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

J. R. BARRA

À propos d'un résultat de Brailovsky concernant une probabilité d'erreur en analyse discriminante

Annales de l'I. H. P., section B, tome 17, n° 1 (1981), p. 21-29

<http://www.numdam.org/item?id=AIHPB_1981__17_1_21_0>

© Gauthier-Villars, 1981, 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/

Ann. Inst. Henri Poincaré, Section B :
Vol. XVII, n° 1, 1981, p. 21-29. *Calcul des Probabilités et Statistique.*

A propos d'un résultat de Brailovsky concernant une probabilité d'erreur en analyse discriminante

par

J. R. BARRA

IRMA B. P. 53, 38041 Grenoble Cedex

Dans un formatisme plus rigoureux et sous des hypothèses plus générales, on retrouve par une méthode différente un résultat de [4]; de plus on propose une explication à la situation apparemment paradoxale mise en évidence par V. Brailovsky dans ce même article.

1. INTRODUCTION

Considérons la structure statistique $\mathscr{S}_0$:

$$[\mathbb{R}^k, \mathscr{B}_{\mathbb{R}_k}; N(\mu_0, \Sigma), N(\mu_1, \Sigma)]$$

et la statistique, appelée fonction discriminante linéaire d'Anderson (cf. [1]) :

$$x \in \mathbb{R}^k \quad W(x) = 2{}^t(\mu_1 - \mu_0)\Sigma^{-1}x + {}^t\mu_0\Sigma^{-1}\mu_0 - {}^t\mu_1\Sigma^{-1}\mu_1 .$$

Si les deux hypothèses $\{ \mu_0 \}$ et $\{ \mu_1 \}$ ont des probabilités *a priori* respectives p_0 et p_1, le test de l'hypothèse $\{ \mu_0 \}$ contre l'hypothèse $\{ \mu_1 \}$ défini par la région critique :

$$W(x) > -2 \log \frac{p_1}{p_0}$$

© Gauthier-Villars

rend minimum la probabilité d'erreur, qui vaut alors :

$$P_0 = p_0 N\left(-\frac{D}{2} + c\right) + p_1 N\left(-\frac{D}{2} - c\right) \tag{1}$$

où :

$$D^2 = {}^t(\mu_1 - \mu_0)\Sigma^{-1}(\mu_1 - \mu_0),\ c = \frac{1}{D}\log\frac{p_1}{p_0}. \tag{2}$$

Soit maintenant la structure statistique $\mathscr{S}$

$$\left[\mathbb{R}^{3k}, \mathscr{B}_{\mathbb{R}^{3k}}, N\left(\mu_0, \frac{1}{n}\Lambda\right) \times N\left(\mu_1, \frac{1}{n'}\Lambda\right) \times N((1-\varepsilon)\mu_0 + \varepsilon\mu_1, \Sigma),\ \mu_0 \in \mathbb{R}^k,\ \mu_1 \in \mathbb{R}^k,\ \varepsilon = 0,1\right];$$

en Analyse Discriminante, on considère le test de l'hypothèse $\{\varepsilon = 0\}$, de probabilité *a priori* p_0, contre l'hypothèse $\{\varepsilon = 1\}$, de probabilité *a priori* p_1, défini par la région critique $\left\{H > -2\log\frac{p_1}{p_0}\right\}$ où la statistique H, directement déduite de W, est l'application

$$\hat{\mu}_0 \in \mathbb{R}^k,\ \hat{\mu}_1 \in \mathbb{R}^k, x \in \mathbb{R}^k \to 2{}^t(\hat{\mu}_1 - \hat{\mu}_0)\Sigma^{-1}x + {}^t\hat{\mu}_0\Sigma^{-1}\hat{\mu}_0 - {}^t\hat{\mu}_1\Sigma^{-1}\hat{\mu}_1.$$

Dans [4] Brailovsky étudie le cas particulier $p_0 = p_1$, $n = n' = N$, et calcule les deux premiers termes du développement limité en $\frac{1}{N}$ de la probabilité d'erreur correspondante au test ci-dessus; en particulier si $\Lambda = \Sigma$, il retrouve un cas particulier d'un résultat d'Okamoto [3], à savoir que la probabilité d'erreur est égale à :

$$N\left(-\frac{D}{2}\right) + \frac{1}{4N\sqrt{2\pi}} e^{-\frac{D^2}{8}}\left(\frac{D}{2} + \frac{2}{D}(k-1)\right) + 0(N^{-3/2}). \tag{3}$$

Il étudie ensuite cette probabilité dans certains cas particuliers où la matrice Λ diffère de Σ.

