ON THE METHOD OF H. CHERNOFF IN CLASSIFICATORY PROBLEM BETWEEN TWO CATEGORIES WITH SPECIFIED $ 2^k $-th MOMENT ($ k = 0, 1, 2 $)

Nishi, Akihiro
Department of Mathematics, Faculty of Education, Saga University

https://doi.org/10.5109/13130

出版情報：統計数理研究. 18 (3/4), pp.51-59, 1979-03. Research Association of Statistical Sciences

ON THE METHOD OF H. CHERNOFF IN CLASSIFICATORY PROBLEM BETWEEN TWO CATEGORIES WITH SPECIFIED 2^k-th MOMENT ($k=0, 1, 2$)

By

Akihiro NISHI*

(Received December 1, 1977)

0. Summary

In this paper we shall seek the upper bound of probabilities (erorr probabilities) of misclassification in the Bayes classification procedures between two categories whose moments of order 2^k ($k=0, 1, 2$) are assumed to be known, while the distribution of each category may well be continuous or discrete.

We shall also give a sufficient condition, which is an improvement of Becher-Chernoff's upper bound. This becomes also a necessary condition in some cases of interest.

1. Introduction

We shall consider throughout this paper, classificatory problem between two categories with known apriori probabilities.

Let $F=(F_1, F_2)$, where F_i is the distribution of the category C_i ($i=1, 2$), and $\Pi=(\Pi_1, \Pi_2)$ ($\Pi_i>0$, $\Pi_1+\Pi_2=1$), where Π_i is the apriori probability of C_i. Let $\phi(x)=(\phi_1(x), \phi_2(x))$ (Lebesgue measurable, $\phi_i(x)\geqq 0$, $\phi_1(x)+\phi_2(x)\equiv 1$) be a (randomized classification) procedure such that an observation x is classified into C_i with probability $\phi_i(x)$, $\Phi=\{\phi(x)\}$ be the collection of all such ϕs and $\Phi^L=\{\phi^L(x)\}$ be the collection of all linear procedures, namely $\phi_i^L(x)$ is an indicator function of some half space.

The conditional error probability of classifying an observation, which is actually taken from C_i, into C_{3-i} ($i=1, 2$) becomes

(1-1) $$e_i(\phi, F)=\int(1-\phi_i(x))dF_i(x), \qquad i=1, 2$$

and the expected error probability with respect to the apriori probabilities Π becomes

(1-2) $$e_\Pi(\phi, F)=\Pi_1 e_1(\phi, F)+\Pi_2 e_2(\phi, F).$$

When the distribution functions $F_1(x)$ and $F_2(x)$ have their density functions (say $f_1(x)$, $f_2(x)$, respectively), it is well known that the $\inf_\phi e_\Pi(\phi, F)$ is attained, assuming

* Department of Mathematics, Faculty of Education, Saga University, Saga 840, Japan.

the equality of the two kinds of cost due to misclassification, by the Bayes classification procedure

(1-3) $$\phi_1^B(x)=\begin{cases} 1, & \text{when } \Pi_1 f_1(x)\geqq \Pi_2 f_2(x) \\ 0, & \text{when } \Pi_1 f_1(x)<\Pi_2 f_2(x). \end{cases}$$

Suppose the true distribution function $F_i(x)$ of the category C_i is not known to us, but it is known to belong to a family $\mathcal{F}_i$. Assuming that $\mathcal{F}_i$ is the family of all possible distribution functions with specified mean $\mu_i(\mu_1 \neq \mu_2)$ and variance σ_i^2, Becher [1] conjectured and Chernoff [2] verified that

(1-4) $$\sup_F \inf_\phi e_{\Pi^0}(\phi, F)=\sup_F \inf_{\phi L} e_{\Pi^0}(\phi^L, F)=\frac{1}{2}(1+S^2)^{-1},$$

where $\Pi^0=(1/2, 1/2)$ and $S=|\mu_1-\mu_2|/(\sigma_1+\sigma_2)$, by two excellent methods (one is analytical, the other geometrical) which, however, seem difficult to be extended to the multivariate case or unequal apriori probabilities case. We shall, however, use Chernoff's geometrical method to evaluate $\sup_F \inf_\phi e_{\Pi^0}(\phi, F)$ in the case of $\mathcal{F}_i$ with specified 2^k-th moment ($k=0, 1, 2$) in section 2. On the other hand, Isii and Taga [3], from the viewpoint of mathematical programming, studied the multivariate case in full generality, and Nishi [4] evaluated $\inf_{\phi L} \sup_F e_\Pi(\phi^L, F)$ and $\inf_{\phi L} \sup_F \max_{i=1,2} e_i(\phi^L, F)$ in detail.

