BAB IV

DASAR-DASAR DISKRIMINAN

Misalkan bahwa kita mempunyai k populasi, yaitu π_1, π_2, ... , π_k. Kita akan mengklasifikasikan suatu individu dengan observasi $\underset{\sim}{x} = (\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots , \underset{\sim}{x}_p)'$ atau suatu group dari n individu dengan observasi $\underset{\sim}{x}_j = (\underset{\sim}{x}_{j1}, \ldots , \underset{\sim}{x}_{jp})'$, dimana j = 1, ... , n, dalam p karakter yang berbeda, kedalam salah satu dari $\pi_1, \pi_2, \ldots , \pi_k$.

Misal k populasi dituliskan dengan :

π_1 , π_2 , ... , π_k

$(\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p)'$, $(\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_p)'$, ... , $(\underset{\sim}{z}_1, \underset{\sim}{z}_2, \ldots, \underset{\sim}{z}_p)'$

atau :

$$\begin{pmatrix} x_{11} & x_{21} & \cdots x_{p1} \\ x_{12} & x_{22} & \cdots x_{p2} \\ \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & \cdots x_{pn} \end{pmatrix}, \quad \begin{pmatrix} y_{11} & y_{21} & \cdots y_{p1} \\ y_{12} & y_{22} & \cdots y_{p2} \\ \vdots & \vdots & \vdots \\ y_{1n} & y_{2n} & \cdots y_{pn} \end{pmatrix}, \quad \ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots, \quad \begin{pmatrix} z_{11} & z_{21} & \cdots z_{p1} \\ z_{12} & z_{22} & \cdots z_{p2} \\ \vdots & \vdots & \vdots \\ z_{1n} & z_{2n} & \cdots z_{pn} \end{pmatrix},$$

Selanjutnya kita asumsikan bahwa setiap π_i mempunyai fungsi distribusi F_i dari vektor random $\underset{\sim}{x} = (\underset{\sim}{x}_1, \ldots , \underset{\sim}{x}_p)'$. Kita asumsikan bahwa bentuk fungsi F_i, untuk setiap i diketahui dan F_i berbeda untuk nilai i yang berlainan, walaupun parameter dari F_i diketahui maupun tidak diketahui.

Misalkan E^p adalah ruang euclid p dimensi, kita akan membagi E^p kedalam daerah yang saling asing yaitu R_1, R_2, ... , R_k , sehingga jika $\underset{\sim}{x}$ atau ($\underset{\sim}{x}_j$, j = 1, ... , n) berada dalam R_i, kita dapat menunjukkan individu atau suatu group dalam π_i. Tetapi, dalam menggunakan aturan klasifikasi, kita mungkin membuat suatu kesalahan, dengan suatu kesalahan klasifikasi individu untuk π_i apabila sebenarnya individu itu untuk π_j ($i \neq j$).

Misalkan bahwa harga kesalahan klasifikasi individu untuk π_j apabila sebenarnya individu tersebut untuk π_i adalah C(j/i). Umumnya C(j/i) tidak semuanya sama, tergantung pada kesalahan relatif yang penting. Selanjutnya diasumsikan bahwa tidak ada penyimpangan untuk klasifikasi yang benar. Dengan kata lain C(i/i) = 0, untuk semua i.

Kita akan mengklasifikasikan suatu individu tunggal dengan observasi $\underset{\sim}{x}$ kedalam salah satu π_i (i = 1, 2,...,k). Misalkan $\underset{\sim}{r}$ = (R_1, ... , R_k)'. Kita akan menunjukkan pembagian aturan klasifikasi untuk ruang E dalam keadaan terpisah, menjadi R_1, ... , R_k dalam $\underset{\sim}{r}$. Probabilitas kesalahan klasifikasi individu dengan observasi $\underset{\sim}{x}$ pada π_i tetapi dimasukkan dalam π_j adalah :

$$P(j/i,\underset{\sim}{r}) = \int_{R_j} f_i(\underset{\sim}{x})\, d\underset{\sim}{x}, \qquad \text{..... (4.1)}$$

dimana :

$$d\underset{\sim}{x} = \prod_{i=1}^{p} dx_i$$

Harapan harga kesalahan klasifikasi pada observasi π_i diatas adalah diberikan dengan :

$$r_i(\underset{\sim}{r}) = \sum_{j=1,\, j \neq i}^{k} C(j/i)\, P(j/i, \underset{\sim}{r}) \quad , \; i = 1, 2, \ldots, k \qquad \ldots\ldots\ldots\ldots \; (4.2)$$

4.1. ATURAN KLASIFIKASI

DEFINISI 4.1.1.

