'écrit alors :

$$\int I^X(\underline{x}) I^{\Theta}(\underline{\theta}') d\pi_{\Theta X} = \pi(\underline{\theta}', \underline{x})$$

Comme  $\pi_{\Theta}(\underline{x})$  est porté par  $\underline{\theta}$ ,  $\pi(\underline{\theta}', \underline{x}) = \pi(\underline{\theta} \cap \underline{\theta}', \underline{x})$ . Or  $\pi_X(\underline{\theta} \cap \underline{\theta}')$  est porté par  $\underline{x}$  puisque majorée par  $\pi_X(\underline{\theta})$  portée par  $\underline{x}$ , donc

$$\pi(\underline{\theta} \cap \underline{\theta}', \underline{x}) = \pi(\underline{\theta} \cap \underline{\theta}', X) = \mu(\underline{\theta} \cap \underline{\theta}')$$

D'où le résultat puisque

$$\int_{\underline{\theta}'} I^{\Theta}(\underline{\theta}) d\mu_{\Theta} = \mu(\underline{\theta} \cap \underline{\theta}')$$

$b) \Rightarrow a)$  : si  $b)$  est vérifiée, alors :

$$(\forall \underline{\theta}), \int Q^X(\underline{\theta}^c) d\pi_X(\underline{\theta}) = 0$$

donc :

$$Q^X(\underline{\theta}) = 1 \pi_X(\underline{\theta})\text{-p. p.}$$

et

$$Q^X(\underline{\theta}) = 0 \pi_X(\underline{\theta}^c)\text{-p. p.}$$

ce qui démontre  $a)$ .

$c)$  : si  $P_X^{\Theta}$  est une transition, la condition  $b)$  est équivalente à :

$$(\forall \underline{\theta}), (\forall_{\mu} \underline{\theta}), \int Q^X(\underline{\theta}) dP_X(\underline{\theta}) = I(\underline{\theta}, \underline{\theta})$$

d'où le résultat.

(III-9) COROLLAIRE. — Soit  $(\mu_{\Theta}, P_X^{\Theta})$  un modèle bayésien tel que la tribu soit dénombrablement engendrée et que  $P_X^{\Theta}$  et  $Q_{\Theta}^X$  soient des transitions. Alors le modèle est séparé si et seulement si il existe une fonction mesurable  $f$  de  $X$  dans  $\Theta$  telle que :

$$(\forall \underline{\theta}), (\forall_{\mu} \underline{\theta} \in \underline{\theta}), P(\underline{\theta}, f^{-1}(\underline{\theta})) = 1$$

Alors, pour  $\alpha > 0$ , la fonction de concentration de seuil  $\alpha$  est donnée par :

$$C_\alpha^X = I^X(f^{-1}(\underline{\theta}_\alpha^c)) \quad P^\mu\text{-p. p.}$$

où  $\underline{\theta}_\alpha$  est la réunion des atomes de  $\mu_\Theta$  de mesures strictement supérieure à  $\alpha$ .

*Démonstration.* — On vérifie immédiatement que la relation précédente entraîne la séparation. Inversement toute transition de  $X$  dans  $\Theta$  à valeur 0 ou 1 étant déterministe [I], c'est-à-dire définie par une fonction mesurable, d'après (III-8) la séparation du modèle entraîne l'existence d'une fonction mesurable  $f$  de  $X$  dans  $\Theta$  telle que  $Q_\Theta^X = f_\Theta^X P^\mu\text{-p. p.}$  D'où l'équivalence.

D'autre part si  $\mu(\underline{\theta}) \leq \alpha$ , nécessairement  $\underline{\theta} \cap \underline{\theta}_\alpha = \emptyset$ , donc :

$$Q^X(\underline{\theta}) = I^X(f^{-1}(\underline{\theta}_\alpha^c)) \quad P^\mu\text{-p. p.}$$

De plus si  $x \in f^{-1}(\underline{\theta}_\alpha^c)$ , alors il existe une partie mesurable  $\underline{\theta}$  de  $\Theta$  de mesure  $\mu(\underline{\theta}) \leq \alpha$ , contenant  $f(x)$ . La deuxième partie du corollaire s'en déduit.

Remarquons que, si la tribu  $\mathcal{H}$  de  $\Theta$  contient les points, la condition du corollaire est vérifiée si :

$$(\forall_\mu \theta), \quad P(\theta, f^{-1}(\theta)) = 1$$

Donc la séparation du modèle entraîne l'existence de supports mesurables disjoints de  $\mu$ -presque toutes les probabilités  $P_X(\theta)$ .

(III-10) DÉFINITION. — On dit qu'un modèle bayésien asymptotique  $((\mu_\Theta, P_X^\Theta), \{X_i, i \in I\})$  est séparé si le modèle bayésien  $(\mu_\Theta, P_X^\Theta)$  est séparé.

(III-11) PROPOSITION. — Tout modèle bayésien asymptotique produit (défini dans III-2), où  $\Theta$  est un espace L. C. D. muni de sa tribu borélienne et  $\mu_\Theta$  une mesure de Radon, la tribu  $\mathcal{Y}$  est dénombrablement engendrée,  $P_Y^\Theta$  est une transition et l'application  $\theta \mapsto P_Y(\theta)$  est injective, est  $\mu$ -séparé.