Here we shall review briefly Chernoff's geometrical method. Let $F_i(x)\in\mathcal{F}_i=\mathcal{F}_i(\mu_i, \sigma_i^2)$ $(i=1, 2)$ and $H(x)\equiv F_1(x)+F_2(x)$. Since F_1 and F_2 are absolutely continuous with respect to H, they have densities f_1 and f_2, respectively. From (1-2) and (1-3) we have

(1-5) $$\begin{aligned} e_{\Pi^0}(\phi^B, F)&=\frac{1}{2}\left[\int_{\{x|f_1(x)<f_2(x)\}} f_1(x)dH(x)+\int_{\{x|f_1(x)\geqq f_2(x)\}} f_2(x)dH(x)\right] \\ &=\frac{1}{2}\int \min[f_1(x), f_2(x)]dH(x) \end{aligned}$$

Let us exclude the trivial cases where $e_{\Pi^0}(\phi^B, F)=0$ (i.e., supports of F_i are disjoint) or $e_{\Pi^0}(\phi^B, F)=1/2$ (i.e., $F_1=F_2$). Put

(1-6) $$p_0=\int \min[f_1(x), f_2(x)]dH(x), \qquad 0<p_0<1,$$

and for $i=1, 2$

(1-7) $$p_i=\int \{f_i(x)-\min[f_1(x), f_2(x)]\}dH(x)=1-p_0, \qquad 0<p_i<1.$$

Thus we may define the densities

(1-8) $$g_0(x)=\min[f_1(x), f_2(x)]/p_0$$

and for $i=1, 2$

(1-9) $$g_i(x)=\{f_i(x)-\min[f_1(x), f_2(x)]\}/p_i.$$

The first two moments corresponding to f_1, f_2, g_0, g_1, g_2 are given by $P_i=(\mu_i, \nu_i)$, $i=1, 2$ $(\nu_i=\mu_i^2+\sigma_i^2)$ which are the fixed points determined by $\mathcal{F}_i=\mathcal{F}_i(\mu_i, \nu_i)$, and $P_i^*=(\mu_i^*, \nu_i^*)$, $i=0, 1, 2$, and these five points satisfy

$$P_i=p_0P_0^*+(1-p_0)P_i^*, \qquad i=1, 2, \tag{1-10}$$

where $p_0/2$ represents the value $e_{\Pi 0}(\phi^B, F)$. Conversely given any three points P_i^* $(i=0, 1, 2)$ on or above the parabola $\eta=\xi^2$ which satisfy (1-10), it is possible to construct F_1, F_2 which will yield these three points. Thus to evaluate $\sup_F e_{\Pi 0}(\phi^B, F)$ is equivalent to find a configuration P_0^*, P_1^*, P_2^* of points on or above the parabola $\eta=\xi^2$ which satisfy (1-10) for a maximal value of p_0.

We see from Fig. 1-1 that we may restrict P_0^* to be on the parabola $\eta=\xi^2$ (i.e., $P_0^*=(\xi, \xi^2)$). Define $p_{0i}(\xi)$ as follows

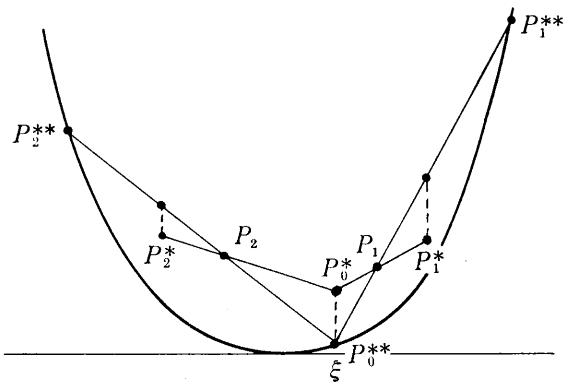

Fig. 1-1.

