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

G. ROMIER

Introduction à la statistique mathématique. I. Modèle d'expérimentation statistique

Annales de l'I. H. P., section B, tome 5, nº 4 (1969), p. 275-288
<http://www.numdam.org/item?id=AIHPB_1969__5_4_275_0>

Ann. Inst. Henri Poincaré, Section B :
Vol. V, n° 4, 1969, p. 275-288. *Calcul des Probabilités et Statistique.*

Introduction à la statistique mathématique.

I. Modèle d'expérimentation statistique

par

G. ROMIER
Maître de Conférences I. U. T. de Grenoble.

I. LES MODÈLES

Préliminaire. Le langage.

1° Dans la suite nous nous sommes efforcés de donner une écriture assez générale des modèles statistiques, dans la direction des études « séquentielles » largement développées par Wald, en acceptant aussi bien le choix entre diverses possibilités de prolonger l'expérimentation que la possibilité d'expérimentation continue.

Pour décrire les « états d'expérimentation » nous avons du choisir un langage adapté. Une première possibilité, souvent adoptée consistait à retenir le langage des graphes. Il nous est apparu en définitive que celui des catégories mettaient mieux en lumière les propriétés du modèle. C'est celui que nous avons retenu. S'il a parfois guidé l'étude de propriétés, il restera dans tout ce travail un langage, sans que les résultats de la théorie soient explicitement utilisés. Nous pensons toutefois que sa raison d'être est plus profonde et éclairera certains domaines comme l'exhaustivité, la comparaison d'expérimentation, etc. (cf. [*2*]).

2° $(\Theta, \mathscr{H})$, $(\Omega, \mathscr{A})$ étant deux espaces mesurables, p_Ω^Θ désigne, sauf indications contraires, une probabilité de transition de Θ vers Ω.

p_Ω^θ est la probabilité sur $\mathscr{A}$ pour la valeur θ de Θ.

$p_{\underline{A}}^\theta$ est la valeur de cette probabilité pour la partie mesurable A de

p_A^θ est la restriction de la probabilité p_Ω^θ à $A \cap \mathscr{A}$ $(A \in \mathscr{A})$.

1. GÉNÉRALITÉS SUR L'EXPÉRIMENTATION STATISTIQUE

1.1. Notations.

Nous introduirons une famille de catégories:

$\mathcal{M}$ CATÉGORIE dont

les objets sont les espaces mesurables

les morphismes sont identifiés aux applications mesurables, les unités étant les applications identiques, la composition celle des applications.

$\mathcal{C}$ CATÉGORIE dont

les objets sont les triplets (Θ, X, p_x^Θ) où Θ désigne un espace mesurable, X un espace mesurable p_x^Θ une transition (probabilité).

Les tribus $\mathcal{T}$ de Θ et $\mathcal{X}$ de X sont sous-entendus.

les morphismes $M : (\Theta, X, p_x^\Theta) \to (H, Y, P_Y^H)$ sont identifiés aux couples (v, u) où:

v est une application mesurable surjective de Θ sur H

u est une application mesurable de X dans Y

$$(\forall \theta \in \Theta)(u(P_X^\theta) = P_Y^{v(\theta)}).$$

Les unités sont les couples d'applications identiques. On pose

$$(v_1, u_1) \circ (v_2, u_2) = (v_1 \circ v_2, u_1 \circ u_2).$$

Si Θ est un espace mesurable donné, on pose:

$\mathcal{C}_\Theta$ SOUS-CATÉGORIE de $\mathcal{C}$ dont

les objets sont les triplets (Θ, X, p_X^Θ), où Θ est l'espace donné

les morphismes sont les morphismes de $\mathcal{C}'(v, u)$, tels que v soit l'application identique de Θ sur lui-même.

Nous élargirons ces définitions en conservant les mêmes objets mais en acceptant pour morphismes des transitions. Soit:

$\overline{\mathcal{C}}$ CATÉGORIE dont les objets sont les triplets (Θ, X, P_X^Θ), les morphismes $[(\Theta, X, P_X^\Theta) \to (H, Y, P_Y^\Theta)]$ étant identifiés aux couples (v, u_Y^X), v application surjective mesurable de Θ dans H, $u_{\mathrm{Y}}^{\mathrm{X}}$ transition de X vers Y telle que:

$$(\forall \theta \in \Theta)(p_{\mathrm{Y}}^{v(\theta)} = u_{\mathrm{Y}}^{\mathrm{X}} \circ p_{\mathrm{X}}^{\theta})$$
$$(v_1, u_1{}_{\mathrm{Z}}^{\mathrm{Y}}) \circ (v_2, u_2{}_{\mathrm{Y}}^{\mathrm{X}}) = (v_1 \circ v_2, u_1{}_{\mathrm{Z}}^{\mathrm{Y}} \circ u_2{}_{\mathrm{Y}}^{\mathrm{X}}).$$

$\overline{\mathscr{C}}_{\Theta}$ est la sous-catégorie obtenue, comme précédemment en fixant l'espace Θ.

