

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

G. ROMIER

Introduction à la statistique mathématique.

III. Décision statistique

Annales de l'I. H. P., section B, tome 5, n° 4 (1969), p. 323-355

<http://www.numdam.org/item?id=AIHPB_1969__5_4_323_0>

© Gauthier-Villars, 1969, 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/>

Introduction à la statistique mathématique.

III. Décision statistique

par

G. ROMIER

Maitre de Conférences I. U. T. de Grenoble.

1. INTRODUCTION

La théorie développée dans le présent chapitre a sa source dans le travail essentiel de Abraham Wald [13] : *Statistical Decision Functions* (J. Wiley, 1950). Il fut complété par la suite par Le Cam [5].

La mise au point présentée ci-après, hormi quelques améliorations techniques et extensions, n'apporte rien de nouveau du point de vue conceptuel. L'essentiel de cette théorie demeure le fait que les stratégies « bayésiennes » approchent bien — dans le cas d'un problème avec « fonction de coût » — les solutions optimales. Les applications en demeurent toujours limitées dans la pratique.

Au paragraphe 2 sont rappelées les principales définitions. Au numéro 3 nous énonçons quelques résultats techniques connus (relèvement, Dunford-Pettis, désintégration) et certaines généralisations utiles par la suite et dans d'autres contextes. Les paragraphes 4 et 5 sont consacrés au théorème de Wald et à sa démonstration. Le paragraphe 6 à l'importante question de l'existence de stratégies « bayésiennes » et à leur caractérisation.

2. RAPPEL DES DÉFINITIONS

2.1. Notations.

Soit $(\Theta, \mathcal{H})$ un espace mesurable. Nous désignons par $\mathcal{C}_\Theta$ la catégorie dont :

— les objets sont les triplets (Θ, X, P_X^Θ) , $(X, \mathcal{X})$ espaces mesurables, P_X^Θ sont les transitions,

— les morphismes $u : (\Theta, X, P_X^\Theta) \rightarrow (\Theta, Y, P_Y^\Theta)$ sont identifiés aux applications mesurables de X dans Y , telles que $u(P_X^\Theta) = P_Y^\Theta$.

Soit I une catégorie, ordonnée par factorisation $(i \leq j) \Leftrightarrow (\text{Hom}(i, j) \neq \emptyset)$ O un élément de I tel que, pour tout $i \in \text{Ob} I$, $\text{Hom}(O, i) \neq \emptyset$, ξ un facteur de I dans $\mathcal{C}_\Theta$ (contravariant).

On désignera par :

γ le facteur de $\mathcal{C}_\Theta$ dans la catégorie des espaces mesurables :

$$\begin{aligned} (\Theta, X, P_X^\Theta) &\rightarrow (X, \mathcal{X}) \\ u &\rightarrow u \end{aligned}$$

α et β les applications source et but de FII dans $\text{Ob} I$

$$V = \text{Hom}(O, I)$$

$$V_i = \{u \in V \mid (\exists v \in FII)(\beta(v_0 u) = i)\} \quad (i \in I)$$

$$A_u = \{v \in V \mid (\exists w \in FII)(v = w_0 u)\} \quad (u \in v)$$

$$B_u = \{v \in V \mid (\exists w \in FII)(u = w_0 v)\} \quad (u \in V)$$

$$B_u^0 = B_u \setminus \{u\}$$

$$(\Omega, \mathcal{A}, \zeta_i)_{i \in I} = \lim_{\leftarrow B_u^0} \gamma_0 \xi$$

$$(\Omega_i, \mathcal{A}_i) = \gamma_0 \xi(i)$$

$$\mathcal{A}_i = \zeta_i^{-1}(\mathcal{A}_i)$$

$$(\Omega_u^-, \mathcal{A}_u^-, \zeta_{ui}^-) = \lim_{\leftarrow} \gamma_0 \xi \text{ (restriction du facteur, aux morphismes de } B_u^0 \text{)}.$$

$$\zeta_u^- \text{ application canonique de } (\Omega, \mathcal{A}) \text{ dans } (\Omega_u^-, \mathcal{A}_u^-)$$

$$\overline{\mathcal{A}}_u = (\zeta_u^-)^{-1}(\mathcal{A}_u^-).$$

Soit $D \subset \prod_{i \in I} D^i$ un espace mesurable, muni de la trace de la tribu

produit. On notera D_j l'espace D muni de la trace de la tribu des cylindres

à base dans $\prod_{i \leq j} D^i$.

Rappelons enfin que pour une transition P_Y^X on note, pour B partie mesurable de Y :

P_B^X transition restriction de P_Y^X à B .

P_B^x fonction mesurable qui associe à x élément de X , la valeur de P_Y^X sur B .

2.2. Définitions.

Soit $\mathcal{E}$ catégorie « structure expérimentale », sous-catégorie de $\mathcal{C}_\Theta$ telle qu'elle a été définie au chapitre I.

DÉFINITION 1. — *Un plan d'expérimentation* est la donnée :

- (1) — d'un espace mesurable $(\theta, \mathcal{H})$ (paramètres)
 — d'un espace mesurable $(D, \mathcal{D})$ (décisions)
 — d'une sous-catégorie $\mathcal{E}$ de $\mathcal{C}_\theta$ (structure expérimentale)
- (2) — d'une catégorie I (états d'expérimentation), ordonnée par factorisation et possédant un élément 0 tel que :

$$(\forall i \in \text{Ob} I) \quad (\text{Hom}(0, i) \neq \emptyset)$$

— d'un foncteur contravariant ξ de I dans $\mathcal{E}$

- (3) — d'une transition P_Ω^θ ou Ω est l'espace associé à $\lim_{\leftarrow} \gamma_0 \xi$, et telle que, pour tout i de I , $\zeta_i(P_\Omega^\theta) = P_{\Omega_i}^\theta$

- (4) — d'une fonction l , mesurable, de $\Theta \Omega \text{VD}$ dans $\mathbb{R}_+$.

DÉFINITION 2. — Une *stratégie* ou fonction de décision, est une transition S_{VD}^θ telle que, sauf sur un ensemble P_Ω^θ -négligeable quel que soit θ :

$$(\delta_1) \quad (\forall i \in I) \quad (S_{V_i D_i}^\Omega \text{ factorise à travers } \zeta_i)$$

$$(\delta_2) \quad (\forall u \in V) \quad (S_{A_u D}^\Omega \text{ factorise à travers } \zeta_u^-)$$

Cette définition des stratégies est trop restrictive pour donner, hors du cas I dénombrable, des résultats suffisamment forts. Nous considérons donc des plans d'expérimentation satisfaisant la condition suivante :

(a) Soit, pour

$$i \in I, J(i) = \{j \in I \mid (j > i) \text{ et } (\forall k \in I)(i < k < j)(\exists l \in I)(i < l < k)\}.$$

Pour tout i appartenant à I :

ou bien $J(i)$ est vide et nous poserons $\overline{\mathcal{A}}_{i+} = \overline{\mathcal{A}}_i$,

ou bien $J(i)$ est non vide et il existe une tribu $\overline{\mathcal{A}}_{i+}$ contenant $\overline{\mathcal{A}}_i$, telle que, pour toute suite i_n décroissante vers i ,

$$i_n \in J, \quad \bigcap_n \overline{\mathcal{A}}_{i_n} = \overline{\mathcal{A}}_{i+}.$$

Nous poserons alors :

DÉFINITION 2b. — Une *stratégie ouverte* est une transition de probabilité S_{VD}^Ω , telle que, sauf sur un ensemble négligeable quel que soit θ :

$$(\delta'_1) \quad (\forall i \in I) \quad (S_{V_i D_i}^\Omega \text{ est } \overline{\mathcal{A}}_{i+}\text{-mesurable})$$

$$(\delta_2) \quad (\forall u \in V) \quad (S_{A_u D}^\Omega \text{ factorise dans } \zeta_u^-).$$

La condition (a) vise à introduire dans le cas d'un processus à temps $\mathbb{R}_+$, la famille $\overline{\mathcal{A}}_{t+}$, continue à droite, associée classiquement à $\overline{\mathcal{A}}_t$. On peut aussi

la construire, par exemple, dans le cas où l'observateur peut changer à des instants $t_1, t_2, \dots, t_m, \dots$, fixés à l'avance, le processus observé, ces derniers étant toujours avec pour temps $\mathbb{R}_+$.

La classe des stratégies ouvertes peut être considérée de deux points de vue :

Soit comme modèle du type suivant : considérons le cas $I = \overline{\mathbb{R}}_+$,

$$V = \{[0, t], \quad t \in \overline{\mathbb{R}}_+\}, \quad \overline{\mathcal{A}}_{t+} = \bigcap_{s>t} \overline{\mathcal{A}}_s.$$

On notera alors que (δ'_1) et (δ_2) sont équivalentes. $\overline{\mathcal{A}}_{t+}$ fait intervenir plus que le processus à l'instant t , sans le considérer aux instants $t + \varepsilon$, $\varepsilon > 0$. Cette tribu résume donc la « tendance » du processus à l'instant t . Si la décision d'arrêt ne peut être considérée comme instantanée, mais exige un intervalle de temps $[t, t + \varepsilon]$, on peut tenir compte de cette « tendance » dans la stratégie d'arrêt.

Soit — et ceci est généralement plus intéressant — comme une approximation de la classe des stratégies, celle des stratégies ouvertes se prêtant mieux à l'étude de l'optimalité. On constate par exemple, que sous des hypothèses assez générales, et dans le cas continu décrit ci-dessus, une stratégie bayésienne naturellement construite dans la classe des stratégies ouvertes, est en fait une stratégie ordinaire.

Remarques.

1. Le point (3) de la définition 1 est en fait une hypothèse sur le plan d'expérimentation. Nous la désignerons dans la suite par (H1). Elle affirme l'existence de l'objet de $\mathcal{C}_\Theta$, qui, sans appartenir à $\mathcal{C}$, donc sans être observable, contient toute l'information du plan d'expérimentation. Il faut noter qu'un tel P_Ω^Θ n'est pas nécessairement unique (sauf hypothèses complémentaires : voir paragraphe 4) et que sa donnée est donc une information complémentaire.

2. La définition de stratégie adoptée sera suffisante, du fait des hypothèses faites ultérieurement dans les théorèmes énoncés (voir aussi chapitre I, exemples de plans d'expérimentation). Il n'est pas évident qu'elle soit suffisamment précise pour recouvrir, dans tous les problèmes, la notion intuitive de stratégie. Chaque cas particulier — ne relevant pas des hypothèses étudiées plus loin — nécessiterait donc vérification.

Dans l'ensemble des stratégies nous considérerons le préordre (π) défini comme suit :

Pour toute stratégie S , posons

$$\theta(S) = \int_{\Omega} \int_{VD} l dS_{VD}^{\omega} dP_{\Omega}^{\theta}.$$

et

$$(\pi): (S_1 < S_2) \Leftrightarrow (\forall \theta \in \Theta)[\theta(S_1) \geq \theta(S_2)] \wedge (\exists \theta \in \Theta)[\theta(S_1) > \theta(S_2)].$$

Il est induit par le produit des ordres duaux de $\mathbb{R}$.

On en étudiera les éléments maximaux et parties cofinales. Nous utiliserons pour cela :

DÉFINITION 3.

a) Soit μ une probabilité sur Θ . Une stratégie S_μ sera dite μ -stratégie, ou stratégie bayésienne relative à μ si :

$$\int_{\Theta} \theta(S_\mu) d\mu(\theta) = \inf_s \int_{\Theta} \theta(S) d\mu(\theta)$$

b) Soit $(\mu) = \{\mu_n\}$ une suite de probabilités sur Θ . Une stratégie $S_{(\mu)}$ sera dite (μ) -stratégie, ou stratégie asymptotiquement bayésienne relative à $\{\mu_n\}$ si :

$$\lim_{n \rightarrow +\infty} \left[\int_{\Theta} \theta(S_{(\mu)}) d\mu_n(\theta) - \inf_s \int_{\Theta} \theta(S) d\mu_n \right] = 0.$$

Pour l'énoncé des théorèmes ultérieurs nous introduirons enfin les définitions techniques suivantes :

DÉFINITION 4. — Le foncteur ξ de I dans $\mathcal{C}_\Theta$ est dit *de type produit* s'il existe $(Y_i, \mathcal{Y}_i)_{i \in I}$ famille d'espaces mesurables, et pour tout $i \in I$, une partie T_i de T , telles que :

$$\Omega_i = \prod_{t \in T_i} Y_t \quad \mathcal{A}_i = \bigotimes_{t \in T_i} \mathcal{Y}_t.$$

Si $u \in \text{Hom}(i, j)$, $T_j \supset T_i$ et ξ_u est la projection de Ω_j sur Ω_i .