$$P_i=p_{0i}(\xi)P_0^{**}+(1-p_{0i}(\xi))P_i^{**}, \qquad i=1, 2, \tag{1-11}$$

where P_i^{**} and P_0^{**} lie on the parabola $\eta=\xi^2$. We have easily

$$p_{0i}(\xi)=\sigma_i^2/\{(\mu_i-\xi)^2+\sigma_i^2\}, \qquad i=1, 2. \tag{1-12}$$

Obviously $p_0(\xi)\equiv\min[p_{01}(\xi), p_{02}(\xi)]$ is the attainable maximum value of p_0 at ξ. We must seek $\sup_\xi p_0(\xi)$ that is equal to $2\sup_F e_{\Pi 0}(\phi^B, F)$. After some calculations we have

$$\sup_F \inf_\phi e_{\pi 0}(\phi, F)=\frac{1}{2}\sup_\xi p_0(\xi)=\frac{1}{2}(1+S^2)^{-1}, \tag{1-13}$$

where $\sup_\xi p_0(\xi)$ is attained at

$$\xi^*=(\sigma_2\mu_1+\sigma_1\mu_2)/(\sigma_1+\sigma_2). \tag{1-14}$$

Moreover the distributions F_1 and F_2 that attain $\sup_F e_{\Pi^0}(\phi^B, F)$ are given as follows. For $i=1, 2$, F_i takes x_0 with probability $(1+S^2)^{-1}$ and x_i with probability $1-(1+S^2)^{-1}$, where $x_0=\mu_1-\sigma_1 S$, $x_1=\mu_1+\sigma_1/S$ and $x_2=\mu_2-\sigma_2/S$ provided that $\mu_1>\mu_2$.

2. Improvement on Becker-Chernoff's bound

Suppose $F(x)$ is a univariate distribution function which has a finite 4-th moment. Denote

(2-1) $$\mu(F)=\int x\,dF(x)\,,\qquad \nu(F)=\int x^2 dF(x)\,,\qquad \tau(F)=\int x^4 dF(x)\,,$$

then Schwarz's inequality asserts that

(2-2) $$\mu(F)^2\leqq\nu(F)\,,\qquad \nu(F)^2\leqq\tau(F)\,.$$

Put

(2-3) $$X=\left\{(\mu(F),\ \nu(F),\ \tau(F))\,\middle|\,\int x^4 dF(x)<\infty\right\}.$$

LEMMA 2-1.

(i) $(\mu, \mu^2, \mu^4)\in X$ for all $-\infty<\mu<\infty$.

(ii) If $(\mu, \nu, \tau)\in X$ and $\nu=\mu^2$, then $\tau=\nu^2$.

(iii) If $\nu>\mu^2$, then $(\mu, \nu, \tau)\in X$ for all $\tau\geqq\nu^2$.

PROOF. It is clear that (i) and (ii) are valid. Now let us prove (iii) by constructing a distribution F which yields (μ, ν, τ) where $\nu>\mu^2$ and $\tau\geqq\nu^2$. Let π be the plane which goes through two points (μ, μ^2, μ^4) and (μ, ν, τ) but does not intersect the ξ-axis (i.e., π and ξ-axis are disjoint or π contains ξ-axis entirely). The intersecting points of the plane π and the curve $L\rightleftarrows\{(\xi, \eta, \zeta)\,|\,\xi=t,\ \eta=t^2,\ \zeta=t^4,\ -\infty<t<\infty\}$ are $(\pm\mu, \mu^2, \mu^4)$ and $(\pm\sqrt{\tau-\nu\mu^2}/\sigma,\ (\tau-\nu\mu^2)/\sigma^2,\ (\tau-\nu\mu^2)^2/\sigma^4)$ (put $\sigma^2=\nu-\mu^2>0$). Let F_0, F_1 and F_2 be one-point distributions which take $x=\mu$, $x=\sqrt{\tau-\nu\mu^2}/\sigma$ and $x=-\sqrt{\tau-\nu\mu^2}/\sigma$, respectively. Define $F_3=pF_1+(1-p)F_2$, where $p=\left(\mu+\frac{\sqrt{\tau-\nu\mu^2}}{\sigma}\right)\Big/\frac{2\sqrt{\tau-\nu\mu^2}}{\sigma}$. Note here that p satisfies $0<p<1$ because of $\tau\geqq\nu^2$ and $\nu>\mu^2$. Then $F=qF_0+(1-q)F_3$, where $q=\rho^2/(\rho^2+\sigma^4)$ (put $\rho^2=\tau-\nu^2$), is a desired one. Finally note that F is obtained as a three-points distribution provided $\tau>\nu^2$. When $\tau=\nu^2$ (i.e. $\rho^2=0$), F is uniquely determined as a two-points distribution which takes $\sqrt{\nu}$ with probability $(\sqrt{\nu}+\mu)/2\sqrt{\nu}$ and $-\sqrt{\nu}$ with probability $(\sqrt{\nu}-\mu)/2\sqrt{\nu}$.

