

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

J. P. FLORENS

Expériences bayésiennes invariantes

Annales de l'I. H. P., section B, tome 18, n° 4 (1982), p. 305-317

http://www.numdam.org/item?id=AIHPB_1982__18_4_305_0

© Gauthier-Villars, 1982, 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/>

Expériences bayésiennes invariantes

par

J. P. FLORENS

Ceramisme. Université d'Aix-Marseille.

RÉSUMÉ. — Une mesure est invariante par une transformation si et seulement si elle est marginalement et conditionnellement invariante. Ce résultat est utilisé pour construire et étudier les expériences bayésiennes invariantes, c'est-à-dire définies par une mesure sur le produit des espaces des paramètres et des échantillons invariante par un couple de transformation de ces deux espaces.

Mots-clés : Expérience bayésienne, invariance, invariance marginale et conditionnelle, exhaustivité ancillarité et invariance, choix de mesure *a priori*.

I. INTRODUCTION

En statistique bayésienne, la nécessité de compléter la description du processus d'échantillonnage par une mesure *a priori* a conduit à s'intéresser à des expériences dans lesquelles la mesure *a priori* était déterminée grâce à une propriété d'invariance relative à une famille de fonctions. De nombreux travaux ont été consacrés aux « *a priori* de Haar », par exemple ceux de Hartigan (1964), Villegas (1971, 1972). Plusieurs autres méthodes de choix d'une *a priori* « non informative » ont été présentées et nous renvoyons pour cela le lecteur à notre article (1978) ou à celui de Bernardo (1979) par exemple.

Dans de nombreux cas la famille de fonctions par lesquelles la mesure *a priori* est invariante est liée à une famille de fonctions opérant sur l'espace des échantillons : l'image d'une probabilité d'échantillonnage par l'une

d'entre elles est une probabilité d'échantillonnage déterminée par la transformation du paramètre par l'une des fonctions opérant sur l'espace paramétrique.

On constate alors que la probabilité sur le produit des espaces paramétrique et d'échantillonnage obtenue par le produit de la mesure *a priori* et de la transition d'échantillonnage est elle-même invariante ainsi que la probabilité marginale sur l'espace des échantillons. La transition *a posteriori* (dont on supposera toujours l'existence dans ce texte) possède alors une propriété analogue à la transition d'échantillonnage.

Plus généralement nous nous proposons de montrer que cette construction d'une expérience bayésienne invariante repose sur une équivalence entre l'invariance d'une probabilité et une double invariance marginale et conditionnelle de celle-ci. Cette présentation plus générale fera l'objet de la section 2 et son application aux expériences bayésiennes sera exposée dans la section 3.

Dans la section 4 nous nous intéresserons à quelques problèmes d'exhaustivité et d'ancillarité dans des expériences bayésiennes invariantes (voir à ce sujet Hall, Wijsmann et Gosh (1965)). Enfin dans la section 5 sera étudiée l'invariance des expériences conditionnelles et nous y développerons l'exemple du modèle de régression.

II. INVARIANCE ET INVARIANCE CONDITIONNELLE

Dans cette section on notera (E, \mathcal{E}, m) un espace mesuré constitué d'un ensemble E , d'une tribu de sous ensembles \mathcal{E} et d'une mesure positive m . ϕ étant une application mesurable de E dans E on notera $\phi.m$ l'image de m par ϕ . Rappelons que $\phi.m$ est définie par

$$(2.1) \quad \phi.m(X) = m(\phi^{-1}(X)) \quad \forall X \in \mathcal{E}$$

$\phi.m$ (notée aussi parfois $m \circ \phi^{-1}$) est une mesure sur (E, \mathcal{E}) et on a (voir Métivier (1972)) :

$$(2.2) \quad \int_X \xi d\phi.m = \int_{\phi^{-1}(X)} \xi \circ \phi dm$$

pour toute variable aléatoire ξ définie sur (E, \mathcal{E})

2.1 DÉFINITION. — m est invariante par ϕ (resp. k -invariante $k \in \mathbb{R}$, $k > 0$) si $m = \phi.m$ (resp. $m = \frac{1}{k} \phi.m$).

On suppose souvent que ϕ est *inversible*, c'est-à-dire bijective et telle que ϕ^{-1} soit mesurable (voir Billingsley (1965)) mais nous nous passerons le plus possible de cette hypothèse. Le résultat suivant donne une propriété intéressante dans le cas où ϕ n'est pas inversible.

2.2 PROPOSITION. — Si ϕ est telle que $\phi(X) \in \mathcal{E}$ pour tout $X \in \mathcal{E}$ et si m est invariante par ϕ , \mathcal{E} est m -p. s. égale à la tribu engendrée par l'ensemble des $\phi[\phi^{-1}(X)]$ quand X parcourt \mathcal{E} .