Nous utiliserons enfin les définitions générales suivantes:

* Si J est une catégorie, FlJ désignant l'ensemble des morphismes, ObJ l'ensemble des objets de J, Hom (i, j), où $i \in \mathrm{O}b\mathrm{J}$, $j \in \mathrm{O}b\mathrm{J}$, l'ensemble des morphismes de i vers j, nous noterons:

α l'application source de FlJ dans ObJ telle que:

$$(\forall (i, j) \in (\mathrm{O}b\mathrm{J})^2)(\forall u \in \mathrm{Hom}\ (i, j))(\alpha(u) = i)$$

β l'application but de FlJ dans ObJ telle que:

$$(\forall (i, j) \in (\mathrm{O}b\mathrm{J})^2)(\forall u \in \mathrm{Hom}\ (i, j))(\beta(u) = j).$$

* J sera dite ordonnée par factorisation si la relation:

$$(i \leqq j) \Leftrightarrow (i \in \mathrm{O}b\mathrm{J})(j \in \mathrm{O}b\mathrm{J})(\exists u \in \mathrm{F}l\mathrm{J})(i = \alpha(u))(j = \beta(u))$$

est une relation d'ordre partiel sur ObJ.

On notera que dans ces conditions toute partie V de FlJ est partiellement ordonnée par la relation:

$$(u \leqq v) \Leftrightarrow (\exists w \in \mathrm{F}l\mathrm{J})(v = w_0 u).$$

Ces conditions interdisent les « circuits » dans FlJ. Si ObJ est alors un ensemble ordonné, la catégorie J n'est pas nécessairement celle qui lui est, classiquement, associée. Nous en verrons un exemple sur l'ensemble des parties d'un ensemble, ordonné par inclusion.

* F désignera le foncteur « ensemble sous-jacent » de $\mathscr{M}$ dans la catégorie des ensembles.

γ) désignera le foncteur de $\mathscr{C}$ sur $\mathscr{M}$ qui associe à tout $(\Theta, \mathrm{X}, \mathrm{P}_{\mathrm{X}}^{\Theta})$ l'ensemble mesurable X, et à tout (v, u), l'application mesurable u.

1.2. Expérimentation statistique.

La définition d'un problème statistique comprend:

La description des expérimentations possibles (structure expérimentale).

Les éléments de résolution du problème (décisions, coûts, critères, etc.).

Nous considérons successivement.

La structure expérimentale.

Des indications sur l'étude « intrinsèque » de cette structure.

Des exemples de problèmes, qui seront développés aux paragraphes suivants.

DÉFINITION 1. — *La structure expérimentale* d'un problème statistique est définie par la donnée d'une sous-catégorie $\mathscr{E}$ de $\mathscr{C}$ satisfaisant aux deux conditions

$$(\mathscr{E}1) \qquad (\forall u \in Fl\mathscr{E})[(\alpha(u) \in \mathscr{E}) \Rightarrow (u \in Fl\mathscr{E}) \wedge (\beta(u) \in Ob\mathscr{E})]$$

($\mathscr{E}2$). Soit λ un foncteur d'une catégorie I dans $\mathscr{E}$ et soit

$$(X, \xi_i)_{i \in ObJ} = \lim_{\leftarrow \mathscr{M}} \gamma_0 \lambda.$$

S'il existe dans $\mathscr{E}$ un système projectif $(A, \xi_i)_{i \in ObJ}$ associé à λ dont l'image par γ soit (X, ξ_i), alors $(A, \xi_i) = \lim\limits_{\leftarrow \xi} \lambda$.

Par abus de langage nous dirons que $\mathscr{E}$ est la (catégorie) structure expérimentale.

La donnée de $\mathscr{E}$ décrit les expérimentations possibles et les relations fonctionnelles qui les relient. La donnée d'un morphisme (v, u) :

$$[(\Theta, X, P_X^\Theta) \rightarrow (H, Y, P_Y^\Theta)]$$

indique que l'expérimentation décrite par l'espace des observations X et la famille P_X^Θ prolonge, en apportant plus de renseignements, celle décrite par Y et P_Y^H. Il en est bien ainsi puisque si $x \in X$ est « réalisé », alors $u(x) \in Y$ est aussi réalisé.

La condition ($\mathscr{E}1$) indique, que, si nous sommes capables d'observer $x \in X$, pour tout u défini sur (X, P_X^Θ) nous sommes capables d'observer par le fait même $u(x)$ défini sur $(u(X), u(P_X^\Theta))$.