Diberikan dua aturan klasifikasi $\underset{\sim}{r}$ dan $\underset{\sim}{r}^*$, kita katakan bahwa $\underset{\sim}{r}$ lebih baik atau sama dengan $\underset{\sim}{r}^*$, jika $r_i(\underset{\sim}{r}) \leq r_i(\underset{\sim}{r}^*)$ untuk semua i. Dan $\underset{\sim}{r}$ lebih baik dari $\underset{\sim}{r}^*$, apabila $r_i(\underset{\sim}{r}) < r_i(\underset{\sim}{r}^*)$.

DEFINISI 4.1.2.

Aturan klasifikasi $\underset{\sim}{r}$ dikatakan Admissible jika tidak ada aturan klasifikasi $\underset{\sim}{r}^*$ yang lebih baik daripada $\underset{\sim}{r}$.

DEFINISI 4.1.3.

Suatu klass dari aturan klasifikasi dikatakan komplit, jika untuk semua aturan $\underset{\sim}{r}^*$ diluar klass ini, kita dapat menentukan aturan $\underset{\sim}{r}$ didalam klass yang lebih baik dari $\underset{\sim}{r}^*$.

DEFINISI 4.1.4.

Aturan klasifikasi $\underset{\sim}{r}^*$ dikatakan minimax diantara klass dari semua aturan $\underset{\sim}{r}$ jika :

$$\max_i r_i(\underset{\sim}{r}^*) = \min_{\underset{\sim}{r}} \max_i r_i(\underset{\sim}{r})$$

Misalkan bahwa p_i menunjukkan proporsi dari π_i dalam populasi, i = 1, 2, ... , k . Jika p_i diketahui, kita dapat menentukan mean dari kesalahan klasifikasi individu dengan menggunakan aturan $\underset{\sim}{r}$. Karena probabilitas pengambilan observasi dari π_i adalah p_i, maka probabilitas pengambilan observasi dari π_i dan secara tepat klasifikasi

itu kedalam π_i dengan pertolongan aturan $\underset{\sim}{r}$, diberikan dengan :

$$p_i \; P(i/i,\underset{\sim}{r}) \quad , \; i = 1, 2, \ldots , k$$

Dengan cara sama, probabilitas pengambilan observasi π_i dari kesalahan klasifikasi ke π_j dimana $i \neq j$ adalah :

$$p_i \; P(j/i,\underset{\sim}{r})$$

Jadi harga :

$$\sum_{i=1}^{k} p_i \sum_{j=1, j\neq i}^{k} C(j/i) \; P(j/i,\underset{\sim}{r}) \qquad \ldots\ldots\ldots \qquad (4.4)$$

adalah mean harga kesalahan klasifikasi untuk aturan $\underset{\sim}{r}$ terhadap prior probabilitas $\underset{\sim}{p} = (p_1, p_2, \ldots , p_k)$.

DEFINISI 4.1.5.

Diberikan $\underset{\sim}{p}$, aturan klasifikasi $\underset{\sim}{r}$ yang meminimumkan mean harga kesalahan klasifikasi disebut Aturan Bayes terhadap $\underset{\sim}{p}$.

Dapat disimpulkan bahwa aturan Bayes dapat menghasilkan probabilitas kesalahan klasifikasi yang besar, disini ada beberapa usaha untuk mengatasi kesulitan ini. Dalam kasus dimana prior probabilitas p_i diketahui, aturan Bayes adalah optimum, dalam arti aturan ini meminimumkan mean harga harapan.

Kita sekarang akan mengevaluasi bentuk eksplisit dari aturan Bayes ini, dalam kasus dimana setiap π_i diambil dari fungsi density probabilitas f_i, $i = 1, \ldots, k$. Kita asumsikan bahwa semua prosedur klasifikasi dianggap sama, jika hal itu hanya berbeda pada kumpulan dari ukuran probabilitas berharga nol.

4.2. CONTOH

Misalkan bahwa :

$$f_i(\underset{\sim}{x}) = \begin{cases} \beta_i^{-1} \exp(-\underset{\sim}{x} / \beta_i) & , \quad 0 < \underset{\sim}{x} < \infty \\ 0 \text{ untuk yang lain} \end{cases}$$

$i = 1, 2, \ldots, k$ dan $\beta_1 < \ldots < \beta_k$ adalah parameter yang tidak diketahui, dan misalkan bahwa $p_i = i/k$, $i = 1, 2, \ldots, k$. Jika $\underset{\sim}{x}$ adalah observasi, aturan Bayes dengan $C(i/j)$ sama, menghendaki kita untuk mangklasifikasikan $\underset{\sim}{x}$ pada π_i apabila :

$$p_i f_i(\underset{\sim}{x}) \geq \max_{(j \neq i)} p_j f_j(\underset{\sim}{x}) , \quad \ldots\ldots\ldots\ldots\ldots \quad (4.5)$$

dengan perkataan lain, untuk $i < j$ apabila :

$$\beta_i^{-1} \exp(-\underset{\sim}{x}/\beta_i) \geq \beta_j^{-1} \exp(-\underset{\sim}{x}/\beta_j) \text{ dipenuhi ,}$$

jika dan hanya jika :

$$\underset{\sim}{x} \leq \frac{\beta_i \beta_j}{\beta_j - \beta_i} \log(\beta_j - \beta_i)$$