DÉFINITION 5. — Le foncteur ξ de I dans $\mathcal{C}_\Theta$, de type produit, est dit (λ) *semi-indépendant* dans les conditions suivantes :

Soit $\{\alpha\}_{\alpha \in A}$ le treillis des parties de T engendré par $\{T_i, i \in I\}$, $(\lambda) = (\lambda_\alpha)_{\alpha \in A}$ un système projectif de probabilités sur

$$\left\{ \mathcal{A}_\alpha = \bigotimes_{t \in \alpha} \mathcal{Y}_t \right\}_{\alpha \in A}.$$

Alors pour tout couple $(\alpha_1, \alpha_2) \in A^2$, $\mathcal{A}_{\alpha_1}$ et $\mathcal{A}_{\alpha_2}$ sont $\lambda_{\alpha_1 \cup \alpha_2}$ conditionnellement indépendantes, par rapport à $\mathcal{A}_{\alpha_1 \cap \alpha_2}$.

DÉFINITION 6. — Les conditions étant celles de la définition 5, nous dirons que le foncteur ξ_1 est (λ) complété du foncteur ξ si :

$$(\forall i \in I)[(\gamma_0 \xi(i) = (\Omega_i, \mathcal{A}_i)) \wedge (\lambda_0 \xi_1(i) = (\Omega_i, \mathcal{A}_i, \lambda_i))].$$

(Rappelons que $\mathcal{A}_p$ désigne la complétée de $\mathcal{A}$ par rapport à P). Dans le cas d'un tel foncteur on prendra, par convention $\mathcal{A}_{1u}$ λ complétée de $\mathcal{A}_u$.

3. RÉSULTATS TECHNIQUES

3.1. Rappels.

THÉORÈME DE RELÈVEMENT (cf. [3], [8]). — Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité complet. $\mathcal{A}$ possède la propriété de relèvement, c'est-à-dire qu'il existe une application $h: L_\infty(\Omega, \mathcal{A}, P) \rightarrow \mathcal{L}_\infty(\Omega, \mathcal{A})$ telle que :

$$h(x) \in x.$$

h est linéaire, isométrique (pour la norme ess. sup et la norme de la convergence uniforme).

h est positive, conserve les constantes et les indicatrices d'ensembles.

h est multiplicative.

COROLLAIRE. — Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, f une application mesurable dans $(E, \mathcal{E})$, telle que $(E, \mathcal{E}, f(P))$ soit complet. Alors $(\Omega, f^{-1}(\mathcal{E}), P)$ possède la propriété de relèvement.

En effet l'application $Y \rightsquigarrow X = Y_0 f$ est une bijection linéaire, isométrique entre $L_\infty(\Omega, f^{-1}(\mathcal{E}), P)$ et $L_\infty(E, \mathcal{E}, f(P))$.

THÉORÈME DE DUALITÉ (cf. [3], [11]). — Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité complet, E un espace de Banach, E' son dual. Le dual de $L'_E(\Omega, \mathcal{A}, P)$ est $L^\infty_E(\Omega, \mathcal{A}, P)$. Si E est séparable il n'est pas nécessaire que $(\Omega, \mathcal{A}, P)$ soit complet.

THÉORÈME DE DÉSINTÉGRATION (cf. [4], [8]). — Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité. Soit f et g deux applications mesurables à valeurs dans $(X, \mathcal{X})$ et $(Y, \mathcal{Y})$. Nous supposons que X est localement compact, $\mathcal{X}$ étant sa tribu borélienne, et que $\mathcal{Y}$ est complète pour la mesure image.

Alors il existe une transition $P_X^\mathcal{Y}$, telle que :

$$(\forall A \in \mathcal{X})(\forall B \in \mathcal{Y})(P[f^{-1}(A) \cap g^{-1}(B)] = \int_B P_A^\mathcal{Y} dg(P).$$

3.2. Extensions.

PROPOSITION 1. — Soit $(Y_t, \mathcal{Y}_t)_{t \in T}$ une famille d'espaces mesurables. Soit sur

$$\left(X = \prod_{t \in T} Y_t, \bigotimes_{t \in T} \mathcal{Y}_t \right)$$

une probabilité P et notons $\mathcal{X}$ la complétée pour P de la tribu produit. Soit I une famille dénombrable de parties T . Notons

$$X_i = \prod_{t \in T} Y_t, \quad \mathcal{X}_i = \left(\bigotimes_{t \in i} \mathcal{Y}_t \right)_P,$$

$\overline{\mathcal{X}}_i$ la tribu des cylindres de X à bases dans $\mathcal{X}_i$.

Supposons que la famille $X_i, \bigotimes_{t \in i} \mathcal{Y}_t$ soit P semi-indépendante au sens de la définition 5. Alors il existe un relèvement h de $L_\infty(X, \mathcal{X}, P)$ dans $\mathcal{L}_\infty(X, \mathcal{X})$ tel que :

$$(\forall i \in I)(\forall \Phi \in M_\infty(X, \mathcal{X}, P))[(\Phi \cap \mathcal{L}_\infty(X, \mathcal{X}_i) \neq 0) \Rightarrow h(\Phi) \in \mathcal{L}_\infty(X, \mathcal{X}_i)].$$

Nous dirons que le relèvement h est *adapté* à la famille $(\overline{\mathcal{X}}_i)_{i \in I}$. Le problème des relèvements adaptés se présente comme suit :

Soit $(X, \mathcal{X}, P)$ un espace probabilisé complet, $\mathcal{X}_1$ une sous-tribu de $\mathcal{X}$. S'il existe un relèvement de $L_\infty(X, \mathcal{X}_1, P)$, alors il se prolonge, selon la méthode de démonstration de Ionescu-Tulcea en un relèvement, qui est donc adapté de $L_\infty(X, \mathcal{X}, P)$.

Pour deux sous-tribus $\mathcal{X}_1, \mathcal{X}_2$ le relèvement doit être aussi adapté avec $\mathcal{X}_3 = \mathcal{X}_1 \cap \mathcal{X}_2$ et $\mathcal{X}_1 \vee \mathcal{X}_2$. Nous verrons sous quelles conditions ci-dessous.

Pour plus de deux tribus, nous devons introduire non seulement les intersections, mais le treillis des tribus engendré, stable par $\cap$ et $\vee$. Même pour une famille finie, ce treillis peut ne pas l'être. En général on ne pourra donc procéder par récurrence. Par contre ceci sera possible pour un système de type produit.

La démonstration se fera en plusieurs étapes marquées par les lemmes suivants.

LEMME 1. — Soit $(X, \mathcal{X}, P)$ un espace probabilisé. Une condition nécessaire et suffisante pour que deux sous-tribus $\mathcal{X}_1$ et $\mathcal{X}_2$ de $\mathcal{X}$, soient indépendantes conditionnellement à $\mathcal{X}_1 \cap \mathcal{X}_2$, est que :

$$(\forall f \in \mathcal{L}_\infty(\Omega, \mathcal{X}_1))[E^2(f) \cap \mathcal{L}_\infty(\Omega, \mathcal{X}_1 \cap \mathcal{X}_2) + \emptyset]$$

(Nous dirons que $\mathcal{X}_1$ et $\mathcal{X}_2$ sont semi-indépendantes).

LEMME 2. — Soit $(X, \mathcal{X}, P)$ un espace probabilisé complet,

$$\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3 = \mathcal{X}_1 \cap \mathcal{X}_2$$

des sous-tribus de $\mathcal{X}$, images réciproques, par des applications ζ_i ($i = 1, 2, 3$), de tribus contenant les points et complètes pour $\zeta_i(P)$. Si $\mathcal{X}_1$ et $\mathcal{X}_2$ sont semi-indépendantes, il existe un relèvement adapté.

Démonstration. — Les grandes lignes suivent celle du théorème de relèvement ([8]).

a) Il existe un relèvement adapté avec $\mathcal{X}_1$ et $\mathcal{X}_3$.

On ne modifie pas $\mathcal{X}_1$ en remplaçant la tribu et l'espace image de ζ_1 par $\zeta_1(X)$ et la trace de la tribu sur cet ensemble. Alors il existe α mesurable telle que $\zeta_3 = \alpha_0 \zeta_1 \dots$. Et il existe un relèvement de $\alpha^{-1}(\mathcal{X}_3)$. Soit sur $\zeta_1(X)$, $\mathcal{N}_1$ l'ensemble des parties de mesure nulle. Nous prolongeons le relèvement à $\alpha^{-1}(\mathcal{X}_3) \vee \mathcal{N}_1$ en prenant :

$$h(A_3 \triangle N) = h(A_3) \quad A_3 \alpha^{-1}(\mathcal{X}_3), \quad N \in \mathcal{N}_1$$

Supposons que $A_3 \triangle N = A'_3 \triangle N'$. Alors $A_3 \triangle A'_3 \subset N \cup N'$ est $\alpha^{-1}(\mathcal{X}_3)$ mesurable et donc de mesure nulle, par suite $h(A_3) = h(A'_3)$ et le relèvement est bien défini. Nous le prolongeons alors comme dans la démonstration du théorème de relèvement en un relèvement de tribu de $\zeta(\mathcal{X})$ (Considérer le système inductif des tribus contenant $\alpha^{-1}(\mathcal{X}_3) \vee \mathcal{N}_1$ et relèvements prolongeant h). Il en résulte un relèvement de $\mathcal{X}_1$ cohérent avec $\mathcal{X}_3$, qui se prolonge en un relèvement cohérent de $\mathcal{X}$.

Remarque. — Cette démonstration s'étend à une suite finie croissante de tribus.

(1) Soit $\mathcal{N}$ l'ensemble des éléments de mesure nulle de $\mathcal{X}$ et $\tilde{\mathcal{X}}_i = \mathcal{X}_i \vee \mathcal{N}$. Alors $\mathcal{X}_1 \cap \mathcal{X}_2 = \mathcal{X}_3$.

Soit $A_1 \in \mathcal{X}_1, A_2 \in \mathcal{X}_2, N_1 \in \mathcal{N}, N_2 \in \mathcal{N}$. Supposons :

$$A_1 \triangle N_1 = A_2 \triangle N_2.$$

Avec les indicatrices.

$$1_{A_1} + 1_{N_1} - 21_{A_1}1_{N_1} = 1_{A_2} + 1_{N_2} - 21_{A_2}1_{N_2}.$$

Soit

$$E^{\mathcal{X}_2}(1_{A_1}) + E^{\mathcal{X}_2}(1_{N_1}) - 2E^{\mathcal{X}_2}(1_{A_1}1_{N_1}) \stackrel{\text{p.s.}}{=} 1_{A_2} + E^{\mathcal{X}_2}(1_{N_2} - 21_{A_2}1_{N_2}).$$

D'où :

$$E^{\mathcal{X}_2}(1_{A_1}) = 1_{A_2} \quad (\text{presque sûrement})$$

$E^{\mathcal{X}_2}(1_{A_1})$ contient par hypothèse de semi-indépendance une fonction $\mathcal{X}_3$ mesurable, A_2 ne diffère donc d'un ensemble de $\mathcal{X}_3$ que pour un ensemble négligeable et par suite $A_2 \triangle N_2$ appartient à $\mathcal{X}_3$.

c) $\tilde{\mathcal{X}}_1$ et $\mathcal{X}_2$ sont semi-indépendantes (et par suite $\mathcal{X}_1$ et $\mathcal{X}_2$ le sont). Soit A une partie de $\mathcal{X}_1$, $AA_1 \triangle N_1$ avec $A_1 \in \mathcal{X}_1$, $N_2 \in \mathcal{N}$.

$$1_A = 1_{A_1} + 1_{N_1} - 21_{A_1}1_{N_1} \quad \text{d'où} \quad E^{\mathcal{X}_2}(1_A) \stackrel{\text{p.s.}}{=} E^{\mathcal{X}_2}(1_{A_1}) = g.$$

Soit $B = B_2 \triangle N_2$, $B_2 \in \mathcal{X}_2$, $N_2 \in \mathcal{N}$:

$$\int_{B_2} g dP = \int_{B_2 \triangle N_2} g dP = P(A_1 \cap B) = P(A_1 \cap B_2)$$

car $A_1 \cap B = (A_1 \cap B_2) \triangle (A_1 \cap N_2)$.

