

共変量判別問題における漸近展開  
による誤確率

阪大 基礎工学部 山口光代

$q$ 次元の共変量を持つ、2つの、 $p+q$ 次元正規分布

$$\pi_1: N\left[\begin{bmatrix} \mu_1 \\ \eta \end{bmatrix}, \Sigma\right], \quad \pi_2: N\left[\begin{bmatrix} \mu_2 \\ \eta \end{bmatrix}, \Sigma\right] \quad \text{のどちらか一方から、}$$

とり出されたことがわかっている観測値  $\begin{bmatrix} x \\ y \end{bmatrix}$  と、 $\pi_1, \pi_2$  から  
とり出された、大きさ  $N_1, N_2$  の2つの標本

$$\begin{bmatrix} x_{11} \\ y_{11} \end{bmatrix}, \dots, \begin{bmatrix} x_{1N_1} \\ y_{1N_1} \end{bmatrix}; \quad \begin{bmatrix} x_{21} \\ y_{21} \end{bmatrix}, \dots, \begin{bmatrix} x_{2N_2} \\ y_{2N_2} \end{bmatrix}$$

に基づいて、どちらかに判別する問題と考える。ここには、

$p$ 次元平均ベクトル  $\mu_1, \mu_2$ , 共通の  $q$ 次元平均ベクトル  $\eta$ ,  
共通の nonsingular な共分散行列  $\Sigma$  は、すべて未知パラメー  
タである。この判別問題に対して、与えられる2つの  
判別統計量を用いる。1つは、Cochran and Bliss [1] による

$$W^* = (\bar{x}_1^* - \bar{x}_2^*)' S_{11 \cdot 2}^{-1} \left[ x^* - \frac{1}{2}(\bar{x}_1^* + \bar{x}_2^*) \right]$$

他の1つは、Fujikoshi and Karazawa [2] による

$$Z^* = N_1(N_1+1+l_1/k_1)^{-1} (x^* - \bar{x}_1^*)' S_{11 \cdot 2}^{-1} (x^* - \bar{x}_1^*) \\ - N_2(N_2+1+l_2/k_2)^{-1} (x^* - \bar{x}_2^*)' S_{11 \cdot 2}^{-1} (x^* - \bar{x}_2^*)$$

である。  $x^* = x - S_{12}S_{22}^{-1}y$ ,  $\bar{x}_i^* = \bar{x}_i - S_{12}S_{22}^{-1}\bar{y}_i$ ,  $\bar{x}_i, \bar{y}_i$ : 標本平均,  
 $S_{11.2} = S_{11} - S_{12}S_{22}^{-1}S_{21}$ ,  $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$  は  $\Sigma$  の最良不偏推定量,  
 $l_i = (y - \bar{y}_i)' S_{22}^{-1} (y - \bar{y}_i)$ ,  $k_i = n/N_i$ ,  $n = N_1 + N_2 - 2$ .

$W^*$  による判別方式,  $Z^*$  による判別方式は, それぞれ,

$$W^* \geq C_1 \Rightarrow [\bar{y}] \in \pi_1 \text{ に判別}, W^* < C_1 \Rightarrow [\bar{y}] \in \pi_2 \text{ に判別} \text{ せよ.}$$

$$Z^* < C_2 \Rightarrow [\bar{y}] \in \pi_1 \text{ に判別}, Z^* \geq C_2 \Rightarrow [\bar{y}] \in \pi_2 \text{ に判別} \text{ せよ.}$$

となり,  $C_1, C_2$  は, 定数, 又は,  $\bar{x}_i, \bar{y}_i, S$  の関数である。  
 この判別方式によつて,  $[\bar{y}]$  が判別するとき, 二種類の誤判別  
 が起こるが, その中の一つの誤判別の確率, 即ち,  $[\bar{y}]$  が実  
 際  $\pi_1$  に属するのにも,  $\pi_2$  に属すると判別する確率

$$P\{W^* < C_1 \mid [\bar{y}] \in \pi_1\}, P\{Z^* \geq C_2 \mid [\bar{y}] \in \pi_1\}$$