*Démonstration.* — Pour  $n \in \mathbb{N}$ , soit  $R_n$  la transition de probabilité « empirique » de  $X_n$  dans  $Y$  définie, pour  $x = (y_1, \dots, y_n, \dots) \in \underline{X}$  et  $\underline{y} \in \mathcal{Y}$ , par :

$$R_n(x, \underline{y}) = \frac{1}{n} \sum_{i=1}^n I(y_i, \underline{y})$$

La loi faible des grands nombres permet d'affirmer que :

$$(\forall \theta), \quad (\forall \underline{y}), \quad (\forall_{P(\theta)} x), \quad R_n(x, \underline{y}) \rightarrow P(\theta, \underline{y})$$

Donc, si  $\mathcal{Y}^0$  est une algèbre dénombrable engendrant, on a :

$$(\forall \theta), \quad (\forall_{P(\theta)} x), \quad (\forall \underline{y} \in \mathcal{Y}^0), \quad R_n(x, \underline{y}) \rightarrow P(\theta, \underline{y})$$

Par suite, pour tout  $\theta$ , l'ensemble  $\underline{x}_\theta$  des points  $x$  de  $\underline{X}$  tels que, pour tout  $\underline{y} \in \mathcal{Y}^0$ ; la suite  $R_n(x, \underline{y})$  converge vers  $P(\theta, \underline{y})$ , est mesurable. De plus les ensembles  $\underline{x}_\theta$  sont des supports deux à deux disjoints des  $P_x(\theta)$ .

Il reste à montrer (cf. III-8) qu'il existe une application mesurable  $f$  de  $\underline{X}$  dans  $\Theta$  telle que :

$$(\forall \mu \theta), \quad f^{-1}(\{\theta\}) = \underline{x}_\theta \quad P_x(\theta)\text{-p. s.}$$

L'ensemble  $\underline{X}^0 = \{x : (\forall \underline{y} \in \mathcal{Y}^0), R_n(x, \underline{y}) \rightarrow \}$  est mesurable et, pour tout  $\theta$ , de probabilité  $P(\theta, \underline{X}^0) = 1$ .

Il suffit évidemment de définir  $f$  sur l'ensemble  $\underline{X}^0$ . Notons  $R$  l'application mesurable de  $\underline{X}^0$  dans  $\mathbb{R}^{\mathcal{Y}^0}$  définie par :

$$(\forall x \in \underline{X}^0), \quad R(x) = \{ \lim_n R_n(x, \underline{y}); \underline{y} \in \mathcal{Y}^0 \}$$

et  $S$  l'application mesurable de  $\Theta$  dans  $\mathbb{R}^{\mathcal{Y}^0}$  définie par :

$$(\forall \theta), \quad S(\theta) = \{ P(\theta, \underline{y}); \underline{y} \in \mathcal{Y}^0 \}$$

L'application  $S$  est de plus injective par hypothèse.

Supposons que  $\mu_\Theta$  soit une probabilité. On peut appliquer le théorème de Lusin aux variables aléatoires réelles  $P^\Theta(\underline{y})$ , pour  $\underline{y} \in \mathcal{Y}^0$  : pour tout  $\varepsilon > 0$ , on peut trouver un compact  $K_{\underline{y}, \varepsilon}$  de  $\Theta$  de probabilité  $\mu(K_{\underline{y}, \varepsilon}) \geq 1 - \varepsilon$ , tel que la restriction de  $P^\Theta(\underline{y})$  à  $K_{\underline{y}, \varepsilon}$  soit une application continue de  $K_{\underline{y}, \varepsilon}$  dans  $\mathbb{R}$ . Soit  $\varepsilon > 0$ . Soit  $\{\varepsilon_{\underline{y}}, \underline{y} \in \mathcal{Y}^0\}$  une suite de nombres réels positifs de somme  $\varepsilon$ . Alors  $K_\varepsilon = \bigcap_{\underline{y} \in \mathcal{Y}^0} K_{\underline{y}, \varepsilon_{\underline{y}}}$  est un compact de  $\Theta$  de probabilité  $\mu(K_\varepsilon) \geq 1 - \varepsilon$ , et la restriction de l'application  $S$  à ce compact est continue. Par suite l'image  $S(K_\varepsilon)$  est compacte donc mesurable dans  $\mathbb{R}^{\mathcal{Y}^0}$ , et, comme  $S$  est injective, on en déduit qu'elle admet une inverse mesurable  $T_\varepsilon$  de  $S(K_\varepsilon)$  dans  $K_\varepsilon$ . Maintenant si on prend une suite  $\{\varepsilon_n\}$  décroissant vers 0, la réunion  $\underline{\Theta}^0$  des  $K_{\varepsilon_n}$  est mesurable et de probabilité  $\mu(\underline{\Theta}^0) = 1$ , et la restriction de l'application  $S$  à  $\underline{\Theta}^0$  admet donc une inverse mesurable  $T$  de  $S(\underline{\Theta}^0)$  dans  $\underline{\Theta}^0$ .