Dans le cas général, nous étudions ci-après, en menant les calculs différemment que V. Brailovsky, la probabilité d'erreur P associée au test ci-dessus de $\{\varepsilon = 0\}$ contre $\{\varepsilon = 1\}$ où on constate aisément que P est une fonction de la seule différence $\mu = \mu_1 - \mu_0$.

On remarquera que la même méthode s'applique si les covariances de $\hat{\mu}_0$ et $\hat{\mu}_1$ sont différentes, mais les formules finales sont moins intéressantes.

2. DÉVELOPPEMENT LIMITÉ A L'ORDRE 2 DE LA PROBABILITÉ D'ERREUR

Soit

$$\sigma^2 = {}^t(\hat{\mu}_1 - \hat{\mu}_0)\Sigma^{-1}(\hat{\mu}_1 - \hat{\mu}_0)$$
$$F = 2cD + {}^t\hat{\mu}_0\Sigma^{-1}\hat{\mu}_0 - {}^t\hat{\mu}_1\Sigma^{-1}\hat{\mu}_1$$
$$G = -2cD + {}^t\hat{\mu}_1\Sigma^{-1}\hat{\mu}_1 - {}^t\hat{\mu}_0\Sigma^{-1}\hat{\mu}_0 + 2{}^t\mu\Sigma^{-1}(\hat{\mu}_0 - \hat{\mu}_1).$$

La probabilité d'erreur conditionnelle à $(\hat{\mu}_0, \hat{\mu}_1)$ est égale à :

$$p_0 N\left(\frac{F}{2\sigma}\right) + p_1 N\left(\frac{G}{2\sigma}\right). \tag{4}$$

Soit A une matrice telle que $A{}^tA = \Lambda$, on peut écrire :

$$\hat{\mu}_0 = \frac{1}{\sqrt{n}} AZ, \quad \hat{\mu}_1 = \mu + \frac{1}{\sqrt{n'}} AZ',$$

où Z et Z' sont des vecteurs aléatoires indépendants et de loi $N(0, \mathbb{1}_k)$. Soit alors

$$\alpha D^2 = {}^tA\Sigma^{-1}\mu$$

et soit Q la forme quadratique sur $\mathbb{R}^k$ définie par la matrice

$$\frac{1}{2D^2} {}^tA\Sigma^{-1}A$$

on a :

$$\frac{F}{2D} = -\frac{D}{2} + c - \frac{D}{\sqrt{n'}} {}^t\alpha Z' + \frac{D}{n} Q(Z) - \frac{D}{n'} Q(Z')$$
$$\frac{G}{2D} = -\frac{D}{2} - c + \frac{D}{\sqrt{n}} {}^t\alpha Z + \frac{D}{n'} Q(Z') - \frac{D}{n} Q(Z)$$
$$\frac{\sigma^2}{D^2} = 1 + 2{}^t\alpha\left(\frac{Z'}{\sqrt{n'}} - \frac{Z}{\sqrt{n}}\right) + 2Q\left(\frac{Z'}{\sqrt{n'}} - \frac{Z}{\sqrt{n}}\right).$$

D'où :

$$\frac{D}{\sigma} = 1 + {}^t\alpha\left(\frac{Z}{\sqrt{n}} - \frac{Z'}{\sqrt{n'}}\right) - Q\left(\frac{Z'}{\sqrt{n'}} - \frac{Z}{\sqrt{n}}\right) + \frac{3}{2}\left[{}^t\alpha\left(\frac{Z'}{\sqrt{n'}} - \frac{Z}{\sqrt{n}}\right)\right]^2 + 0\left(\frac{1}{n} + \frac{1}{n'}\right)$$

où, comme dans toute la suite, $0\left(\frac{1}{n}+\frac{1}{n'}\right)$ désigne un infiniment petit dont le produit avec n et le produit avec n' tendent vers 0 quand n et n' tendent vers l'infini. On en déduit que :