La condition ($\mathscr{E}2$) signifie de son côté que si l'on considère une famille d'expérimentation (X_i) et s'il existe une « plus petite » expérimentation possible, les prolongeant, elle est unique. Ainsi, si $(X_1, P_{X_1}^\Theta)$ décrit une expérimentation, $(X_2, P_{X_2}^\Theta)$ une autre, et si nous possédons, parmi les expérimentations possibles, une troisième $(X_1X_2, P_{X_1X_2}^\Theta)$, alors cette dernière, prolonge X_1 et X_2 et est unique.

Ces deux propriétés sont nécessaires pour la cohérence du système, mais nous ne les utiliserons pas pratiquement.

Nous complèterons ces définitions par deux données, qui n'interviendront pas dans la suite.

DÉFINITION 2. — $\overline{\mathscr{E}}$ est la plus petite sous-catégorie pleine de $\overline{\mathscr{C}}$ contenant $\mathscr{E}$ ([1]).

Cette définition est justifiée par le fait que $\mathscr{E} \subset \mathscr{C} \subset \overline{\mathscr{C}}$.

La catégorie $\overline{\mathscr{E}}$ contient toute « l'information » que l'on peut tirer de la structure expérimentale. En introduisant les espaces qui sont liés aux expérimentations par des transitions, elle décrit les « non-observables », sur lesquels l'expérimentation apporte des renseignements, dans le cadre du modèle probabiliste et statistique.

DÉFINITION 3. — Le problème sera dit *paramétrique* s'il existe un espace mesurable Θ fixe, tel que $\mathscr{E}$ soit une sous-catégorie de $\mathscr{C}_\Theta$. Sinon il sera dit *non paramétrique.*

Remarques.

1. Il ne faut pas confondre problème non paramétrique, au sens de la définition ci-dessus, avec *méthodes de traitement statistique non paramétriques.*

Ceci nous paraît éclairer le faux débat sur la « taille » de Θ à partir de laquelle un problème serait non paramétrique. Dans ce débat la catégorie $\mathscr{E}$ contient en général un « plus grand » objet (! c'est-à-dire que l'on ne dispose que d'une expérimentation possible et de ses conséquences). Le problème est alors, par nature paramétrique.

Ceci dit les méthodes de traitement faisant jouer un rôle important à l'espace des paramètres, sont souvent limitées au cas où $\Theta = \mathbb{R}^n$. Dans ce débat la « taille » de Θ importe en fin de compte moins que les renseignements fournis par l'expérience (Pour ce débat, cf. [*4*]).

2. La condition pour qu'un problème soit paramétrique est la suivante : considérons le foncteur de $\mathscr{E}$ dans $\mathscr{M}$: $P_X^\Theta \to \Theta$ et $(v, u) \to v$. S'il admet une limite projective avec applications canoniques *surjectives*, le système est paramétrique. On trouvera dans [5] des conditions suffisantes pour qu'il en soit ainsi.

1.3. **Indications sur l'étude intrinsèque.**

Nous signalerons deux directions sans les aborder plus avant :

a) *Équivalences naturelles* (Isomorphismes au sens de la catégorie). On peut s'interroger sur les isomorphismes des diverses catégories en jeu :

([1]) Une sous-catégorie pleine J_0 de J est une sous-catégorie telle que si $i \in J_0$, $j \in J_0$, et si $u \in FlJ$ avec $\alpha(u) = i$, $\beta(u) = j$, alors $u \in FlJ_0$.

*a*1) Dans $\mathscr{E}$ l'étude est celle des bijections bimesurables et de ce fait peu intéressante.

*a*2) Dans $\overline{\mathscr{E}}$ on est conduit à l'étude des transitions inversibles, c'est-à-dire des transitions Q_Y^X telles qu'il existe R_X^Y avec

$$Q_Y^X \circ R_X^Y = \delta_Y^Y \qquad \text{et} \qquad R_X^Y \circ Q_Y^X = \delta_X^X$$

(applications identiques).

*a*3) Une autre équivalence est introduite dans [*2*].

Nous nous plaçons dans le cas d'un problème paramétrique et dans $\overline{\mathscr{E}}$. Deux objets sont dits équivalents, $P_X^\Theta \sim P_Y^\Theta$ s'il existe Q_Y^X et R_X^Y telles que

$$Q_Y^X P_X^\Theta = P_Y^\Theta \qquad \text{et} \qquad R_X^Y P_Y^\Theta = P_X^\Theta.$$

Ceci revient à considérer comme équivalents deux morphismes de Hom (i, j) et à substituer à $\overline{\mathscr{E}}$ la catégorie ayant les mêmes objets et pour morphismes les classes d'équivalences.

b) *Équivalences logiques.*

Parmi les diverses équivalences généralement acceptées nous en retiendrons une :

Soit dans le cadre paramétrique :

$$u : [(\Theta, X, P_X^\Theta) \to (\Theta, Y, P_Y^\Theta)].$$

DÉFINITION 4. — On dira que *u est exhaustif* et que (Θ, Y, P_Y^Θ) est un *objet exhaustif* pour (Θ, X, P_X^Θ) si $\mathscr{X}$ désignant la tribu de X, $\mathscr{Y}$ celle de Y :

$$u^{-1}(\mathscr{Y}) \qquad \text{est exhaustive pour} \qquad (X, \mathscr{X}, P_X^\Theta) \ (^1)$$

On introduit la relation d'équivalence fermeture transitive de la relation « posséder un objet exhaustif commun ».

