ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria III: MATEMATYKA STOSOWANA XXIX (1987)

Ministerstwo Nauki i Szkolnictwa Wyższego

WIESŁAW PASEWICZ (Szczecin)

# Estymatory kwadratowej funkcji dyskryminacyjnej w przypadku specjalnych struktur macierzy kowariancji

(*Praca wpłynęła do Redakcji 1985.03.27*)

**1. Wstęp.** Załóżmy, że w populacji $\pi_i$ ($i = 1, 2$) wektor losowy $Z_i$ ma symetryczny rozkład normalny [8], tj. rozkład normalny z zerowym wektorem wartości oczekiwanych i macierzą kowariancji $\Sigma_i$ postaci

$$\Sigma_i = \sigma_i^2 [(1-\varrho_i) I + \varrho_i E], \tag{1}$$

gdzie $\sigma_i^2$ jest wspólną wariancją składowych wektora losowego $Z_i$, $\varrho_i$ jest wspólnym współczynnikiem korelacji dla wszystkich par składowych wektora losowego $Z_i$, $I_{p\times p}$ jest macierzą jednostkową, a $E_{p\times p}$ jest macierzą, której wszystkie elementy są jedynkami. Niech $f_i(z)$ będzie funkcją gęstości wektora losowego $Z_i$, $i = 1, 2$. Problem polega na znalezieniu estymatora kwadratowej funkcji dyskryminacyjnej postaci

$$U(z) = \ln \frac{f_1(z)}{f_2(z)}. \tag{2}$$

W pracy [7] przedstawiony został estymator bayesowski funkcji $f_i(z)$ dla $i = 1, 2$. Podstawiając w miejsce nieznanych funkcji $f_1(z)$ i $f_2(z)$ ich odpowiednie estymatory, otrzymamy natychmiast estymator funkcji $U(z)$.

W niniejszej pracy dyskutowany będzie semibayesowski estymator funkcji $U(z)$ oraz estymator klasyczny tej funkcji, w przypadku gdy parametry $\sigma_i$ i $\varrho_i$ nie są znane.

**2. Estymacja semibayesowska kwadratowej funkcji dyskryminacyjnej z symetrycznym rozkładem normalnym.** Załóżmy, że $Z_{ijk}$ jest $k$-tą obserwacją, $j$-tej cechy z $i$-tej populacji, gdzie $k = 1, \ldots, N_i$; $j = 1, \ldots, p$; $i = 1, 2$. Wprowadźmy następujące oznaczenia:

$$A_i = \sum_k \sum_j z_{ijk}^2 = pN_i S_i^2, \tag{3}$$

$$A_i + B_i = \sum_k \Big(\sum_j z_{ijk}\Big)^2 = pN_i S_i^2 [1 + (p-1) r_i], \tag{4}$$

$$a = \sum_j z_j^2, \qquad a + b = \Big(\sum_j z_j\Big)^2. \tag{5}$$

Załóżmy dalej ([2], [7]), że łączny rozkład a priori parametrów $\sigma_i$, $\varrho_i$ ma następującą gęstość:

$$g(\sigma_i, \varrho_i) \propto \{\sigma_i(1-\varrho_i)[1+(p-1)\varrho_i]\}^{-1}. \tag{6}$$

Niech $T(z)$ będzie semibayesowskim estymatorem kwadratowej funkcji dyskryminacyjnej $U(z)$. Estymator ten znajdziemy jako wartość oczekiwaną funkcji $U(z)$ [6] postaci

$$U(z) = \frac{1}{2}p\ln\frac{\sigma_2^2}{\sigma_1^2} + \frac{1}{2}(p-1)\ln\frac{1-\varrho_2}{1-\varrho_1} + \frac{1}{2}\ln\frac{1+(p-1)\varrho_2}{1+(p-1)\varrho_1} + \tag{7}$$