Slightly modifying the above method to obtain a distribution which yields $P=(\mu, \nu, \tau)\in X$, we have a following lemma.

LEMMA 2-2. *Any finite distinct points of X, which satisfy $\tau>\nu^2$ except at most one point, can be generated by distributions whose supports are disjoint to each other, and each of these distributions is given except at most one distribution by a three-points or a four-points distribution.*

Let $\mathcal{F}_i=\mathcal{F}_i(\mu_i, \nu_i, \tau_i)$ be a family of all possible distributions with specified mean μ_i, second moment ν_i, and fourth moment τ_i ($i=1, 2$). Put $P_1=(\mu_1, \nu_1, \tau_1)$ and $P_2=(\mu_2, \nu_2, \tau_2)$. Then for each pair $(F_1, F_2)\in\mathcal{F}_1\times\mathcal{F}_2$, similarly to (1-10), there correspond three points P_0^*, P_1^* and P_2^* of X which satisfy

(2-4) $$P_i=p_0P_0^*+(1-p_0)P_i^*, \quad i=1, 2, \qquad \text{where } p_0=2e_{\Pi 0}(\phi^B, F).$$

Conversely given any three points P_i^* ($i=0, 1, 2$) of X which satisfy (2-4), it is possible to construct F_1, F_2 which will yield these three points according to lemma (2-2). Thus we may restrict P_0^* to be on the surface

$$M\equiv\{(\xi, \eta, \zeta)\,|\,\xi=s(-t)+(1-s)t,\ \eta=t^2,\ \zeta=t^4,\ 0<s<1,\ 0<t<\infty\}$$

according to lemma (2-1) (iii).

Let us put

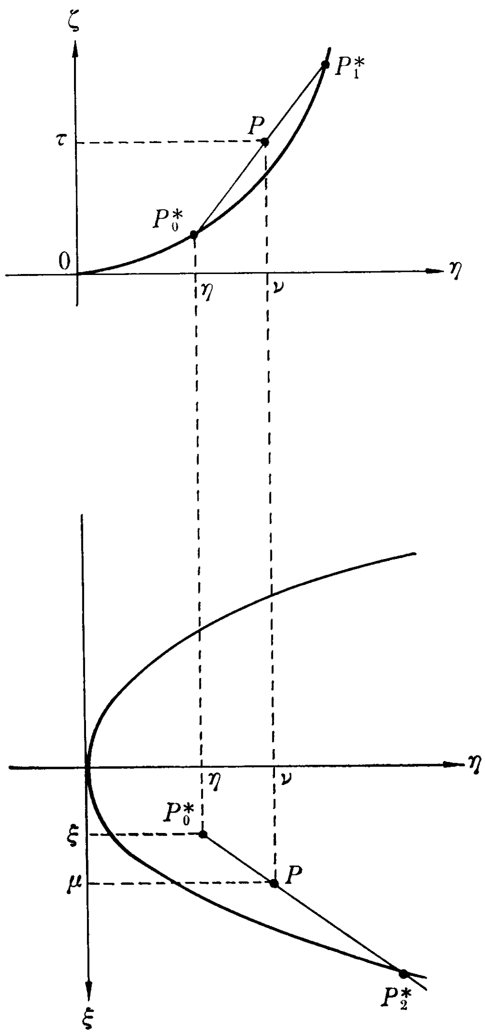

Fig. 2 1.

$$Y=\{(\xi,\eta,\zeta)\mid-\infty<\xi<\infty,\ \eta>0,\ \zeta>\eta^2\}$$
(2-5)
$$Z=\{(\xi,\eta,\zeta)\mid-\infty<\xi<\infty,\ \eta>\xi^2,\ -\infty<\zeta<\infty\}\ ,$$

then the set X defined by (2-3) becomes

(2-6) $$X=(Y\cap Z)\cup L\cup M\,.$$

We shall not consider the trivial cases of $\sigma^2=0$ or $\rho^2=0$ hereafter. We omit the subscript i in μ_i, ν_i, τ_i and P_i for simplicity for a while.