Démonstration. — Soit \mathcal{F} cette tribu. Par hypothèse $\phi[\phi^{-1}(X)] \in \mathcal{E}$ pour tout x et donc $\mathcal{F} \subset \mathcal{E}$. De plus soit X quelconque de \mathcal{E} . On a alors :

$$\begin{aligned} m(\phi[\phi^{-1}(X)]\Delta X) &= m(\phi^{-1}(\phi[\phi^{-1}(X)]\Delta X)) \\ &= m(\phi^{-1} \circ \phi \circ \phi^{-1}(X)\Delta\phi^{-1}(X)) \\ &= m(\phi^{-1}(X)\Delta\phi^{-1}(X)) \\ &= 0 \end{aligned}$$

(Δ désignant la différence symétrique).

d'où l'on déduit $\mathcal{E} \subset \mathcal{F}$ m p. s. □

Notons en particulier que si m est une probabilité une vérification analogue montrerait que $m(\phi(E)) = 1$,

Soit \mathcal{F} une sous tribu de \mathcal{E} nous noterons \mathcal{F}^ϕ la sous tribu des $\phi^{-1}(X)$, $X \in \mathcal{F}$. L'indépendance de deux sous tribus sera notée $\mathcal{F}_1 \perp \mathcal{F}_2$ et l'indépendance de ces deux sous tribus conditionnellement à une troisième sera notée $\mathcal{F}_1 \perp \mathcal{F}_2/\mathcal{F}_3$ (soulignons que cette indépendance n'est définie que si la restriction de m à \mathcal{F}_3 est σ -finie, ce qui implique que m est σ -finie). (Des précisions sur l'indépendance conditionnelle peuvent être trouvées dans Mouchart-Rolin (1978)).

2.3 PROPOSITION. — Soient \mathcal{F}_1 et \mathcal{F}_2 deux sous tribus de \mathcal{E} telles que $\mathcal{F}_1 \perp \mathcal{F}_2$ et supposons que m soit invariante par ϕ . On a alors $\mathcal{F}_1^\phi \perp \mathcal{F}_2^\phi$.

Démonstration. — Soient $X_1 \in \mathcal{F}_1^\phi$ et $X_2 \in \mathcal{F}_2^\phi$. Il existe alors $Y_1 \in \mathcal{F}_1$ et $Y_2 \in \mathcal{F}_2$ tels que $X_1 = \phi^{-1}(Y_1)$ et $X_2 = \phi^{-1}(Y_2)$. On a alors :

$$\begin{aligned} m(X_1 \cap X_2) &= m(\phi^{-1}(Y_1) \cap \phi^{-1}(Y_2)) = m(\phi^{-1}(Y_1 \cap Y_2)) \\ &= m(Y_1 \cap Y_2) = m(Y_1)m(Y_2) = m(X_1)m(X_2) \quad \square \end{aligned}$$

Remarquons que ce résultat ne serait pas exact dans le cas d'une k -invariance de m par ϕ si $k \neq 1$ et que la transposition a l'indépendance conditionnelle exige une hypothèse supplémentaire mais permet un assouplissement de la notion d'invariance (voir proposition 2.6).

Soit \mathcal{F} une sous tribu de \mathcal{E} on dira que \mathcal{F} est *stable* par ϕ si $\mathcal{F}^\phi \subset \mathcal{F}$ et que \mathcal{F} est *invariante* par ϕ si $\mathcal{F}^\phi = \mathcal{F}$.

2.4 DÉFINITION. — Soient (E, \mathcal{E}, m) un espace mesuré, ϕ une application mesurable de E dans E et \mathcal{F} une sous tribu de \mathcal{E} . On dira que m est *invariante par ϕ conditionnellement à \mathcal{F}* si

- a) la restriction de m à \mathcal{F}^ϕ est σ -finie,
- b) \mathcal{F} est stable par ϕ ,
- c) $\forall \xi$ v. a. intégrable définie sur (E, \mathcal{E}) ,

$$E(\xi \circ \phi \mid \mathcal{F}^\phi) = E(\xi \mid \mathcal{F}^\phi) \circ \phi^* \quad \text{p. s.}$$

ϕ^* désignant l'application ϕ considérée comme une application de (E, \mathcal{F}^ϕ) dans (E, \mathcal{F}^ϕ) mesurable (car $\mathcal{F}^\phi \subset \mathcal{F}$ implique $(\mathcal{F}^\phi)^\phi \subset \mathcal{F}^\phi$) et $E(\xi \mid \mathcal{F}^\phi)$ désignant l'espérance conditionnelle de ξ par \mathcal{F}^ϕ pour la mesure m .

Exemple. — Soit $\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left(\begin{pmatrix} m_x \\ m_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}\right)$ et $\phi\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + a \\ y + a \end{pmatrix}$. Cette probabilité est invariante conditionnellement à x si le coefficient de régression $\frac{\sigma_{xy}}{\sigma_y^2}$ est égal à 1.

Remarques. — 1) L'invariance conditionnelle par rapport à la tribu triviale $\{\phi, E\}$ est équivalente à l'invariance sous l'hypothèse $m(E)$ fini.

De plus toute mesure σ -finie m est conditionnellement invariante par rapport à \mathcal{E} pour toute fonction ϕ .

2) La *k-invariance conditionnelle* peut être définie en remplaçant la condition c de la définition par la condition

$$E(\xi \circ \phi \mid \mathcal{F}^\phi) = \frac{1}{k} E(\xi \mid \mathcal{F}^\phi) \circ \phi^* \quad \text{p. s.}$$

Le résultat suivant reliant invariance et invariance conditionnelle est le résultat essentiel pour l'utilisation de cette notion.

