Problems

⊕ *Section 2.1*

1. In the two-category case, under the Bayes' decision rule the conditional error is given by Eq. 7. Even if the posterior densities are continuous, this form of the conditional error virtually always leads to a discontinuous integrand when calculating the full error by Eq. 5.

(a) Show that for arbitrary densities, we can replace Eq. 7 by $P(error|x) = 2P(\omega_1|x)P(\omega_2|x)$ in the integral and get an upper bound on the full error.

(b) Show that if we use $P(error|x) = \alpha P(\omega_1|x)P(\omega_2|x)$ for $\alpha < 2$, then we are not guaranteed that the integral gives an upper bound on the error.

(c) Analogously, show that we can use instead $P(error|x) = P(\omega_1|x)P(\omega_2|x)$ and get a lower bound on the full error.

(d) Show that if we use $P(error|x) = \beta P(\omega_1|x)P(\omega_2|x)$ for $\beta > 1$, then we are not guaranteed that the integral gives an lower bound on the error.

⊕ *Section 2.2*

2. Consider minimax criterion for the zero-one loss function, i.e., $\lambda_{11} = \lambda_{22} = 0$ and $\lambda_{12} = \lambda_{21} = 1$.

(a) Prove that in this case the decision regions will satisfy

$$\int\limits_{\mathcal{R}_2} p(\mathbf{x}|\omega_1)d\mathbf{x} = \int\limits_{\mathcal{R}_1} p(\mathbf{x}|\omega_2)d\mathbf{x}$$

(b) Is this solution always unique? If not, construct a simple counterexample.

3. Consider the minimax criterion for a two-category classification problem.

(a) Fill in the steps of the derivation of Eq. 22.

(b) Explain why the overall Bayes risk must be concave down as a function of the prior $P(\omega_1)$, as shown in Fig. 2.4.

(c) Assume we have one-dimensional Gaussian distributions $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$ but completely unknown prior probabilities. Use the minimax criterion to find the optimal decision point x^* in terms of μ_i and σ_i under a zero-one risk.

(d) For the decision point x^* you found in (**??**), what is the overall minimax risk? Express this risk in terms of an error function $\text{erf}(\cdot)$.

(e) Assume $p(x|\omega_1) \sim N(0, 1)$ and $p(x|\omega_2) \sim N(1/2, 1/4)$, under a zero-one loss. Find x^* and the overall minimax loss.

(f) Assume $p(x|\omega_1) \sim N(5, 1)$ and $p(x|\omega_2) \sim N(6, 1)$. Without performing any explicit calculations, determine x^* for the minimax criterion. Explain your reasoning.

4. Generalize the minimax decision rule in order to classify patterns from three categories having triangle densities as follows:

$$p(x|\omega_i) = T(\mu_i, \delta_i) \equiv \begin{cases} (\delta_i - |x - \mu_i|)/\delta_i^2 & \text{for } |x - \mu_i| < \delta_i \\ 0 & \text{otherwise,} \end{cases}$$

where $\delta_i > 0$ is the half-width of the distribution ($i = 1, 2, 3$). Assume for convenience that $\mu_1 < \mu_2 < \mu_3$, and make some minor simplifying assumptions about the δ_i's as needed, to answer the following:

(a) In terms of the priors $P(\omega_i)$, means and half-widths, find the optimal decision points x_1^* and x_2^* under a zero-one (categorization) loss.

(b) Generalize the minimax decision rule to *two* decision points, x_1^* and x_2^* for such triangular distributions.

(c) Let $\{\mu_i, \delta_i\} = \{0, 1\}, \{.5, .5\}$, and $\{1, 1\}$. Find the minimax decision rule (i.e., x_1^* and x_2^*) for this case.

(d) What is the minimax risk?

5. Consider the Neyman-Pearson criterion for two univariate normal distributions: $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$ and $P(\omega_i) = 1/2$ for $i = 1, 2$. Assume a zero-one error loss, and for convenience $\mu_2 > \mu_1$.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $p(x|\omega_1) \sim N(-1, 1)$ and $p(x|\omega_2) \sim N(1, 1)$ and $E_1 = 0.05$.

(e) Compare your result to the Bayes error rate (i.e., without the Neyman-Pearson conditions).

6. Consider Neyman-Pearson criteria for two Cauchy distributions in one dimension

$$p(x|\omega_i) = \frac{1}{\pi b} \cdot \frac{1}{1 + \left(\frac{x - a_i}{b}\right)^2}, \qquad i = 1, 2.$$

Assume a zero-one error loss, and for simplicity $a_2 > a_1$, the same "width" b, and equal priors.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $b = 1$ and $a_1 = -1$, $a_2 = 1$ and $E_1 = 0.1$.