On peut montrer :

* Si les espaces X sont finis, cette relation d'équivalence est identique à la relation (*a*3) définie ci-dessus.

* Si pour tout $(\Theta, X, P_X^\Theta) \in Ob\mathscr{E}$, il existe une probabilité λ sur (X, x), telle que toute P_X^Θ soit absolument continue par rapport à λ alors :

La relation « avoir un objet exhaustif commun » est transitive (cf. ci-dessous chap. II).

Nous ignorons si les résultats s'étendent (une définition plus large de l'exhaustivité semble nécessaire cf. [*3*]).

(1) Rappelons que la sous-tribu $\mathscr{B}$ de $\mathscr{X}$ est dite exhaustive pour $(X, \mathscr{X}, P_X^\Theta)$, si, quelle que soit f, $\mathscr{X}$ mesurable, réelle, bornée, il existe g appartenant à $E_\Theta^{\mathscr{B}} f$ quel que soit θ. On a montré depuis [*6*] que la relation d'équivalence par exhaustivité est identique à la relation (*a* 3) si les familles P_X^θ sont dominées et les espaces X localement compacts.

2. MODÈLE DE LA DÉCISION STATISTIQUE

2.1. Définitions et notations générales.

DÉFINITION 5. — Un plan d'expérimentation statistique est la donnée

* D'un foncteur ξ d'une catégorie I dans la structure expérimentale $\mathscr{E}$.
* D'une transition P_X^Θ, où X est l'espace associé à

$$\lim_{\leftarrow \mathscr{M}} \gamma_0 \xi = (X, \zeta_i)_{i \in I},$$

telle que pour tout $i \in I$

$$\xi(i) = (\Theta, X_i, \zeta_i(P_X^\Theta)).$$

La catégorie I sera appelée catégorie des « états d'expérimentation ». Les éléments de I décrivent le stade ou est parvenu l'expérimentation. On admettra toujours qu'elle contient un « premier élément » O, c'est-à-dire un élément tel que:

$$\mathrm{Hom}\,(0, i) \neq \varnothing, \qquad \forall i \in I.$$

On supposera en outre qu'elle est ordonnée par la factorisation ($\leqq$).

Nous utiliserons les notations suivantes:

F est l'ensemble des parties maximales de I totalement ordonnée par factorisation.

$V = \mathrm{Hom}\,(0, I)$

$V_i = \{ u \in V \mid (\exists\, v \in F/I)(\beta(v_0 u) = i) \}$ $\qquad (i \in I)$

$A_u = \{ v \in V \mid (\exists w \in F/I) v = w_0 u) \}$ $\qquad (u \in V)$

$B_u = \{ v \in V \mid (\exists w \in F/I)(u = w_0 v) \}$ $\qquad (u \in V)$.

$\mathring{B}_u = B_u / \{ u \}$ (facteurs stricts de u).

X_u^- = espace associé à la limite projective de la restriction de $\gamma_0 \xi$ à $\mathring{B}_u$.

ζ_u^- application canonique de X dans X_u^-.

Dans la suite nous supposerons que V est muni d'un tribu $\mathscr{V}$, rendant mesurables les divers sous-ensembles de V introduits.

X_φ = espace associé à la limite projective de la restriction de $\gamma_0 \xi$ à I_φ sous-catégorie pleine de I, telle que $ObI_\varphi = \{ i \in f \mid f \in \varphi \}$ $(\varphi \subset F)$.

ζ_φ = application canonique de X dans X_φ.

Remarque sur la définition 5. — L'existence, et le choix de P_X^Θ, signifie que l'on est capable de définir une transition rassemblant toute l'information du plan d'expérience, cette transition n'étant pas en général, dans $\mathscr{E}$, c'est-à-dire ne correspondant pas à une expérience effectivement réalisable.

2.2. Définition du modèle.

Dans ce qui suit nous considérons une famille $\{D^i\}_{i\in I}$ d'espaces mesurables. Dans $\prod_{i\in I} D^i$ nous considérons une partie D mesurable pour la tribu produit, dont les projections D_i sur $\prod_{j\in V_i} D^j$ sont mesurables. Nous dirons que D est un espace de décision. Un élément de D^i peut être considéré comme une décision prise à la suite de l'observation correspondant à i.