$$\frac{\mathrm{F}}{2\sigma} = x + y + z + 0\left(\frac{1}{n}+\frac{1}{n'}\right)$$

où

$$x = c - \frac{\mathrm{D}}{2},\quad y = -{}^t\alpha\left[\left(\frac{\mathrm{D}}{2}-c\right)\frac{\mathrm{Z}}{\sqrt{n}} + \left(\frac{\mathrm{D}}{2}+c\right)\frac{\mathrm{Z}'}{\sqrt{n'}}\right].$$

$$z = \left(\frac{3\mathrm{D}}{2}-c\right)\cdot\frac{\mathrm{Q}(\mathrm{Z})}{n} - \left(\frac{\mathrm{D}}{2}+c\right)\cdot\frac{\mathrm{Q}(\mathrm{Z}')}{n'}$$

$$+\ \frac{1}{2}{}^t\alpha\left(\frac{\mathrm{Z}'}{\sqrt{n'}}-\frac{\mathrm{Z}}{\sqrt{n}}\right)^t\left[3\left(\frac{\mathrm{D}}{2}-c\right)\frac{\mathrm{Z}}{\sqrt{n}} + \left(\frac{\mathrm{D}}{2}+3c\right)\frac{\mathrm{Z}'}{\sqrt{n'}}\right]\alpha + \frac{2c-\mathrm{D}}{\sqrt{nn'}}q(\mathrm{Z},\mathrm{Z}')$$

où q est la forme bilinéaire à Q.

D'autre part on a :

$$\mathrm{N}\left(x+y+z+0\left(\frac{1}{n}+\frac{1}{n'}\right)\right) = \mathrm{N}(x) + \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\left(y+z-\frac{x}{2}y^2\right) + 0\left(\frac{1}{n}+\frac{1}{n'}\right),$$

et en prenant l'espérance mathématique en Z et Z′ il vient :

$$\mathrm{E}(y) = 0,\quad \mathrm{E}(y^2) = {}^t\alpha\alpha\left(\frac{1}{n'}\left(c+\frac{\mathrm{D}}{2}\right)^2 + \frac{1}{n}\left(c-\frac{\mathrm{D}}{2}\right)^2\right),$$

$$\mathrm{E}(z) = \frac{{}^t\alpha\alpha}{2}\left(\left(\frac{\mathrm{D}}{2}+3c\right)\frac{1}{n'} + 3\left(c-\frac{\mathrm{D}}{2}\right)\frac{1}{n}\right) - \frac{\mathrm{Tr}(\Lambda\Sigma^{-1})}{2\mathrm{D}^2}\left[\frac{1}{n'}\left(c+\frac{\mathrm{D}}{2}\right)+\frac{1}{n}\left(c-\frac{3\mathrm{D}}{2}\right)\right],$$

puisque ([2], p. 90) on sait que

$$\mathrm{E}(\mathrm{Q}(\mathrm{Z})) = \mathrm{E}(\mathrm{Q}(\mathrm{Z}')) = \frac{1}{2\mathrm{D}^2}\mathrm{Tr}({}^t\mathrm{A}\Sigma^{-1}\mathrm{A}) = \frac{\mathrm{Tr}(\Lambda\Sigma^{-1})}{2\mathrm{D}^2}.$$

D'autre part on a :

$${}^t\alpha\alpha = \frac{1}{\mathrm{D}^4}{}^t\mu\Sigma^{-1}\mathrm{A}{}^t\mathrm{A}\Sigma^{-1}\mu = \frac{1}{\mathrm{D}^4}{}^t\mu\Sigma^{-1}\Lambda\Sigma^{-1}\mu.$$

On a de même :

$$\frac{G}{2\sigma} = x' + y' + z' + 0\left(\frac{1}{n} + \frac{1}{n'}\right)$$

où x', y', z' sont respectivement déduits de x, y, z par le changement :

$$c \to -c,\ \alpha \to -\alpha,\ n \to n',\ n' \to n.$$

En tenant compte de (1) et de (4) il vient :

$$P = P_0 + \frac{p_0}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\left[E(z) - \frac{x}{2}E(y^2)\right] + \frac{p_1}{\sqrt{2\pi}} e^{-\frac{x'^2}{2}}\left[E(z') - \frac{x'^2}{2}E(y'^2)\right] + 0\left(\frac{1}{n} + \frac{1}{n'}\right)$$

en posant :

$$K = \sqrt{\frac{p_0 p_1}{2\pi}} \exp\left\{-\frac{D^2}{8} - \frac{c^2}{2}\right\}$$

et en utilisant (2) on obtient :

$$P = P_0 + K\left[E(z) + E(z') - \frac{x}{2}E(y^2) - \frac{x'}{2}E(y'^2)\right] + 0\left(\frac{1}{n} + \frac{1}{n'}\right).$$