Par suite g est presque sûrement égale (dans $\mathcal{X}$) à $E^{\mathcal{X}_2}(1_A)$, soit donc à une fonction $\mathcal{X}_3$ mesurable g est donc presque sûrement égale à une fonction $\mathcal{X}_3$ mesurable. Par les opérations classiques on passera sans difficultés aux fonctions intégrables quelconques.

Le second point résulte alors de ce que $\tilde{\mathcal{X}}_1 \cap \mathcal{X}_2$ contient les ensembles négligeables de $\mathcal{X}_2$ et avec $A_3 \triangle N$ ou

$$A_3 \in \mathcal{X}_3, \quad N \in \mathcal{N}, \quad E^{\mathcal{X}_2}(1_{A_3 \triangle N}) \stackrel{\text{p.s.}}{=} 1_{A_3} = E^{\mathcal{X}_2}(g)$$

g fonction négligeable, dont l'espérance conditionnée par $\mathcal{X}_2$ est négligeable donc $\mathcal{X}_1 \wedge \mathcal{X}_2$ mesurable.

d) Nous considérons maintenant les tribus $\mathcal{X}_i$ contenant $\tilde{\mathcal{X}}_1$, semi-indépendantes de $\mathcal{X}_2$ et possédant un relèvement h_i adapté à $\mathcal{X}_3$, $\mathcal{X}_2$ et $\mathcal{X}_i \cap \mathcal{X}_2$. D'après le point (a) il existe un tel relèvement h_0 associé à $\tilde{\mathcal{X}}_1$.

Nous ordonnons les couples $(\mathcal{X}_i, h_i)$ par :

$$((\mathcal{X}_i, h_i) < (\mathcal{X}_j, h_j)) = (\mathcal{X}_i \subset \mathcal{X}_2 \text{ et } h_j \text{ prolonge } h_i).$$

Montrons que l'ensemble ainsi ordonné est inductif. Et soit pour cela une famille totalement ordonnée $((\mathcal{X}_i, h_i))_{i \in \mathcal{I}}$.

d1) La famille n'admet pas de partie cofinale dénombrable, toute fonction mesurable de l'union des $\mathcal{X}_i$ est mesurable d'une des composantes. Le prolongement des h_i est alors immédiat. Si f est $\mathcal{X}_2$ mesurable, comme

$$(U\mathcal{X}_i) \cap \mathcal{X}_2 = U(\mathcal{X}_i \cap \mathcal{X}_2)$$

est une tribu, le même résultat est valable pour la restriction aux intersections avec $\mathcal{X}_2$.

Enfin, comme si f est $\mathcal{X}_i$ mesurable, $E^{\mathcal{X}_2}f$ est $\mathcal{X}_i \cap \mathcal{X}_2$ mesurable, donc $U(\mathcal{X}_i \cap \mathcal{X}_2)$ mesurable, la semi-indépendance demeure vraie.

d2) La famille admet une partie cofinale dénombrable. Nous en déduisons une suite croissante de tribus $\mathcal{X}_n$ ($n > 3$) et les relèvements h_n associés. Soit $\vee \mathcal{X}_n$ la tribu engendrée par $\mathcal{X}_n$ ($n > 3$).

— $\vee \mathcal{X}_n$ et $\mathcal{X}_2$ sont semi-indépendantes. Soit f une fonction $\vee \mathcal{X}_n$ mesurable $f = \lim E^{\mathcal{X}_n} f$ (presque sûrement) et $E^{\mathcal{X}_2} f = E^{\mathcal{X}_2} [\lim E^{\mathcal{X}_1} f]$.

Si f est bornée, il en est de même de $E^{\mathcal{X}_n} f$ (et par le même nombre). Nous pouvons intervertir limite et $E^{\mathcal{X}_2}$

$$E^{\mathcal{X}_2} f = \lim (E^{\mathcal{X}_2} E^{\mathcal{X}_n} f)$$

et comme $E^{\mathcal{X}_2} E^{\mathcal{X}_n} f$ est $\mathcal{X}_2 \cap \mathcal{X}_n$ mesurable, $E^{\mathcal{X}_2} f$ est $\mathcal{X}_2 \cap (\vee \mathcal{X}_n)$ mesurable.

— Si f est $\mathcal{X}_2$ mesurable, le relèvement est $\mathcal{X}_2$ mesurable. Or $E^{\mathcal{X}_n} f$ est $\mathcal{X}_2 \cap \mathcal{X}_n$ mesurable, par hypothèse $h_n(E^{\mathcal{X}_n} f)$ est $\mathcal{X}_2$ -mesurable. Par suite, nous placant comme au point (a) dans $\xi_2(X)$ la limite des $h_n(E^{\mathcal{X}_n} f)$ suivant un ultra-filtre plus fin que le filtre de Frachet sur N nous définit une fonction $\mathcal{X}_2$ mesurable puisque dans $\xi_2(X)$ la tribu est complète (cf. la démonstration de Ionescu-Tulcea).

e) Il nous faut enfin montrer que si $(\mathcal{X}_0, h_0)$ est un élément maximal, $\mathcal{X}$ contient $\mathcal{X}_2$.

Supposons qu'il existe $A_2 \in \mathcal{X}_2 \notin \mathcal{X}$. La tribu $\mathcal{X} \vee \{A_2\}$ est formée d'éléments $(X \cap A_2) \cup (Y \cap A_2^c)$, $X \in \mathcal{X}$, $Y \in \mathcal{X}$.

$$e1) (\mathcal{X} \vee \{A_2\}) \cap \mathcal{X}_2 = (\mathcal{X} \cap \mathcal{X}_2) \vee \{A_2\}.$$

Soit

$$1_X 1_{A_2} + 1_Y 1_{A_2^c} = 1_B, \quad B \in \mathcal{X}_2$$

ou $F^{\mathcal{X}_2}(1_X) 1_{A_2} + E^{\mathcal{X}_2}(1_Y) 1_{A_2^c} = 1_B$ (presque sûrement $\mathcal{X}_2$). $\mathcal{X}$ et $\mathcal{X}_2$ étant semi-indépendantes, $E^{\mathcal{X}_2}(1_X)$ et $E^{\mathcal{X}_2}(1_Y)$ sont $\mathcal{X} \cap \mathcal{X}_2$ mesurables. Donc

$$B \in (\mathcal{X} \cap \mathcal{X}_2) \vee \{A_2\}.$$

e2) $\mathcal{X} \vee \{A_2\}$ et $\mathcal{X}_2$ sont semi-indépendantes.

Comme ci-dessus, si $E \in \mathcal{X} \vee \{A_2\}$, $E^{\mathcal{X}_2}(1_B) = E^{\mathcal{X}_2}(1_X) 1_{A_2} + E^{\mathcal{X}_2}(1_Y) 1_{A_2^c}$ dont nous venons de voir qu'il est $(\mathcal{X} \vee \{A_2\}) \cap \mathcal{X}_2$ mesurable.

e3) Il reste à prolonger le relèvement. Nous utiliserons le schéma de la démonstration donnée par Meyer (*Probabilité et Potentiel*, VII, T [8]).

Soit A , appartenant à $\mathcal{X}$, contenant presque sûrement A_2 et de probabilité minimum. A^c est contenu presque sûrement dans A_2^c soit $P(A^c \cap A_2) = 0$, et de probabilité maximum. Nous construirons B , contenant presque sûrement A^c , appartenant à $\mathcal{X} \cap \mathcal{X}_2$ et tel que $P(B \cap A_2) = 0$. Il sera encore de probabilité maximum. On fera alors le prolongement de h à l'aide de B : il sera bien adapté à $(\mathcal{X} \vee \{A_2\}) \cap \mathcal{X}_2 = (\mathcal{X} \cap \mathcal{X}_2) \vee \{A_2\}$, puisque l'on pourrait travailler seulement dans cet ensemble.

Soit $f_A = E^{\mathcal{X}_2}(1_{A^c})$ qui est $\mathcal{X} \cap \mathcal{X}_2$ mesurable. Pour tout B appartenant à $\mathcal{X} \cap \mathcal{X}_2$:

$$\begin{aligned}\int_{A_2 \cap G} f_A dP &= \int_{A_2 \cap G} 1_A dP = 0 \\ \int_{A_2^c \cap G} f_A dP &= \int_{A_2^c \cap G} 1_A dP = P(A \cap G).\end{aligned}$$

A_2 et G sont de la forme $\zeta_2^{-1}(A_2^{(2)})$ et $\zeta_2^{-1}(G^{(2)})$ où nous pouvons encore supposer ζ_2 surjective. Alors à f_A correspond de la même manière $f_A^{(2)}$ et la première relation s'écrit :

$$\int_{A^{(2)} \cap G^{(2)}} f_A^{(2)} dP^{(2)} = 0.$$

Donc sur $A_2^{(2)}$, $f_A^{(2)} = 0$, sauf sur un ensemble de mesure nulle dont l'image réciproque est dans $\mathcal{X} \cap \mathcal{X}_2$ (car $\mathcal{X}$ est complète).

Par suite sur A_2 , $f_A = 0$ sauf sur la trace d'un ensemble de mesures nulles de $\mathcal{X} \cap \mathcal{A}_2$. Soit alors $B = [f_A \neq 0]$, $B \in \mathcal{X} \cap \mathcal{X}_2$ car $f_A = E^{\mathcal{X}_2}(1_{A^c})$ et $\mathcal{X} \cap \mathcal{X}_2$ mesurable, et de ce qui précède résulte que $P(6_2 \cap B) = 0$.

Enfin

$$P(A^c \cap B) = \int_{A_2^c \cap B} f_A dP = \int_{A_2^c} f_A dP = \int_{A_2^c \cap X} f_A dP = P(A^c)$$

d'où B contient A^c presque sûrement.

Ceci achève la démonstration.

LEMME 3. — Soit $(X, \mathcal{X}, P)$ un espace probabilisé complet, $(\mathcal{X}_i)_{i \in I}$ une famille dénombrable de sous-tribus images réciproques de tribus complètes et contenant les points. Si pour toute partie finie de I , il existe un relèvement h_j adopte ou $(\mathcal{X}_i)_{i \in J}$, alors il existe un relèvement adapté où $(\mathcal{X}_i)_{i \in I}$.

Démonstration. — Soit I_n une suite de parties finies de I , croissante vers I . A chacune de ces parties est associé un relèvement h_n , et si f est $\mathcal{X}_i$ mesurable, pour n assez grand, tous les $h_n(f)$ sont $\mathcal{X}_i$ mesurables.

Posons $h_n(f) = f_n$ et pour $\mathcal{X} \in X$, $\mathcal{X}_n(f) = f_n(\mathcal{X})$. $\mathcal{X}_n$ est une forme linéaire continue de norme 1 sur $F = L_\infty(X, \mathcal{X}, P)$ et appartient donc à la boule unité du dual E' .