$$+\frac{1}{2}\left\{\left[\frac{1}{\sigma_2^2(1-\varrho_2)} - \frac{1}{\sigma_1^2(1-\varrho_1)}\right]a - \left[\frac{\varrho_2}{(1-\varrho_2)[1+(p-1)\varrho_2]\sigma_2^2} - \right.\right.$$

$$\left.\left. - \frac{\varrho_1}{(1-\varrho_1)[1+(p-1)\varrho_1]\sigma_1^2}\right](a+b)\right\},$$

względem łącznego rozkładu a posteriori parametrów $\sigma_i$ i $\varrho_i$ [3] dla $i = 1, 2$ lub (co jest równoważne) jako wartość oczekiwaną funkcji danej wzorem

$$U(z) = \frac{1}{2}\ln\frac{\alpha_2}{\alpha_1} + \frac{1}{2}(p-1)\ln\frac{\theta_2}{\theta_1} + \frac{a+b}{2p}\left(\frac{1}{\alpha_2} - \frac{1}{\alpha_1}\right) + \frac{(p-1)a-b}{2p}\left(\frac{1}{\theta_2} - \frac{1}{\theta_1}\right) \tag{8}$$

względem łącznego rozkładu a posteriori parametrów $\alpha_i$, $\theta_i$ ($i = 1, 2$) określonych wzorami [2]

$$[1+(p-1)\varrho_i]\sigma_i^2 = \alpha_i, \quad (1-\varrho_i)\sigma_i^2 = \theta_i. \tag{9}$$

Wiadomo [7], że zmienne losowe $\alpha_i$, $\theta_i$ są niezależne, a ich gęstości, tj.

$$P(\alpha_i|A_i, B_i) = (2p)^{-\frac{1}{2}N_i}\frac{(A_i+B_i)^{\frac{1}{2}N_i}}{\alpha_i^{\frac{1}{2}N_i+1}\Gamma\left(\frac{N_i}{2}\right)}\exp\left(-\frac{1}{2p}\frac{A_i+B_i}{\alpha_i}\right), \tag{10}$$

$$P(\theta_i|A_i, B_i) = \tag{11}$$

$$= (2p)^{-\frac{1}{2}N_i(p-1)}\frac{[(p-1)A_i - B_i]^{\frac{1}{2}(p-1)N_i}}{\theta_i^{\frac{1}{2}(p-1)N_i+1}\Gamma[\frac{1}{2}(p-1)N_i]}\exp\left[-\frac{1}{2p}\frac{(p-1)A_i - B_i}{\theta_i}\right]$$

dla $i = 1, 2$, są gęstościami uogólnionego rozkładu gamma [5] postaci

$$f(x) = \begin{cases} 0 & \text{dla } x \leqslant 0 \\ \dfrac{|\alpha|\lambda^{p/\alpha}}{\Gamma(p/\alpha)}x^{p-1}\exp(-\lambda x^{\alpha}) & \text{dla } x > 0,\ \alpha p > 0,\ \lambda > 0. \end{cases} \tag{12}$$

Udowodnimy teraz następujące lematy.

LEMAT 1. *Jeżeli zmienne losowe* $\alpha_i$ *i* $\theta_i$ *mają uogólnione rozkłady gamma odpowiednio postaci* (10) *i* (11), *to istnieją wartości oczekiwane zmiennych losowych* $\ln\alpha_i$ *i* $\ln\theta_i$, *przy czym*

$$E(\ln\alpha_i) = \ln\frac{A_i+B_i}{2p} - \psi\left(\frac{N_i}{2}\right) \tag{13}$$

*oraz*

$$E(\ln\theta_i) = \ln\frac{(p-1)A_i - B_i}{2p} - \psi\left[\frac{(p-1)N_i}{2}\right] \tag{14}$$

*dla* $i = 1, 2$ $(\psi(x) = \Gamma'(x)\Gamma^{-1}(x))$.