Now, we assume that P_0^* lies on the surface $M_0\equiv\{(\xi,\eta,\zeta)\mid-\infty<\xi<\infty,\ -\infty<\eta<\infty,\ \zeta=\eta^2\}$ containing M for convenience. Let P_1^* and P_2^* be the intersection points of the line P_0^*P and the surface ∂X or ∂Y, respectively, where ∂X is the boundary of the set X, and the three points P_0^*, P, and P_i^* lie on the line in this order. For arbitrarily fixed ξ, we define a ratio as $q^{(1)}(\eta\mid\xi)=\overline{PP_1^*}/\overline{P_0^*P_1^*}$ (see Fig. 2-1), then some calculations yield

(2-7) $$q^{(1)}(\eta\mid\xi)=\begin{cases}\rho^2/\{(\nu-\eta)^2+\rho^2\}\ , & \text{when } \eta>\tau/\nu \text{ or } \eta<\nu\\ \nu/\eta\ , & \text{when } \nu\leqq\eta\leqq\tau/\nu\,.\end{cases}$$

Thus $q^{(1)}(\eta\mid\xi)$ is independent of ξ, continuous and monotonic increasing for $\eta<\nu$, monotonic decreasing for $\eta\geqq\nu$. Obviously $q^{(1)}(\eta\mid\xi)$ satisfies $\lim_{|\eta|\to\infty}q^{(1)}(\eta\mid\xi)=0$ and $q^{(1)}(\nu\mid\xi)=1$. Similarly to $q^{(1)}(\eta\mid\xi)$ we define a ratio as $q^{(2)}(\eta\mid\xi)=\overline{PP_2^*}/\overline{P_0^*P_2^*}$ (see Fig. 2-1). When $\xi\neq\mu$, $q^{(2)}(\eta\mid\xi)$ is continous, monotonic decreasing, $\lim_{\eta\to+\infty}q^{(2)}(\eta\mid\xi)=0$, $\lim_{\eta\to-\infty}q^{(2)}(\eta\mid\xi)=1$ and

(2-8) $$q^{(2)}(\xi^2\mid\xi)=\sigma^2/\{(\mu-\xi)^2+\sigma^2\}\,.$$

When $\xi=\mu$, we have obviously

(2-9) $$q^{(2)}(\eta\mid\mu)=\begin{cases}1\ , & \text{when } \eta<\nu\\ \sigma^2/(\eta-\mu^2)\ , & \text{when } \eta\geqq\nu\,.\end{cases}$$

LEMMA 2-3. *For any $\xi\neq\mu$ there exist unique $\eta(\xi)$ which satisfies*

(2-10) $$\sup_{-\infty<\eta<\infty}\ \min_{i=1,2}q^{(i)}(\eta\mid\xi)=q^{(i)}(\eta(\xi)\mid\xi)=\rho^2/\{\omega^+(\xi)^2+\rho^2\}\ ,$$

where $\omega^+(\xi)$ is the unique positive solution of the following equation

(2-11) $$\sigma^2\omega^4+\rho^2\omega^3-2\mu\rho^2(\mu-\xi)\omega^2-(\mu-\xi)^2\rho^4=0\,.$$

When $\xi=\mu$, we have

(2-12) $$\sup_{-\infty<\eta<\infty}\ \min_{i=1,2}q^{(i)}(\eta\mid\xi)=1\,.$$

PROOF. For arbitrarily fixed $\xi\neq\mu$ the equation $q^{(1)}(\eta\mid\xi)=q^{(2)}(\eta\mid\xi)$ relative to η has only one solntion $\eta(\xi)$ (in the region of $\eta<\nu$) which satisfies

(2-13) $$\sup_{-\infty<\eta<\infty}\ \min_{i=1,2}q^{(i)}(\eta\mid\xi)=q^{(i)}(\eta(\xi)\mid\xi)\,,$$

because $q^{(1)}(\eta\mid\xi)$ is monotonic increasing for $\eta<\nu$, monotonic decreasing for $\eta\geqq\nu$ and

$q^{(2)}(\eta|\xi)$ is monotonic decreasing.