Soit $\mathcal{U}$ un ultra-filtre de N , plus fin que le filtre de Fréchet, et soit dans E' muni de sa topologie faible :

$$x_0 = \lim_{\mathcal{U}} x_n$$

Définissons $h(f)$ par :

$$h(f)(x) = x_0(f).$$

Lorsque $\lim_n f_n(x)$ existe, $x_0(f)$ coïncide avec cette limite, et $h(f)$ est $\mathcal{X}$ mesurable, puisque cette tribu est complète et que f_n converge presque sûrement. .

Soit maintenant f une classe de fonctions $\mathcal{X}_i$ mesurables. $\mathcal{X}_i$ est image réciproque d'une tribu complète par une application ζ_i que nous pouvons ici encore supposer surjective à valeurs dans (X_i, Y_i, P_i) . Posons

$$F_i = L_\infty(X_i, \mathcal{X}_i, P_i).$$

Si x et x' ont même image par ζ_i , x_n et x'_n sont égales comme formes linéaires sur F_i et il en est de même des limites x_0 et x'_0 .

Par suite $x_0(f)$ est définie à partir de l'image de f dans F_i (en se limitant à l'ensemble des indices n , tels que $i \in I_n$). On définit ainsi une fonction $h^i(f)$, sur X_i qui est $\mathcal{Y}_i$ mesurable, puisque cette tribu est complète, et telle que

$$h(f) = h^i(f) \circ \zeta_i$$

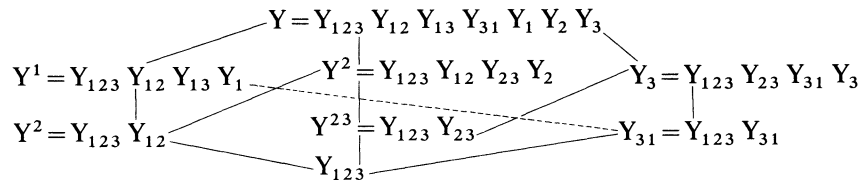
$h(f)$ est donc bien $\mathcal{X}_i$ mesurable.

Démonstration de la proposition 1. — Nous noterons d'abord la petite difficulté suivante :

$\overline{\mathcal{X}}_{i \cap j} \subset \overline{\mathcal{X}}_i \cap \overline{\mathcal{X}}_j$ et en général l'égalité n'est pas vraie. On définira donc pour appliquer le lemme 2 un relèvement de $\overline{\mathcal{X}}_{i \cap j}$ que l'on prolongera à $\overline{\mathcal{X}}_i \cap \overline{\mathcal{X}}_j$. Il est immédiat toutefois que ces deux tribus ne diffèrent que par des éléments de mesure nulle de $\mathcal{X}$, et le relèvement pourra être choisi à valeurs dans $\mathcal{L}_\infty(X, \overline{\mathcal{X}}_{i \cap j})$.

Par la définition de la semi-indépendance, on peut supposer I filtrant, et par le lemme 3 I fini. Pour $n = 2$ le résultat est d'après la remarque que nous venons de faire celui du lemme 2.

Prenons le cas de $n = 3$. On a le schéma ci-dessous, avec des notations évidentes :



Nous opérons comme suite: il nous faut successivement l'existence d'un relèvement :

de $Y_{123} Y_{12} Y_{29} Y_{31} Y_2 Y_3$ adapté à $Y_{123} Y_{12} Y_{31}, Y^2$ et Y^3

(alors le prolongement à Y adapté à Y^1, Y^2, Y^3 résulte du lemme 2 et de la remarque faite au début de la démonstration),

de $Y_{123} Y_{12} Y_{23} Y_{31} Y_3$ adapté à $Y_{123} Y_{12} Y_{31}, Y_{123} Y_{12} Y_{23}$ et Y^3

(d'où le relèvement précédent de la même manière),

de $Y_{123} Y_{12} Y_{23} U_{31}$ adapté à $Y_{123} Y_{12} Y_{31}, Y_{123} Y_{12} Y_{23}, Y_{123} Y_{23} Y_{31}$

(même opération),

de $Y_{123} Y_{23} Y_{31}$ adapté à $Y_{123} Y_{31}$ et $Y_{123} Y_{23}$

(même opération).

Le lemme 2 et la remarque du début achève la démonstration. L'extension à n quelconque se fait selon le même procédé.

PROPOSITION 2. — Soit $(\Omega, \mathcal{A}, \mathcal{P})$ un espace probabilisé complet, $\{\mathcal{A}_i, i \in I\}$ une famille de sous-tribus, h un relèvement adapté. Soit A un espace localement compact, $(A_i)_{i \in I}$ une famille d'ouverts de A , E l'espace des fonctions continues sur A , à support compact. E' son dual (mesures de Radon bornées). Dans la boule unité de $L_E^\infty(\Omega, \mathcal{A}, \mathcal{P})$, l'ensemble des éléments admettant un représentant S_A^Ω dont les restrictions aux A_i soient $\mathcal{A}_i$ mesurables, est fermé pour la topologie faible $\sigma(L_E^\infty, L_E^1)$.

Démonstration. — A chaque élément de L_E^∞ , $(\Omega, \mathcal{A}, \mathcal{P})$ est associée, par le théorème de dualité même, une transition à valeurs dans A , muni de sa tribu de Baire. Nous nous contenterons ici de cette transition, la proposition suivante étudiant le passage à la tribu de Borel.

Pour u forme linéaire sur $L_1 \oplus E$, on posera $w: e \rightarrow u(\cdot, e)$, $w = h_0 w$ (h relèvement). $e \xrightarrow{w} u(\cdot, e) \xrightarrow{h} g_e = \langle G, e \rangle \in \mathcal{L}_\infty(\Omega, \mathcal{A})$. G est ici une fonction de Ω dans les mesures de Radon bornées, scalairement mesurables, définissant donc une transition de Baire.

Soit u_α un ultrafiltre de telles formes, appartenant à la boule unité forte et convergent simplement vers u_0 . A u_0 , nous associons g_{02} et S_{0A}^Ω . Pour tout f et tout e , $u_0(f, e) = \lim u_\alpha(f, e)$, soit dans $L_\infty(\Omega, \mathcal{A}, \mathcal{P})$:

$$u_0(\cdot, e) = \lim u_\alpha(\cdot, e)$$

$u_\alpha(\cdot, e)$ sont, pour e à support dans A_i , des éléments de $L_\infty(\Omega, \mathcal{A}_i, \mathcal{P})$.

Or la boule unité de $L_\infty(\Omega, \mathcal{A}_i, P)$ est fermée dans $L_\infty(\Omega, \mathcal{A}, P)$ pour la topologie faible de ce dernier (cf. remarque 1 ci-dessous). Donc

$$g_{0,e} \in \mathcal{L}_\infty(\Omega, \mathcal{A}_i).$$

Pour un compact G_δ de A_i , soit C , il existe une suite décroissante de e_n , à support compact, contenu dans A_i , décroissante vers 1_C (Tout compact G_δ de A_i est un compact G_δ de A contenu dans A_i et réciproquement). D'où :

$$S_{0_C}^\Omega = \lim \downarrow g_{0,e_n}$$

et par suite est $\mathcal{A}_i$ mesurable.

Remarques.

1. Si $\mathcal{B} \subset \mathcal{A}$, la boule unité de $L_\infty(\Omega, \mathcal{B}, P)$ est fermée dans $L_\infty(\Omega, \mathcal{A}, P)$ pour sa topologie faible. En effet la topologie induite par la faible de $L_\infty(\Omega, \mathcal{A}, P)$ est la faible de $L_\infty(\Omega, \mathcal{B}, P)$ car pour $f \in L_1(\Omega, \mathcal{A}, P)$, $g \in L_\infty(\Omega, \mathcal{B}, P)$:

$$E_P(fg) = E_P[E^{\mathcal{B}}(f)g]$$

2. La proposition 2 démontre en fait que toute forme u dont la restriction aux A_i appartient à $L^\infty(\Omega, \mathcal{A}_i, P)$, admet un représentant dont la restriction aux A_i est $\mathcal{A}_i$ mesurable.

Nous achèverons la liste des résultats techniques en indiquant les liens entre probabilités de transition et éléments de $L_E^\infty(\Omega, \mathcal{A}, P)$.

PROPOSITION 3. — Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité complet, $(\mathcal{A}_i)_{i \in I}$ une famille de sous-tribus de $\mathcal{A}$, possédant un relèvement adapté. Soit A un espace localement compact, $(A_i)_{i \in I}$ une famille d'ouverts de A . Toute transition de Baire S_A^Ω , telle que les restrictions aux A_i soient $\mathcal{A}_i$ mesurables, est scalairement presque sûrement égale à une transition à valeurs dans les mesures de Borel régulières satisfaisant aux mêmes conditions (Une transition sera dite de Baire si la mesure S_A^ω est considérée sur la tribu de Baire de A).

Démonstration. — Nous verrons au cours de celle-ci que l'on peut se limiter, à démontrer le prolongement de S_A^Ω .

A la transition de Baire S_A^Ω s'attache une forme bilinéaire continue sur $L_1 \times E$ (E fonctions continues à support compact) que nous relevons en une transition de Baire S_A^Ω , scalairement presque sûrement égale à S_A^Ω , et que nous allons prolonger (démonstration de Ionescu-Tulcea dans [8]).

Soit U un ouvert de A . Pour g fonction continue à support compact, nous noterons

$$S_g^\Omega = \int_A g dS_A^\Omega.$$

Pour chaque ω , la mesure de Borel régulière se définit par

$$S_v^\omega = \sup_{0 \leq g \leq 1} S_g^\omega.$$

Rappelons que S_g^Ω est image par h , et si g est à support dans A_i est $\mathcal{A}_i$ mesurable. Soit

$$H = \text{Ess Sup}_{0 \leq g \leq 1} S_g^\Omega h(H) \geq h(S_g^\Omega) = S_g^\Omega.$$

Donc $h(H) \geq S_v^\Omega$ et par suite $H \geq S_v^\Omega$ presque sûrement.

Si on se rapporte à la construction de Ess Sup , on constate que $S_v^\Omega \geq H$ presque sûrement. Comme $\mathcal{A}$ est complète (ou $\mathcal{A}_i$ image réciproque de tribu complète) $S_v^\Omega \in H$ et par suite est $\mathcal{A}$ mesurable ($\mathcal{A}_i$ mesurable). Le passage à la tribu est immédiat.

4. LE THÉORÈME FONDAMENTAL DE LA DÉCISION STATISTIQUE

4.1. Énoncé des principaux résultats.

THÉORÈME 1. — Soit ξ un plan d'expérimentation, I une fonction du coût tels que :

H1) Il existe, et il a été choisi, une transition P_Ω^Θ , telle que, pour tout

$$i \in I, \quad P_{\Omega_i}^\Theta = \zeta_i(P_\Omega^\Theta).$$

H2) I possède une partie cofinale dénombrable totalement ordonnée J , et il existe un système projectif de probabilité $(\lambda) = (\lambda_i)_{i \in J}$ sur X_i , tel que, pour tout i appartenant à J :

$$P_{\Omega_i}^\Theta \ll \lambda_i.$$

H3) V et D sont localement compacts. Pour tout u , et pour tout i appartenant à J , $A_u \cap V_i$ est compact.

H4) (a) Pour tout i de J et pour tout $(u, v) \in V_i^2$, il existe $u \wedge v \in V_i$, facteur maximal commun unique de u et v .

(b) Pour l'ordre induit par la factorisation et la topologie la plus fine rendant continue l'application de V_i dans B_u , $v \rightarrow u \wedge v$, B_u est isomorphe à un compact de $\mathbb{R}$.

H5) (a) Pour tout i appartenant à J , il existe une partie dénombrable V_{0i} de V_i telle que pour tout $u \in V_i$,

$$A_u \cap V_i = \bigcap_{v \in V_{0i}} (A_v \cap V_i)$$

v facteur strict de u .

