Problem 1: (22 pts) Let the conditional densities for a two-category one-dimensional problem be given by the following Cauchy distribution:

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

1. (6 pts) By explicit integration, check that the distribution are indeed normalized.

2. (9 pts) Assuming \( P(\omega_1) = P(\omega_2), \) show that \( P(\omega_1|x) = P(\omega_2|x) \) if \( x = \frac{a_1 + a_2}{2}, \) that is, the minimum error decision boundary is a point midway between the peaks of the two distributions, regardless of \( b. \)

3. (7 pts) Show that the minimum probability of error is given by

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

Problem 2: (21 pts) Let \( \omega_{\text{max}}(x) \) be the state of nature for which \( P(\omega_{\text{max}}|x) \geq P(\omega_i|x) \) for all \( i, i = 1, \ldots, c. \)

1. (7 pts) Show that \( P(\omega_{\text{max}}|x) \geq \frac{1}{c}. \)

2. (7 pts) Show that for the minimum-error-rate decision rule the average probability of error is given by

\[ P(\text{error}) = 1 - \int P(\omega_{\text{max}}|x)p(x)dx. \]

3. (7 pts) Show that \( P(\text{error}) \leq \frac{1}{c}. \)

Problem 3: (22 pts) In many machine learning applications, 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

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

where \( \lambda_r \) is the loss incurred for choosing the \((c + 1)\)th action, rejection, and \( \lambda_s \) is the loss incurred for making any substitution error.

1. (10 pts) Please derive the decision rule with the minimum risk.

2. (6 pts) What happens if \( \lambda_r = 0? \)

3. (6 pts) What happens if \( \lambda_r > \lambda_s? \)

Problem 4: (12 pts + 10 extra points) Let the components of the vector \( x = [x_1, \ldots, x_d]^T \) be binary-valued (0 or 1), and let \( P(\omega_j) \) be the prior probability for the state of nature \( \omega_j \) and \( j = 1, \ldots, c. \) We define

\[ p_{ij} = P(x_i = 1|\omega_j), \quad i = 1, \ldots, d, j = 1, \ldots, c, \]

with the components of \( x_i \) being statistically independent for all \( x \) in \( \omega_j. \)

1. (12 pts) Show that the minimum probability of error is achieved by the following decision rule:

\[ \text{Decide } \omega_k \text{ if } g_k(x) \geq g_j(x) \text{ for all } j \text{ and } k, \]

where

\[ g_j(x) = \sum_{i=1}^{d} x_i \ln \frac{p_{ij}}{1 - p_{ij}} + \sum_{i=1}^{d} \ln(1 - p_{ij}) + \ln P(\omega_j). \]
2. **(10 extra pts)** If the components of \( x \) are ternary valued (1, 0, or \(-1\)), show that a minimum probability of error decision rule can be derived that involves discriminant functions \( g_j(x) \) that are quadratic function of the components \( x_i \).

**Question 5: (23 pts)** Suppose we have three categories with prior probabilities \( P(\omega_1) = 0.5, P(\omega_2) = P(\omega_3) = 0.25 \) and the class conditional probability distributions

\[
p(x|\omega_1) \sim N(0, 1) \\
p(x|\omega_2) \sim N(0.5, 1) \\
p(x|\omega_3) \sim N(1, 1)
\]

where \( N(\mu, \sigma^2) \) represents the normal distribution with density function

\[
p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.
\]

We sample the following sequence of four points: \( x = 0.6, 0.1, 0.9, 1.1 \).

1. (9 pts) Calculate explicitly the probability that the sequence actually came from \( \omega_1, \omega_3, \omega_3, \omega_2 \).
2. (6 pts) Repeat for the sequence \( \omega_1, \omega_2, \omega_2, \omega_3 \).
3. (8 pts) Find the sequence of states having the maximum probability.