が一定の値,  $\alpha$ , になるように, 分離点  $C_1, C_2$  を定める場合,  
 を考へる。Studentized- $W^*$ , Studentized- $Z^*$  の分母関数の  
 漸近展開を用いると,  $C_1, C_2$  は,  $(N_1, N_2, n)^{-2}$  次の項を無視可  
 ると, 次の様に与えられる。  $\Delta^2 = (\mu_1 - \mu_2)' \Sigma_{11.2}^{-1} (\mu_1 - \mu_2)$  とする。

$$C_1 = C_1(W^*) = A(W^*)D + \frac{1}{2}D^2, \quad C_2 = C_2(Z^*) = 2A(Z^*)D - D^2,$$

$$A(W^*) = u_0 - (2N_1\Delta)^{-1} \{ \Delta u_0 - 2(p-1) \} - (4n)^{-1} \{ u_0^2 + (4p+4q-3)u_0 \},$$

$$A(Z^*) = -u_0 - (2N_1\Delta)^{-1} \{ u_0^2 + \Delta u_0 - (p-1) \} + (2N_2\Delta)^{-1} [ u_0^2 + 2\Delta u_0 + \Delta^2 + p-1 ]$$

$$- (4n)^{-1} \{ u_0^2 + (4p-3)u_0 - 2q\Delta \}. \quad D^2 = (\bar{x}_1^* - \bar{x}_2^*)' S_{11.2}^{-1} (\bar{x}_1^* - \bar{x}_2^*).$$

これらの分離点に対し, もう一方の誤判別の確率, 即ち,  
 $[\bar{y}]$  が実際は,  $\pi_2$  に属するのにも,  $\pi_1$  に属すると判別する

確率は、それぞれ、

$$Pr\{W^* \geq C_\alpha(W^*) | [\frac{r}{y}] \in \Pi_2\} = 1 - \Phi(u_0 + \Delta) + \phi(u_0 + \Delta) [(2N_1\Delta)^{-1}(\Delta^2 + p - 1) + (2N_2\Delta)^{-1}\{4\Delta(u_0 + \Delta) - 3(p - 1)\} + (4\pi)^{-1}\Delta\{u_0^2 + 2(p + q - 1)\} + O_2],$$

$$Pr\{Z^* < C_\alpha(Z^*) | [\frac{r}{y}] \in \Pi_2\} = 1 - \Phi(u_0 + \Delta) + \phi(u_0 + \Delta) [-(2N_1\Delta)^{-1}\{2u_0^2 + 4\Delta u_0 + 3\Delta^2 + p - 1\} + (2N_2\Delta)^{-1}\{2(u_0 + \Delta)(u_0 + 3\Delta) - (p - 1)\} + (4\pi)^{-1}\Delta\{u_0^2 + 2(p - q - 1)\} + O_2]$$

$\Phi(u)$ ,  $\phi(u)$  は  $N[0, 1]$  の cdf, pdf,  $u_0$  は  $\Phi(u)$  の  $100\alpha\%$  点, とする。 小さな確率  $\varepsilon$  与える統計量は、

$\Delta$	$R_2 \leq R_1$	$R_1 < R_2 < 2R_1 + q$	$2R_1 + q = R_2$	$2R_1 + q < R_2$
0				$u_0 < u^{**}$ $u^{**} \leq u_0$
↓ k	$Z^*$	$W^*$	$W^*$	$W^*$
		$\Delta_4$	$\Delta_0$	$\Delta_4$ $W^*$
		$Z^*$	$Z^*$	$\Delta_3$ $W^*$

$$\Delta_0 = -(u_0^2 + p - 1) / 2u_0$$

$$\Delta_3 = \{-(R_2 - R_1)u_0 + \sqrt{D_2}\} / (R_2 - 2R_1 - q)$$

$$\Delta_4 = \{-(R_2 - R_1)u_0 - \sqrt{D_2}\} / (R_2 - 2R_1 - q)$$

$$D_2 = (R_2 - R_1)\{(R_1 + q)u_0^2 - (R_2 - 2R_1 - q)(p - 1)\}$$

$$u^{**} = -\{(R_2 - 2R_1 - q)(p - 1) / (R_1 + q)\}^{1/2}$$

$N_1=100, N_2=30, \alpha=0.05$   $V_{.95} \sim 2$









