Czechoslovak Mathematical Journal

Albert Pérez

Transformation ou σ-algèbre suffisante et minimum de la probabilité d'erreur

Czechoslovak Mathematical Journal, Vol. 7 (1957), No. 1, 115–123

Persistent URL: `http://dml.cz/dmlcz/100235`

Чехословацкий математический журнал, т. 7 (82) 1957, Прага

TRANSFORMATION OU σ-ALGÈBRE SUFFISANTE ET MINIMUM DE LA PROBABILITÉ D'ERREUR

ALBERT PÉREZ, Prague.

(Reçu le 29. novembre 1955.)

Après avoir caractérisé la classe des fonctions de décision donnant le minimum de la probabilité d'erreur, on étudie la convergence de ce minimum par rapport à une suite croissante de σ-algèbres de notre espace d'épreuves, ainsi que son rapport très étroit avec les transformations ou σ-algèbres suffisantes.

Dans ce qui suit nous considérons un espace mesurable $(X, \mathbf{S})$ et un système $\mathcal{M} = (\mu_1, \ldots, \mu_n)$ de n mesures de probabilité sur $\mathbf{S}$ avec les probabilités a priori respectives $p_1, \ldots, p_n, \sum_{i=1}^{n} p_i = 1$, dont le système est noté par $\mathcal{P}$. Soit $\mathbf{B} = (B_1, \ldots, B_n)$ une partition de $(X, \mathbf{S})$, c'est-à-dire, un système de n ensembles disjoints de $\mathbf{S}$, dont l'union est égale à X. La probabilité d'erreur que nous commettons en décidant que la mesure vraie est μ_i quand $x \in B_i$, $i = 1, \ldots, n$, est donnée par:

$$e_{\mathbf{B}}(\mathcal{M}, \mathcal{P}) = 1 - \sum_{i=1}^{n} p_i \mu_i(B_i)\,. \tag{1}$$

Il s'agit, tout d'abord, de déterminer la classe des partitions de $(X, \mathbf{S})$, dont la probabilité d'erreur respective obtient son minimum pour les $\mathcal{M}$ et $\mathcal{P}$ donnés.

Il s'agit, ensuite, d'étudier la suite des minima des probabilités d'erreur relatifs à une suite $\{\mathbf{S}_k\}$ croissante de sous-σ-algèbres de $\mathbf{S}$, et de donner, en particulier, des conditions pour que la suite ci-dessus des minima converge vers le minimum de la probabilité d'erreur relatif à la σ-algèbre $\mathbf{S}$.

Il s'agit, enfin, d'examiner s'il n'existe pas de sous-σ-algèbres de $\mathbf{S}$ telles, que les minima des probabilités d'erreur respectifs soient égaux au minimum de la probabilité d'erreur relatif à $\mathbf{S}$, indépendamment de la distribution a priori $\mathcal{P}$. Nous verrons que ce sont les sous-σ-algèbres de $\mathbf{S}$ suffisantes par rapport au système $\mathcal{M}$ de mesures considéré. L'intérêt qu'il y a à choisir parmi ces σ-algèbres la plus petite possible est évident.

I. La classe des partitions donnant le minimum de la probabilité d'erreur

Il s'agit de déterminer la classe des partitions de $(X, \mathbf{S})$ minimisant la probabilité d'erreur (1), le système de mesures de probabilité $\mathcal{M}$ ainsi que leur distribution a priori $\mathcal{P}$ étant donnés.

Remarquons que, malgré certaines analogies, ni dans le travail de Liapounoff[1]) ni dans aucune autre publication nous n'avons pas trouvé la réponse au problème posé ci-dessus.