Geometrical interpretation of the existence of $\eta(\xi)$ satisfying (2-13) asserts that the points P_1^* and P_2^* of Fig. 2-1 are coincident (thus $P_1^*=P_2^*=(t(\xi), t(\xi)^2, t(\xi)^4)$ for an appropriate $t(\xi)$ and the three points P_0^*, P, and P_1^* lie on some line in this order. Hence we have

(2-14) $$(t(\xi)-\mu)/(\mu-\xi)=(t(\xi)^2-\nu)/(\nu-\eta(\xi))=(t(\xi)^4-\tau)/(\tau-\eta(\xi)^2)=c(\xi)\,,$$

where $c(\xi)$ is a positive constant and

(2-15) $$\sup_{-\infty<\eta<\infty}\min_{i=1,2} q^{(i)}(\eta|\xi)=q^{(i)}(\eta(\xi)|\xi)=c(\xi)/(1+c(\xi))\,.$$

Solving the equations (2-14) for $c(\xi)$, we have

(2-16) $$c(\xi)=\rho^2/(\nu-\eta(\xi))^2\,.$$

Substituting (2-16) for (2-14), $\eta(\xi)$ must satisfy the equation

(2-17) $$\sigma^2(\nu-\eta)^4+\rho^2(\nu-\eta)^3-2\mu\rho^2(\mu-\xi)(\nu-\eta)^2-(\mu-\xi)^2\rho^4=0\,,$$

which has only one solution for $\eta<\nu$. Therefore if we put $\omega=\nu-\eta$, then $\omega^+(\xi)=\nu-\eta(\xi)$ is the unique positive solution for the equation (2-11).

Thus we obtain (2-10) from (2-15) and (2-16).

When $\xi=\mu$, (2-7) and (2-9) show that $\sup_{-\infty<\eta<\infty}\min_{i=1,2} q^{(i)}(\eta|\mu)=q^{(i)}(\nu|\mu)=1$.

Let us put

(2-18) $$r(\xi)=\sup_{\eta>\xi^2}\min_{i=1,2} q^{(i)}(\eta|\xi)\,,$$

(2-19) $$\psi(\xi)=\sigma^2\xi^4-(2\nu\sigma^2+\rho^2)\xi^2+2\mu\rho^2\xi+(\sigma^2\nu^2-\rho^2\mu^2)\,.$$

LEMMA 2-4.

(i) *For* $|\xi|<\sqrt{\nu}$

(2-20) $$r(\xi)=\begin{cases}\rho^2/\{\omega^+(\xi)^2+\rho^2\}\,, & \text{when } \psi(\xi)>0\\ \sigma^2/\{(\mu-\xi)^2+\sigma^2\}\,, & \text{when } \psi(\xi)\leqq 0\,.\end{cases}$$

(ii) *For* $\sqrt{\nu}\leqq|\xi|\leqq\sqrt{\tau/\nu}$

(2-21) $$r(\xi)=\sigma^2/\{(\mu-\xi)^2+\sigma^2\}\,.$$

(iii) *For* $|\xi|>\sqrt{\tau/\nu}$

(2-22) $$r(\xi)=\begin{cases}\rho^2/\{(\nu-\xi^2)^2+\rho^2\}\,, & \text{when } \psi(\xi)>0\\ \sigma^2/\{(\mu-\xi)^2+\sigma^2\}\,, & \text{when } \psi(\xi)\leqq 0\,.\end{cases}$$

PROOF.