(b) Le plan d'expérimentation satisfait à la condition (a) (cf. 2.2) et il existe une partie dénombrable I_0 de I et pour tout i une suite i_n d'éléments de I_0 telle que :

$$V_i \subset \overset{\circ}{V}_{i_n} \text{ (intérieur de } V_{i_n} \text{ dans } V)$$

$\overline{\mathcal{A}}_{i+} = \xi_i^{-1}(\mathcal{A}_i)$ est la tribu des événements asymptotiques de

$$\overline{\mathcal{A}}_{in} = \xi_{in}^{-1}(\mathcal{A}_{in})$$

H6) Pour tout i de J , $v \in V$, il existe $\tau_i(v) \in V_i$, facteur maximal unique de v contenu dans V_i . L'application τ_i de V dans V_i est continue. Pour tout i de I , il existe j dans J tel que :

$$\tau_i^{-1}(V_i) = V_j.$$

H7) Ou bien V et D sont à base dénombrable,

ou bien ξ est (λ) complète d'un foncteur de type produit, (λ) semi-indépendant.

H8) 1 est fonction mesurable de $\Theta\Omega V D$ dans $\mathbb{R}_+$ est croissante de $i = \beta(v)$; sa restriction à $V_i D$ est limitée croissante d'une suite de fonctions de $\mathcal{L}_{\mathcal{B}(V_i D)}^1(\Omega_i, \mathcal{A}_i, P_{\Omega_i}^\Theta)$ l tend vers $+\infty$ à l'infini de D , pour tout (θ, ω, v) , et vers $+\infty$ à l'infini de V pour tout (θ, ω, d) .

Alors, pour tout préordre de type (π) sur l'ensemble des stratégies ouvertes

W1) La classe des éléments maximaux est cofinale.

W2) La classe des (μ_n) stratégies, (μ_n) suite de probabilité de support fini sur Θ , est cofinale.

W3) Il existe une partie compacte Σ d'un espace vectoriel topologique et une application surjective de l'ensemble des stratégies sur Σ , telle que toute classe de stratégies ayant pour image la fermeture dans Σ de l'image des μ -stratégies, μ de support fini, est cofinale.

Avant de discuter les hypothèses du théorème, nous donnerons trois corollaires qui l'éclairent et en montrent les limites.

COROLLAIRE 1. — Soit $(\Omega, \mathcal{A}, P_\Omega^\Theta)$ un espace muni d'une transition λ une probabilité sur $\mathcal{A}$ dominant P. Soit D un espace localement compact, 1 une fonction de $\Theta\Omega D$ dans $\mathbb{R}_+$ limite d'une suite croissante de fonctions de $\mathcal{L}_{\mathcal{C}(D)}^1(\Omega, \mathcal{A}, P_\Omega^\Theta)$. Supposons que 1 tend vers $+\infty$ à l'infini de D, quels que soient θ et ω . Supposons $\mathcal{A}$ complète pour λ .

Alors les conclusions du théorème 1 sont exactes.

COROLLAIRE 2 (Wald [13]). — Soit :

— D' un espace localement compact à base dénombrable.
 $(Y_n, \mathcal{Y}_n)$ une famille, dénombrable d'espaces mesurables, P_Ω^Θ une transition sur

$$\Omega = \prod_{n=1}^{+\infty} Y_n, \quad \mathcal{A} = \bigotimes_{n=1}^{+\infty} \mathcal{Y}_n,$$

telle qu'il existe pour toute partie finie i de N une probabilité λ_i sur

$$\Omega_i = \prod_{n \in i} Y_n, \quad \mathcal{A}_i = \bigotimes_{n \in i} \mathcal{Y}_n,$$

dominant $P_{\Omega_i}^\Theta$, les λ_i formant un système projectif.

I une famille de parties finies de N, V une famille de suites croissantes : $(i_1 \subset i_2 \subset \dots \subset i_k)$ d'éléments de I, telles qu'il existe une suite croissante d'entiers c_k , avec $i_k \subset \{1, 2, \dots, c_k\}$.

1 une fonction de $\Theta\Omega V D$ dans $\mathbb{R}_+$ satisfaisant H_8 (topologie discrète sur V).

Alors les conclusions du théorème 1 sont exactes pour l'ensemble des stratégies vraies.

Remarque. — Même conclusion avec D localement compact et les systèmes semi-indépendants.

COROLLAIRE 3. — Soit : D un espace localement compact.

— $(Y_t, \mathcal{Y}_t)_{t \in \mathbb{R}_+}$, une famille d'espaces mesurables, P_Ω^Θ une transition sur

$$\Omega = \prod_{t \in \mathbb{R}_+} Y_t, \quad \mathcal{A} = \bigotimes_{t \in \mathbb{R}_+} \mathcal{Y}_t,$$

— λ_i un système projectif de probabilités sur

$$\Omega_i = \prod_{t \leq i} Y_t, \quad \mathcal{A}_i = \bigotimes_{t \leq i} \mathcal{Y}_t$$

tel que $P_{\Omega_i}^\Theta$ soit absolument continu par rapport à λ_i .

Soit

$$\mathcal{A}_i = \left(\bigotimes_{t \leq i} y_t \right)_{\lambda_i} \quad V = \{ [0, i], i \in \mathbb{R}_+ = I \},$$

— l une fonction de $\Theta\Omega\text{VD}$ dans $\overline{\mathbb{R}}_+$ satisfaisant (H8).

Alors les conclusions du théorème 1 sont exactes pour l'ensemble des stratégies ouvertes.

4.2. Remarques sur les énoncés précédents.

1) L'hypothèse (H1) ne soulève guère de difficultés au niveau des applications. Si ξ est de type produit et si les $(Y_i, \mathcal{Y}_i)$ sont par exemple des espaces polonais munis de leurs tribus boréliens et si I est filtrante elle est automatiquement satisfaite (cf. par exemple Neveu [9]).

2) Les hypothèses (H1), (H5) et (H6) caractérisent essentiellement la catégorie I des états d'expérimentation. (H1) met en correspondance un « chemin » avec par exemple un intervalle de $\mathbb{R}$ ou de $\mathbb{N}$.

(H5) est une hypothèse de séparabilité et de continuité à droite. La première partie (H5a) est très restrictive : les « embranchements » possibles sont au plus dénombrable.

La seconde partie implique la continuité à droite du processus.

(H6) enfin est analogue à la limitation de Wald [13] : à chaque étape (cas discret) le nombre des choix possibles, pour la suite est fini (cf. corollaire 2). Elle semble toutefois moins limitative (cf. I, exemple 3).

3) L'hypothèse de démonstration (H2) est très restrictive dès que l'on s'écarte des schémas discrets (cf. corollaire 3 par exemple). Elle exige que l'ensemble $\{P_{\Omega}^{\theta}, \theta \in \Theta\}$ soit relativement compact pour la topologie affaiblie du Banach $\mathcal{M}(\Omega_i)$.

Dans la perspective des énoncés du théorème 1, les efforts faits pour l'éliminer (cf. Le Cam [6]) conduisent à une notion de stratégie qui nous paraît sans intérêt pratique (formes linéaires sur $L_1(\Omega, \mathcal{A}, P_{\Omega}^{\theta} \theta \in \Theta) \otimes \mathcal{C}(D)$, elles n'ont pas de « représentations » du type transitions S_{Ω}^{Ω}).

5) Les stratégies considérées sont telles que les dimensions dépendent non seulement de l'état d'expérimentation mais du chemin suivi pour y parvenir. A l'exception du cas discret (I et V dénombrables), l'écriture nécessaire pour restreindre la notion de stratégies est assez lourde. Nous ne la considérerons pas. Dans le cas discret il est facile de voir que les résultats ne sont pas affectés.

4.3. Démonstration du théorème 1.

Elle résultera d'une série de lemmes.

LEMME 1 (Le Cam [5]). — Soit Σ une partie compacte convexe d'un espace vectoriel topologique, Θ une famille d'applications linéaires, semi-continues inférieurement de Σ dans $\mathbb{R}_+$, un préordre de type π sur Σ . Alors :

- 1) La classe des éléments maximaux est cofinale.
- 2) L'ensemble des (μ_n) -éléments, (μ_n) parcourant les suites de probabilités de supports finis sur Θ , est cofinal.
- 3) La fermeture B dans Σ de la classe des μ -éléments, μ parcourant les probabilités de support fini sur Θ , est cofinale.

Démonstration. — Elle diffère de celle de Le Cam en évitant toute référence au minima. Les points essentiels sont néanmoins ceux de [5].

1) Soit $x_0 \in \Sigma$ et soit $\mathcal{F}$ ensemble des parties F de Σ qui satisfont :

F est fermée.

Tout élément de F est supérieur ou égal (π) à x_0 .

Un élément de F n'est strictement dominé par aucun élément non contenu dans F .

$\mathcal{F}$ est non vide $\{x \geq x_0\}$ est fermé et pour cet ensemble le troisième point résulte de la transitivité du préordre.

$\mathcal{F}$ est inductif décroissant pour l'inclusion toute intersection décroissante est fermée (compacte de Σ), formée d'éléments supérieurs ou égaux à x_0 . Et s'il existe x dominant strictement un élément y de F , il appartient avec y à tous les F_i de la famille décroissante considérée. Il est donc dans F .

Soit F_0 un élément minimal, $y \in F_0$ et $x > y$. Alors $\{z > y\} \subset F_0$ donc $x \in F_0$. Mais $\{z \geq x\}$ est un élément de F contenu dans F_0 et qui ne contient pas y . F_0 n'est pas minimal. Il y a contradiction et tout élément de F_0 est maximal et domine x_0 .

2a) Nous utiliserons le résultat suivant :

Soit A un convexe $\mathbb{R}^n$ contenu dans $\mathbb{R}_+^n$. Soit $M \in A$, $M = (x_1, x_2, \dots, x_n)$. Pour toute mesure de probabilité sur $\{1, 2, \dots, n\}$ définie par $(\mu_1, \mu_2, \dots, \mu_n) = \mu$, posons

$$\mu(M) = \sum_{i=1}^n \mu_i x_i.$$

Supposons que pour toute telle mesure, avec $\mu_i > 0$, quel que soit i ,

$$\mu(M_0) - \inf_{M \in A} \mu(M) \geq \varepsilon \quad (M_0 \in A).$$

Alors il existe M_0'' dans A tel que pour tout $i = 1, 2, \dots, n$, on ait

$$x_i(M_0) \geq x_i(M_0'') + \frac{\varepsilon}{3}.$$

En effet : soit A_1 la fermeture de A , elle aussi contenue dans $\mathbb{R}_+^n$, et C le cône intersection des demi-espaces fermés contenant A_1 , et limités par des hyperplans parallèles aux axes de coordonnées. La variété linéaire engendrée par A est $\mathbb{R}^n$ tout entier. En effet si cette variété est contenue dans un hyperplan qui ne soit parallèle à aucun axe de coordonnée, pour cet hyperplan $\mu(M)$ est constant sur A et ceci contredit la propriété énoncée.

Supposons alors que les éléments de A satisfassent $x_n = 0$ et qu'il existe $v_1, v_2, \dots, v_{n-1}$, strictement positifs et de somme 1 avec,

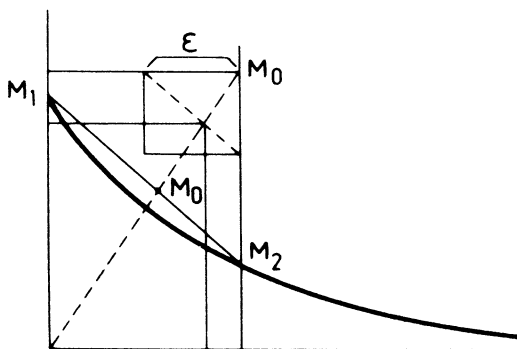
$$(v_1 x_{01} + \dots + v_{n-1} x_{0n-1}) - \inf_{M \in A} (v_1 x_1 + \dots + v_{n-1} x_{n-1}) = a.$$

Soit alors $0 < \mu_n < 1$, et posons $\mu_i = v_i(1 - \mu_n)$ pour $i = 1, \dots, n-1$. Alors

$$(\mu_1 x_{01} + \dots + \mu_n x_{0n}) - \inf_{M \in A} (\mu_1 x_1 + \dots + \mu_n x_n) = a(1 - \mu_n)$$

qui peut être rendu $< \varepsilon$, pour $1 - \mu_n$ assez petit. Ainsi $x_n = 0$ (ou constant) sur tout A est impossible, ce qui prouve que A engendre $\mathbb{R}^n$.