Maka sangatlah mudah untuk menunjukkan bahwa ini adalah fungsi naik dari β_j untuk $\beta_i < \beta_j$ tertentu dan fungsi naik dari β_i untuk $\beta_j < \beta_i$ tertentu. Karena $f_i(\underset{\sim}{x})$ menurun dalam $\underset{\sim}{x}$ untuk $\underset{\sim}{x} > 0$, berarti bahwa kita mengklasifikasikan $\underset{\sim}{x}$ pada π_i apabila :

$$\underset{\sim}{x}_{i-1} \leq \underset{\sim}{x} \leq \underset{\sim}{x}_i ,$$

dimana : $\underset{\sim}{x}_0 = 0$, $\underset{\sim}{x}_k = \infty$ dan

$$\underset{\sim}{x}_i = \frac{\beta_i \beta_{i+1}}{\beta_{i+1} - \beta_i} (\log \beta_{i+1} - \log \beta_i)$$

Ini penting sebagai catatan bahwa jika p_i sebanding pada

β_i, maka aturan Bayes tidak membuat observasi pada individu dan akan selau mengklasifikasikan pada π_k.

Sebagai contoh yang lain adalah : misalkan bahwa

$$f_i(X) = \begin{cases} \dfrac{1}{(2\pi)^{1/2}} \exp\{-1/2(X-\mu_i)^2\} & -\infty < X < \infty \\ 0 \quad \text{untuk yang lain} \end{cases}$$

dimana μ_i adalah parameter yang tidak diketahui, dan misalkan $p_j = i/k$, $i = 1, 2, \ldots, k$. Aturan Bayes dengan $C(i/j)$ sama, menghendaki kita untuk mengklasifikasikan observasi X pada π_j apabila :

$$(X - \mu_j)^2 \geq \max_{i, i \neq j} \{(X - \mu_i)^2\} , \quad \ldots\ldots\ldots\ldots \quad (4.6)$$

Untuk kasus khusus $k = 2$, aturan klasifikasi Bayes lawan priór (p_1, p_2) diberikan dengan menunjukkan X pada :

$$\begin{cases} \pi_1, \text{ apabila} & \dfrac{f_1(X)}{f_2(X)} > \dfrac{C(1/2)p_2}{C(2/1)p_1} \\ \pi_2, \text{ apabila} & \dfrac{f_1(X)}{f_2(X)} < \dfrac{C(1/2)p_2}{C(2/1)p_1} \\ \pi_1 \text{ atau } \pi_2, \text{ apabila} & \dfrac{f_1(X)}{f_2(X)} = \dfrac{C(1/2)p_2}{C(2/1)p_1} \end{cases} \quad \ldots\ldots \quad (4.7)$$

Meskipun demikian, jika berada dalam π_i, $i = 1,2$

$$P\left\{ \frac{f_1(X)}{f_2(X)} = \frac{C(1/2)p_2}{C(2/1)p_1} \,\middle|\, \pi_i \right\} = 0 \quad \ldots\ldots\ldots\ldots \quad (4.8)$$

maka aturan Bayes adalah tunggal kecuali untuk kumpulan dari ukuran probabilitas berharga nol.

4.3. ATURAN LIKELIHOOD

Rasio likelihood aturan klasifikasi $\underset{\sim}{r} = (R_1, R_2, \ldots, R_k)$ ditentukan dengan:

$$R_j : C_j f_j(\underset{\sim}{x}) > \max_{i,i\neq j} C_i f_i(\underset{\sim}{x}) \quad \ldots\ldots\ldots\ldots \quad (4.9)$$

untuk konstanta positip $C_1, \ldots, C_k$. Khusus, jika C_i semua sama, maka aturan klasifikasinya disebut Aturan Likelihood Maximum.

Jika distribusi F_i tidak lengkap diketahui (mean, variansinya tidak diketahui), pelengkap informasi pada F_i atau parameter didalamnya diperoleh langsung dari sampel dalam populasi yang bersesuaian. Maka asumsi lengkap diketahui dari F_i, aturan klasifikasi yang baik apabila $\underset{\sim}{R} = (R_1, R_2, \ldots, R_k)$ antara lain memenuhi Bayes, Minimax dan Aturan Rasio Likelihood.