Soit λ une mesure sur $\mathbf{S}$ dominant $\mathcal{M}$, $\mathcal{M} \ll \lambda$, et $f_1(x), \ldots, f_n(x)$ les densités de Radon-Nikodym respectives de $\mu_1, \ldots, \mu_n$ par rapport à λ [1]. Soit, pour $i, j = 1, 2, \ldots, n$,

$$M_{ij} = M_{ji} = \{x : p_i f_i(x) = p_j f_j(x)\} . \tag{2}$$

Appelons $K_{\mathbf{B}}$ la classe de partitions de $(X, \mathbf{S})$ contenant la partition $\mathbf{B} = (B_1, \ldots, B_n)$ et toute partition $\mathbf{C} = (C_1, \ldots, C_n)$ telle, que les relations suivantes soient valables modulo λ,

$$C_i \cap B_j \subset M_{ij} , \quad i \neq j , \quad i, j = 1, 2, \ldots, n .$$

A toute partition de cette classe correspond une et même valeur de la probabilité d'erreur. En effet, si la partition $\mathbf{C} \in K_{\mathbf{B}}$,

$$\sum_{i=1}^{n} p_i \mu_i(C_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i \mu_i(C_i \cap B_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_j \mu_j(C_i \cap B_j) = \sum_{j=1}^{n} p_j \mu_j(B_j) ,$$

puisqu'alors, pour $i \neq j$, $[\lambda]$ $C_j \cap B_j \subset M_{ij}$ et pour $x \in M_{ij}$ nous avons $p_i f_i(x) = p_j f_j(x)$. Par conséquent, $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{B}}(\mathcal{M}, \mathcal{P})$, d'où le théorème suivant

Théorème 1. *À toute partition de la classe $K_{\mathbf{B}}$ engendrée par la partition $\mathbf{B}$ correspond une probabilité d'erreur égale à $e_{\mathbf{B}}(\mathcal{M}, \mathcal{P})$.*

Soit, pour $i \neq j$,

$$N_{ij} = \{x : p_i f_i(x) > p_j f_j(x)\} , \quad i, j = 1, 2, \ldots, n , \tag{3}$$

$$L_{ij} = M_{ij} \quad \text{pour} \quad i < j , \quad L_{ij} = \emptyset \quad \text{pour} \quad i > j . \tag{4}$$

Soit $\mathbf{A} = (A_1, \ldots, A_n)$ la partition de $(X, \mathbf{S})$ définie par:

$$A_i = \bigcap_{\substack{j=1 \\ j \neq i}}^{n} (N_{ij} \cup L_{ij}) , \quad i = 1, 2, \ldots, n . \tag{5}$$

D'après le Théorème 1., à toute partition de la classe $K_{\mathbf{A}}$ engendrée par la partition $\mathbf{A}$ correspond une probabilité d'erreur égale à $e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$. Nous allons maintenant montrer le théorème suivant:

Théorème 2. *À la classe $K_{\mathbf{A}}$ correspond le minimum de la probabilité d'erreur.*

---

[1]) *А. А. Ляпунов:* О выборе из конечного числа законов распределения. Усп. мат. наук, том *6* (1951).

Démonstration. Soit $\mathbf{C} = (C_1, \ldots, C_n)$ une partition quelconque de $(X, \mathbf{S})$ et $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P})$ la probabilité d'erreur correspondante. Nous voulons montrer que $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) \geqq e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$. Pour cela remarquons que

$$e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) = 1 - \sum_{j,i=1}^{n} p_i \mu_i(A_j \cap C_i) \, , \quad e_{\mathbf{A}}(\mathcal{M}, \mathcal{P}) = 1 - \sum_{i,j=1}^{n} p_j \mu_j(A_j \cap C_i) \quad (6)$$

et que, d'après la définition de la partition $\mathbf{A}$,

$$p_i \mu_i(A_j \cap C_i) \leqq p_j \mu_j(A_j \cap C_i) \, , \quad i, j = 1, 2, \ldots, n \, . \quad (7)$$

En comparant les deux formules (6) à l'aide des relations (7) on obtient facilement le résultat voulu, c. q. f. d.

Théorème 3. *Toute partition de* $(X, \mathbf{S})$, $\mathbf{C} = (C_1, \ldots, C_n)$, *pour laquelle* $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$, *appartient à* $K_{\mathbf{A}}$.

Démonstration. En se reportant à la définition de la classe $K_{\mathbf{A}}$, nous constatons que toute partition $\mathbf{B} = (B_1, \ldots, B_n)$ de $K_{\mathbf{A}}$ est caractérisée uniquement par le fait que chacun de ses ensembles B_i satisfait modulo λ à la relation: $B_i \cap A_j \subset M_{ij}$ ou à la relation équivalente: $\lambda(B_i \cap N_{ji} \cap \bigcap_{i \neq k \neq j} L_{jk}) = 0$.

