PUBLICATIONS MATHÉMATIQUES ET INFORMATIQUES DE RENNES

J. R. BARRA

Quelques théorèmes généraux de statistique mathématique

Publications des séminaires de mathématiques et informatique de Rennes, 1967-1968
« Publications des séminaires du département de mathématiques », , exp. nº 3, p. 1-9
<http://www.numdam.org/item?id=PSMIR_1967-1968____A3_0>

QUELQUES THEOREMES GENERAUX DE STATISTIQUE MATHEMATIQUE

par

J.R. BARRA

Il s'agit moins de présenter des résultats très originaux, qu'un formalisme de base pour la statistique mathématique : ce formalisme est tout simplement celui du Calcul des Probabilités élémentaires moderne, mais il suffit à placer, nous semble-t-il, les résultats essentiels dans leur cadre mathématique naturel.

1. - Soit $\mathcal{P} = \{P_\theta \; \theta \; \Xi\}$ une famille de lois de probabilité sur l'espace mesurable $(\Omega,\mathcal{Q})$; on appelle <u>structure statistique</u> le triplet $(\Omega,\mathcal{Q},\mathcal{P})$. Cette notion joue le rôle de l'espace probabilisé en C.des P. et de le préciser éclaircit bien des démonstrations de statistique.

Deux notions sont importantes :

a) celle bien classique de <u>tribu exhaustive sur $(\Omega,\mathcal{Q},\mathcal{P})$:</u>

$$\mathcal{B} \subset \mathcal{Q} \quad \forall A \in \mathcal{Q} \quad , \exists \, P(A/\mathcal{B}) \text{ indépendant de } P \in \mathcal{P}$$

b) celle, moins clairement mise en évidence de <u>tribu libre</u> sur $(\Omega,\mathcal{Q},\mathcal{P})$

$$\mathcal{C} \subset \mathcal{Q} \quad (\Omega,\mathcal{C},\mathcal{P}) \text{ est un espace de probabilité !!!}$$

(ce qui signifie, si $\mathcal{C}$ est induit par une statistique T, que la loi de probabilité de T est indépendante de P dans $\mathcal{P}$).

Ces deux notions sont complementaires, non seulement au point de vue de la théorie de l'information, mais aussi de l'indépendance stochastique.

Soit $\mathcal{B}$ une tribu exhaustive sur $(\Omega, \mathcal{A}, \mathcal{P})$

a) si $(\Omega, \mathcal{B}, \mathcal{P})$ est quasi-complète, $\mathcal{B}$ est indépendante de toute tribu libre sur $(\Omega, \mathcal{A}, \mathcal{P})$, quel que soit $P \in \mathcal{P}$.

b) si $\forall$ P, $P' \in \mathcal{P}$, $\nexists B \in \mathcal{B}$ tel que

$P(B) = 1$, $P'(B) = 0$

alors toute tribu indépendante de $\mathcal{B}$ quel que soit $P \in \mathcal{P}$ est libre sur $(\Omega, \mathcal{A}, \mathcal{P})$

La partie a) découle d'une ext. du th de Lehmann-Scheffé et Linnik ((cf. Le cours de Statistique analytique chez G.V.) en a donné de jolis exemples d'applications ; la démonstration de b) se trouve également dans ce livre.

Remarquons que cette notion de liberté, et celle de liberté en moyenne : "T statistique, vectorielle, sommable, est libre en moyenne sur $(\Omega, \mathcal{A}, \mathcal{P})$ si E(T) est constante sur $\mathcal{P}$", contiennent les notions de similar tests, tests invariants selon Linnik, distribution par methods,...

De même, la notion si importante, de liberté par rapport à un paramètre fantôme μ , si $\theta = (\lambda,\mu)$, signifie que pour tout λ fixé, on a liberté en moyenne relativement à μ.

3 - Considérons une structure paramétrée par un paramètre $\theta \in (\Xi, \mathcal{C})$ telle que : P_θ soit une probabilité de transition sur $\Omega \times \Xi$.