(i) Since $\xi^2<\nu$, (2-7) and (2-8) show that

$$\begin{aligned}q^{(2)}(\xi^2|\xi)-q^{(1)}(\xi^2|\xi)&=\sigma^2/\{(\mu-\xi)^2+\sigma^2\}-\rho^2/\{(\nu-\xi^2)^2+\rho^2\}\\&=\psi(\xi)/\{(\mu-\xi)^2+\sigma^2\}\{(\nu-\xi^2)^2+\rho^2\}\end{aligned}$$

provided $\xi\neq\mu$. It is easily seen that $\eta(\xi)>\xi^2$ is equivalent to $\psi(\xi)>0$. Thus (2-20) follows immediately from (2-10) and the graphs of $q^{(i)}(\eta|\xi)$ $(i=1, 2)$. When $\xi=\mu$, (2-7) and (2-9) show that $r(\mu)=q^{(i)}(\nu|\mu)=1$, which is contained in (2-20) because of $\omega^{+}(\mu)=0$ and $\psi(\mu)=\sigma^{4}>0$.

(ii) Since $\nu\leqq\xi^{2}\leqq\tau/\nu$, (2-7) and (2-8) show that

$$q^{(2)}(\xi^{2}|\xi)-q^{(1)}(\xi^{2}|\xi)=\sigma^{2}/\{(\mu-\xi)^{2}+\sigma^{2}\}-\nu/\xi^{2}$$
$$=-(\nu-\mu\xi)^{2}/\xi^{2}\{(\mu-\xi)^{2}+\sigma^{2}\}\leqq 0\,.$$

Since $q^{(1)}(\eta|\xi)$ and $q^{(2)}(\eta|\xi)$ are monotonic decreasing for

$$\eta>\xi^{2}(\geqq\nu)\,,\qquad r(\xi)=\min_{i=1,2}q^{(i)}(\xi^{2}|\xi)\,.$$

(iii) Similarly to the case (ii), $r(\xi)=\min_{i=1,2}q^{(i)}(\xi^{2}|\xi)$.

Since $\xi^{2}>\tau/\nu$, (2-7) and (2-8) show that

$$q^{(2)}(\xi^{2}|\xi)-q^{(1)}(\xi^{2}|\xi)=\psi(\xi)/\{(\mu-\xi)^{2}+\sigma^{2}\}\{(\nu-\xi^{2})^{2}+\rho^{2}\}\,.$$

Let $q_j^{(1)}(\xi)$, $q_j^{(2)}(\xi)$, $\psi_j(\xi)$, $\eta_j(\xi)$ and $r_j(\xi)$ be those that correspond to the points $P_j=(\mu_j,\nu_j,\tau_j)$ $(j=1,2)$. We may assume $\mu_1>\mu_2$ without loss of generality.

Let $p_0(\xi)$ be the attainable maximum value of p_0 at ξ (see (2-4)), then

$$p_0(\xi)=\sup_{\eta>\xi^2}\min_{j=1,2}\min_{i=1,2}q_j^{(i)}(\eta|\xi)\,. \tag{2-23}$$

Thus we have

$$\sup_F\inf_\phi e_{\Pi 0}(\phi,F)=\frac{1}{2}\sup_{-\infty<\xi<\infty}\sup_{\eta>\xi^2}\min_{j=1,2}\min_{i=1,2}q_j^{(i)}(\eta|\xi)\,, \tag{2-24}$$

though it seems difficult to represent the right hand side of (2-24) explicitly.

THEOREM 2-5. *A sufficient condition for* $\sup_F\inf_\phi e_{\Pi 0}(\phi,F)<\frac{1}{2}(1+S^{2})^{-1}$ *is*

$$\begin{cases}|\xi^*|<\sqrt{\nu} \quad \text{and} \quad \psi_j(\xi^*)>0\\ \text{or}\\ |\xi^*|>\sqrt{\tau_j/\nu_j} \quad \text{and} \quad \psi_j(\xi^*)>0\end{cases} \tag{2-25}$$

for $j=1$ *or* 2, *where* $\xi^*=(\sigma_1\mu_2+\sigma_2\mu_1)/(\sigma_1+\sigma_2)$.

PROOF. Since $\min_{j=1,2}r_j(\xi)=\min_{j=1,2}\sup_{\eta>\xi^2}\min_{i=1,2}q_j^{(i)}(\eta|\xi)$, we have

$$\sup_{\eta>\xi^2}\min_{j=1,2}\min_{i=12}q_j^{(i)}(\eta|\xi)\leqq\min_{j=1,2}r_j(\xi)\,. \tag{2-26}$$

From the definition of the ratios $q_j^{(i)}(\eta|\xi)$, we have clearly

$$\sup_{\eta>\xi^2}\min_{j=1,2}\min_{i=1,2}q_j^{(i)}(\eta|\xi)\leqq\min[p_{01}(\xi),\,p_{02}(\xi)] \qquad \text{(see (1-12))}. \tag{2-27}$$

Hence the conclusion follows from Lemma 2-4, (2-26) and (2-27).