Si donc la partition $\mathbf{C}$ n'appartenait pas à $K_{\mathbf{A}}$, il y aurait au moins deux indices i_0 et j_0 différents tels, que $\lambda(C_{i_0} \cap D) > 0$ pour $D = N_{j_0 i_0} \cap \bigcap_{i_0 \neq k \neq j_0} L_{j_0 k}$ de sorte que

$$\sum_{i=1}^{n} p_i \mu_i(A_i \cap D) = p_{j_0} \mu_{j_0}(D) > \sum_{i=1}^{n} p_i \mu_i(C_i \cap D) \, .$$

En effet, compte tenu de ce que $p_{j_0} f_{j_0}(x) > p_{i_0} f_{i_0}(x)$ pour $x \in D$ nous avons $p_{i_0} \mu_{i_0}(C_{i_0} \cap D) < p_{j_0} \mu_{j_0}(C_{i_0} \cap D)$. Mais cela suffit pour que $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) < < e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$, en contradiction avec notre hypothèse que $e_{\mathbf{C}}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$. Donc $\mathbf{C}$ appartient nécessairement à la classe $K_{\mathbf{A}}$ c. q. f. d.

Corollaire. *La classe* $K_{\mathbf{A}}$ *coïncide avec la classe des partitions donnant, pour les* $\mathcal{M}$ *et* $\mathcal{P}$ *donnés, le minimum de la probabilité d'erreur.*

Démonstration. Elle résulte immédiatement des Théorèmes 2. et 3.

II. Sur la convergence de la suite des minima de la probabilité d'erreur relatifs à une suite croissante de σ-algèbres

Dans la section I, toutes les partitions utilisées étaient constituées d'ensembles de la σ-algèbre $\mathbf{S}$. Ceci, en particulier, était le cas pour la classe des partitions $K_{\mathbf{A}}$ donnant le minimum de la probabilité d'erreur correspondant aux $\mathcal{M}$ et $\mathcal{P}$ donnés.

Nous pouvons donc écrire pour cette dernière $e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$ au lieu de $e_{\mathbf{A}}(\mathcal{M}, \mathcal{P})$, ce que nous ferons par la suite.

Considérons maintenant une suite croissante $\{\mathbf{S}_k\}$ de sous-σ-algèbres de $\mathbf{S}$ et la suite $\{e_k(\mathcal{M}, \mathcal{P})\}$ des minima respectifs de la probabilité d'erreur.

Nous avons, évidemment,

$$e_1(\mathcal{M}, \mathcal{P}) \geqq e_2(\mathcal{M}, \mathcal{P}) \geqq \ldots \geqq e_{\mathbf{S}}(\mathcal{M}, \mathcal{P}) , \tag{8}$$

puisque, par hypothèse, $\mathbf{S}_1 \subset \mathbf{S}_2 \subset \ldots \subset \mathbf{S}$.

La question qui se pose maintenant est de savoir dans quelles conditions la limite de la suite (8) est égale à $e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$.

Soit $f_i^{(k)}(x)$ la densité de Radon-Nikodym de μ_i par rapport à λ mesurable par rapport à $\mathbf{S}_k$, $k = 1, 2, \ldots, i = 1, 2, \ldots, n$.

On s'aperçoit alors facilement (voir la définition de la classe $K_{\mathbf{A}}$ et, en particulier, la définition de la partition $\mathbf{A}$ dans (5) écrite à l'aide des densités mesurables par rapport à $\mathbf{S}_k$) que

$$e_k(\mathcal{M}, \mathcal{P}) = 1 - \int \max_i (p_i f_i^{(k)}(x))\, d\lambda \tag{9}$$

et que, si $k < s$, pour tout ensemble E de $\mathbf{S}_k$ nous avons

$$\int_E \max_i (p_i f_i^{(k)}(x))\, d\lambda \leqq \int_E \max_i (p_i f_i^{(s)}(x))\, d\lambda .$$

Il en résulte que le processus $\{\max_i(p_i f_i^{(k)}(x)), k \geqq 1\}$ est une semi-martingale (voir [2], chap. VII, 1.2s). Suivant le théorème 4.1s de [2], et puisque la limite de la suite (8) est finie, nous avons (partie (i) du théorème) que la limite de la suite $\{\max_i (p_i f_i^{(k)}(x))\}$, soit $g_\infty(x)$, existe presque sûrement $[\lambda]$.

