1 Chapter 7 Discrimination and Detection

Problem 7.7 *We want to choose our criterion k for a decision-making task such that the probability of being correct is maximized. Starting from equation (7.20), derive an expression for the criterion k.*

Answer

Here is the situation, assume we are in a binary classification setting with two categories $C = -1$ and $C = +1$. The observer sees a stimulus s and has noisy measurements.

Based on these, they choose a criterion k such that they respond:

$$\hat{C} = \begin{cases} +1 & \text{if } x > k \\ -1 & \text{if } x < k \end{cases}$$

Let $p(C)$ denote the prior over class and $p(x \mid C)$ denote the class-conditional stimulus distribution (assume Gaussian with mean μ_C, shared variance σ^2).

From Eq. (7.20), the probability of being correct is:

$$p(\text{correct}) = p(C = +1) \cdot \Phi\left(\frac{\mu_1 - k}{\sigma}\right) + p(C = -1) \cdot \Phi\left(\frac{k - \mu_{-1}}{\sigma}\right)$$

Our goal is to maximize this with respect to k. Differentiate with respect to k:

$$\frac{d}{dk} p(\text{correct}) = -p(C = +1) \cdot \phi\left(\frac{\mu_1 - k}{\sigma}\right) \cdot \left(\frac{1}{\sigma}\right) + p(C = -1) \cdot \phi\left(\frac{k - \mu_{-1}}{\sigma}\right) \cdot \left(\frac{1}{\sigma}\right)$$

Maximizing so derivative should be set to zero:

$$-p(C = +1) \cdot \phi\left(\frac{\mu_1 - k}{\sigma}\right) + p(C = -1) \cdot \phi\left(\frac{k - \mu_{-1}}{\sigma}\right) = 0$$

Now use the fact that $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$:

$$p(C = +1) \cdot \exp\left(-\frac{(\mu_1 - k)^2}{2\sigma^2}\right) = p(C = -1) \cdot \exp\left(-\frac{(k - \mu_{-1})^2}{2\sigma^2}\right)$$

Take log of both sides:

$$\log p(C = +1) - \frac{(\mu_1 - k)^2}{2\sigma^2} = \log p(C = -1) - \frac{(k - \mu_{-1})^2}{2\sigma^2}$$

Multiply both sides by $2\sigma^2$ and rearrange:

$$(\mu_1 - k)^2 - (k - \mu_{-1})^2 = 2\sigma^2 \log\left(\frac{p(C = +1)}{p(C = -1)}\right)$$

Solve the left-hand side:

$(\mu_1 - k)^2 - (k - \mu_{-1})^2 = (\mu_1^2 - 2k\mu_1 + k^2) - (k^2 - 2k\mu_{-1} + \mu_{-1}^2) = \mu_1^2 - 2k\mu_1 - k^2 + 2k\mu_{-1} - \mu_{-1}^2$

Simplify:

$$(\mu_1^2 - \mu_{-1}^2) + 2k(\mu_{-1} - \mu_1) = 2\sigma^2 \log\left(\frac{p(C=+1)}{p(C=-1)}\right)$$

Solve for k:

$$k = \frac{1}{2(\mu_{-1} - \mu_1)}\left[2\sigma^2 \log\left(\frac{p(C=+1)}{p(C=-1)}\right) - (\mu_1^2 - \mu_{-1}^2)\right]$$

This is the optimal criterion k that maximizes the probability of being correct.

■