On note alors :

$$S = \frac{1}{D^2}\,{}^t\mu\Sigma^{-1}\Lambda\Sigma^{-1}\mu,\ R = \mathrm{Tr}\,(\Lambda\Sigma^{-1}) - S$$

$$K' = \frac{D}{2} + \frac{2c^2}{D} + 2c\,\frac{n - n'}{n + n'},$$

et on obtient la formule finale

$$P = P_0 + \left(\frac{1}{n} + \frac{1}{n'}\right)\frac{K}{4}\left[\frac{2}{D}R + K'S\right] + 0\left(\frac{1}{n} + \frac{1}{n'}\right). \qquad (5)$$

On remarque que R et S sont des fonctions linéaires de Λ et que K' est toujours positif, P est donc une fonction croissante de R et S.

3. LES RÉSULTATS DE BRAILOVSKY

Soit Σ_d la matrice diagonale qui a même diagonale que Σ; Brailovsky établit la proposition suivante dont la démonstration est reprise dans [5].

THÉORÈME. — « *Si* ${}^t\mu = (1, 0, \ldots, 0)$ *et si* $\Lambda = \Sigma_d + \beta\Sigma_e (\beta \leq 1)$ *alors* P *est supérieure à l'expression* (3) ».

Brailovsky considère alors trois situations pratiques intéressantes et relevant de ce théorème.

a) Si à propos de la structure statistique $\mathscr{S}_0$ on estime *a priori* successivement et indépendamment les diverses composantes de μ_0 et de μ_1 par des échantillons de tailles respectives n et n', alors on obtient $\mathscr{S}$ avec $\Lambda = \Sigma_d$.

b) Soit maintenant λ un rationnel supérieur à 1 et tel que $\lambda n, \lambda n', \dfrac{n}{n}, \dfrac{n'}{\lambda}$ soient entiers; Brailovsky considère le cas où dans les échantillons respectifs $[\{ {}^0X_j^i, j = 1, \ldots, \lambda_n, i = 1, \ldots, k \}, \{ {}^1X_j^i, j = 1, \ldots, \lambda n', i = 1, \ldots, k \}]$ des lois $N(\mu_0, \Sigma)$ et $N(\mu_1, \Sigma)$, il manque des observations de la façon suivante : pour tout $i = 1, \ldots, k$, la composante d'indice j c'est-à-dire ${}^0X_j^i$ (resp. ${}^1X_j^i$) n'est observée que si $j \in I_0^i$ (resp. $j \in I_1^i$) où les parties I_0^i (resp. I_1^i) de $\{ 1, \ldots, \lambda_n \}$, (resp. $\{ 1, \ldots, \lambda n' \}$) sont telles que :

$$\forall i = 1, \ldots, k, \text{ card } I_0^i = n, \text{ card } I_1^i = n'$$

$$\forall i \neq i', = 1, \ldots, k, \text{ card } (I_0^i \cap I_0^{i'}) = \frac{n}{\lambda}, \text{ card } (I_1^i \cap I_1^{i'}) = \frac{n'}{\lambda}.$$

Brailovsky estime alors μ_0 et μ_1 comme suit :

$$\forall i = 1, \ldots, k \quad \hat{\mu}_0^i = \frac{1}{n} \sum_{j \in I_0^i} {}^0X_j^i \qquad \hat{\mu}_1^i = \frac{1}{n'} \sum_{j \in I_1^i} {}^1X_j^i,$$

et il obtient encore $\mathscr{S}$ avec :

$$\Lambda = \Sigma_d + \frac{1}{\lambda}\Sigma_e.$$

c) Brailovsky suppose enfin que pour estimer μ_0 (resp. μ_1) on dispose d'un échantillon de taille n de la loi $N(\mu_0, \Sigma)$ (resp. $N(\mu_1, \Sigma)$ et, en plus, d'un échantillon indépendant du précédent et de taille γn (resp. $\gamma n'$) de la loi $N(\mu_0, \Sigma_d)$ (resp. $N(\mu_1, \Sigma_d)$), où γ est un rationnel donné tel que γn et $\gamma n'$ soient entiers.