Dowód. Znajdźmy najpierw drugą pochodną funkcji charakterystycznej zmiennej losowej $\ln\alpha_i$ w punkcie zero. Mamy

$$\varphi(t) = E(e^{it\ln\alpha_i}) = \int_0^\infty \alpha_i^{it} P(\alpha_i|A_i, B_i)\,d\alpha_i = \frac{\Gamma\left(\frac{N_i}{2} - it\right)}{\Gamma\left(\frac{N_i}{2}\right)}\left(\frac{A_i+B_i}{2p}\right)^{it}.$$

Stąd

$$\varphi''(0) = 2\psi\left(\frac{N_i}{2}\right)\ln\frac{A_i+B_i}{2} - \left(\ln\frac{A_i+B_i}{2p}\right)^2 - \frac{\Gamma''\left(\frac{N_i}{2}\right)}{\Gamma\left(\frac{N_i}{2}\right)} < \infty.$$

Wiadomo [1], [4], że jeśli funkcja charakterystyczna $\varphi(t)$ zmiennej losowej $X$ ma skończoną pochodną parzystą $\varphi^{(2l)}(0)$, to istnieją momenty rzędów $1, 2, \ldots, 2l$ zmiennej losowej $X$.

Z powyższego wynika, że istnieje wartość oczekiwana zmiennej losowej $\ln\alpha_i$. Ażeby uzyskać równość (13), wystarczy teraz zauważyć, że $E(\ln\alpha_i) = \varphi'(0)/i$.

Podobnie można dowieść, że istnieje $E(\ln\theta_i)$ i jest równa prawej stronie równości (14).

LEMAT 2. *Jeżeli zmienna losowa* $X$ *ma uogólniony rozkład gamma postaci* (12), *to*

$$E(X^{-1}) = \begin{cases} -\dfrac{p}{\lambda}, & \text{gdy } \alpha = -1 \text{ i } p < 0, \\ \dfrac{j}{p-1}, & \text{gdy } \alpha = 1 \text{ i } p > 0. \end{cases} \tag{15}$$

Dowód. Zauważmy, że $E(X^{-1}) = E(|X^{-1}|)$. Wystarczy zatem wykazać równość (15).

W przypadku gdy $\alpha = -1$ i $p < 0$, mamy

$$E(X^{-1}) = \frac{\lambda^{-p}}{\Gamma(-p)} \int_0^\infty x^{p-2} \exp\left(-\frac{\lambda}{x}\right) dx = \frac{\Gamma(1-p)}{\lambda\Gamma(-p)} = -\frac{p}{\lambda}.$$

Podobnie można dowieść drugą część tezy.

Korzystając z lematu 2, otrzymujemy

$$E(\alpha_i^{-1}) = \frac{pN_i}{A_i + B_i} \tag{16}$$

oraz

$$E(\theta_i^{-1}) = \frac{p(p-1)N_i}{(p-1)A_i - B_i}. \tag{17}$$

Z lematu 1 i równości (16) i (17), przy uwzględnieniu związku (4) i związku postaci

$$(p-1)A_i - B_i = p(p-1)N_i S_i^2(1-r_i) \tag{18}$$

wynika bezpośrednio następujące twierdzenie.

TWIERDZENIE 1. *Dla ustalonej obserwacji z, semibayesowski estymator* $T(z)$ *ma postać*

$$\begin{aligned} T(z) = {} & \frac{1}{2}\ln\frac{S_2^2[1+(p-1)r_2]}{S_1^2[1+(p-1)r_1]} + \frac{1}{2}(p-1)\ln\frac{S_2^2(1-r_2)}{S_1^2(1-r_1)} + \\ & + \frac{a+b}{2p}\left\{\frac{1}{S_2^2[1+(p-1)r_2]} - \frac{1}{S_1^2[1+(p-1)r_1]}\right\} + \\ & + \frac{(p-1)a-b}{2p}\left[\frac{1}{S_2^2(1-r_2)} - \frac{1}{S_1^2(1-r_1)}\right] + \frac{1}{2}p\ln\frac{N_2}{N_1} + K, \end{aligned} \tag{19}$$

gdzie $S_i$, $r_i$ (określone wzorami (3) i (4)) są estymatorami parametrów odpowiednio $\sigma_i$, $\varrho_i$ dla $i = 1, 2$ oraz

$$K = \frac{1}{2}\left[\psi\left(\frac{N_1}{2}\right) - \psi\left(\frac{N_2}{2}\right)\right] + \frac{p-1}{2}\left\{\psi\left[\frac{1}{2}(p-1)N_1\right] - \psi\left[\frac{1}{2}(p-1)N_2\right]\right\}. \tag{20}$$