Soit ϕ une statistique réelle

$$(\Omega,\mathcal{A},\mathcal{P}) \xrightarrow{\phi} (\mathbb{R},\mathcal{B})$$

Beaucoup de problèmes de test ou d'estimation consistent à étudier l'application β :

$$\phi \longrightarrow E(\phi) = \beta_{\phi}(\Theta)$$

Considérons cette application β comme une application de

(1) $\quad \mathcal{L}_\infty(\Omega,\mathcal{A}) \xrightarrow{\beta} \mathcal{L}_\infty(\Xi,\mathcal{C})$

On sait qu'à toute mesure bornée m sur $(\Xi,\mathcal{C})$ correspond une mesure bornée β^*_m sur $(\Omega,\mathcal{A})$ définie par :

$$\forall\, A\in\mathcal{A} \quad \beta^*_m(A) = \int_\Xi P_\Theta(A)\, dm$$

et telle alors que $\displaystyle\int_\Xi \beta_\phi(\Theta)\, dm = \int_\Omega \phi\, d\beta^*_m$

D'où l'application duale

(2) $\quad \mathcal{M}(\Xi,\mathcal{C}) \xrightarrow{\beta} \mathcal{M}(\Omega,\mathcal{A})$

Si la structure est dominée par P^*, le schéma se simplifie en

(3) $\quad L_\infty(\Omega,\mathcal{A},P^*) \xrightarrow{\beta} \mathcal{L}_\infty(\Xi,\mathcal{C})$

(4) $\quad L_1(\Omega,\mathcal{A},P^*) \xleftarrow[\beta^*]{} \mathcal{M}(\Xi,\mathcal{C})$

Désignons par T le sous-ensemble de $L_\infty(\Omega,\mathcal{A},P^*)$

$$\|2\phi-1\|_\infty \leq 1$$

et par T' le sous-ensemble de $\mathcal{L}_\infty(\Omega,\mathcal{A})$:

$$0 \leq \phi \leq 1$$

L'image par β de T' (ou de T dans le cas dominé) est un <u>sous-ensemble</u> convexe R de $\mathcal{L}_\infty(\Xi,\mathcal{C})$ que nous dirons <u>associé à la structure</u> $(\Omega,\mathcal{A},\mathcal{P})$

Or, on sait (Banach-Alaoglu) que si $L_\infty(\Omega,\mathcal{A},P^*)$ est muni de la topologie $\sigma\{L_\infty,L_1\}$, T est compact ; d'autre part si $\mathcal{L}_\infty(\Xi,\mathcal{C})$ est muni de la topologie $\sigma\{\mathcal{L}_\infty\ ,\mathcal{M}\}$ il est facile de constater que β ((3) est une application continue, d'où finalement :

"<u>Le domaine associé à une structure dominée est convexe et compact</u>".

Plusieurs théorèmes généraux de théorie de tests s'appuyent sur ce résultat et nous en citons quelques uns :

4 - <u>Soit</u> $(\Omega,\mathcal{A},\{P_\Theta,(\Xi,\mathcal{C})\})$ <u>une structure dominée quels que soient $\Xi_o \subset \Xi$, $\Theta_1 \in \Xi - \Xi_o$, $o \leq \alpha \leq 1$, il existe un test U.M.P. de Ξ_o contre Θ_1 de niveau α</u>".

Ce théorème est donné avec une condition inutile ($\mathcal{A}$ à base dénombrable) dans Lehman, mais la démonstration est la même :

Le sous-ensemble de T défini par

$$\beta_\phi(\Theta) \leq \alpha \qquad \forall\ \Theta \in \Xi_o$$

est fermé dans T, donc compact et la fonction de $\phi,\beta_\phi\ (\Theta_1)$, continue, atteint son maximum.

<u>Conjecture</u>. "Si la structure n'est pas dominée, $\exists$ une loi P' (combinaison convexe des P_Θ) telle qu'il n'existe pas de test U.M.P. de Ξ contre P'.

5 - On peut de plus situer les tests optimaux sur R :

"Un point β_ϕ de R appartient à l'intérieur de R si quel que soient Ξ_o, Ξ_1 parties disjointes de Ξ, ϕ n'est pas admissible pour tester Ξ_o contre Ξ_1".