Soient $\bar{x}_0$, (resp. $\bar{x}_1$), $\bar{x}'_0$, (resp. $\bar{x}'_1$) les moyennes empiriques respectives de ces 4 échantillons, Brailovsky estime μ_0 et μ_1, respectivement par :

$$\hat{\mu}_0 = \frac{\bar{x}_0 + \gamma\bar{x}'_0}{1 + \gamma} \qquad \hat{\mu}_1 = \frac{\bar{x}_1 + \gamma\bar{x}'_1}{1 + \gamma},$$

et obtient encore $\mathscr{S}$ avec :

$$\Lambda = \frac{1}{1+\gamma}\Sigma_d + \frac{1}{(1+\gamma)^2}\Sigma_e .$$

En particulier pour $k = 2$, en désignant par ρ le coefficient de corrélation associé à Σ, on obtient :

$$\mathscr{S}(\Sigma_d) = \frac{1+\rho^2}{1-\rho^2},\ \mathrm{R}(\Sigma_d) = 1 ,$$

et donc si ${}^t\mu = (1, 0)$, si $c = 0$ (probabilités *a priori* égales) et si

$$0 \leq \gamma \leq \frac{2\mathrm{D}^2\rho^2}{(\mathrm{D}^2+4)(1-\rho^2)} - 1 ,$$

la probabilité d'erreur est supérieure à celle obtenue en utilisant seulement $\bar{x}_0$ et $\bar{x}_1$ comme estimateurs de μ_0 et μ_1. Et Brailovsky cite pour conclure les valeurs numériques :

$$c = 0,\quad \mathrm{D} = 3,\quad \mathrm{P}_0 = 7\ \%,\quad \rho^2 = 0{,}64,\quad \gamma \leq 1{,}47$$

qui illustrent un cas où il vaudrait mieux selon lui ne pas utiliser toutes les observations dont on dispose.

4. DISCUSSION

Les résultats de Brailovsky ne sont pas valables quelque soit μ, ce qui en limite beaucoup la portée puisque μ est un paramètre inconnu. Mais la formule (5) permet une comparaison systématique de P à la valeur obtenue pour $\Lambda = \Sigma$; plus précisément μ n'intervient que par l'intermédiaire de D et de S et Λ apparaît de plus par l'intermédiaire de $\mathrm{T} = \mathrm{Tr}\,(\Lambda\Sigma^{-1})$.

En posant $\beta = \log\dfrac{p_1}{p_0}$ on est conduit à étudier le signe de

$$\mathrm{T} - k + \left(\frac{\mathrm{D}^2}{4} - \beta\frac{n-n'}{n+n'} + \frac{\beta^2}{\mathrm{D}^2} - 1\right)(\mathrm{S}-1) \qquad (6)$$

Or il existe ([2], p. 86) une matrice B telle que :

$${}^t\mathrm{B}\dot{\mathrm{B}} = \Sigma^{-1},\quad \mathrm{B}\Lambda{}^t\mathrm{B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

où $\lambda_1 \leq, \ldots, \leq \lambda_k$ sont les valeurs propres ordonnées de $\Lambda\Sigma^{-1}$; alors en posant

$$z = \frac{B\mu}{\| B\mu \|}$$

on a :

$$T = \lambda_1 + \ldots + \lambda_k \; ; \; S = \sum_1^k \lambda_i z_i^2 , \quad (\lambda_1 \leq \ldots \leq \lambda_k \; ; \; z_1^2 + \ldots + z_k^2 = 1).$$

On constate que S est fonction de la seule position « relative » de la direction de μ par rapport à Σ et Λ, et varie de λ_1 à λ_k.

Une étude élémentaire montre qu'en général, pour une matrice Λ donnée, le signe de (6) n'est pas constant pour tout μ, c'est-à-dire pour tout couple (S, D); il n'y a donc pas de conclusion générale possible sur le choix de la matrice Λ.