If the equality holds in (2-26), then the equation (2-24) becomes

$$\sup_F\inf_\phi e_{\Pi 0}(\phi,F)=\frac{1}{2}\sup_{-\infty<\xi<\infty}\min_{j=1,2}r_j(\xi)\,. \tag{2-28}$$

Thus Lemma (2-4) and (2-27) show that the condition (2-25) is also necessary.

When $\nu_1=\nu_2$ and $\tau_1=\tau_2$, the inequality (2-26) reduces to a equality because of $q_1^{(1)}(\eta|\xi)=q_1^{(2)}(\eta|\xi)$.

Hence we obtain the following theorem.

THEOREM 2-6. *When* $\nu_1=\nu_2$ *and* $\tau_1=\tau_2$, *the condition* (2-25) *is a necessary and sufficient condition for* $\sup_F \inf_\varphi e_{II^0}(\phi, F)<\frac{1}{2}(1+S^2)^{-1}$.

COROLLARY 2-7. *Wnen* $\mu_1=\mu$, $\mu_2=-\mu$ $(\mu>0)$, $\nu_1=\nu_2=\nu$ *and* $\tau_1=\tau_2=\tau$, $\sup_F \inf_\varphi e_{II^0}(\phi, F)$ $<\frac{1}{2}(1+S^2)^{-1}$ *holds if and only if* $\tau<\nu^3/\mu^2$. *Moreover, in this case, we have*

$$\sup_F \inf_\varphi e_{II^0}(\phi, F)=\frac{1}{2}\rho^2/\{\omega^+(0)^2+\rho^2\}\ , \tag{2-29}$$

where $\omega^+(0)$ *is the unique positive solution of the equation*

$$\sigma^2\omega^4+\rho^2\omega^3-2\mu^2\rho^2\omega^2-\mu^2\rho^4=0\,. \tag{2-30}$$

PROOF. Since $\sigma_1=\sigma_2$ and $\mu_2=-\mu_1$, we have $\xi^*=0$ which satisfy $|\xi^*|<\sqrt{\nu}$. Thus the condition (2-25) is equivalent to $\psi_1(0)=\psi_2(0)=\sigma^2\nu^2-\rho^2\mu^2=\nu^3-\tau\mu^2>0$.

Let $\omega_1^+(\xi)$ and $\omega_2^+(\xi)$ be the positive solutions of the equations $\sigma^2\omega^4+\rho^2\omega^3-2\mu\rho^2(\mu-\xi)\omega^2-(\mu-\xi)^2\rho^4=0$ and $\sigma^2\omega^4+\rho^2\omega^3-2\mu\rho^2(\mu+\xi)\omega^2-(\mu+\xi)^2\rho^4=0$, respectively, then $\omega_1^+(\xi)$ is monotonic decreasing for $\xi<\mu$ and $\lim_{\xi\uparrow\mu}\omega_1^+(\xi)=\omega_1^+(\mu)=0$, similarly $\omega_2^+(\xi)$ is monotonic increasing for $\xi>-\mu$ and $\lim_{\xi\downarrow-\mu}\omega_2^+(\xi)=\omega_2^+(-\mu)=0$. It is obvious that $\omega_1^+(0)$ $=\omega_2^+(0)=\omega^+(0)$. Thus (2-29) follows from Lemma 2-4 and (2-28).

References

[1] BECKER, P., Recognition of Patterns. Polyteknisk Forlag, Copenhagen, 1968.

[2] CHERNOFF, H., *A bound on the classification error for discriminating between populations with specified means and variances.* Studi probabilità statistica e ricerca operativa in onore di Giuseppe Pompilj (1971), 203-211.

[3] ISII, K. and TAGA, Y., *Bound on the classification error for discriminating between multivariate populations with specified means and covariance matrices.* Ann. Statist. **6** (1978), 132-141.

[4] NISHI, A., *On linear classification procedures between two categories with known mean vectors and covariance matrices.* Ann. Inst. Statist. Math. **29** (1977), Part A, 433-444.