On sait que ϕ admissible comme test de Ξ_o contre Ξ_1 signifie maximal pour le préordre

$$\phi' \geq \phi \Longleftrightarrow \begin{cases} \beta_{\phi'}(\Theta) \leq \beta_\phi(\Theta) & \Theta \in \Xi_o \\ \beta_{\phi'}(\Theta) \geq \beta_\phi(\Theta) & \Theta \in \Xi_1 \end{cases}$$

que l'on distinguera de la notion de quasi-admissibilité souvent confondue.

Démonstration.

Si R est un voisinage de β_ϕ pour $\sigma\{\mathcal{L}_\infty , \mathcal{L}_1\}$, $\exists\ \varepsilon > o$ et $\Pi_1,\dots,\Pi_N \in \mathcal{M}(\Xi,\mathcal{C})$ telles que R contienne

$$\left|\int (f-\beta_\phi)\, d\Pi_i\right| \leq \varepsilon \ \forall\ i = 1 \dots N$$

Quels que soient Ξ_o et Ξ disjointes la fonction

$$f = \begin{cases} \beta_\phi + \varepsilon/K & \Theta \in \Xi \\ \beta_\phi - \varepsilon/K & \Theta \in \Xi \qquad K = \sup \|\Pi_i\| \\ \beta_\phi & \text{sinon.} \end{cases}$$

est donc dans R et le test ϕ' qui l'admet comme fonction puissance $\beta_{\phi'} = f$ est strictement meilleur que ϕ comme test de Ξ_o contre Ξ_1.

6 - Les points frontières de R sont donc particulièrement intéressants et on peut généraliser comme suit un résultat donné par Lehman et qui est une sorte de réciproque du lemme de Neyman et Pearson :

"Si la structure est dominée par P, l'intérieur $\overset{\circ}{R}$ de R étant non vide, alors si β_ϕ est point frontière de R, $\exists$ $m \in \mathcal{M}(\Xi, \mathcal{C})$ telle que si

$$p^*(\omega) = \frac{d}{dP} \left(\int_\Xi P_\Theta \, dm\right)$$

on ait P- presque partout

$$p^*(\omega) > o \implies \phi = 1$$

$$p^*(\omega) < o \implies \phi = 0 \text{ ".}$$

Démonstration : R est ici fermé, donc β_ϕ appartient à R , mais non à R^o qui est convexe, donc $\exists$ (Hahn-Banach) une f-linéaire continue.

Donc $m \in \mathcal{M}(\Xi, \mathcal{C})$ telle que

$$\int_\Xi \beta_\psi \, dm \leq \int_\Xi \beta_\psi dm \qquad \forall \text{ le test } \psi$$

soit encore $\int_\Omega \psi \, d\beta^*_m \leq \int_\Omega \phi \, d\beta^*_m$ "

ou enfin $\int_\Omega (\phi-\psi) p^* dp \geq o$ "

D'autre part si $p^* = \begin{cases} 1 & p^* > o \\ 0 & p^* \leq o \end{cases}$

on a également $\int_\Omega (\phi^*-\psi) p^* dp \geq o$ "

donc $\phi = \phi^*$ P-presque partout.

7 - L'application du théorème de KREIN-MILLMAN permet souvent de trouver des tests déterministes optimaux, comme points extrémaux d'un sous-ensemble convexe de R. Soit donc $\mathcal{D}$ le sous-ensemble de R, correspondant aux tests déterministes, c'est-à-dire aux fonctions

$$P_\Theta(A) \qquad \forall A \in \mathcal{A}.$$

"Tout point extrémal de R appartient à $\mathcal{D}$ ".

En effet si ϕ^* est un point extrémal, non déterministe

$\exists\ A \in \mathcal{A},\ \varepsilon > o\ ,\ p \in \mathcal{P}$ tels que

$$\begin{cases} \varepsilon \leq \phi^* \leq 1-\varepsilon & \text{sur } A \\ p(A) > o \end{cases}$$

les deux tests $\phi^* + \varepsilon\chi_A$ et $\phi^* - \varepsilon\chi_A$ fournissent alors la contradiction.

(Lyapounov) "Si Ξ est fini, les lois P_ϕ étant non atomiques, alors R $\equiv \mathcal{D}$".