En particulier dans le cas $\Lambda = \Sigma_d$ mis en évidence par Brailovsky, il est facile de voir que $\lambda_1, \ldots, \lambda_k$ sont les inverses des valeurs propres de la *matrice de corrélation* associée à Σ, c'est-à-dire.

$$\Sigma_d^{1/2} \Sigma^{-1} \Sigma_d^{1/2} ,$$

et donc seule cette matrice intervient dans le signe de (6). On constate encore que ce signe dépend de μ.

Enfin la situation, paradoxale selon Brailovsky, que nous avons mise en évidence à la fin du § 3, résulte selon nous, seulement de ce que dans les cas *b*) et *c*) il existe un estimateur de variance moindre que celui utilisé par Brailovsky ce qui explique la perte d'information artificielle.

En effet dans le cas *b*) on peut appliquer les techniques générales des modèles linéaires sur plan d'expériences quelconque (cf. [*2*], p. 140); plus précisément soit $E^* \subset [1, k] \times [1, N]$ (N est la taille de l'échantillon empirique considéré) l'ensemble des couples (i, j) pour lesquels on dispose d'une observation $X_{i,j}$ l'élément aléatoire $X = \{ X_{i,j}, (i,j) \in E^* \}$ est un élément aléatoire gaussien à valeurs dans l'espace Ω^* des applications de E^* dans $\mathbb{R}$, dont la loi est définie par :

$$E(X_{i,j}) = m_i \qquad \mathrm{Cov}(X_{i,j}, X_{i',j'}) = \begin{cases} \rho_{i,i'}\sigma_i\sigma_{i'} & \text{si } j = j' \\ 0 & \text{si } j \neq j' \end{cases}$$

où $\Sigma = \{ \rho_{i,i'}\sigma_i\sigma_{i'} \}$ et où ${}^t m = (m_1, \ldots, m_k)$ est soit μ_0 soit μ_1. Soit V le sous-espace vectoriel de Ω^* formé des applications de E^* dans $\mathbb{R}$ qui ne dépendent pas de j; alors, en munissant Ω^* de la norme définie par l'inverse de la covariance de X, laquelle ne dépend que de Σ et est donc connue,

la projection de X sur V est un estimateur de m qui est sans biais et de variance minimum, donc meilleur que celui proposé par Brailovsky.

Enfin dans le cas c) un calcul simple montre que l'estimateur de μ_0 (résultat identique pour μ_1) défini par :

$$\bar{x}_0 \frac{1}{\gamma}\Sigma_d\left(\Sigma + \frac{1}{\gamma}\Sigma_d\right)^{-1} + \bar{x}_0'\Sigma\left(\Sigma + \frac{1}{\gamma}\Sigma_d\right)^{-1}$$

est sans biais, de covariance

$$\frac{1}{\gamma}\Sigma_d\left(\Sigma + \frac{1}{\gamma}\Sigma_d\right)^{-1}\Sigma \tag{7}$$

laquelle est inférieure à celle de l'estimateur choisi par Brailovsky. Il reste à vérifier qu'en remplaçant Λ par (7), la situation paradoxale mise en évidence par Brailovsky disparaît !

En conclusion, les précédentes réserves ne doivent pas faire oublier l'intérêt du travail de Brailovsky lequel a très bien montré la nécessité théorique et pratique d'étudier le comportement de la probabilité P quand Λ est différent de Σ.

BIBLIOGRAPHIE

[1] T. W., ANDERSON, *An introduction to Multivariate Statistical Analysis*, J. Wiley, New York.

[2] J. R. BARRA, *Notions fondamentales de statistique mathématique*, Dunod, Paris, 1971.

[3] M. OKAMOTO, An asymptotic expansion for the distribution of the linear discriminant function *A. M. S.*, t. **34**, 4, 1963, p. 1286-1301.

[4] V. BRAILOVSKY. On influence of sample set structures on Decision Rule. Quality for the case of linear Discriminant Function *(à paraître)*.

[5] J. R. BARRA, *Influence de l'estimation des paramètres sur la probabilité d'erreur en Analyse Discriminante* (selon V. Brailovsky), 1979. Séminaire de l'École Polytechnique, novembre 1979.

(Manuscrit reçu le 11 septembre 1980).