Nous pouvons alors sans limiter la généralité, supposons que $C = \mathbb{R}_+^n$. Considérons alors pour toute droite du cône $\mathbb{R}_+^n$, le point de A_1 le plus proche de 0 et soit B l'ensemble de ces points. Choisissons, pour chaque axe de coordonnée, le point M_α qui est, s'il existe l'intersection de B et de la parallèle à l'axe considérée, et sinon le point de contact de B et de l'hyperplan de coordonnée orthogonal à l'axe. Le point de coordonnées $x_i(M) - \varepsilon/2$ est dans le demi-espace fermé limité à l'hyperplan défini par les points M_α ,



qui ne contient pas l'origine. Il en résulte que le convexe engendrée par les M_α , qui est contenu dans A_1 , contient un point M'_0 tel que

$$x_i(M) - x_i(M'_0) \geq \frac{\varepsilon}{2}.$$

Considérons alors l'ouvert

$$|x_i(M) - x_i(M'_0)| < \frac{\varepsilon}{6},$$

quel que soit i . Il contient un point M''_0 de A qui satisfait aux conditions imposées.

2b) Soit d'abord Θ fini, et μ parcourt l'ensemble des probabilités sur Θ de poids positif en tout point. On désigne encore par

$$\mu(x) = \sum_{\theta} \mu(\Theta)\theta(x).$$

Soit x_0 tel que

$$\mu(x_0) - \inf_{x \in \Sigma} \mu(x) \geq \varepsilon$$

pour tous les μ considérés. Alors il existe $x_1 \in \Sigma$, tel que

$$\mu(x_1) \leq \mu(x_0) = \frac{\varepsilon}{3}$$

quel que soit μ . Si l'un des $\Theta_i(x_0)$ est infini il n'y a rien à démontrer. Sinon c'est le résultat (2a). On notera que la compacité de Σ n'intervient pas.

Soit alors Σ compact, et x_0 tel que :

$$\inf_{\mu} [\mu(x_0) - \inf_{x \in \Sigma} \mu(x)] \geq \varepsilon > 0,$$

μ parcourant les probabilités sur Θ , de support fini.

Pour chaque partie finie T de Θ , l'ensemble des x_1 tels que pour tout μ de support T ,

$$\mu(x_1) \leq \mu(x_0) - \frac{\varepsilon}{3}$$

est non vide. Par suite nous constituons un filtre de parties de Σ , image du filtre des parties finies de Θ . Il possède un point adhérent x_1 qui est tel que, pour toute mesure μ de support fini.

Par suite : ou bien $\inf_{\mu} [\mu(x_0) - \inf_x \mu(x)] = 0$ (en prenant $\infty - \infty = 0$) et x_0 est une (μ_n) -élément,

ou bien x_0 est majoré pour le préordre (π) .

3) Soit d'abord Θ fini. Considérons H famille finie de probabilité sur Θ , strictement positives en tout point de Θ . Alors $\mu(x) = +\infty$ quel que soit μ dans H ou bien $\mu(x) < +\infty$ quel que soit μ dans H . Pour le problème associé à H (au lieu de Θ), tout élément x_0 est majoré par un élément maximal d'après le point (1). Si celui-ci est infini pour tout μ de H , il est majoré par n'importe quelle v -élément (v mesure sur H). Sinon l'admissible est une v -stratégie, v probabilité sur H (cf. ci-dessous lemme 2). Il existe donc x_1 , qui est un η -élément du problème original, η mesure induite sur Θ par v , et tel que pour toute mesure μ combinaison des mesures de H , $\mu(x_1) \leq \mu(x)$. Associons à H l'ensemble A_H de ces x_1 . Lorsque H parcourt l'ensemble des familles finies de probabilités de poids strictement positif en tout point de Θ , les A_H parcourent un filtre de parties de Σ , qui admet donc un point adhérent x_1 dans B . On a alors pour tout x sur Θ , $\mu(x_1) \leq \mu(x_0)$.

Si Θ est quelconque on réédite l'opération avec les parties finies de Θ . Et le même raisonnement conduit à la même conclusion.

Nous avons utilisé au cours de la démonstration le :

LEMME 2. — Soit Σ partie convexe d'un espace vectoriel topologique, Θ famille d'applications linéaires de Σ dans $\mathbb{R}_+$. Supposons que pour tout x , la fonction $x : \theta \rightsquigarrow \theta(x)$ soit continue et finie sur Θ topologique pseudo-compact (Weierstrassein). Alors tout élément maximal pour un préordre π est un μ -élément.

Démonstration. — $x \rightsquigarrow \bar{x}$ est une application linéaire dans $\mathcal{C}(\theta, \mathbb{R})$. L'image de Σ , soit Σ est convexe, et le préordre se transporte, les éléments maximaux se correspondant.

Soit $\bar{x}_0$ un élément maximal, C_0 le cône translaté du cône des fonctions strictement négatives, et de sommet $\bar{x}_0$. $\bar{x}_0$ étant maximal on a $\bar{\Sigma} \cap C_0 = \emptyset$.

Comme C_0 est ouverte pour la topologie de Banach (toute fonction continue atteint son minimum) il existe une forme linéaire continue positive qui sépare Σ et C_0 . Soit μ la mesure de Radon associée, $\bar{x}_0$ est un μ -élément.

LEMME 3. — Soit ξ un plan d'expérimentation satisfaisant :

(H1)

(H2 α) Il existe une probabilité λ sur Ω telle que $P_\Omega^\theta \ll \lambda$.

(H3 α) V et D sont compacts. Pour tout u , A_u est compact.

(H4 α) H4 est satisfaite pour V .

(H5 α) H5 est satisfaite pour V .

(H7)

Alors l'image dans $L_{\mathcal{M}(VD)}^\infty(\Omega, \mathcal{A}, \lambda)$ des stratégies S_{VD}^Ω est compacte pour $\sigma(L_{\mathcal{M}(VD)}^\infty L_{\mathcal{C}(VD)}^1)$.

Démonstration. — A toute stratégie S_{VD}^Ω on associe évidemment un élément de L^∞ . Nous savons que réciproquement à un élément de L^∞ on peut associer une transition S_{VD}^Ω mais qui ne satisfait pas nécessairement aux conditions imposées aux stratégies :

($\delta'1$) ($\forall i \in I$) $S_{V_i D_i}^\Omega$ est $\bar{x}_{i+}$ mesurable.

($\delta 2$) ($\forall u \in V$) $S_{A_u D}^\Omega$ factorise à travers ζ_u^- .

(sauf sur un ensemble de mesure nulle quel que soit θ).

Pour que ($\delta'1$) soit satisfaite il suffit d'après l'hypothèse (H5) que :

($\forall i \in I_0$) $S_{V_i D_i}^\Omega$ factorise à travers ζ_i .

Pour que ($\delta 2$) soit satisfaite il suffit qu'elle le soit pour $u \in V_0$. En effet par l'hypothèse (H5a) $A_u D$ est intersection dénombrable de $A_0 D$, $v \in V_0$, A_v décroissant et ainsi $S_{A_u D}^\Omega$ factorisera à travers la limite de ζ_v^- donc de ζ_u^- qui en est la limite.

Nous modifierons encore cette condition. Pour chaque u appartenant à V_0 on considère une suite décroissante vers $1_{\{u\}}$ de fonctions φ_n^u continues en B_n (ensemble des facteurs de u) pour la topologie la plus fine rendant $v \rightarrow u \wedge v$ continue (hypothèse H4). Soit alors $h_n^u(v, d) = \varphi_n^u(u \wedge v)$. Il suffit que ($\delta 2$) soit satisfaite pour que les

$$\int_{VD} h_n^u dS_{VD}^\Omega.$$

On peut considérer les transitions S_{VD}^Ω convenables, comme des formes linéaires u sur $L_1(\Omega, \mathcal{A}, \lambda) \otimes \mathcal{G}(VD)$ qui sont telles que :

$$u(\cdot, a) \in L_\infty(\Omega, \mathcal{Y}_a, \lambda)$$

dans les cas suivants :

— a à support dans $\mathring{V}_i D$, ne dépendant que de D_i , $i \in I_0$, $\mathcal{Y}_a = \bar{\mathcal{A}}_i = \zeta_i^{-1}(\mathcal{A}_i)$,

— $a = h_n^u$, $u \in V_0$, $\mathcal{Y}_a = \bar{\mathcal{A}}_u$,

— $a = 1_{VD}$, $\mathcal{Y}_a = \mathcal{A}$ et $u(f, a) = \int f d\lambda$.

Ces formes sont définies par :

$$u(\cdot, a): f \rightsquigarrow \int_X f \int_{VD} a dS_{VD}^\omega d\lambda(\omega)$$

Elles sont dans des boules fortes, dont le rayon dépend de a , du L_∞ considéré. En outre elles sont positives. Les formes linéaires satisfaisant à ces diverses conditions forment un ensemble faiblement fermé de $L_{\mathcal{M}(VD)}^\infty$. Il suffit de

le vérifier pour chaque condition. Pour $a = 1_{VD}$, ceci résulte de ce que VD est compact. Pour les autres, se reporter à la démonstration de la proposition 2.

Ainsi nous définissons une application de l'ensemble des stratégies dans un ensemble compact Σ . Il reste à montrer que cette application est surjective.

Si ξ est λ complété d'un système de type produit, λ semi-indépendant les conditions imposées étant en infinité dénombrable, la possibilité d'associer à toute forme linéaire une stratégie résulte de II 2.1 lemme 5. Il en est de même pour un plan d'expérience borélien.

Si V et D sont à base dénombrable, on considère un relèvement quelconque, et une algèbre dénombrable, engendrant la tribu de VD et contenant les A_u , $u \in V_0$ et les bases de V_i , $i \in I_0$. La forme linéaire définit par ces intermédiaires une transition S_{VD}^Ω qui sauf sur un ensemble de λ mesure nulle satisfait aux conditions requises.

LEMME 8. — Soit ξ un plan d'expérimentation pour lequel D est compact et qui satisfait (H1), (H2), (H3), (H4), (H5), (H6), (H7).

Il existe alors une partie compacte Σ d'un espace vectoriel topologique et une application surjective σ de l'ensemble des stratégies sur Σ , telles que :

Pour toute fonction l mesurable de ΩVD dans $\mathbb{R}$, croissante de $\delta = \beta(v)$, et dont la restriction à $V_i D$ est limite croissante d'une suite de fonctions de $\mathcal{L}_{\mathcal{E}(V_i D)}^1(\Omega, \mathcal{A}, \lambda_i)$, l'application :

$$S_{VD}^\Omega \rightsquigarrow \lim_{i \in J} \int_{\Omega} \int_{V_i D} l dS_{V_i D} d\lambda_i,$$

factorise à travers σ , le second facteur étant linéaire et semi-continue inférieurement sur Σ .

Démonstration. — L'application τ_j restreinte à V_i (avec $u: j \rightarrow j$) est l'identité sur V_j . Elle définit une application de $L_{\mathcal{M}(V_i D)}^\infty$ (mesures images), et si g_i est continue sur $V_i D$, $\tau_j \circ g_i$ est continue sur $V_i D$. L'application ainsi induite de $L_{\mathcal{M}(V_i D)}^\infty$ dans $L_{\mathcal{M}(V_j D)}^\infty$ est donc continue pour les topologies faibles. Elle conserve la propriété ($\delta 2$). La propriété ($\delta' 1$) est conservée, sauf pour les $i' \in I$, buts de flèches u telles que $\tau_j^{-1}(u) \neq u$. On considérera donc dans $L_{\mathcal{M}(V_i D)}^\infty$ l'ensemble des éléments satisfaisants ($\delta 2$) et ($\delta 1$) sauf peut-être pour les i' définies comme ci-dessus. En se restreignant alors aux boules unités (ce qui est partout implicite) les τ_i définissent alors un système projectif de compacts. Soit Σ sa limite projective. Il est clair qu'à chaque S_{VD} satisfaisant (δ) on associe naturellement un élément de Σ .

