1 Consider a two-class, two-dimensional classification task, where the feature vectors in each of the classes ω_1, ω_2 are distributed according to

$$p(\boldsymbol{x}|\omega_1) = \frac{1}{\left(\sqrt{2\pi\sigma_1^2}\right)^2} \exp\left(-\frac{1}{2\sigma_1^2}(\boldsymbol{x}-\boldsymbol{\mu}_1)^T(\boldsymbol{x}-\boldsymbol{\mu}_1)\right)$$

$$p(\boldsymbol{x}|\omega_2) = \frac{1}{\left(\sqrt{2\pi\sigma_2^2}\right)^2} \exp\left(-\frac{1}{2\sigma_2^2}(\boldsymbol{x}-\boldsymbol{\mu}_2)^T(\boldsymbol{x}-\boldsymbol{\mu}_2)\right)$$

with

$$\boldsymbol{\mu}_1 = [1, 1]^T, \quad \boldsymbol{\mu}_2 = [1.5, 1.5]^T, \quad \sigma_1^2 = \sigma_2^2 = 0.2$$

Assume that $P(\omega_1) = P(\omega_2)$ and design a Bayesian classifier

(a) that minimizes the error probability

(b) that minimizes the average risk with loss matrix

$$\Lambda = \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix}$$

Using a pseudorandom number generator, produce 100 feature vectors from each class, according to the preceding pdfs. Use the classifiers designed to classify the generated vectors. What is the percentage error for each case? Repeat the experiments for $\boldsymbol{\mu}_2 = [3.0, 3.0]^T$.

pdfs = probability density functions

2 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?

3 In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0 & i = j \quad i, j = 1, ..., c \\ \lambda_r & i = c + 1 \\ \lambda_s & \text{otherwise,} \end{cases}$$

where λ_r is the loss incurred for choosing the $(c+1)$th action, rejection, and λ_s is the loss incurred for making a substitution error. Show that the minimum risk is obtained if we decide ω_i if $P(\omega_i|\mathbf{x}) \geq P(\omega_j|\mathbf{x})$ for all j and if $P(\omega_i|\mathbf{x}) \geq 1 - \lambda_r/\lambda_s$, and reject otherwise. What happens if $\lambda_r = 0$? What happens if $\lambda_r > \lambda_s$?

4 Consider the classification problem with rejection option.

(a) Use the results of Problem 3 to show that the following discriminant functions are optimal for such problems:

$$g_i(\mathbf{x}) = \begin{cases} p(\mathbf{x}|\omega_i)P(\omega_i) & i = 1, ..., c \\ \frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} p(\mathbf{x}|\omega_j)P(\omega_j) & i = c + 1. \end{cases}$$

(b) Plot these discriminant functions and the decision regions for the two-category one-dimensional case having

- $p(x|\omega_1) \sim N(1, 1)$,
- $p(x|\omega_2) \sim N(-1, 1)$,
- $P(\omega_1) = P(\omega_2) = 1/2$, and
- $\lambda_r/\lambda_s = 1/4$.

(c) Describe qualitatively what happens as λ_r/λ_s is increased from 0 to 1.

(d) Repeat for the case having

- $p(x|\omega_1) \sim N(1, 1)$,
- $p(x|\omega_2) \sim N(0, 1/4)$,
- $P(\omega_1) = 1/3, P(\omega_2) = 2/3$, and
- $\lambda_r/\lambda_s = 1/2$.