Wyniki w postaci wzoru (19) uzyskali również autorzy pracy [2]. Ponieważ funkcję daną wzorem (7) można zapisać w postaci (8), więc łatwo dostrzec, że $\lim_{N_1, N_2 \to \infty} T(z) = U(z)$. Zatem estymator $T(z)$ jest estymatorem zgodnym kwadratowej funkcji dyskryminacyjnej $U(z)$.

Zauważmy, że iloczyn prawych stron równości (10) i (11) jest proporcjonalny do iloczynu gęstości uogólnionych rozkładów gamma zmiennych losowych $A_i + B_i$ i $(p-1)A_i - B_i$. Wynika stąd, że zmienne te są niezależne o następujących gęstościach:

$$(21)\qquad P(A_i+B_i|\alpha_i)=\frac{(A_i+B_i)^{\frac{1}{2}N_i}}{(2p\alpha_i)^{\frac{1}{2}N_i}\,\Gamma(\frac{1}{2}N_i+1)}\exp\left(-\frac{A_i+B_i}{2p\alpha_i}\right),$$

$$(22)\qquad P[(p-1)A_i-B_i|\theta_i]=\frac{[(p-1)A_i-B_i]^{\frac{1}{2}(p-1)N_i}}{(2p\theta_i)^{\frac{1}{2}(p-1)N_i+1}\,\Gamma[\frac{1}{2}(p-1)N_i+1]}\times$$
$$\times\exp\left[-\frac{(p-1)A_i-B_i}{2p\theta_i}\right]$$

dla $i=1, 2$.

Jeśli zapiszemy funkcję $T(z)$ daną wzorem (19) w postaci

$$(23)\qquad T(z)=\frac{1}{2}\ln\frac{A_2+B_2}{A_1+B_1}+\frac{1}{2}(p-1)\ln\frac{(p-1)A_2-B_2}{(p-1)A_1-B_1}+$$
$$+\frac{a+b}{2p}\left[\frac{pN_2}{A_2+B_2}-\frac{pN_1}{A_1+B_1}\right]+$$
$$+\frac{(p-1)a-b}{2p}\left[\frac{p(p-1)N_2}{(p-1)A_2-B_2}-\frac{p(p-1)N_1}{(p-1)A_1-B_1}\right]+K,$$

to wyliczenie wartości oczekiwanej tej funkcji sprowadzi się do wyliczenia wartości oczekiwanych funkcji zmiennych $A_i+B_i$ oraz $(p-1)A_i-B_i$, tj. do wyliczenia wartości wyrażeń $E[\ln(A_i+B_i)]$ i $E\{\ln[(p-1)A_i-B_i]\}$ oraz $E[(A_i+B_i)^{-1}]$ i $E\{[(p-1)A_i-B_i]^{-1}\}$ (o ile istnieją), względem rozkładów prawdopodobieństwa zmiennych losowych odpowiednio $A_i+B_i$ oraz $(p-1)A_i-B_i$.

