THÉORIE DES PROBABILITÉS. – *Un problème de minimum pour le couplage de deux probabilités.* Note (*) de **Jean Moulin Ollagnier** et **Didier Pinchon,** transmise par M. Robert Fortet.

Le théorème principal de cette Note caractérise le minimum de l'intégrale d'une fonction continue sur le produit de deux espaces compacts pour les probabilités ayant des marginales données. On retrouve comme cas particuliers, l'expression de différentes distances sur les probabilités d'un même espace.

The main theorem of this Note characterizes the infimum of the mean value of a continuous function on the product of two compact spaces for probabilities with given marginals. As a particular application, one can obtain the known expressions of some distances on the set of all probabilities on a space.

Soient X et Y deux espaces compacts, μ et ν deux probabilités définies respectivement sur X et Y. Soit f une fonction continue sur le produit $X \times Y$. On désigne par $I(f)$ la borne inférieure des nombres $\lambda(f)$ où λ décrit l'ensemble des probabilités sur $X \times Y$ de marginales μ et ν.

Une probabilité sur $X \times Y$ étant une forme linéaire positive de masse totale 1 sur l'espace $C(X \times Y)$ des fonctions continues sur $X \times Y$, le nombre $I(f)$ est caractérisé à l'aide du théorème de Hahn-Banach par

$$I(f) = \sup_{\varphi,\, \psi} J(\varphi, \psi),$$

où φ et ψ décrivent respectivement les espaces $C(X)$ et $C(Y)$ et

$$(1) \qquad J(\varphi, \psi) = \inf_{(x,\, y) \in X \times Y} [f(x, y) + \varphi(x) - \mu(\varphi) + \psi(y) - \nu(\psi)].$$

Le but de cette étude est de restreindre le domaine du couple (φ, ψ) pour obtenir une expression simplifiée de (1).

Nous étudions d'abord le cas où X et Y sont deux ensembles finis : le théorème 1 montre que $I(f)$ est atteint pour certains couples (φ, ψ) possédant nécessairement une propriété d'association. Dans le cas général, il est suffisant de se restreindre aux couples vérifiant cette propriété (th. 2).

ÉTUDE DU CAS FINI. – Soient X et Y deux ensembles finis, μ et ν deux probabilités sur X et Y respectivement, chargeant tous les points. Soit f une fonction sur $X \times Y$.

Il est clair que $I(f) = \sup_{\varphi,\, \psi} J(\varphi, \psi)$, où φ et ψ sont des fonctions sur X et Y vérifiant $\mu(\varphi) = \nu(\psi) = 0$ et

$$J(\varphi, \psi) \geqq J(0, 0) = \inf_{x,\, y} f(x, y).$$

De telles fonctions vérifient alors, pour tout x et tout y

$$\varphi(x) + \psi(y) \geqq \inf_{x,\, y} f(x, y) - \sup_{x,\, y} f(x, y) = -M.$$

Puisque φ et ψ sont d'intégrales nulles, elles prennent des valeurs négatives ou nulles, ce qui permet d'obtenir séparément

$$(\forall x \in X), \quad \varphi(x) \geqq -M \qquad \text{et} \qquad (\forall y \in Y), \quad \psi(y) \geqq -M.$$

Utilisant encore la condition d'intégrale nulle, on obtient :

$$(\forall x \in \mathrm{X}), \quad \varphi(x) \leqq \mathrm{M} \sup_{t \in \mathrm{X}} \frac{1-\mu(\{t\})}{\mu(\{t\})}$$

et

$$(\forall y \in \mathrm{Y}), \quad \psi(y) \leqq \mathrm{M} \sup_{z \in \mathrm{Y}} \frac{1-\nu(\{z\})}{\nu(\{z\})}.$$

On peut donc se restreindre à prendre le maximum sur un compact de $\mathrm{R}^{\mathrm{X}} \times \mathrm{R}^{\mathrm{Y}}$ et $\sup_{\varphi, \psi} \mathrm{J}(\varphi, \psi)$ est atteint.