2.5 PROPOSITION. — Soient (E, \mathcal{E}, m) un espace mesuré, ϕ une application mesurable de E dans E et \mathcal{F} une sous tribu de \mathcal{E} stable par ϕ et telle que la restriction de m à \mathcal{F} soit σ -finie. Les deux propositions suivantes sont alors équivalentes :

- a) m est k -invariante par ϕ ,
- b) la restriction de m à \mathcal{F}^ϕ est k -invariante par ϕ^* (cf. définition 2.4) et m est invariante par ϕ conditionnellement à \mathcal{F} .

Démonstration. — Vérifions tout d'abord que a implique b .

La k -invariance de m implique évidemment la k -invariance de sa restric-

tion à \mathcal{F}^ϕ qui est bien définie car $\phi^{-1}(X) \in \mathcal{F}^\phi$ si $X \in \mathcal{F}^\phi$. On a de plus, pour toute v. a. intégrale ξ de (E, \mathcal{E}) et tout $X \in \mathcal{F}$

$$\int_X \xi dm = \frac{1}{k} \int_X \xi d\phi \cdot m = \frac{1}{k} \int_{\phi^{-1}(X)} \xi \circ \phi dm = \frac{1}{k} \int_{\phi^{-1}(X)} E(\xi \circ \phi | \mathcal{F}^\phi) dm_{\mathcal{F}^\phi}$$

la première égalité résulte de l'invariance, la seconde de (2.2) et la troisième de la définition de l'espérance conditionnelle. $m_{\mathcal{F}^\phi}$ désigne la restriction de m à \mathcal{F}^ϕ . Mais on peut écrire aussi :

$$\begin{aligned} \int_X \xi dm &= \int_X E(\xi | \mathcal{F}^\phi) dm_{\mathcal{F}^\phi} = \frac{1}{k} \int_X E(\xi | \mathcal{F}^\phi) d\phi^* \cdot m_{\mathcal{F}^\phi} \\ &= \frac{1}{k} \int_{\phi^{-1}(X)} E(\xi | \mathcal{F}^\phi) \circ \phi^* dm_{\mathcal{F}^\phi} \end{aligned}$$

en utilisant les mêmes propriétés dans un ordre différent. Le rapprochement des deux calculs de l'espérance de ξ donne alors le résultat.

Réciproquement on peut écrire :

$$\begin{aligned} m(X) &= \int_E E(X | \mathcal{F}^\phi) dm_{\mathcal{F}^\phi} = \frac{1}{k} \int_E E(X | \mathcal{F}^\phi) d\phi^* \cdot m_{\mathcal{F}^\phi} \\ &= \frac{1}{k} \int_E E(X | \mathcal{F}^\phi) \circ \phi^* dm_{\mathcal{F}^\phi} = \frac{1}{k} \int_E E(\phi^{-1}(X) | \mathcal{F}^\phi) dm_{\mathcal{F}^\phi} = \frac{1}{k} m(\phi^{-1}(X)). \end{aligned}$$

$E(X | \mathcal{F}^\phi)$ désignant l'espérance conditionnelle de la fonction caractéristique de X . La première et la dernière égalités se déduisent de la définition de l'espérance conditionnelle, la deuxième de l'invariance de $m_{\mathcal{F}^\phi}$ et la quatrième de l'invariance conditionnelle. \square

Ce résultat montre qu'une mesure k -invariante est alors *marginale* et conditionnellement invariante pour toute tribu stable par ϕ . Ce résultat donne aussi le moyen de construire une mesure invariante par des mesures marginalement et conditionnellement invariante. Ce point sera illustré dans la section suivante.

La proposition suivante donne une application du concept d'invariance conditionnelle qui sera elle aussi utilisée dans la section suivante.

2.6 PROPOSITION. — Soient $\mathcal{E}_1, \mathcal{E}_2$ et \mathcal{F} trois sous tribus de \mathcal{E} et m sur (E, \mathcal{E}) invariante par ϕ conditionnellement à \mathcal{F} . On a alors l'implication :

$$\mathcal{E}_1 \perp\!\!\!\perp \mathcal{E}_2 / \mathcal{F}^\phi \Rightarrow \mathcal{E}_1^\phi \perp\!\!\!\perp \mathcal{E}_2^\phi / \mathcal{F}^\phi$$

(ces indépendances conditionnelles sont entendues relativement à la mesure m).

Démonstration. — Soient ξ_1 et ξ_2 deux variables aléatoires intégrables

respectivement \mathcal{F}_1^ϕ et \mathcal{F}_2^ϕ mesurable. On sait (voir par exemple Métivier (1972)) qu'il existe alors deux v. a. intégrables \mathcal{F}_1 et \mathcal{F}_2 mesurables, notées ξ'_1 et ξ'_2 telles que $\xi_1 = \xi'_1 \circ \phi$ et $\xi_2 = \xi'_2 \circ \phi$. On a alors :

$$\begin{aligned} E(\xi_1 \xi_2 | \mathcal{F}^\phi) &= E[(\xi'_1 \circ \phi)(\xi'_2 \circ \phi)] = E(\xi'_1 \cdot \xi'_2 \circ \phi | \mathcal{F}^\phi) \\ &= E(\xi'_1 \cdot \xi'_2 | \mathcal{F}^\phi) \circ \phi^* \quad \text{p. s.} \\ &= E(\xi'_1 | \mathcal{F}^\phi) E(\xi'_2 | \mathcal{F}^\phi) \circ \phi^* \quad \text{p. s.} \\ &= [E(\xi'_1 | \mathcal{F}^\phi) \circ \phi^*] [E(\xi'_2 | \mathcal{F}^\phi) \circ \phi^*] \quad \text{p. s.} \\ &= E(\xi_1 | \mathcal{F}^\phi) E(\xi_2 | \mathcal{F}^\phi) \quad \text{p. s.} \end{aligned}$$

en utilisant l'invariance conditionnelle puis l'hypothèse d'indépendance et de nouveau l'invariance conditionnelle.