D'autre part, pour chaque $i = 1, 2, \ldots, n$, la suite $\{f_i^{(k)}(x)\}$ constitue une martingale, puisque si $k < s$ nous avons pour tout ensemble E de $\mathbf{S}_k$

$$\int_E f_i^{(k)}(x)\, d\lambda = \int_E f_i^{(s)}(x)\, d\lambda = \int_E f_i(x)\, d\lambda .$$

Soit $f_i^{(\infty)}(x)$ la densité de μ_i par rapport à λ mesurable par rapport à la σ-algèbre engendrée par le système d'ensembles $\bigcup_{k=1}^{\infty} \mathbf{S}_k$, que nous désignons par $\mathbf{S}_\infty$. Évidemment, $\mathbf{S}_\infty \subset \mathbf{S}$.

D'après le théorème (4.4) (ii) de Doob [2] sur les martingales, nous avons

$$\lim_{k\to\infty} f_i^{(k)}(x) = f_i^{(\infty)}(x) \quad [\lambda] \tag{10}$$

et d'après cette rélation

$$\lim_{k\to\infty} \max_i(p_i f_i^{(k)}(x)) = g_\infty(x) = \max_i(p_i f_i^{(\infty)}(x)) \quad [\lambda] . \tag{11}$$

Mais alors aussi, puisque d'une part les $g_k(x) = \max_i(p_i f_i^{(k)}(x))$ sont non-négatifs et d'autre part le processus $\{g_k(x), \mathbf{S}_k, 1 \leqq k \leqq \infty\}$ est une semi-martingale, il résulte (voir théorème 4.1s de Doob cité précédemment, partie (ii,b)) que

$$\lim_{k\to\infty} \int g_k(x)\,\mathrm{d}\lambda = \int g_\infty(x)\,\mathrm{d}\lambda\,. \tag{12}$$

D'où le théorème suivant:

Théorème 4. *Nous avons toujours*

$$\lim_{k\to\infty} e_k(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}_\infty}(\mathcal{M}, \mathcal{P})\,.$$

Par la suite nous ferons usage du théorème 1 (p. 233) de [1], adapté dans le langage de σ-algèbres suffisantes par Bahadur [3]; voir aussi [4], spécialement le paragraphe *24.4. Formulons ce théorème sous la forme suivante:

Lemme 1. *Une condition nécessaire et suffisante pour que la sous-σ-algèbre $\mathbf{S}_0$ de $\mathbf{S}$ soit suffisante pour un système dominé $\mathcal{M}$ de mesures sur $\mathbf{S}$ est qu'il existe une mesure λ sur $\mathbf{S}$ équivalente par rapport à $\mathcal{M}$ et que la densité de μ par rapport à λ soit presque sûrement $[\lambda]$ égale à une fonction mesurable par rapport à $\mathbf{S}_0$, pour chaque μ de $\mathcal{M}$.*

Dans notre cas il est toujours possible de choisir la mesure λ équivalente par rapport à notre système de mesures $\mathcal{M}$. Rappelons la définition: λ est équi valente par rapport à $\mathcal{M}$, si elle domine $\mathcal{M}$ et si elle s'annule pour chaque ensemble de $\mathbf{S}$, sur lequel toutes les mesures de $\mathcal{M}$ s'annulent.