Si  $\mu_\Theta$  est une mesure de Radon, comme  $\Theta$  est dénombrable à l'infini, il existe une suite  $\{\theta_n\}$  croissant vers  $\Theta$  de compacts de  $\Theta$  de mesures  $\mu(\theta_n) < +\infty$ . Pour tout  $n$ , soit  $\mu'_n$  la probabilité définie par :

$$(\forall \theta), \quad \mu'_n(\theta) = \frac{\mu(\theta \cap \theta_n)}{\mu(\theta_n)}$$

La construction précédente fournit, pour tout entier  $n$ , un ensemble mesu-

nable  $\theta_n^0$  de  $\Theta$  contenu dans  $\theta_n$ , tel que  $\mu(\theta_n - \theta_n^0) = 0$ . Donc, pour  $\underline{\Theta}^0 = \bigcup_n \theta_n^0$ , on a  $\mu(\underline{\Theta}^{0c}) = 0$ , et le résultat précédent reste vrai.

On a donc :

$$\underline{\Theta}^0 \xrightarrow[S]{S} S(\underline{\Theta}^0) \subset \mathbb{R}^{\mathcal{X}^0} \xleftarrow{R} \underline{X}^0$$

L'ensemble mesurable  $\underline{X}^0 = R^{-1}(S(\underline{\Theta}^0))$  de  $X$  est la réunion des ensembles  $\underline{x}_\theta$  pour  $\theta \in \underline{\Theta}^0$ , donc :

$$(\forall \theta \in \underline{\Theta}^0), \quad P(\theta, \underline{x}^0) = 1$$

Par suite si  $f$  est une application mesurable de  $X$  dans  $\Theta$ , égale à  $T \circ R$  sur  $\underline{x}^0$ , elle vérifie :

$$(\forall \theta \in \underline{\Theta}^0), \quad f^{-1}(\{\theta\}) = \underline{x}_\theta$$

d'où la séparation.

(III-12) PROPOSITION. — Soit  $((\mu_\Theta, P_X^\Theta), \{X_i, i \in I\})$  un modèle bayésien asymptotique séparé. Alors si  $P_X^\Theta$  est une transition, on a :

$$(\forall \theta: \mu(\theta) < +\infty), \quad (\forall_\mu \theta), \quad Q_n^{X_n}(\theta) \rightarrow I(\theta, \theta) \quad P_X(\theta)\text{-p. s.}$$

Si de plus  $\Theta$  est un espace L. C. D. muni de sa tribu borélienne, et si  $\mu_\Theta$  est une mesure de Radon, alors pour toute version des  $Q_n$ , on a :

$$(\exists \underline{x}^0: (\forall_\mu \theta), P(\theta, \underline{x}^0) = 0), \quad (\forall x \notin \underline{x}^0), \quad Q_{n\Theta}(x) \rightarrow I_\Theta(\theta)$$

Enfin si  $\mu_\Theta$  est sans atomes, on a :

$$(\forall \alpha > 0), \quad C_{n\alpha}^X \rightarrow 1 \quad P^\mu\text{-p. p.}$$

*Démonstration.* — La première partie de la proposition résulte directement de la première remarque (III-4) et de la proposition (III-8). D'autre part il résulte de (III-5) et (III-9) que l'on a :

$$(\exists \underline{x}^0: (\forall_\mu \theta), P(\theta, \underline{x}^0) = 0), \quad (\forall x \notin \underline{x}^0), \quad Q_{n\Theta}(x) \rightarrow I_\Theta(f(x))$$

La deuxième partie de la proposition s'en déduit puisque :

$$(\forall_\mu \theta), \quad f = \theta \quad P_X(\theta)\text{-p. s.}$$

La dernière partie résulte de (III-6) et de (III-8).

### BIBLIOGRAPHIE

- [1] Ph. CAILLOT, Étude algébrique et topologique des transitions. Thèse de 3<sup>e</sup> cycle. Paris, 1969.
- [2] DUNFORD et SCHWARTZ, *Linear operators* (part I). Interscience Publishers John Wiley and Sons. New York, 1963.

- [3] P. R. HALMOS, *Measure Theory*. Van Nostrand, New York, 1950.
- [4] HALMOS et SAVAGE, Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *A. M. S.*, vol. **20**, 1949, p. 225-241.
- [5] G. A. HUNT, *Martingales et processus de Markov*. Monographie de la société mathématique de France. Dunod, 1966.
- [6] LE CAM, Sufficiency and approximate sufficiency. *A. M. S.*, 1964.
- [7] MARTIN-PETIT-LITTAYE, Comparaison des expériences. *Ann. Inst. Henri Poincaré*, 1971.
- [8] NEVEU, *Bases mathématiques du calcul des probabilités*. Masson et Cie, 1964.

*Manuscrit reçu le 15 décembre 1971.*

---

*Directeur de la publication : GUY DE DAMPIERRE.*

IMPRIMÉ EN FRANCE

DÉPOT LÉGAL ÉD. N° 1859b.

IMPRIMERIE BARNÉOUD S. A. LAVAL, N° 6362. 4-1972.