L'égalité de l'espérance conditionnelle du produit avec le produit des espérances conditionnelles implique le résultat. \square

Ce résultat est particulièrement utile dans le cas où $\mathcal{F}^\phi = \mathcal{F}$. Ces propositions 2.5 et 2.6 montrent en particulier que si m est k -invariante par ϕ et si \mathcal{F} satisfait les conditions a et b de la définition 2.4 l'indépendance $\mathcal{E}_1 \perp\!\!\!\perp \mathcal{E}_2 / \mathcal{F}^\phi$ implique $\mathcal{E}_1^\phi \perp\!\!\!\perp \mathcal{E}_2^\phi / \mathcal{F}^\phi$.

III. EXPÉRIENCE BAYÉSIENNE INVARIANTE

Dans cette section une *expérience bayésienne* sera définie par une mesure π sur un espace mesurable produit $(A \times S, \mathcal{A} \otimes \mathcal{S})$ dont les restrictions à \mathcal{A} et \mathcal{S} sont toutes deux σ -finies. L'interprétation de cette définition est la suivante : A et S sont respectivement l'espace paramétrique et l'espace des échantillons, les restrictions à \mathcal{A} et \mathcal{S} définissent les mesures *a priori* et prédictive et celles-ci étant σ -finies on peut définir les espérances conditionnelles à \mathcal{A} et \mathcal{S} qui sont respectivement les espérances d'échantillonnage et *a posteriori*.

On appellera *transformation* de l'expérience $(A \times S, \mathcal{A} \otimes \mathcal{S}, \pi)$ tout couple de deux applications f et g respectivement de A dans A , \mathcal{A} -mesurable et de S dans S , \mathcal{S} -mesurable et on notera ϕ l'application mesurable définie sur $A \times S$ par $\phi(a, s) = (f(a), g(s))$.

3.1 DÉFINITION. — On appellera *expérience bayésienne k -invariante* la donnée d'une expérience bayésienne $(A \times S, \mathcal{A} \otimes \mathcal{S}, \pi)$ et d'une transformation ϕ telle que π soit k -invariante par ϕ .

En remarquant que par hypothèse les restrictions de π à \mathcal{A} et à \mathcal{S} sont σ -finies et que ϕ est définie de telle façon que $\phi^{-1}(\mathcal{A}) = f^{-1}(\mathcal{A}) \subset \mathcal{A}$ et $\phi^{-1}(\mathcal{S}) = g^{-1}(\mathcal{S}) \subset \mathcal{S}$ on déduit de la proposition 2.6 le résultat suivant :

3.2 PROPOSITION. — Les trois propositions suivantes sont équivalentes :

a) $(A \times S, \mathcal{A} \otimes \mathcal{S}, \pi)$ est k -invariante par ϕ .

b) La mesure *a priori* restreinte à \mathcal{A}^f est k -invariante par f et π est invariante par ϕ conditionnellement à \mathcal{A} .

c) La mesure prédictive restreinte à \mathcal{S}^g est k -invariante par g et π est invariante par ϕ conditionnellement à \mathcal{S} .

Commentaire. — Pour faciliter notre commentaire nous nous limiterons au cas où $f^{-1}(\mathcal{A}) = \mathcal{A}$ et $g^{-1}(\mathcal{S}) = \mathcal{S}$ (ce qui est réalisé en particulier si f et g sont inversibles).

Une expérience bayésienne est souvent construite progressivement de la manière suivante : on considère tout d'abord une transition d'échantillonnage notée par exemple $P^a(X)$ ($a \in A, X \in \mathcal{S}$) une statistique g à valeurs dans S telle que pour tout a de A il existe a' de A satisfaisant la condition $g.P^a = P^{a'}$. On définit alors l'application f de A dans A par $f(a) = a'$ ou encore $g.P^a = P^{f(a)}$, en supposant la mesurabilité de f vérifiée. Pour toute distribution μ *a priori* sur (A, \mathcal{A}) la mesure produit π sur $\mathcal{A} \otimes \mathcal{S}$ de μ et de P^a (voir Neveu (1964)) est alors invariante par $\phi = (f, g)$ conditionnellement à \mathcal{A} . Si de plus μ est choisie k -invariante (et σ -finie) on satisfait alors les conditions *b* de la proposition 3.2 et on en déduit que π est k -invariante par ϕ .