Soit σ l'application linéaire ainsi définie. Il faut montrer qu'elle est surjective.

A un élément de Σ , correspond une suite de probabilités $S_{iV_iD}^\Omega$, $i \in I$, images les unes des autres par les applications τ_i . Cette correspondance est entre classes et donc presque sûre seulement. Ceci ne présente pas de difficultés dans le cas où V et D sont à la base dénombrable. Dans les autres cas on aura en utilisant un relèvement adapté, correspondance exacte entre les $S_{iV_iD}^\Omega$ (cf. remarque 1 ci-dessous).

Soit alors $i_0 \in J$, a un sous-ensemble mesurable de $V_{i_0}D$. Pour $i \geq i_0$, S_{ia}^ω est défini, et si $j > i$, $S_{ia}^\omega \leq S_{ia}^\omega = S_{i\tau_j^{-1}(a)}^\omega$ ($a \cap V_i$, et $\tau_j^{-1}(a) \cap V_i \supset a$).

On a donc une suite décroissante de nombres positifs, qui admet une limite par Vitali-Hahn-Saks (cf. [2]), celle-ci définit une mesure, positive, de masse ≤ 1 sur V_iD . Les propriétés de mesurabilité sont satisfaites sans restriction, puisque par (H6b), ($\delta 2$) est satisfaite, à partir d'un certain rang par $S_{ia}^\omega \subset V_iD$.

On construira donc ainsi une suite de mesures $S_{iV_iD}^\Omega$ restrictions, les unes des autres. Nous les considérons sur les tribus de Baire. Pour tout élément de la tribu de Baire de VD nous définissons $\bar{S}^\Omega$ comme limite des $\bar{S}_i^\Omega$ pour les intersections avec les V_iD . On définit un contenu additif, approché par les compacts puisque approché par les V_iD et dans ceux-ci par les compacts. Nous définissons bien ainsi une mesure pour chaque x (la trace de la tribu de Baire et un localement compact sur un compact qu'il contient est la tribu de Baire). La mesurabilité en ω résulte d'un calcul classique.

Ceci montre la surjectivité de σ (que toute transivité de Baire se prolonge en une transition de Borel résulte de II, lemme 8, en utilisant un relèvement adapté). Il nous reste à voir la factorisation.

Désignons par l_i la restriction de l à $\Theta\Omega V_iD$. On a :

$$\int l_i d\bar{S}_i \leq \int l_i dS_i = \int (l_i \circ \tau_i) dS \leq \int l dS.$$

et en passant à la limite les deux extrémités sont égales.

Or $\int_X \int_{V_iD} l_i dS_{iV_iD}^\Omega d\lambda_i$ définit une forme linéaire semi-continue inférieurement. Ceci demeure par passage à la limite croissante.

Remarque. — Nous avons indiqué l'utilisation d'un relèvement adapté. Or il n'existe pas nécessairement de mesure λ sur X prolongeant les λ_i . Pour ceci on définira

$$L_\infty(\Omega, \mathcal{A}_i, \lambda_i, i \in I) = \bigcup_{i \in J} L_\infty(\Omega, \mathcal{A}_i, \lambda_i)$$

en considérant, par un abus de notation évident,

$$L_{\infty}(\Omega, \mathcal{A}_i, \lambda_i) \subset L_{\infty}(\Omega, \mathcal{A}_j, \lambda_j)$$

pour $i \leq j$ (en fait on considère la limite inductive).

Il est alors facile de construire comme au 3 un relèvement adapté de $L_{\infty}(\Omega, \mathcal{A}_i, \lambda_i, i \in I)$ dans $\mathcal{L}_{\infty}(\Omega, \mathcal{A})$, et sous les mêmes hypothèses. C'est celui que nous utiliserons ci-dessus.

DÉMONSTRATION DU THÉORÈME 1. — Elle est immédiate en prenant le compact fini d'Alexandroff de D . Les hypothèses sur l évitent que des stratégies ayant avec une probabilité non nulle des masses à l'infini de V ou D , soient optimales.

5. EXISTENCE ET CARACTÉRISATION DES μ -STRATÉGIES

L'existence de μ -stratégies est un corollaire du théorème 1 :

THÉORÈME 2. — Soit ξ un plan d'expérimentation, μ une probabilité sur Θ . Soit ν la mesure sur $\Theta\Omega$ composée de P_{Ω}^{Θ} et de μ . Et supposons que ξ satisfasse les hypothèses du théorème 1 en remplaçant Ω par $\Theta\Omega$, Ω_i par $\Theta\Omega_i$, et (λ) et P_{Ω}^{Θ} par ν .

Alors il existe une μ -stratégie ouverte.

Nous consacrerons l'essentiel de ce paragraphe à l'étude des μ -stratégies dans le cas où I est l'ensemble ordonné N .

5.1. Résultat technique préalable.

PROPOSITION 2. — Soit D un espace compact, $(\Omega, \mathcal{A}, P)$ un espace probabilisé complet, $\mathcal{B}$ une sous-tribu de $\mathcal{A}$, possédant la propriété de relèvement, l limite d'une suite croissante de fonctions bornées positives de $\mathcal{L}_{\mathcal{C}(D)}^1(\Omega, \mathcal{A}, P)$, S_D^{Ω} une transition $\mathcal{B}$ mesurable. Alors :

1) $E^{\mathcal{B}} f$ peut être choisie pour chaque ω , semi-continue inférieurement en d .

$$2) \int_D (E^{\mathcal{B}} f)(\omega, d) dS_D^{\omega} = E^{\mathcal{B}} \left[\int_D f(\omega, d) dS_D^{\omega} \right] \quad [P]$$

3) Si D est métrisable, les hypothèses de complétion et de relèvement sont inutiles.

Démonstration. — Il suffit évidemment d'opérer avec une fonction bornée à valeurs continues.

1) Pour tout $d > 0$, il existe un ensemble mesurable Ω_α de Ω tel que $P(\Omega_\alpha) \geq 1 - \alpha$, et une suite f_n de fonctions étagées, indépendantes de α , convergent uniformément, en norme de $\mathcal{C}(D)$ vers f , sur Ω_α .

Soit h un relèvement de Ω , prolongeant un relèvement de $\mathcal{B}$, et considérons

$$(E^{\mathcal{B}} f_n)(\omega, d) \cdot f_n(\omega, d) = \sum_{m=1}^k 1_{A_{n,m}}(\omega) g_{n,m}(d),$$

où les fonctions $g_{n,m}$ soient continues. Alors

$$(E^{\mathcal{B}} f_n)(\omega, d) = \sum_{m=1}^k E^{\mathcal{B}}(1_{A_{n,m}})(\omega) g_{n,m}(d) \quad (\text{p. s.})$$

et

$$h[(E^{\mathcal{B}} 1_{\Omega_\alpha} f_n)(\cdot, d)] = \sum_{m=1}^k h[E^{\mathcal{B}} 1_{\Omega_\alpha} 1_{A_{n,m}}] g_{n,m}(d).$$

L'opération $E^{\mathcal{B}}$ étant contractante et h une isométrie, pour $n \geq N(\varepsilon)$

$$(\forall d \in D)(\forall \omega \in \Omega_\alpha)(|f(\omega, d) - f_n(\omega, d)| \leq \varepsilon)$$

soit

$$(\forall d \in D)(\|E^{\mathcal{B}}(1_{\Omega_\alpha} f)(\cdot, d) - E^{\mathcal{B}}(1_{\Omega_\alpha} f_n)(\cdot, d)\|_{L_\infty} \leq \varepsilon).$$

soit

$$(\forall d \in D)(\|h(E^{\mathcal{B}}(1_{\Omega_\alpha} f)(\cdot, d)) - h(E^{\mathcal{B}}(1_{\Omega_\alpha} f_n)(\cdot, d))\|_{L_\infty} \leq \varepsilon)$$

et par suite, quel que soit $\omega \in \Omega_\alpha$, $h(E^{\mathcal{B}} 1_{\Omega_\alpha} f)(\omega, \cdot)$ est continue sur D .

Soit maintenant α_m suite décroissante vers 0, et Ω_{α_m} suite croissante

$$\begin{aligned} P(\Omega_m \Omega_\alpha) &= 1, & \Omega_\infty &= \lim \Omega_{\alpha_m} \\ 1_{\Omega_{\alpha_m}} f &\uparrow 1_{\Omega_\infty} f. \\ E^{\mathcal{B}} 1_{\Omega_{\alpha_m}} f &\uparrow E^{\mathcal{B}} 1_{\Omega_\infty} f \end{aligned}$$

presque sûrement pour chaque d .

Par suite $h[E^{\mathcal{B}} 1_{\Omega_{\alpha_m}} f](\cdot, d)$ est une suite croissante vers un élément de $(E^{\mathcal{B}} 1_{\Omega_\infty} f)(\cdot, d) = (E^{\mathcal{B}} f)(\cdot, d)$.

D'où dans $(E^{\mathcal{B}} f)$ un élément semi-continu inférieurement pour chaque ω .

2) Pour f étagée, $f = \sum_{i=1}^k 1_{A_i} g_i$, on a pour $A \in \mathcal{B}$:

$$I = \int_A \int_D \sum_{i=1}^k 1_{A_i}(\omega) g_i(d) dS_D^\omega dP(\omega) = \sum_{i=1}^k \int_A 1_{A_i}(\omega) \int_D g_i(d) dS_D^\omega dP(\omega).$$

La seconde intégrale est une fonction $\mathcal{B}$ mesurable, donc :

$$I = \sum_{i=1}^k \int_A E^{\mathcal{B}} 1_{A_i} \int_D g_i(d) dS_D^\omega dP(\omega) = \int_A \int_D (E^{\mathcal{B}} f)(\omega, d) dS_D^\omega dP(\omega).$$

On a ensuite pour les f_n étagées et $n \geq N(\varepsilon)$

$$(\forall d \in D) \quad \|(E^{\mathcal{B}} 1_{\Omega_x} f)(\cdot, d) - (E^{\mathcal{B}} 1_{\Omega_x} f_n)(\cdot, d)\|_{L_x} \leq \varepsilon.$$

et

$$\|(1_{\Omega_x} f)(\cdot, d) - (1_{\Omega_x} f_n)(\cdot, d)\|_{\mathcal{L}_x} \leq \varepsilon.$$

ce qui étend la propriété $1_{\Omega_x} f$. Par passage aux limites croissantes le résultat est vrai pour f .

3) Seul le point (1) est à redémontrer.

f mesurable sur $\Omega \times D$ et continue sur D est encore limite uniforme d'une suite de fonctions mesurables de Ω dans $\mathcal{C}(D)$. Alors pour tout $d \in D$ et tout $\omega \in \Omega$, $|f_n(\omega, d) - f(\omega, d)| \leq \varepsilon$, pour n assez grand. D_0 étant un ensemble dénombrable dense dans D , il existe un ensemble de mesure nulle M , tel que pour tout $d \in D_0$ et pour tout ω de $\Omega \setminus M$, pour une version donnée de l'espérance :

$$|E^{\mathcal{B}} f_n - E^{\mathcal{B}} f| \leq \varepsilon.$$

Donc, pour tout $\omega \in \Omega$, $E^{\mathcal{B}} f$ est limite uniforme de fonctions uniformément continues sur D_0 , et donc uniformément continue sur D_0 . Par convergence dominée (f bornée), son prolongement à D est encore dans $E^{\mathcal{B}} f$ et est continue pour chaque ω .