Sous cette hypothèse, nous pouvons formuler le théorème suivant,

Théorème 5. *Pour que* $\lim_{k\to\infty} e_k(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$ *il suffit que, pour tout i, la suite $\{f_i^{(k)}\}$ converge en λ-mesure vers $f_i(x)$, ou, ce qui revient au même, que $f_i^{(\infty)} = f_i\ [\lambda]$,* ce *qui de sa part est équivalent à ce que la σ-algèbre $\mathbf{S}_\infty$ soit suffisante par rapport au système de mesures $\mathcal{M}$.*

Démonstration. Le fait que la première condition est ici équivalente à la seconde résulte de la rélation (10). Le fait que la seconde condition est équivalente à la troisième est une conséquence immédiate du Lemme 1. Partant de la condition $f_i^{(\infty)} = f_i[\lambda]$, $i = 1, \ldots, n$, nous obtenons $e_{\mathbf{S}}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}_\infty}(\mathcal{M}, \mathcal{P})$, ce qui, en combinaison avec la Théorème 4, nous donne le résultat voulu, c. q. f. d.

La question qui se pose maintenant est de savoir s'il existe une suite décroissante $\{\mathbf{P}_k\}$ de partitions finies de $(X, \mathbf{S})$ telle, que la suite (croissante) $\{\mathbf{S}_k\}$ des σ-algèbres, dont chacune est engendrée par le système d'ensembles de la partition respective, nous assure les conditions du Théorème 5. Rappelons qu'une suite de partitions est décroissante, si chacun des ensembles de chaque partition de la suite est contenu dans un ensemble de la partition précédente.

Lemme 2. *Il existe toujours une suite décroissante $\{\mathbf{P}_k\}$ de partitions de $(X, \mathbf{S})$ telle, que les conditions du Théorème 5 soient assurées, les densités étant prises par rapport aux σ-algèbres engendrées par les $\mathbf{P}_k$ respectifs.*

Démonstration. Pour tout $i = 1, \ldots, n$, et pour $k = 1, 2, \ldots$, écrivons $E_{kj}^{(i)} = \{x : (j-1)/2^k \leqq f_i(x) < j/2^k\}$, $j = 1, 2, \ldots, k \,.\, 2^k$, et de même $E_k^{(i)} = \{x : f_i(x) \geq k\}$. Le système d'ensembles $\{E_{kj}^{(i)}, E_k^{(i)}\}$ constitue pour chaque i et k une partition $\mathbf{P}_k^{(i)}$ de $(X, \mathbf{S})$ et la suite $\{\mathbf{P}_k^{(i)}\}$ est, pour chaque i, décroissante. En prenant comme $\mathbf{P}_k$ l'intersection (produit) de tous les $\mathbf{P}_k^{(i)}$, $i = 1, \ldots, n$, la suite $\{\mathbf{P}_k\}$ nous assure la convergence en λ-mesure de chaque suite $\{f_i^{(k)}\}$ vers le f_i correspondant. En effet, pour chaque i, $\lim_{k \to \infty} \lambda(E_k^{(i)}) = 0$ et, pour $x \in X - \bigcup_{i=1}^{n} E_k^{(i)}$, nous avons $|f_i^{(k)}(x) - f_i(x)| < \frac{1}{2^k}$ $[\lambda]$, c. q. f. d.

Remarque. La construction de la suite $\{\mathbf{P}_k\}$ est telle, que la σ-algèbre $\mathbf{S}_\infty$ résultante est la plus petite σ-algèbre par rapport à laquelle tous les $f_i(x)$ sont mesurables. $\mathbf{S}_\infty$ est alors une σ-algèbre suffisante minimale pour le système de mesures $\mathcal{M}$. (Voir, par exemple, [4], paragraphe *24.4.)

Partant du Théorème 5 et du Lemme 2 nous obtenons le

Théorème 6. *Il existe toujours une suite décroissante de partitions finies de $(X, \mathbf{S})$ telle, que la suite des minima respectifs de la probabilité d'erreur converge vers le minimum de la probabilité d'erreur correspondant à $\mathbf{S}$.*

Remarquons que la dite suite de partitions peut dans certains cas ne pas dépendre des mesures (et densités) impliquées. Ceci est le cas, par exemple, d'un expace $(X, \mathbf{S})$ métrique séparable. Toute suite régulière de partitions de cet espace nous assure alors la convergence, dont il est question dans le Théorème 6.

Rappelons que par suite régulière de partitions d'un espace métrique nous entendons une suite décroissante de partitions $\{\mathbf{P}_k\}$ telle, que le supremum des diamètres des ensembles constituant $\mathbf{P}_k$ tend vers zéro, quand k tend vers l'infini.