Mais de plus si la σ -finitude de la distribution prédictive P (c'est-à-dire la restriction de π à \mathcal{S}) est satisfaite la proposition 3.2 montre que P est aussi k -invariante et que π est invariante par ϕ conditionnellement à \mathcal{S} . Plus précisément, dans le cas où l'existence de versions régulières des probabilités conditionnelles est vérifiée, cela signifie que l'on peut choisir une transition *a posteriori* notée $\mu^s(E)$ ($s \in S, E \in \mathcal{A}$) satisfaisant $f . \mu^s = \mu^{g(s)}$.

Exemple. — Soient $S = \mathbb{R}^n$ et $s = (s_1, \dots, s_n)$ tel que les s_i sont distribués normalement de moyenne $a \in A = \mathbb{R}$ et de variance 1. La distribution *a priori* sur A est la mesure de Lebesgue. Considérons un couple de fonctions (f, g) respectivement de \mathbb{R} dans \mathbb{R} et de \mathbb{R}^n dans \mathbb{R}^n défini par $f(a) = a + b$ et $g(s_1, \dots, s_n) = (s_1 + b, \dots, s_n + b)$. La mesure *a priori* est invariante par f et la probabilité d'échantillonnage satisfait la propriété d'invariance conditionnelle. On en déduit donc que la mesure produit dont la densité par rapport à la mesure de Lebesgue dans \mathbb{R}^{n+1} est

$$\pi(a, s_1, \dots, s_n) = (2\pi)^{-\frac{n}{2}} \exp -\frac{1}{2} \sum_{i=1}^n (s_i - a)^2$$

satisfait l'invariance par (f, g) .

De plus la proposition 3.2 permet de conclure que la mesure prédictive

est invariante par g et que la probabilité *a posteriori* est conditionnellement invariante.

Dans notre exemple la mesure prédictive admet la densité suivante par rapport à la mesure de Lebesgue dans \mathbb{R}^n :

$$p(s) = (2\pi)^{-\frac{n-1}{2}} n^{-\frac{1}{2}} \exp -\frac{1}{2} \Sigma (s_i - \bar{s})^2$$

$$\left(\bar{s} = \frac{1}{n} \Sigma s_i \right)$$

Il est facile de vérifier l'invariance par g de cette mesure. La probabilité *a posteriori* est caractérisée par la densité conditionnelle suivante :

$$\mu(a/s) = (2\pi)^{-\frac{1}{2}} n^{\frac{1}{2}} \exp -\frac{1}{2n} (a - \bar{s})^2$$

On vérifie aisément que l'image par f de cette probabilité admet la densité $\mu(a/g(s))$.

IV. INVARIANCE ET RÉDUCTIONS ADMISSIBLES

Par étude des réductions admissibles d'un modèle bayésien nous entendons essentiellement la recherche des tribus exhaustives et ancillaires sur l'espace paramétrique et sur l'espace des observations. On sait en effet que de telles tribus permettent de « réduire » sans perte d'information l'expérience par marginalisation (dans le cas de l'exhaustivité) ou par conditionnalisation (dans le cas de l'ancillarité). Une présentation synthétique de ces notions a été donnée dans Florens-Mouchart (1977) et ce qui suit peut être considéré comme un complément de ce travail. Des détails sur les notions que nous allons employer peuvent y être trouvés.

Rappelons tout d'abord une sous tribu \mathcal{C} de \mathcal{S} (resp. \mathcal{B} de \mathcal{A}) est *exhaustive* si elle vérifie $\mathcal{A} \perp\!\!\!\perp \mathcal{S}/\mathcal{C}$ (resp. $\mathcal{A} \perp\!\!\!\perp \mathcal{S}/\mathcal{B}$) et que cette propriété justifie l'intérêt pour l'expérience marginale définie par la restriction de π à $\mathcal{A} \otimes \mathcal{C}$ (resp. $\mathcal{B} \otimes \mathcal{S}$). Cette restriction est notée $\pi_{\mathcal{A} \otimes \mathcal{C}}$ (resp. $\pi_{\mathcal{B} \otimes \mathcal{S}}$). Si π est k -invariante par une application ϕ on ne peut en général pas définir une propriété d'invariance pour l'expérience marginale sauf si la tribu exhaustive est stable par ϕ , c'est-à-dire si $\mathcal{C}^g \subset \mathcal{C}$ dans le cas d'une sous tribu de \mathcal{S} ou si $\mathcal{B}^f \subset \mathcal{B}$ dans le cas d'une sous tribu de \mathcal{A} .

Si, par exemple, \mathcal{C} est stable par ϕ on peut alors définir ϕ^* comme l'application mesurable de $(A \times S, \mathcal{A} \otimes \mathcal{C})$ dans lui-même égale à ϕ en tout point et la proposition implique que $\pi_{\mathcal{A} \otimes \mathcal{C}}$ est k -invariante par ϕ^* .

Si $\bar{\mathcal{C}}$ est exhaustive et stable la réduction par marginalisation préserve alors les informations contenues dans l'expérience complète et garde la propriété d'invariance. On peut s'interroger sur l'existence d'une plus petite tribu exhaustive stable et la réponse à ce problème fait l'objet de la proposition suivante :

4.1 PROPOSITION. — Soit une expérience bayésienne $(A \times S, \mathcal{A} \otimes \mathcal{S}, \pi)$ et $\phi = (f, g)$ un couple de fonctions mesurables de A sur lui-même et de S sur lui-même. Il existe une tribu $\bar{\mathcal{C}}^* \subset \mathcal{S}$ telle que

- $\bar{\mathcal{C}}^*$ est exhaustive et stable par g ,
- $\forall \bar{\mathcal{C}}$ exhaustive et stable par g , $\bar{\mathcal{C}}^* \subset \bar{\mathcal{C}}$ p. s.