Można wykazać, że zachodzą następujące związki:

$$(24)\qquad E[\ln(A_i+B_i)]=\ln(2p\alpha_i)+\psi(\tfrac{1}{2}N_i+1),$$

$$(25)\qquad E\{\ln[(p-1)A_i-B_i]\}=\ln(2p\theta_i)+\psi[\tfrac{1}{2}(p-1)N_i+1],$$

$$(26)\qquad E[(A_i+B_i)^{-1}]=(pN_i\alpha_i)^{-1},$$

$$(27)\qquad E\{[(p-1)A_i-B_i]^{-1}\}=[p(p-1)N_i\theta_i]^{-1}$$

(dla $i=1, 2$) oraz

$$(28)\qquad \psi(x+1)=\psi(x)+\frac{1}{x}.$$

Stąd (dla ustalonej obserwacji $z$) wartość oczekiwana estymatora kwadratowej funkcji dyskryminacyjnej $T(z)$ będzie postaci

$$(29)\qquad E[T(z)]=\frac{1}{2}\ln\frac{\alpha_2}{\alpha_1}+\frac{1}{2}(p-1)\ln\frac{\theta_2}{\theta_1}+\frac{a+b}{2p}\left(\frac{1}{\alpha_2}-\frac{1}{\alpha_1}\right)+$$
$$+\frac{(p-1)a-b}{2p}\left(\frac{1}{\theta_2}-\frac{1}{\theta_1}\right)+2\left(\frac{1}{N_2}-\frac{1}{N_1}\right).$$

Z powyższej równości wynika, że estymator $T(z)$ jest estymatorem obciążonym kwadratowej funkcji dyskryminacyjnej $U(z)$ postaci (8) lub (7).

**3. Estymacja klasyczna kwadratowej funkcji dyskryminacyjnej z symetrycznym rozkładem normalnym.** Stosując podejście klasyczne do estymacji, można postąpić w dwojaki sposób.

Estymować najpierw funkcję gęstości, a następnie znaleźć estymator kwadratowej funkcji dyskryminacyjnej jako iloraz wyestymowanych funkcyj gęstości lub estymować bezpośrednio kwadratową funkcję dyskryminacyjną.

Oczywiście, w przypadku estymacji klasycznej (w przeciwieństwie do estymacji bayesowskiej i semibayesowskiej) estymator funkcji dyskryminacyjnej nie zmieni się.

Oznaczmy przez $V(z)$ estymator klasyczny kwadratowej funkcji dyskryminacyjnej $U(z)$ postaci (7) lub (8). Jeśli w funkcji (7) nieznane parametry $\sigma_i$, $\varrho_i$ $(i = 1, 2)$ zastąpimy ich odpowiednimi estymatorami $S_i$ i $r_i$ określonymi przez równości (3) i (4), to otrzymamy klasyczny estymator funkcji $U(z)$ postaci

$$V(z) = \frac{1}{2}p\ln\frac{S_2^2}{S_1^2} + \frac{1}{2}(p-1)\ln\frac{1-r_2}{1-r_1} + \frac{1}{2}\ln\frac{1+(p-1)r_2}{1+(p-1)r_1} + \tag{30}$$

$$+\frac{1}{2}\left\{\left[\frac{1}{S_2^2(1-r_2)} - \frac{1}{S_1^2(1-r_1)}\right]a - \right.$$

$$\left. -\left[\frac{r_2}{(1-r_2)[1+(p-1)r_2]S_2^2} - \frac{r_1}{(1-r_1)[1+(p-1)r_1]S_1^2}\right](a+b)\right\}.$$

Korzystając z tożsamości postaci

$$\frac{a}{1-r_i} - \frac{r_i(a+b)}{(1-r_i)[1+(p-1)r_i]} \equiv \frac{1}{p}\left[\frac{(p-1)a-b}{1-r_i} + \frac{a+b}{1+(p-1)r_i}\right] \quad (i = 1, 2), \tag{31}$$

otrzymujemy estymator

$$V(z) = \frac{1}{2}\ln\frac{[1+(p-1)r_2]S_2^2}{[1+(p-1)r_1]S_1^2} + \frac{1}{2}(p-1)\ln\frac{(1-r_2)S_2^2}{(1-r_1)S_1^2} + \tag{32}$$

$$+\frac{a+b}{2p}\left\{\frac{1}{[1+(p-1)r_2]S_2^2} - \frac{1}{[1+(p-1)r_1]S_1^2}\right\} +$$

$$+\frac{(p-1)a-b}{2p}\left[\frac{1}{(1-r_2)S_2^2} - \frac{1}{(1-r_1)S_1^2}\right].$$