Un autre cas important, où sont assurées les conditions du Théorème 5, c'est le cas, où $(X, \mathbf{S})$ est un produit cartésien infini, $(X, \mathbf{S}) = \prod_{i=1}^{\infty} (X_i, \mathcal{S}_i)$. La suite croissante de σ-algèbres $\{\mathbf{S}_k = \prod_{i=1}^{k} \mathcal{S}_i\}$ engendre alors une σ-algèbre $\mathbf{S}_\infty$ qui, par définition, est égale à $\mathbf{S}$. Dans ce cas, nous aurons alors toujours, indépendamment de $\mathcal{M}$ et $\mathcal{P}$, que $\lim_{k \to \infty} e_{\mathbf{S}_k}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$.

III. Transformation ou σ-algèbre suffisante et minimum de la probabilité d'erreur

Le fait qu'il peut exister, mais pas toujours, de sous-σ-algèbres de $\mathbf{S}$ essentiellement inférieures à $\mathbf{S}$, pour lesquelles le minimum de la probabilité d'erreur n'est pas, comme dans le cas général, supérieur à celui correspondant à $\mathbf{S}$, mais égal, présente une grande importance.

Déjà dans le Lemme 1 et dans le Théorème 5 nous avons rencontré la notion de σ-algèbre suffisante et nous avons constaté que, si $\mathbf{S}_\infty$ est suffisante par rapport au système de mesures $\mathcal{M}$, alors $e_{\mathbf{S}_\infty}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$, quelle que soit la distribution a priori $\mathcal{P}$. La question qui se pose maintenant est de savoir si cette condition est aussi nécessaire.

Pour cela il nous faut encore certains résultats de Halmos et Savage [1] (Théorèmes 2 et 3 de [1]), que nous allons exprimer cependant non pas dans le langage de transformations suffisantes, mais dans celui de σ-algèbres suffisantes.

Lemme 4. *Une condition nécessaire et suffisante pour que la sous-σ-algèbre $\mathbf{S}_0$ de $\mathbf{S}$ soit suffisante pour un système dominé $\mathcal{M}$ de mesures sur $\mathbf{S}$ est que, pour toute paire μ_1 et μ_2 de mesures dans $\mathcal{M}$, le rapport des densités respectives soit égal $[\mu_1 + \mu_2]$ à une fonction mesurable par rapport à $\mathbf{S}_0$.*

Nous sommes maintenant en mesure de montrer le théorème suivant:

Théorème 7. *Pour que, quelle que soit la distribution a priori $\mathcal{P}$, le minimum de la probabilité d'erreur correspondant à $\mathbf{S}_0$, sous-σ-algèbre de $\mathbf{S}$, soit égal à celui correspondant à $\mathbf{S}$, il faut et il suffit que $\mathbf{S}_0$ soit une σ-algèbre suffisante par rapport à $\mathcal{M}$.*

Démonstration. Le „il suffit" est déjà contenu dans le Théorème 5. Pour montrer le „il faut"[2]), nous allons exploiter le fait que, par hypothèse, $e_{\mathbf{S}_0}(\mathcal{M},\mathcal{P}) = e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$, quel que soit $\mathcal{P}$. Nous allons pouvoir ainsi étudier séparement le cas pour chaque paire μ_i, μ_j de mesures de $\mathcal{M}$, en prenant $\mathcal{P} = (0, 0, \ldots, p_i, 0, \ldots, p_j, 0, \ldots, 0)$, et respectant toujours le fait que la somme des probabilités sera dans chaque cas égale à l'unité. Une fois prouvé que, pour toute paire de mesures de $\mathcal{M}$, les conditions du Lemme 4 sont satisfaites, nous allons pouvoir conclure, suivant ce lemme, que $\mathbf{S}_0$ est suffisante pour $\mathcal{M}$. Il suffit de le montrer, par exemple, pour la paire μ_1, μ_2, la numérotation étant arbitraire. C'est ce que nous ferons par la suite.