$\bar{\mathcal{C}}^*$ est appelée tribu exhaustive stable presque sûrement la plus petite. De la même façon il existe une sous tribu \mathcal{B}^* de \mathcal{B} exhaustive stable presque sûrement la plus petite.

Démonstration. — Soit $\bar{\mathcal{C}} \subset \mathcal{S}$. Nous noterons $\bar{\bar{\mathcal{C}}}$ la tribu engendrée par $\bar{\mathcal{C}}$ et les ensembles de mesure nulle. Considérons alors la famille des sous tribus de \mathcal{S} stable par g , exhaustives et telles que $\bar{\mathcal{C}} = \bar{\bar{\mathcal{C}}}$.

Soit $\bar{\mathcal{C}}^*$ l'intersection des tribus de cette famille. $\bar{\mathcal{C}}^*$ est alors exhaustive (voir Mouchard-Rolin (1978)), et stable. Si $\bar{\mathcal{C}}$ est exhaustive et stable, $\bar{\bar{\mathcal{C}}}$ appartient alors à la famille précédemment décrite et $\bar{\mathcal{C}}^* \subset \bar{\bar{\mathcal{C}}}$. Ceci entraîne $\bar{\mathcal{C}}^* \subset \bar{\mathcal{C}}$ p. s. ; l'existence de \mathcal{B}^* se démontre de manière identique. \square

L'invariance d'une expérience bayésienne peut être reliée à ses propriétés d'identification. Rappelons qu'une expérience est identifiée si toute tribu exhaustive de l'espace paramétrique est p. s. égale à \mathcal{A} . Supposons alors qu'une expérience est construite à partir d'une transition d'échantillonnage satisfaisant $g.P^a = P^{f(a)}$ (g peut en particulier être l'identité), que f soit telle que $f(E) \in \mathcal{A}$ si $E \in \mathcal{A}$ et que la tribu \mathcal{A}^* engendrée par les $f(E)$ quand E décrit \mathcal{A}^f soit la plus petite tribu rendant mesurable $P^a(X) \forall X \in \mathcal{S}$. (Par exemple si $A = \mathbb{R}^2$, $S = \mathbb{R}^n$ avec

$$s_i \sim \mathbb{N}(a_1 + a_2, 1)(a = (a_1, a_2)s = (s_1, \dots, s_n))$$

on peut choisir g telle que $g(s_1, \dots, s_n) = (s_1 + \lambda, \dots, s_n + \lambda)$ et $f(a_1, a_2) = (a_1 + a_2 + \lambda, 0)$ et les conditions précédentes sont alors vérifiées.) Si de plus la mesure *a priori* est choisie invariante par f la proposition entraîne l'égalité presque sûre de \mathcal{A} et de \mathcal{A}^* et donc l'identification de l'expérience. Un moyen pour spécifier une mesure *a priori* rendant le modèle identifié est donc de la choisir invariante par une fonction f satisfaisant les conditions précédemment énoncées.

Considérons maintenant une sous tribu $\bar{\mathcal{C}}$ de \mathcal{S} (ou \mathcal{B} de \mathcal{A}) *ancillaire*, c'est-à-dire satisfaisant la condition $\bar{\mathcal{C}} \perp \mathcal{A}$ (resp. $\mathcal{B} \perp \mathcal{S}$). Cette condition

justifie alors l'intérêt pour l'*expérience conditionnelle* définie par la probabilité π sur $\mathcal{A} \otimes \mathcal{S}$ conditionnelle à \mathcal{C} (ou \mathcal{B}) (supposée régulière). Si π est k -invariante il est là encore nécessaire d'imposer une condition supplémentaire pour que cette invariance soit maintenue dans l'expérience conditionnelle : il faut supposer \mathcal{C} (resp. \mathcal{B}) stable par ϕ et supposer de plus que la restriction de π à \mathcal{C} (resp. \mathcal{B}) est σ -finie. Si ces deux conditions sont vérifiées l'expérience conditionnelle est alors conditionnellement invariante comme le montre la proposition 2.5.

Il n'existe évidemment pas de plus grande sous tribu ancillaire stable mais l'utilisation du théorème de Zorn permet de conclure à l'existence d'éléments maximaux dans la famille des sous tribus ancillaires stables. Ce résultat est évident mais ne présente qu'un intérêt pratique réduit.

V. INVARIANCE ET EXPÉRIENCE CONDITIONNELLE

L'étude des propriétés d'invariance des expériences conditionnelles nécessite une extension de la proposition 2.5.