Zauważmy, że zachodzi związek

$$T(z) - \left(\frac{1}{2}p\ln\frac{N_2}{N_1} + K\right) = V(z), \tag{33}$$

gdzie $T(z)$ określone przez równość (19) jest zgodnym i obciążonym estyma-

torem semibayesowskim kwadratowej funkcji dyskryminacyjnej $U(z)$. Powyższy związek implikuje zgodność i obciążoność estymatora $V(z)$, przy czym

$$(34)\quad E[V(z)] = \frac{1}{2}\ln\frac{\alpha_2}{\alpha_1} + \frac{1}{2}(p-1)\ln\frac{\theta_2}{\theta_1} + \frac{a+b}{2p}\left(\frac{1}{\alpha_2} - \frac{1}{\alpha_1}\right) + \frac{(p-1)a-b}{2}\left(\frac{1}{\theta_2} - \frac{1}{\theta_1}\right) + \frac{1}{2}\left[\psi\left(\frac{N_2}{2}\right) - \psi\left(\frac{N_1}{2}\right)\right] + \frac{1}{2}(p-1)\left\{\psi\left[\frac{1}{2}(p-1)N_2\right] - \psi\left[\frac{1}{2}(p-1)N_1\right]\right\} + \frac{1}{2}p\ln\frac{N_1}{N_2} + 2\left(\frac{1}{N_2} - \frac{1}{N_1}\right)$$

gdzie $\alpha_i$ i $\theta_i$ $(i = 1, 2)$ określone są przez równości (9).

Uwaga. Można łatwo znaleźć estymator nieobciążony funkcji $U(z)$ postaci

$$(35)\quad U^*(z) = \frac{1}{2}\ln\frac{S_2^2[1+(p-1)r_2]}{S_1^2[1+(p-1)r_1]} + \frac{1}{2}(p-1)\ln\frac{S_2^2(1-r_2)}{S_1^2(1-r_1)} + \frac{a+b}{2p}\left\{\frac{1}{S_2^2[1+(p-1)r_2]} - \frac{1}{S_1^2[1+(p-1)r_1]}\right\} + \frac{(p-1)a-b}{2p}\left[\frac{1}{S_2^2(1-r_2)} - \frac{1}{S_1^2(1-r_1)}\right] + \frac{1}{2}\left[\psi\left(\frac{N_1}{2}\right) - \psi\left(\frac{N_2}{2}\right)\right] + \frac{p-1}{2}\left\{\psi\left[\frac{1}{2}(p-1)N_1\right] - \psi\left[\frac{1}{2}(p-1)N_2\right]\right\} + \frac{1}{2}p\ln\frac{N_2}{N_1} + 2\left(\frac{1}{N_1} - \frac{1}{N_2}\right).$$

**Literatura cytowana**


[1] H. Cramer (1958), *Metody matematyczne w statystyce*, PWN, Warszawa.

[2] M. M. Desu, S. Geisser (1973), *Methods and applications of equalmean discrimination*, In: *Discriminant Analysis and Applications*, T. Cacoullos, Academic Press, New York, 139-159.

[3] T. S. Ferguson (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.

[4] M. Fisz (1967), *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, Warszawa.

[5] T. Gerstenkorn, T. Śródka (1983), *Kombinatoryka i rachunek prawdopodobieństwa*, PWN, Warszawa.

[6] W. Pasewicz (1978), *Klasyfikacja wielowymiarowych obserwacji w przypadku równych wektorów średnich*, Ósme Colloquium Metodologiczne z Agro-Biometrii, PAN, 388-395.

[7] W. Pasewicz (1985), *Bayesowska metoda dyskryminacyjna w przypadku specjalnej struktury macierzy kowariancji*, Matematyka Stosowana 26, 119-126.

[8] C. R. Rao (1965), *Linear Statistical Inference and Its Applications*, New York.