Démonstration : Soit ϕ un test donné, et soit V_ϕ le sous-ensemble de T formé des tests équivalents à ϕ

$$\psi : \int \phi\, dp_\Theta = \int \psi\, dp_\Theta \qquad \forall\ \Theta$$

V_ϕ est fermé dans T donc compact pour $\sigma(L_\infty, L_1)$, non vide et donc a des points extrémaux, montrons que ceux-ci correspondent à des tests déterministes.

Soit ϕ^* un élément extrémal, non déterministe $\exists\ A \in \mathcal{A},\ \varepsilon > o,\ \Theta \in$

tels que $\begin{cases} \varepsilon \leq \phi^* \leq 1-\varepsilon & \omega \in A \\ \quad P_\Theta(A) > o \end{cases}$

Comme P_Θ est non-atomique, $\exists\ A_1, \ldots, A_{N+1}$ sous-ensembles disjoints de A tels que $P_\Theta(A_j) > o$

alors le système d'équation $\sum_{j=1}^{N} P_\Theta(A_j)x_j = o \quad \forall\ \Theta = 1\ldots N$ admet au moins une solution, à laquelle on peut imposer

$$|x_j| \leq \varepsilon \qquad j = 1\ldots N$$

Alors les deux tests $\phi^* + \sum_j x_j\, \chi_{A_j}$, $\phi^* - \sum_j x_j\, \chi_{A_j}$ appartiennent à V_ϕ, sont distincts de ϕ^* (dans L_∞) et on a la contradiction cherchée.

Romanosky et Sudakoff ont même démontré, de façon analogue, le raffinement suivant de ce théorème :

"Quel que soit le test ϕ sur la structure $|\mathbb{R}^2, B^2, (P_1,\ldots,P_N)|$ dominée par la mesure de Lebesgue , $\exists\ A \in B^2$ tel que

$$P_i(A|x) = E(\phi|x) \qquad x \text{ p-partout}$$

$$P_i(A|y) = E(\phi|y) \qquad y \text{ p-partout}$$

x,y désignant les coordonnées de $\mathbb{R}^2$.

8 - En conclusion, nous considérons, comme question où le formalisme est également facile à préciser, le problème de l'estimation ensembliste (Confidence Régions)

Soit la structure statistique :

$(\Omega, \mathcal{A}, P)$, $P = \{P_\Theta, \ \Theta \in \Xi \subset (\Xi_1 \times \Xi_2 , \mathcal{C}_1 \times \mathcal{C}_2)\}$

Θ_1 est le paramètre à estimer, Θ_2 le paramètre fantôme.

Définition (I) (correspondant aux Confidence Régions. p. ex. CRAMER)

Une estimation ensembliste de Θ_1 est une application

$$E : \Omega \longrightarrow \mathcal{C}_1$$

telle que $\forall\ \Theta_1$ la partie de Ω $E^*_{\Theta_1}$ définie par

$$\Theta_1 \in E(\omega) \qquad \text{appartienne à } \mathcal{A} .$$

Définition II   Légèrement plus forte

Une estimation ensembliste est un ensemble E de $\mathcal{A} \times \mathcal{E}_1$

Le seuil d'une estimation ensembliste est défini par

$$\inf(P_\Theta(E^*_{\Theta_1}) \,|\, \Theta = (\Theta_1, \Theta_2) \in \Xi)$$

l'estimation ensembliste est libre si $P_\Theta(E^*_{\Theta_1})$ est constant sur Ξ

On constate facilement le rapport avec la théorie des tests, mais c'est le cas d'une estimation libre, qui est le plus intéressant, et me semble contenir la notion de probabilité Fiduciaire (Fisher) en effet :

"Si E est une estimation ensembliste (II) libre de seuil α , $\forall$ la loi de probabilité Q sur $(\Xi, \mathcal{E})$, si π est la loi de probabilité sur $(\Omega \times \Xi)$ qui est alors induite par la probabilité de transition P_Θ on a

$$\pi(E) = \alpha \text{ ".}$$

D'où le "renversement" des assertions de la probabilité fiduciaire.