THÉORÈME 1. — *Soit* (g, h) *un couple réalisant* $\mathrm{J}(g, h) = \sup_{\varphi, \psi} \mathrm{J}(\varphi, \psi)$. g *et* h *vérifient les propriétés suivantes :*

$$\inf_{y \in \mathrm{Y}} [f(x, y) + g(x) + h(y)] \quad \textit{est indépendant de } x$$

et

$$\inf_{x \in \mathrm{X}} [f(x, y) + g(x) + h(y)] \quad \textit{est indépendant de } y.$$

Démonstration. — Soit X′ l'ensemble des points de X tels qu'il existe y de Y avec

$$f(x, y) + g(x) + h(y) = \inf_{z, t} [f(z, t) + g(z) + h(t)]$$

et Y′ le sous-ensemble de Y défini de manière semblable.

Soit ε positif; considérons le couple $g_\varepsilon = g + \varepsilon\, 1_{\mathrm{X}'}$, $h_\varepsilon = h + \varepsilon\, 1_{\mathrm{Y}'}$. On peut écrire :

$$\mathrm{J}(g_\varepsilon, h_\varepsilon) \leqq \mathrm{J}(g, h)$$

ce qui donne

$$\inf_{x, y} [f(x, y) + g(x) + h(y) + \varepsilon\, 1_{\mathrm{X}'}(x) + \varepsilon\, 1_{\mathrm{Y}'}(y) - \varepsilon\mu(\mathrm{X}') - \varepsilon\nu(\mathrm{Y}')]$$
$$\leqq \inf_{x, y} [f(x, y) + g(x) + h(y)]$$

Il existe une suite ε_n tendant vers zéro et un couple (x_0, y_0) assurant le minimum du terme de gauche pour tous les ε_n; ce couple assure le minimum du terme de droite et donc pour un certain ε strictement positif on a $\varepsilon [2 - \mu(\mathrm{X}') - \nu(\mathrm{Y}')] \leqq 0$ d'où $\mathrm{X} = \mathrm{X}'$ et $\mathrm{Y} = \mathrm{Y}'$, ce qui donne le résultat.

On peut modifier g d'une constante de façon à obtenir $\inf_{y \in \mathrm{Y}} [f(x, y) + g(x) + h(y)] = 0$ pour tout x, c'est-à-dire choisir $g = h^0$ avec $h^0(x) = \inf_y [f(x, y) + h(y)]$.

D'autre part $\inf_x [f(x, y) + h^0(x) + h(y)]$ est une constante dont on vérifie aisément qu'elle est nulle.

CAS GÉNÉRAL. — Les résultats obtenus dans le cas fini conduisent à poser les définitions suivantes :

DÉFINITION. — Soit h une fonction continue sur Y; on appelle associée de h et on note h^0 la fonction continue sur X définie par

$$h^0(x) = -\inf_y [f(x, y) + h(y)].$$

L'application d'association est continue en norme uniforme. On définit de la même façon l'associée d'une fonction continue sur X.

THÉORÈME 2. – *On a* $I(f)=\sup_{g,h} -[\mu(g)+\nu(h)]$ *où le couple* (g, h) *vérifie* $g=h^0$ *et* $h=g^0$.

Démonstration. – Ceci résulte des deux faits suivants :

ψ étant fixé, $\sup_{\varphi} J(\varphi,\psi)=J(\psi^0, \psi)$ [et de même $\sup_{\psi} J(\varphi, \psi)=J(\varphi, \varphi_0)$] et d'autre part, pour tout φ, $\varphi^{000}=\varphi^0$ (resp. $\psi^{000}=\psi^0$).

On peut donc associer à tout couple (φ, ψ), le couple $(g, h)=(\psi^0, \psi^{00})$ [ou $(\varphi^{00}, \varphi^0)$] tel que $J(\varphi, \psi)\leqq J(g, h)$ et on a $g^0=h$ et $h^0=g$.