4. Generalize the minimax decision rule in order to classify patterns from three categories having triangle densities as follows:

$$p(x|\omega_i) = T(\mu_i, \delta_i) \equiv \begin{cases} (\delta_i - |x - \mu_i|)/\delta_i^2 & \text{for } |x - \mu_i| < \delta_i \\ 0 & \text{otherwise,} \end{cases}$$

where $\delta_i > 0$ is the half-width of the distribution ($i = 1, 2, 3$). Assume for convenience that $\mu_1 < \mu_2 < \mu_3$, and make some minor simplifying assumptions about the δ_i's as needed, to answer the following:

(a) In terms of the priors $P(\omega_i)$, means and half-widths, find the optimal decision points x_1^* and x_2^* under a zero-one (categorization) loss.

(b) Generalize the minimax decision rule to *two* decision points, x_1^* and x_2^* for such triangular distributions.

(c) Let $\{\mu_i, \delta_i\} = \{0, 1\}, \{.5, .5\}$, and $\{1, 1\}$. Find the minimax decision rule (i.e., x_1^* and x_2^*) for this case.

(d) What is the minimax risk?

5. Consider the Neyman-Pearson criterion for two univariate normal distributions: $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$ and $P(\omega_i) = 1/2$ for $i = 1, 2$. Assume a zero-one error loss, and for convenience $\mu_2 > \mu_1$.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $p(x|\omega_1) \sim N(-1, 1)$ and $p(x|\omega_2) \sim N(1, 1)$ and $E_1 = 0.05$.

(e) Compare your result to the Bayes error rate (i.e., without the Neyman-Pearson conditions).

6. Consider Neyman-Pearson criteria for two Cauchy distributions in one dimension

$$p(x|\omega_i) = \frac{1}{\pi b} \cdot \frac{1}{1 + \left(\frac{x - a_i}{b}\right)^2}, \qquad i = 1, 2.$$

Assume a zero-one error loss, and for simplicity $a_2 > a_1$, the same "width" b, and equal priors.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $b = 1$ and $a_1 = -1$, $a_2 = 1$ and $E_1 = 0.1$.

(e) Compare your result to the Bayes error rate (i.e., without the Neyman-Pearson conditions).

⊕*Section 2.4*

7. Let the conditional densities for a two-category one-dimensional problem be given by the Cauchy distribution described in Problem 6.

(a) By explicit integration, check that the distributions are indeed normalized.

(b) Assuming $P(\omega_1) = P(\omega_2)$, show that $P(\omega_1|x) = P(\omega_2|x)$ if $x = (a_1 + a_2)/2$, i.e., the minimum error decision boundary is a point midway between the peaks of the two distributions, regardless of b.

(c) Plot $P(\omega_1|x)$ for the case $a_1 = 3$, $a_2 = 5$ and $b = 1$.

(d) How do $P(\omega_1|x)$ and $P(\omega_2|x)$ behave as $x \rightarrow -\infty$? $x \rightarrow +\infty$? Explain.

8. Use the conditional densities given in Problem 6, and assume equal prior probabilities for the categories.

(a) Show that the minimum probability of error is given by

$$P(\text{error}) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_2 - a_1}{2b}\right|.$$

(b) Plot this as a function of $|a_2 - a_1|/b$.

(c) What is the maximum value of $P(\text{error})$ and under which conditions can this occur? Explain.

9. Consider the following decision rule for a two-category one-dimensional problem: Decide ω_1 if $x > \theta$; otherwise decide ω_2.

(a) Show that the probability of error for this rule is given by

$$P(\text{error}) = P(\omega_1) \int_{-\infty}^{\theta} p(x|\omega_1)\, dx + P(\omega_2) \int_{\theta}^{\infty} p(x|\omega_2)\, dx.$$

(b) By differentiating, show that a necessary condition to minimize $P(\text{error})$ is that θ satisfy

$$p(\theta|\omega_1)P(\omega_1) = p(\theta|\omega_2)P(\omega_2).$$

(c) Does this equation define θ uniquely?

(d) Give an example where a value of θ satisfying the equation actually *maximizes* the probability of error.

10. Consider

(a) True or false: In a two-category one-dimensional problem with continuous feature x, a monotonic transformation of x leave the Bayes error rate unchanged.