Soient $f_1(x)$ et $f_2(x)$ les densités respectives de μ_1 et μ_2 par rapport à $\mu_1 + \mu_2$, mesurables par rapport à $\mathbf{S}$ et $f_1^{(0)}(x)$ et $f_2^{(0)}(x)$ celles mesurables par rapport à $\mathbf{S}_0$.

Soit $K_{\mathbf{A}}$ la classe des partitions de $(X, \mathbf{S})$ donnant le minimum de la probabilité d'erreur $e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$ et $K_{\mathbf{A}^{(0)}}$ celle des partitions de $(X, \mathbf{S}_0)$ donnant $e_{\mathbf{S}_0}(\mathcal{M}, \mathcal{P})$.

D'après le Corollaire de la section I, l'hypothèse de l'égalité $e_{\mathbf{S}_0}(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$ entraîne que $K_{\mathbf{A}^{(0)}} \subset K_{\mathbf{A}}$. En particulier, la partition $\mathbf{A}^{(0)} \in K_{\mathbf{A}}$.

[2]) On remarquera que le „il faut" n'est pas une conséquence des résultats de *Bahadur* ([3], notamment, § 7), puisque dans notre cas nous n'avons nullement besoin, comme Bahadur, de toutes les fonctions de décision possibles, mais seulement de celles qui minimisent la probabilité d'erreur.

Nous pouvons alors facilement se convaincre (voir définition de la classe $K_{\mathbf{A}}$ en rapport avec la démonstration du Théorème 3) que

$$A_1^{(0)} \cap N_{21} = A_2^{(0)} \cap N_{12} = \emptyset \quad [\mu_1 + \mu_2]\,.$$

Autrement dit,

$$\{x : p_1 f_1^{(0)}(x) > p_2 f_2^{(0)}(x)\} \cap \{x : p_1 f_1(x) < p_2 f_2(x)\} =$$
$$= \{x : p_1 f_1^{(0)}(x) < p_2 f_2^{(0)}(x)\} \cap \{x : p_1 f_1(x) > p_2 f_2(x)\} = \emptyset \quad [\mu_1 + \mu_2]\,,$$

ce qui encore s'écrit, en posant $p_1/p_2 = r$ et tenant compte du fait que les rapports $f(x) = f_1(x)/f_2(x)$, $f^{(0)}(x) = f_1^{(0)}(x)/f_2^{(0)}(x)$ existent $[\mu_1 + \mu_2]$,

$$\{x : f^{(0)}(x) > r\} \cap \{x : f(x) < r\} = \{x : f^{(0)}(x) < r\} \cap \{x : f(x) > r\} = \emptyset\, [\mu_1 + \mu_2].$$

Tenant maintenant de nouveau compte du fait que la distribution a priori est arbitraire, laissons r parcourir toutes les valeurs (non négatives) de l'ensemble R des nombres rationnels, qui est, comme on le sait, dénombrable. Nous aurons alors:

$$\bigcup_{r \in R} (\{x : f^{(0)}(x) > r\} \cap \{x : f(x) < r\}) =$$
$$= \bigcup_{r \in R} (\{x : f^{(0)}(x) < r\} \cap \{x : f(x) > r\}) = \emptyset \quad [\mu_1 + \mu_2]\,,$$

ce qui de sa part signifie, R étant partout dense, que $f^{(0)}(x) = f(x)$ $[\mu_1 + \mu_2]$.

Ainsi les conditions du Lemme 4 sont satisfaites, c. q. f. d.

RÉFÉRENCES

[1] *P. Halmos* et *L. Savage:* „Application of the Radon-Nikodym theorem to the theory of sufficient statistics", Ann. Math. Stat., Vol. *20*, 1949, 225—241.

[2] *J. Doob:* „Stochastic Processes", New York, 1953.

[3] *R. Bahadur:* „Sufficiency and Statistical decision functions", Ann. Math. Stat., Vol. *25*, 1954, 423—462.

[4] *M. Loève:* „Probability Theory", New York, 1955.

Резюме

ДОСТАТОЧНОЕ ПРЕОБРАЗОВАНИЕ ИЛИ σ-АЛГЕБРА И МИНИМУМ ВЕРОЯТНОСТИ ОШИБКИ

АЛЬБЕР ПЕРЕЗ (Albert Pérez), Прага.