THÉORÈME 3. – X *et* Y *étant des espaces compacts et f une fonction continue sur* X × Y, $I(f)=\sup_{\varphi,\psi} J(\varphi, \psi)$ *est atteint par un couple de fonctions.*

Démonstration. – En application du théorème 2, il suffit de voir que le maximum est atteint par un certain couple de fonctions associées. On remarque que si (g, h) est un couple de fonctions associées, $(g+a, h-a)$ l'est aussi si a est une constante et que $J(g,h)=J(g+a,h-a)$.

Soit C l'ensemble des fonctions continues g sur X vérifiant $\inf_{x\in X} g(x)=0$ et $g^{00}=g$. On a alors

$$I(f)=\sup_{g\in C} J(g, g^0),$$

$J(g, g^0)$ étant une fonction continue de g, le maximum en est atteint si C est compact dans C(X).

Ceci résulte du théorème d'Ascoli :

En effet $g=g^{00}$ est équivalent à

$$(\forall x\in X),\quad (\exists y\in Y),\quad (\forall z\in X),\qquad g(x)+f(x, y)\leqq g(z)+f(z, y). \tag{2}$$

On en déduit en particulier, pour tout x et z de X :

$$|g(x)-g(z)|\leqq\sup_{y} |f(x, y)-f(z, y)|.$$

Ce qui prouve tout à la fois l'uniforme équicontinuité des fonctions vérifiant $g=g^{00}$ et le fait que pour tout x de X et tout g de C, $0\leqq g(x)\leqq 2\,\|f\|$.

Remarques. – 1. On suppose que X = Y. Une condition nécessaire et suffisante pour que, pour tout h de C(X) tel que $h=h^{00}$, on ait $h^0=-h$ est que f vérifie les deux conditions : $f(x, x)=0$ pour tout x de X et $f(x, y)+f(y, z)\geqq f(x, z)$ pour tout x, y, z de X (3). On obtient alors le résultat :

$$I(f)=\sup_{h=h^{00}} [\nu(h)-\mu(h)].$$

Cela permet de voir en particulier que $I(f)$ comme fonction de μ et de ν est une distance si et seulement si f est une distance sur X. Dans ce cas les fonctions de C sont les fonctions lipschitziennes de rapport 1, comme l'a remarqué Gilles Royer [1].

2. Si $X \subset Y$ et si f est une fonction sur $Y \times Y$ satisfaisant (3), si h est une fonction sur X f-lipschitzienne c'est-à-dire vérifiant, pour tout x et y de X, $h(x)-h(y) \leqq f(x, y)$ alors l'opposé de l'associée de h est un prolongement f-lipschitzien de h à Y tout entier. Ce résultat fait l'objet d'un exercice dans le livre de topologie de G. Choquet.

3. Si $X=Y$ est un espace fini et f l'indicatrice sur $X \times X$ du complémentaire du graphe d'une relation d'ordre, f vérifie (3) et C est constitué par les fonctions croissantes comprises entre 0 et 1.

4. Si X et Y sont des espaces polonais et f une fonction s. c. s. bornée sur $X \times Y$, on peut étendre le résultat du théorème 2 pour obtenir une généralisation des résultats de V. Strassen [2]. Nous ne le faisons pas dans cette Note.

(*) Séance du 22 mai 1978.

[1] G. Royer, *Processus de diffusion associé à certains modèles d'Ising à spins continus* [*Z. Wahrscheinlichkeitstheorie* (à paraître)].

[2] V. Strassen, *Annals Math. Stat.*, 36, 1965, p. 423-439.

J. M. O. : *Département de Mathématiques, Centre scientifique et polytechnique, Université Paris-Nord, avenue Jean-Baptiste-Clément*, 93430 *Villetaneuse;*

D. P. : *Laboratoire de Calcul des Probabilités, Université Pierre-et-Marie-Curie*, Tour 56, 4, *place Jussieu*, 75230 *Paris Cedex* 05.