DÉFINITION 6. — *a*) Un problème de décision statistique est le donné:
D'un plan d'expérimentation ξ.
D'un espace de décision (D, Ø).
D'une fonction mesurable l de ΘXVD dans $\overline{\mathbb{R}}_+$.
b) Une stratégie de décision est une transition S^X_{VD} telle que, sauf sur un ensemble P^θ_X négligeable, quel que soit θ:

$\delta 1)$ $(\forall\, i\in I)(S^X_{V_iD_i}$ factorise dans $\zeta_i)$ (restriction)
$\delta 2)$ $(\forall u\in V)(S^X_{A_uD}$ factorise dans $\zeta_u)$ (fonction mesurable).

Remarque. — La définition de stratégie ainsi adoptée sera suffisante pour les problèmes considérés ultérieurement. Elle ne répond sans doute pas pour tous les exemples concevables à la notion intuitive de stratégie. Une définition plus forte consisterait à substituer à (δ1) et (δ2) des conditions analogues, pour toute partie de I et pour toute partie de V, avec des limites projectives convenables.

La décision statistique se caractérise par la « fonction de risque » associée à toute stratégie S:

$$S(\theta) = \int_X \int_{VD} l(\theta, x, v, d) dS^x_{VD} dP^\theta_X.$$

Soit $\mathcal{S}$ une famille de stratégies. On la munit du préordre

$$(S_1 < S_2) \Leftrightarrow \begin{cases} (\forall\theta\in\Theta)(S_2(\theta) \leqq S_1(\theta)) \\ (\exists\ \theta\in\Theta)(S_2(\theta) < S_1(\theta)) \end{cases}$$

Il est défini comme le produit des ordres duaux de $\overline{\mathbb{R}}$.

On s'intéresse alors, en particulier à l'étude des éléments maximaux de $\mathcal{S}$. Dans cette étude les éléments du type suivant retiendront notre attention.

DÉFINITION 7. — Soit μ une mesure de probabilité sur Θ. Un élément S_μ de $\mathscr{S}$ sera dit μ-stratégie (relativement à $\mathscr{S}$), ou stratégie bayesienne relativement à μ de $\mathscr{S}$ si :

$$(\forall S \in \mathscr{S})\left(\int_\Theta S(\theta)d\mu \geqq \int_\Theta S_\mu(\theta)d\mu\right).$$

Soit $(\mu) = (\mu_n)$ une suite de probabilité sur Θ. Un élément $S_{(\mu)}$ de $\mathscr{S}$ sera dit (μ)-stratégie (relativement à $\mathscr{S}$), ou stratégie asymptotiquement bayesienne relativement à (μ) et $\mathscr{S}$ si :

$$\lim_{n \to +\infty}\left[\int_\Theta S_{(\mu)}(\theta)d\mu_n - \operatorname*{Inf}_{S \in \mathscr{S}} \int_\Theta S(\theta)d\mu_n\right] = 0.$$

Remarque. — Certains problèmes statistique (tests d'hypothèses) peuvent se ramener à des problèmes de décision statistique, en prenant pour préordre sur $\mathscr{S}$ un préordre satisfaisant aux conditions suivantes (S désignant la fonction risque)

$$(\pi 1) \qquad (S_1 \leqq S_2) \Rightarrow (S_1 \leqq S_2)$$
$$(\pi 2) \qquad (\forall \theta \in \Theta)(S_1(\theta) > S_2(\theta)) \Rightarrow (S_1 < S_2)$$
$$(\pi 3) \qquad (\forall S_0)(\exists a_0(\theta)(\{ S \geqq S_0 \} = \{ S : S(\theta) \leqq a_0(\theta) \}.$$

3. EXEMPLES

Ils nous montreront d'une part la notive intuitive de stratégie, d'autre part le passage de l'un à l'autre modèles envisagés au paragraphe 2.

3.1. Échantillonnage dénombrable. Étude des stratégies.

3.1.1. CONSTRUCTION DU MODÈLE

L'expérimentation dispose d'une famille dénombrable d'espaces d'observations $Y_1, Y_2, \ldots, Y_n$. Pour chaque valeur du paramètre θ, la probabilité du système est définie par la donnée, pour toute partie finie i de N de la probabilité P_{X_i} où

$$X_i = \prod_{k \in i} Y_k$$

soit $P^\Theta_{X_i}$. Si $i \supset j$ on définit une projection de X_i sur X_j l'image de $P^\Theta_{X_i}$ étant $P^\Theta_{X_j}$.