5.1 PROPOSITION. — Soient (E, \mathcal{E}, m) un espace mesuré et ϕ une application mesurable de E dans E . On considère deux sous tribus de \mathcal{E} , stables par ϕ , notées \mathcal{F} et \mathcal{G} telles que $\mathcal{F} \subset \mathcal{G}^\phi$. On suppose que la restriction de m à \mathcal{F} est σ -finie. Les deux propriétés suivantes sont alors équivalentes :

- a) m est k -invariante par ϕ conditionnellement à \mathcal{F} ,
- b) la restriction de m à \mathcal{G}^ϕ est k -invariante par ϕ^* (fonction mesurable de (E, \mathcal{G}^ϕ) dans (E, \mathcal{G}^ϕ) égale à ϕ en tout point) conditionnellement à \mathcal{F} et m est invariante par ϕ conditionnellement à \mathcal{G} .

La démonstration de ce résultat est identique à celle de la proposition 2.5. La proposition 2.5 peut être obtenue à partir de la proposition 5.1 en remplaçant \mathcal{F} par $\{\emptyset, E\}$ et \mathcal{G} par \mathcal{F} .

Une implication de la proposition 5.1 est la suivante : si m est invariante par ϕ conditionnellement à \mathcal{F} , m est invariante conditionnellement à toute sous tribu \mathcal{G} telle que $\mathcal{F} \subset \mathcal{G}^\phi \subset \mathcal{G}$.

Si \mathcal{M} est une sous tribu de la tribu produit $\mathcal{A} \otimes \mathcal{S}$ sur laquelle est définie une expérience bayésienne rappelons que l'*expérience conditionnelle* à \mathcal{M} est définie par la probabilité conditionnelle (supposée régulière) à \mathcal{M} sur $\mathcal{A} \otimes \mathcal{S}$. Une expérience conditionnelle est donc en fait une probabilité de transition caractérisée par $\pi^m(E \times X)$ engendrant pour tout m une probabilité sur $\mathcal{A} \otimes \mathcal{S}$ et étant \mathcal{M} -mesurable pour tout $E \times X$.

La définition 2.4 de l'invariance conditionnelle a été donnée en se fondant

sur une probabilité donnée sur tout (E, \mathcal{E}) mais on peut remarquer qu'une définition analogue peut être donnée pour une probabilité de transition $m^f(X)$ ($X \in E, f \in E, m^f(X) \mathcal{F}$ mesurable $\forall X$) en supprimant la condition a (qui ne sert qu'à assurer l'existence des espérances conditionnelles) et en remplaçant l'égalité presque sûre de la condition c par une égalité. Ceci nous conduit à la définition suivante :

5.2 DÉFINITION. — L'expérience conditionnelle $(A \times S, \mathcal{A} \otimes \mathcal{S}, \pi^m)$ est k -invariante par ϕ fonction mesurable de $A \times S$ dans lui-même si π^m est k -invariante par ϕ conditionnellement à \mathcal{M} .

On supposera par la suite que $\phi = (f, g)$ (voir définition 3.1). La proposition suivante est alors une application de la proposition 5.1. Nous nous limiterons ici au cas où $\mathcal{M} = \mathcal{C} \subset \mathcal{S}$. Un résultat analogue serait obtenu en prenant $\mathcal{M} = \mathcal{B} \subset \mathcal{A}$.

5.3 PROPOSITION. — Soit une expérience bayésienne conditionnelle définie par une probabilité de transition π^f sur $(A \times S, \mathcal{A} \otimes \mathcal{S}) \mathcal{C}$ -mesurable et $\phi = (f, g)$ une transformation de cette expérience telle que \mathcal{A} soit invariante par f et \mathcal{C} et \mathcal{S} soient invariantes par g . Les trois résultats suivants sont alors équivalents :

- a) π^f est k -invariante par ϕ conditionnellement à \mathcal{C} .
- b) P^{af} (probabilité d'échantillonnage conditionnelle à \mathcal{C}) est invariante par ϕ conditionnellement à $\mathcal{A} \otimes \mathcal{C}$ et μ^f (probabilité *a priori* conditionnelle à \mathcal{C}) est k -invariante conditionnellement à \mathcal{C} .
- c) μ^s (probabilité *a posteriori*) est invariante conditionnellement à \mathcal{S} et P^f (probabilité prédictive sur \mathcal{S} conditionnelle à \mathcal{C}) est invariante conditionnellement à \mathcal{C} .

Démonstration. — Ce résultat est une application de la proposition 5.1 en faisant tout d'abord $\mathcal{F} = \mathcal{C}$ et $\mathcal{G} = \mathcal{A} \otimes \mathcal{S}$ pour obtenir l'équivalence de a et b puis $\mathcal{F} = \mathcal{C}$ et $\mathcal{G} = \mathcal{S}$ pour obtenir l'équivalence de a et c . \square

Les invariances de $\mathcal{A}, \mathcal{S}, \mathcal{C}$ ont été supposées de façon à obtenir une proposition simple mais le cas où ces tribus sont seulement stables pourrait être traité de façon analogue. L'étude d'expériences conditionnelles à une sous tribu de \mathcal{A} serait tout à fait semblable.

L'étude des expériences conditionnelles invariantes est en partie destinée à fournir un cadre théorique au problème du choix d'une mesure *a priori* dans un modèle conditionnel.

Un exemple d'un tel modèle est fourni par le modèle de régression multidimensionnel pour lequel le problème du choix d'une mesure *a priori* « peu informative » n'a pas été totalement résolu (voir Zellner (1971) et Dreze (1976) par exemple).

