MATERIAL SUPLEMENTAR

O limite ou limiar de discriminação é determinado as probabilidades de uma amostra ser atribuída à classe A e B são iguais: $P(\hat{y}_i|A) = P(\hat{y}_i|B)$. O que é equivalente à equação a seguir:

$$p(Y|A).P(A) = p(Y|B).P(B) \; (1)$$

Onde:

$$p(Y|A) = \frac{1}{s_A\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_A}{s_A}\right)^2} \; (2)$$

$$p(Y|B) = \frac{1}{s_B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_B}{s_B}\right)^2} \; (3)$$

$$P(A) = \frac{I_A}{I_A + I_B} \; (4)$$

$$P(B) = \frac{I_B}{I_A + I_B} \; (5)$$

Onde

- p($\hat{y}_i$|A) e p($\hat{y}_i$|B) são as funções de densidade de probabilidade para as classes A e B, respectivamente;
- s_A e s_B os desvios padrões das estimativas obtidas no conjunto treinamento para as classes A e B, respectivamente;
- $\bar{y}_A$e $\bar{y}_B$ as médias das estimativas de y obtidas no conjunto treinamento para as classes A e B, respectivamente;
- $P(A)$ e $P(B)$ são as probabilidades *a priori* obtidas no conjunto treinamento para as classes A e B, respectivamente;
- I_A e I_B o número de amostras do conjunto de treinamento para as classes A e B, respectivamente;

Para obter a equação que permite determinar o limite de discriminação, iniciamos substituindo as equações 2 e 3 em 1, obtendo:

$$\frac{1}{s_A\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_A}{s_A}\right)^2}.P(A) = \frac{1}{s_B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_B}{s_B}\right)^2}.P(B) \; (6)$$

Multiplicando os dois lados da equação por $\sqrt{2\pi}$, a expressão pode ser simplificada a:

$$\frac{1}{s_A} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_A}{s_A}\right)^2}.P(A) = \frac{1}{s_B} e^{-\frac{1}{2}\left(\frac{y-\bar{y}_B}{s_B}\right)^2}.P(B) \; (7)$$

Aplicando a propriedade $a^{-n} = \frac{1}{a^n}$·, se obtém:

$$\frac{1}{S_A}.\frac{1}{e^{\frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2}}.P(A) = \frac{1}{S_B}.\frac{1}{e^{\frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2}}.P(B) \ (8)$$

$$\frac{P(A)}{S_A}.\frac{1}{e^{\frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2}} = \frac{P(B)}{S_B}.\frac{1}{e^{\frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2}} \ (9)$$

Aplicando a função de logaritmo natural (ln) dos dois lados, temos que:

$$\ln\left[\frac{P(A)}{S_A}\right] + \ln(1) - \ln\left[e^{\frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2}\right] = \ln\left[\frac{P(B)}{S_B}\right] + \ln(1) - \ln\left[e^{\frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2}\right] \ (10)$$

Sabendo que $\ln(1) = 0$ e $\ln(e^x) = x$ e rearranjando a equação temos que:

$$\ln\left[\frac{P(A)}{S_A}\right] - \frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2 = \ln\left[\frac{P(B)}{S_B}\right] - \frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2 (11)$$

$$-\frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2 + \frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2 + ln\left[\frac{P(A)}{S_A}\right] - ln\left[\frac{P(B)}{S_B}\right] = 0 \ (12)$$

$$-\frac{1}{2}\left(\frac{y-\bar{y}_A}{S_A}\right)^2 + \frac{1}{2}\left(\frac{y-\bar{y}_B}{S_B}\right)^2 + ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (13)$$

Multiplicando por (-2) e resolvendo os parênteses, temos que:

$$\left(\frac{y-\bar{y}_A}{S_A}\right)^2 - \left(\frac{y-\bar{y}_B}{S_B}\right)^2 - 2ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (14)$$

$$\frac{(y^2-2y\bar{y}_A+\bar{y}_A^2)}{S_A^2} - \frac{(y^2-2y\bar{y}_B+\bar{y}_B^2)}{S_B^2} - 2ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (15)$$

Fazendo o mínimo múltiplo comum, temos que:

$$S_B^2(y^2 - 2y\bar{y}_A + \bar{y}_A^2) - S_A^2(y^2 - 2y\bar{y}_B + \bar{y}_B^2) - 2S_A^2S_B^2 ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (16)$$

$$S_B^2y^2 - 2S_B^2y\bar{y}_A + S_B^2\bar{y}_A^2 - S_A^2y^2 + 2S_A^2y\bar{y}_B - S_A^2\bar{y}_B^2 - 2S_A^2S_B^2 ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (17)$$

Finalmente, reorganizando a equação, obtemos a seguinte solução:

$$(S_B^2 - S_A^2)y^2 + (2S_B^2\bar{y}_A - 2S_A^2\bar{y}_B)y + S_B^2\bar{y}_A^2 - S_A^2\bar{y}_B^2 - 2S_A^2S_B^2 ln\left[\frac{P(A)}{P(B)}.\frac{S_A}{S_B}\right] = 0 \ (18)$$

Caso as probabilidades de ocorrência das classes A ou B forem iguais (P(A)=P(B)), a solução é simplificada a:

$$(S_B^2 - S_A^2)y^2 + (2S_B^2\bar{y}_A - 2S_A^2\bar{y}_B)y + S_B^2\bar{y}_A^2 - S_A^2\bar{y}_B^2 - 2S_A^2S_B^2 ln\left[\frac{S_A}{S_B}\right] = 0 \ (19)$$

Pode-se observar que as equações 18 e 19 são equações de segundo grau e que a única variável é y. Substituindo os valores experimentais e resolvendo a equação para y, o limiar ou limite de discriminação é determinado por uma das raízes da equação.