On expérimente de la manière suivante: on décide d'une partie j_1 finie de N et on fait une première observation dans X_{i_1} soit x_{i_1}. En fonction de j_1 et de ce premier résultat on décide d'une seconde partie finie j_2 disjointe de j_1. Lors de l'arrêt de l'opération on a choisi une suite croissante de parties finies $i_1 = j_1$, $i_2 = j_1 \cup j_2, \ldots, i_n = j_1 \cup j_2 \ldots \cup j_n$ et observé une suite $x_{i_1}, x_{i_2}, \ldots, x_{i_n}$, telle que x_{i_k} est la projection de x_{i_n}. On remarque en outre que la suite $(i_1, i_2, \ldots, i_n)$ est une fonction de x_{i_n} telle que i_1 ne dépend pas de x, i_2 dépend seulement de la projection x_{i_1}, etc. On formalisera cette situation comme suit:

I est une partie de l'ensemble des parties finies de $\mathbb{N}$,

V est l'ensemble des suites croissantes finies $(i_1 \subset i_2 \subset i_3 \ldots \subset i_n)$ où les i_k appartiennent à I.

$$X = \prod_{k \in \mathbb{N}} Y_k.$$

Une stratégie est une transition (ou une application) de X vers V satisfaisant aux conditions de la définition 6 (§ 2.2). Nous allons montrer que cette définition correspond bien à la notion de stratégie.

3.1.2. Sur la définition globale des stratégies dans le cas discret

α) *Définitions.*

Soit T un ensemble dénombrable, $(Y_t, \mathscr{Y}_t)$ une famille d'ensembles mesurables indexée par 8.

Soit I l'ensemble des parties finies de T, ordonnée par inclusion, et soit $F/I = \{ u = (i_1 \subsetneq i_2 \subsetneq \ldots \subsetneq i_n),\ i_1 \ldots i_n \in I,\ n \in N \}$ avec la composition (non partout définie):

$$(i_1 \subsetneq i_2 \subsetneq \ldots \subsetneq i_n)(i_n \subsetneq i_{n+1} \subsetneq \ldots \subsetneq i_{n+m}) = (i_1 \subsetneq \ldots \subsetneq i_{n+m})$$

(I.F/I) définit une catégorie dite des états d'expérimentation

$$\alpha(i_1, \ldots, i_n) = i_1, \qquad \beta(i_1, \ldots, i_n) = i_n$$

On notera

$$\begin{aligned} V_i &= \{ u \mid \alpha(u) = \varnothing, \quad \exists w, \quad \beta(w_0 u) = i \}. \\ A_u &= \{ v \mid \alpha(v) = \varnothing, \quad \exists w, \quad v = w_0 u \}. \\ V &= \{ u \mid \alpha(u) = \varnothing \}. \end{aligned}$$

$$X_i = \prod_{t \in i} Y_t, \qquad \mathscr{X}_i = \bigotimes_{t \in i} \mathscr{Y}_i.$$

Avec les applications projections, $(X_i, x_i)_{i \in I}$ forme un système projectif image de I par un facteur évident. Soit

$$\left(X = \prod_{t \in T} Y_t, \qquad x = \bigotimes_{t \in T} Y_t, \zeta_i\right)$$

la limite projective du système.

β) Définition constructive d'une stratégie.

Désignons pour tout $u \in V$, par $I_u = \{ i \in I, i \not\supseteq \beta(u) \}$ et donnons-nous pour chaque u une « transition $S_{I_u \cup D}^{X_{\beta(a)}} = S_{uI_u \cup D}$, probabilité de l'étape suivante ». Ceci définit un mode opératoire ou fonction de décision. A cette famille $(S_{uI_u \cup D}^{X}, u \in V)$ nous pouvons associer une transition S_{VD}^{X}. Nous nous poserons les questions suivantes :

— comment définir S_{VD}^{X},

— quelles sont les conditions sur S_{VD}^{X} ainsi définies pour que $(S_u^X, u \in V)$ soit une « bonne fonction de décision »,

— quelles sont, étant donné S_{VD}^{X}, les conditions pour qu'elle provienne d'une fonction de décision.

γ) Construction de S_{VD}^{X}. Conditions nécessaires.

Soit $u \in V$, $d \in \emptyset$ et considérons la décomposition maximale de u :

$$u = u_n \circ u_{n-1} \circ \ldots \circ u_1 \qquad \text{et} \qquad v_k = u_k \circ u_{k-1} \ldots \circ u_1,\ v_n = u.$$

Considérons

$$S_{\{u\}d}^{x} = S_{\emptyset}^{x}(u_1) S_{u_1}^{x}(v_1) \ldots S_{v_{n-1}}^{x}(v_n) S_{v_n}^{x}(d) \tag{1}$$

Cette quantité définit pour chaque x la « probabilité que l'opération conduise en définitive à parcourir le chemin u et à prendre une décision dans d. Si $i = \beta(u)$, elle définit une fonction $\zeta_i^{-1}(x_i) = \bar{x}_i$ mesurable, quel que soit d.