Soient $y_t \in \mathbb{R}^n$, $z_t \in \mathbb{R}^k$, $\pi \in \mathbb{R}^{n+k}$, $\Omega \in \mathbb{R}^{n \times n}$, $t = 1, \dots, T > k$, y_t est engendré conditionnellement à z_t (qui est observé) et aux paramètres π et Ω par une probabilité normale de moyenne πz_t et de variance Ω . Les y_t sont (conditionnellement) indépendants. Soit de plus la famille des transformations $\phi_{B,C}$ définies par :

$$\begin{aligned} \phi_{B,C}(\pi, \Omega, y_1, \dots, y_T, z_1, \dots, z_T) \\ = (B\pi + C, B\Omega B', B y_1 + C z_1, \dots, B y_T + C z_T, z_1, \dots, z_T) \end{aligned}$$

avec $B \in \mathbb{R}^{n \times n}$ régulière et $C \in \mathbb{R}^{n \times k}$.

Pour toutes matrices B et C , $\phi_{B,C}$ satisfait les hypothèses de la proposition 5.3 et la probabilité d'échantillonnage est invariante par $\phi_{B,C}$ conditionnellement aux paramètres et aux z_1, \dots, z_T . Il existe alors une unique mesure invariante sur l'espace paramétrique par le groupe des transformations $(\pi, \Omega) \rightarrow (B\pi + C, B\Omega B')$. Cette mesure admet la densité

$$D(\pi, \Omega) = |\Omega|^{-\frac{n+k+1}{2}}$$

par rapport à la mesure de Lebesgue. Le fait de n'étudier que le modèle conditionnel aux z_t pour estimer π et Ω suppose implicitement que la mesure sur B et Ω conditionnelle aux z_t soit la même que la mesure *a priori* (Voir Florens-Mouchart (1977)). $D(\pi, \Omega)$ peut donc être retenue comme densité *a priori* dans le modèle conditionnel. Cette mesure *a priori* est celle choisie par Dreze (1976) et on peut vérifier que celle choisie par Zellner (1971), de densité $|\Omega|^{-\frac{n+1}{2}}$ n'est que relativement invariante (voir Nachbin (1965)).

La mesure produit de notre mesure *a priori* et de la transition d'échantillonnage est donc invariante par $\phi_{B,C}$ conditionnellement aux z_t . La mesure prédictive est caractérisée par la densité :

$$D(Y | Z) = |Y' M_Z Y|^{-\frac{T}{2}} |Z' Z|^{-\frac{n}{2}}$$

avec $Y = (y'_1, \dots, y'_T)$, $Z = (z'_1, \dots, z'_T)$ et $M_Z = I - Z(Z'Z)^{-1}Z'$.

Cette distribution est (conditionnellement aux z_t) invariante par les transformations définies par

$$(y_t, z_t) \rightarrow (B y_t + C z_t, z_t) \quad \forall t$$

La probabilité *a posteriori* de π conditionnelle à Ω , Y et Z est normale de moyenne $(Z'Z)^{-1}Z'Y$ et de variance $(Z'Z)^{-1} \otimes \Omega$. Ω est *a posteriori* Whishart-inverse de paramètres $Y' M_Z Y$, Y et T . Cette probabilité *a posteriori* est invariante par ϕ_{BC} conditionnellement à Y et Z . Des détails sur les calculs précédents peuvent être trouvés dans Tiao et Zellner (1964).

BIBLIOGRAPHIE

- J. M. BERNARDO, Reference Posterior Distributions for Bayesian Inference (with Discussion). *J. R. Statist. Soc.*, Series B, t. **41**, n° 2, 1979, p. 113-147.
- P. BILLINGSLEY, *Ergodic Theory and Information*, New-York, Wiley, 1965.
- J. DREZE, Bayesian limited Information Analysis of simultaneous Equations Model, *Econometrica*, t. **44**, 1976, p. 1045-1075.
- J. P. FLORENS et M. MOUCHART, *Reduction of bayesian experiments*, Core D. P. 7737, Université de Louvain, 1977.
- J. P. FLORENS, Mesures *a priori* et invariance dans une expérience bayésienne, *Publications de l'ISUP*, t. **XXIII**, Fasc. 1-2, 1978, p. 29-56.
- W. S. HALL, R. A. WIJSMAN et J. K. GOSH, The relationship between sufficiency and invariance with applications in sequential analysis, *Ann. of Math. Statistics*, t. **36**, 1965, p. 575-614.
- J. HARTIGAN, Invariant prior distributions, *Ann. of Math. Statistics*, t. **35**, 1964, p. 836-845.
- M. METIVIER, *Notions fondamentales de la théorie des probabilités*, Paris, Dunod, 1972.
- M. MOUCHART et J. M. ROLIN, *A note on conditional Independence*, mimeo, 1978.
- L. NACHBIN, *The Haar Integral*, Princeton, D. Van Nostrand, 1965.
- J. NEVEU, *Bases mathématiques du calcul des probabilités*, Paris, Masson, 1964.
- G. C. TIAO et A. ZELLNER, On the bayesian estimation of multivariate regression. *J. Roy. Statist. Soc.*, Series B, t. **26**, 1964, p. 277-285.
- A. ZELLNER, *An Introduction to bayesian Inference in Econometrics*, New York, Wiley, 1971.

(Manuscrit reçu le 17 novembre 1981)