(Поступило в редакцию 29/XI 1955 г.)

В работе рассматривается измеримое пространство $(X, \mathbf{S})$ и система $\mathcal{M} = (\mu_1, \ldots, \mu_n)$ n вероятностных мер на $\mathbf{S}$ с априорными вероятностями,

равными соответственно, $p_1, \ldots, p_n$, $\sum_{i=1}^{n} p_i = 1$, причем система этих вероятностей обозначается через $\mathcal{P}$. Пусть $\mathbf{B} = (B_1, \ldots, B_n)$ – разбиение пространства $(X, \mathbf{S})$, т. е. система n непересекающихся множеств из $\mathbf{S}$, соединение которых равно X. Вероятность ошибки, $e_{\mathbf{B}}(\mathcal{M}, \mathcal{P})$, которую мы сделаем, принимая за настоящую меру μ_i при $x \in B_i$, $i = 1, \ldots, n$, дана формулой (1).

Цель отдела I — дать характеристику класса разбиений $\mathbf{B}$, минимализирующих $e_{\mathbf{B}}(\mathcal{M}, \mathcal{P})$. Пусть λ — мера на $\mathbf{S}$ такая, что все $\mu_i \in \mathcal{M}$ абсолютно непрерывны по отношению к λ, пусть $f_1(x), \ldots, f_n(x)$ означают, соответственно, производные мер $\mu_1, \ldots, \mu_n$ по отношению к λ [1].

Определим, далее, множество M_{ij} соотношением (2) и обозначим через $K_{\mathbf{B}}$ класс разбиений пространства $(X, \mathbf{S})$, содержащий $\mathbf{B}$ и всякое разбиение $\mathbf{C} = (C_1, \ldots, C_n)$ такое, что при $i \neq j$, $C_i \cap B_j \subset M_{ij}$ по модулю λ. Каждому разбиению из $K_{\mathbf{B}}$ соответствует вероятность ошибки, равная $e_{\mathbf{B}}(\mathcal{M}, \mathcal{P})$ (Теорема 1), и класс $K_{\mathbf{A}}$, где $\mathbf{A}$ определенно соотношениями (3), (4), (5), совпадает с классом разбиений, дающих для данных $\mathcal{M}$ и $\mathcal{P}$ минимум вероятности ошибки, обозначенный через $e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$ (Теоремы 2, 3).

В отделе II рассматривается возрастающая последовательность $\{\mathbf{S}_k\}$ под-σ-алгебр из $\mathbf{S}$ и соответствующая последовательность $\{e_k(\mathcal{M}, \mathcal{P})\}$ минимумов вероятности ошибки. Справедливы соотношения (8), (9) и (10), где $f_i^{(k)}$ означают производные мер μ_i по отношению к λ, измеримые по отношению к $\mathbf{S}_k$, $k = 1, 2, \ldots, \infty$, причем $\mathbf{S}_\infty$ означает минимальную σ-алгебру над $\bigcup_{k=1}^{\infty} \mathbf{S}_k$. Процесс $\{g_k(x) = \max_i p_i f_i^{(k)}(x),\ 1 \leqq k \leqq \infty\}$, является семимартингалом (смотри [2]), и, вследствие (11), справедливо также (12), откуда $\lim_{k \to \infty} e_k(\mathcal{M}, \mathcal{P}) = e_{\mathbf{S}_\infty}(\mathcal{M}, \mathcal{P})$ (Теорема 4).

Пусть $\mathbf{S}_0$ — под-σ-алгебра из $\mathbf{S}$ (например, $\mathbf{S}_\infty$). В общем случае $e_{\mathbf{S}_0}(\mathcal{M}, \mathcal{P}) \geqq$ $\geqq e_{\mathbf{S}}(\mathcal{M}, \mathcal{P})$; знак равенства имеет для произвольного $\mathcal{P}$ место тогда и только тогда, если $\mathbf{S}_0$ является σ-алгеброй, достаточной для $\mathcal{M}$. (Теорема 7, отдел III). Этот результат не является следствием результатов, полученных Бахадуром [3].