Pour chaque x, nous pouvons alors définir une mesure en VD comme suit. Soit $u_1, u_2, \ldots, u_n, \ldots$, les éléments de V dans un ordre arbitraire, soit $\Delta \subset VD$, nous notons

$$\Delta^i = \{ d \mid (v_i, d) \in \Delta \}$$

et

$$S_{\Delta}^{x} = \sum_{1}^{+\infty} X_{\{v_n\} X \Delta^n}^{x}.$$

Pour une famille dénombrable de Δ_n, deux à deux disjoints, commutant les sommations :

$$\sum_{1}^{n} S^x_{\Delta_n} = \sum_{n=1}^{+\infty}\sum_{i=1}^{+\infty} S^x_{\{v_i\}\Delta^i_n} = \sum_{i=1}^{+\infty}\sum_{n=1}^{+\infty} S^x_{\{v_i\}\Delta^i_n} = \sum_{i-1}^{+\infty} S^x_{\{v_i\}\mathrm{X}\ \Delta^i_n} = S^x_{\bigcup_n \Delta_n}.$$

Ainsi à la stratégie $(S^x_{u\mathrm{l}_u \cup \mathrm{D}}, u \in \mathrm{V})$ nous pouvons associer une transition $S^{\mathrm{X}}_{\mathrm{VD}}$ définie ci-dessus.

La masse totale est inférieure ou égale à 1 quel que soit x. La stratégie est fermée, au sens de Bahadur [*1*] si pour tout x (presque pour tout quand on probabilisera le modèle) $S^{\mathrm{X}}_{\mathrm{VD}} = 1$.

Pour les conditions nécessaires, nous en retenons deux intuitives.

2) Pour tout i, la restriction $S^{\mathrm{X}}_{\mathrm{V}_i\mathrm{D}}$ est une transition $\zeta_i^{-1}(x_i) = \overline{x}_i$ mesurable.

Ceci provient du fait que, quel que soit d, $(2')S^{\mathrm{X}}_{\{u\}\mathrm{X}d}$ est $\overline{x}_{\beta(u)}$ mesurable ce qui résulte de (1). On notera d'ailleurs que $(2') \Leftrightarrow (2)$.

3) Pour tout u $S^{\mathrm{X}}_{\mathrm{A}_u\mathrm{D}}$ est $x_{\beta(v_{n-1})} = x_u$ mesurable où v_{n-1} a le sens défini avant la relation.

Ceci résulte toujours de la relation (1) :

$$S^{\mathrm{X}}_{\mathrm{A}_u v} = S^{\mathrm{X}}_{u_1}(v_2) \ldots S^{\mathrm{X}}_{v_{n-1}}(v_n).$$

δ) *Réciproque. Les conditions sont suffisantes.*

Soit u, et $v = (u, j)\ j \geqq \beta(u)$.

Nous définissons

$$S^x_u(v) = \frac{S^x_{\mathrm{A}_u\mathrm{D}}}{S^x_{\mathrm{A}_u\mathrm{D}}}$$

qui est $\overline{x_{\beta(u)}}$ mesurable par (3) et pour $d \in \mathscr{D}$:

$$S^x_u(d) = \frac{S^x_{\{u\}\mathrm{X}d}}{S^x_{\mathrm{A}_u\mathrm{D}}}$$

qui est $\overline{x_{\beta(u)}}$ mesurable par (2) et (3).

Il est clair que le système de probabilités conditionnelles ainsi défini redonne $S^{\mathrm{X}}_{\mathrm{VD}}$.

Conséquence. — La classe des stratégies de décision est en bijection avec la classe des $S^{\mathrm{X}}_{\mathrm{VD}}$, qui sont positives de norme $\leqq 1$, (1 pour les fermées) et qui satisfont (2) et (3).

3.2. Échantillonnage « continu ».

Soit T un intervalle fermé de $\overline{\mathbb{R}}_+$ ou de $\overline{\mathbb{N}}$. On dispose ici d'une famille d'espaces d'observations $(\mathrm{Y}_t)_{t\in\mathrm{T}}$. 0 désignant l'origine de T, nous prendrons pour I (et V) l'ensemble des intervalles fermés $[0, t] = i$, $t \in \mathrm{T}$ et

$$\mathrm{X}_i = \prod_{s \leqq t} \mathrm{Y}_s.$$

L'existence de la transition $\mathrm{P}_{\mathrm{X}}^{\Theta}$ est ici encore très naturelle. Quant aux stratégies, la propriété ($\delta 1$) suffit à les définir. C'est donc une extension de la notion de temps d'arrêt.

Nous sommes dans un cas particulier du paragraphe 2.5.