5.2. Caractérisation dans le cas d'une seule expérience.

THÉORÈME 3. — Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, μ une probabilité sur Θ , $\nu_{\Theta\Omega}$ la mesure composée de P_Ω^Θ et μ , D la mesure marginale sur Ω . Soit D espace de décision localement compact, l fonction de coût limite d'une suite croissante de fonctions bornées positives de $\mathcal{L}_{\mathcal{B}(D)}^1(\Theta\Omega, \mathcal{A} \oplus \mathcal{H}, \nu_{\Theta\Omega})$, admettant une limite à l'infini de D , l tendant vers $+\infty$ à l'infini de D .

Supposons enfin D à base dénombrable où $\mathcal{A}$ complète pour D . Alors :

Pour qu'une transition S_D^Ω soit une μ -stratégie, il faut et il suffit que η -presque partout, S_D^ω donne la mesure 1 à toute partie mesurable contenant les points où $(E_\eta l)(\omega, \cdot)$ est minimum.

Ce résultat est un corollaire immédiat de la proposition 2.

On remarquera qu'il ne permet pas de définir de manière constructive une μ -stratégie. Toutefois nous noterons certains cas particuliers, où l'on peut en déduire l'existence d'une μ -stratégie certaine (fonction de Ω dans D) (cf. [10]).

1. D est à base dénombrable (en utilisant le résultat de [1] selon lequel toute multiplication mesurable à valeurs dans les compacts d'un compact métrisable possède une section mesurable).

2. D est produit de compacts métrisables, ou D est compact convexe d'un espace $\mathbb{R}^T$, la fonction l étant convexe.

5.3. Caractérisation des μ -stratégies dans le cas séquentiel totalement ordonné.

I est l'ensemble ordonné des entiers. Le résultat est dû à Wald et Wolfowitz ([12], [13]) dans le cas particulier d'indépendance.

Nous opérerons la construction en deux étapes.

1) CONSTRUCTION DE LA DÉCISION $d \in D$ CONDITIONNELLEMENT A L'ARRÊT.

Nous devons rendre minimum :

$$r(S) = \int_{\Theta\Omega} \sum_{n=0}^{+\infty} \int_D l(z, n, d) dS_D^{\omega, n} S_n^\omega dv$$

où $Z = \Theta\Omega$, $S_D^{\Omega n}$ est la probabilité conditionnelle si l'on arrête l'opération, S_n^Ω la probabilité de s'arrêter à l'instant n .

On peut écrire :

$$\begin{aligned} r(S) &= \sum_{n=0}^{+\infty} \int_{\Theta\Omega} \int_D l(z, n, d) dS_D^{\omega, n} S_n^\omega dv = \sum_{n=0}^{+\infty} \int_{\Theta\Omega} E^{\mathcal{A}_n} \left[\int_D l(z, n, d) dS_D^{\omega, n} \right] S_n^\omega dv \\ &= \sum_{n=0}^{+\infty} \int_{\Theta\Omega} \left[\int_D (E^{\mathcal{A}_n} l)(z, n, d) dS_D^{\omega, n} \right] S_n^\omega dv. \end{aligned}$$

D'après la définition des stratégies et la proposition 2

$$r(S) = \int_{\Theta\Omega} \sum_{n=0}^{+\infty} \left[\int_D (E^{\overline{\mathcal{A}}_n} l)(z, n, d) dS_D^{\omega, n} \right] S_n^{\omega} dv.$$

On aura une solution optimale en choisissant :

d'une part, pour tout n , $S_D^{\Omega, n}$, comme au théorème 2,
d'autre part, en posant

$$\lambda(\omega, n) = \inf_{d \in D} (E^{\overline{\mathcal{A}}_n} l)(z, n, d)$$

et en rendant minimum

$$\int_{\Theta\Omega} \sum_{n=0}^{+\infty} \lambda(\omega, n) S_n^{\omega} dv.$$

$\{S_n\}_{n \in \mathbb{N}}$ parcourant l'ensemble des suites adaptées de variables positives, pour lesquelles,

$$\sum_{n=0}^{+\infty} S_n^{\Omega} = 1.$$

On notera que, la fonction $\lambda(\cdot, n)$ est $\overline{\mathcal{A}}_n$ mesurable, si la tribu $\mathcal{A}_n$ est complète ou si D est métrisable, la fonction $E^{\overline{\mathcal{A}}_n} l$ pouvant être choisie continue en d (proposition 2).

2) CARACTÉRISATION DE LA DÉCISION D'ARRÊT.

Soit

$$\Lambda_n \in \text{Ess} \inf_T E^{\overline{\mathcal{A}}_n} \sum_{m=n}^{+\infty} \lambda(\omega, m) T_m$$

où $T = \{T_m\}_{m \geq n}$ est une suite de variables positives, adaptée à $\overline{\mathcal{A}}_n$, telle que

$$\sum_{m=n}^{+\infty} T_m = 1.$$

Et considérons

$$T(\omega) = \inf \{ n \mid \lambda(\omega, n) = \Lambda_n(\omega) \}$$

Dans les hypothèses où $\lambda(., n)$ est $\mathcal{A}_n$ mesurable, T définit un temps d'arrêt. Montrons qu'il minimise

$$\int_{\Theta \Omega} \sum_{n=0}^{+\infty} \lambda(\omega, n) S_n^\omega dv.$$

Soit $\{S_n\}_{n \geq 0}$ une suite de variables positives, adaptée à $(\mathcal{A}_n)$, et telle que

$$\sum_{n=0}^{+\infty} S_n = 1.$$

Posons $\Omega_{m,n} = \{\omega \mid (T(\omega) = m) \wedge (S_n^\omega \neq 0)\}$ et supposons $P(\Omega_{m,n}) \neq 0$.

(a) $n < m$. Alors $\Omega_{m,n}$ est $\mathcal{A}_m$ mesurable.

Par définition de T on peut trouver $(U_k)_{k \geq n}$ suite adaptée de variables positives telle que

$$\sum_{k=n}^{+\infty} U_k = 1, \lambda(\omega, n) > \lambda(\omega, n) U_n + E^{\mathcal{A}_n} \left(\sum_{k=n+1}^{+\infty} \lambda(., k) \right) (\omega)$$

sur une partie A de $\Omega_{m,n}$ de mesure non nulle. Il est clair que A est $\mathcal{A}_n$ mesurable et que l'on peut choisir $U_n = 0$.

Posons alors

$$S'_k(\omega) = S_k(\omega) \quad \text{si} \quad k < n \quad \text{ou} \quad \omega \notin A.$$

$$S'_n(\omega) = 0 \quad \text{si} \quad \omega \in A.$$

$$S'_n(\omega) = S_n(\omega) \quad \text{si} \quad \omega \notin A.$$

$$S'_k(\omega) = S_k(\omega) + U_k(\omega) \quad \text{si} \quad \omega \in A \quad \text{et} \quad k > n.$$

La stratégie définie par $\{S'_n\}_{n \in \mathbb{N}}$ est strictement meilleure que $\{S_n\}_{n \in \mathbb{N}}$.

(b) $m < n$. Soit $\Omega_m = \{\omega \mid T(\omega) = m\}$. Posons

$$U_k(\omega) = \frac{S_k(\omega)}{1 - [S_1(\omega) + S_2(\omega) + \dots + S_{m-1}(\omega)]}$$

pour $k \geq n$. Et soit

$$B = \left\{ \omega \mid \lambda(\omega, m) \leq E^{\mathcal{A}_m} \left[\sum_{k=m}^{+\infty} \lambda(., k) U_k \right] (\omega) \right\}.$$

B contient presque sûrement Ω_m et est de mesure non nulle.

Choisissons

$$\begin{aligned} S'_k(\omega) &= S_k(\omega) & \text{si } \omega \notin B & \quad \text{ou si } k < m \\ S'_m(\omega) &= \sum_{k=m}^{+\infty} (., S_k(\omega)) & \text{si } \omega \in B \\ S'_k(\omega) &= 0 & \text{si } k > m \quad \text{et } \omega \in B \end{aligned}$$

$\{S'_k\}_{k \in \mathbb{N}}$ définit une stratégie meilleure que S (ou équivalente).

Il résulte de là que le temps d'arrêt construit ci-dessus est optimum. Nous pouvons alors énoncer le théorème suivant :

THÉORÈME 4. — Soit $(\Omega, \mathcal{A})$ un espace mesurable, P_Ω^Θ une transition, μ une probabilité sur Θ , ν composée sur $Z = \Theta\Omega$ de μ et de P_Ω^Θ . Soit $(\mathcal{A}_n)_{n \in \mathbb{N}}$ une suite croissante de sous-tribus de $\mathcal{A}$. Soit D un espace localement compact et l limite d'une suite croissante de fonctions positives linéaires de $\mathcal{L}_{\mathcal{B}(D)}^1(Z, \mathcal{L}, \nu)$ admettant une limite à l'infini de D , l tendant vers $+\infty$ à l'infini de D quel que soit $z \in Z$.

Supposons D à base dénombrable ou $\mathcal{A}_n$ α -complète de type produit. On obtient une μ -stratégie selon le procédé suivant :

(a) On arrête l'expérience selon :

$$T(\omega) = \inf \left\{ n \mid \inf_d E^{\bar{\mathcal{A}}_n} l(., d)(\omega) \leq \left(\text{Ess Inf}_S \left(E^{\bar{\mathcal{A}}_n} \int_D l(z, n, d) dS_D^z \right)(\omega) \right) \right\}$$

où S parcourt les transitions à support dans $[n, +\infty)D$.

(b) On choisit une décision $d \in D$, conditionnellement à $T = n$, selon la procédure du théorème 3.

Remarque. — On peut déduire de ce théorème (cf. [7]), le résultat suivant concernant les tests d'hypothèses séquentiels.

Soit $(\Omega, \mathcal{A}, P_i)_{i=1,2}$ un espace probabilisé. Soit $\{X_n\}_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes quel que soit $i = 1, 2$, α_1 et α_2 deux nombres strictement composés entre 0 et 1. Parmi toutes les stratégies séquentielles des risques inférieurs à α_1 et α_2 , il existe une stratégie minimisant l'espérance du nombre d'expériences de la forme suivante :

Il existe deux nombres π_1 et π_2 tel que

L'expérimentation s'arrête au premier instant où

$$\rho = \frac{dP_2}{d(P_1 + P_2)} \bigg/ \frac{dP_1}{d(P_1 + P_2)} \notin]\pi_1, \pi_2[$$

On décide alors P_1 , si $\rho \leq \pi_1$, P_2 si $\rho \geq \pi_2$.

BIBLIOGRAPHIE

- [1] CASTAING, Multi-applications mesurables. Thèse Caen, 1967.
- [2] DUNFORD-SCHWARTZ, *Linear Operators I*. Interscience Publishers, 1953.
- [3] IONESCU-TULCEA, On the lifting property I. *J. Math. Anal. Appl.*, 1961.
- [4] IONESCU-TULCEA, Desintegration of measures. *Annales de l'Institut Fourier*, Grenoble, 1964.
- [5] LE CAM, An extension of Wald's theory of Statistical Decision functions. *A. M. S.*, 1955.
- [6] LE CAM, Approximate Sufficiency. *A. M. S.*, 1964.
- [7] LEHMAN, *Testing Statistical Hypothesis*. J. Wiley, 1959.
- [8] MEYER, *Probabilités et Potentiel*. Hermann, 1966.
- [9] NEVEU, *Bases mathématiques du calcul des probabilités*. Masson, 1964.
- [10] ROMIER, Institut de Statistique de l'Université de Paris (Thèse à paraître).
- [11] L. SCHWARTZ, Produits tensoriels d'E. V. T. Séminaire I. H. P., Paris, 1955.
- [12] WALD-WOLFOWITZ, Bayes solutions of sequential decision functions. *A. M. S.*, 1950.
- [13] WALD, *Statistical decision functions*. J. Wiley, 1950.
- [14] WOLFOWITZ-KIEFER-DWORETZKY, Sequential decision problems for processes with continuous time parameter. *A. M. S.*, 1953.

Manuscrit reçu le 11 juin 1969.