3.3. Modification d'un paramètre inconnu. Introduction à la commande statistique.

Nous supposerons ici que nous sommes capables d'agir sur un paramètre θ, par exemple réel, d'un certain système. On observe par ailleurs une grandeur aléatoire $\mathrm{X}_{\theta,t}$ $t = 1, 2, \ldots, n$, dont la distribution dépend de θ. A un instant t donné l'action consiste à augmenter ou à diminuer θ d'une quantité f_t connue (le paramètre étant lui-même inconnu). Si f_t prend ses valeurs dans un ensemble borné de $\mathbb{R}$ soit F, si les observations sont faites à temps discret, $t = 1, 2, \ldots, n$ et si elles sont indépendantes, chaque état d'observation i est constitué par la suite, réordonnée en grandeur croissante, des nombres f_1, $f_1 + f_2, \ldots, f_1 + f_2 + \ldots + f_n$, soit

$$i = (a_1 \leqq a_2 \leqq \ldots \leqq a_k).$$

On lui associe un espace d'observation $\mathrm{X}(a_1, a_2, \ldots, a_k)$ et une probabilité:

$$\mathrm{P}^{\theta_0 + a_1} \times \mathrm{P}^{\theta_0 + a_2} \times \ldots \times \mathrm{P}^{\theta_0 + a_k}$$

pour chaque valeur initiale θ_0 du paramètre (inconnue). Il est facile de s'assurer que l'ensemble I des états est filtrant, et que, du fait de l'indépendance on peut définir sur la limite X des espaces $\mathrm{X}_{(a_1,\ldots,a_k)}$ une probabilité $\mathrm{P}_{\mathrm{X}}^{\theta_0}$ représentant le système.

Ce problème constitue un cas simple de commande statistique et le modèle se généralise aisément au cas d'un temps continu et où θ est une fonction aléatoire de t.

4. TESTS ET ESTIMATION

Nous introduisons pour finir deux techniques, plus utilisées pratiquement que la décision statistique, et qui seront étudiées en détails aux chapitres IV et V.

La situation est analogue à celle de la décision statistique (ci-dessus § 3). Mais préalablement à la fonction de coût, on introduit une fonction h de Θ dans D. La valeur de $h(\theta)$ peut être considérée comme la « bonne décision » lorsque θ est la « vraie » valeur du paramètre. On distingue :

Le cas où D est fini, $\mathrm{D} = \{d_0, d_2, \ldots, d_k\}$ h induit alors une partition de Θ, finie, $\Theta = \Theta_0 \cup \Theta_1 \ldots \cup \Theta_k$, $\Theta_i = h^{-1}(d_i)$. Les Θ_i sont dits les « hypothèses ». L'essentiel des techniques développées jusqu'ici porte sur le cas où $k = 1$ (Test de deux hypothèses). On introduit en général au niveau des critères une dissymétrie entre Θ_0, hypothèse nulle, celle dont il s'agit de vérifier le bien fondé, et Θ_1, contre hypothèse, ce qui se passe sinon.

Le cas où D est une partie d'un espace vectoriel E. La donnée d'une stratégie $\mathrm{S}_\mathrm{D}^\mathrm{X}$ induit alors une transition $\mathrm{H}_\mathrm{D}^\Theta = \mathrm{S}_\mathrm{D}^\mathrm{X} \circ \mathrm{P}_\mathrm{D}^\Theta$. On peut pour chaque θ s'intéresser au barycentre de la mesure $\mathrm{H}_\mathrm{D}^\theta$ soit $\mathrm{H}(\theta)$. Il est naturel d'introduire des conditions du type $h(\theta) = \mathrm{H}(\theta)$ quel que soit θ (estimation sans biais). Si l'espace vectoriel est normé ($\|.\|$) il est naturel aussi d'introduire une fonction de coût du type :

$$l(\theta, d) = \| d - \mathrm{H}(\theta) \|$$

Le schéma général de la décision statistique conduit alors à une généralisation du critère de la variance minimum.

BIBLIOGRAPHIE

[1] BAHADUR, Sufficiency and Statistical decision functions. *A. M. S.*, vol. **25**, 1954, p. 423-462.

[2] CENSOV, *Méthodes infinitésimales de la Statistique Mathématique*. Institut de mathématiques, Moscou, 1966.

[3] LE CAM, Sufficiency and approximate sufficiency. *A. M. S.*, vol. **35**, 1964, p. 1419-1455.

[4] BELL, Séminaire ISUP, 1967.

[5] BOURBAKI, *Théorie des Ensembles*. Livre I. Hermann, Paris.

[6] ROMIER, Exhaustivité et équivalence des objets statistiques. Note aux *C. R. Acad. Sci.*, *A paraître*.

Manuscrit reçu le 11 juin 1969.