1. [10 points] Consider a two-class problem with the following class-conditional probability density functions (pdfs): 

\[ p(x \mid \omega_1) \sim N(0, \sigma^2) \]

and 

\[ p(x \mid \omega_2) \sim N(1, \sigma^2). \]

Show that the threshold, \( \tau \), minimizing the average risk can be written as 

\[ \tau = \frac{1}{2} - \sigma^2 \ln \left( \frac{\lambda_{12} P(\omega_2)}{\lambda_{21} P(\omega_1)} \right), \]

where we have assumed that \( \lambda_{11} = \lambda_{22} = 0 \).

2. [10 points] In many pattern classification problems, the classifier is allowed to reject an input pattern by not assigning it to any one of the \( c \) classes. If the cost of rejection is not too high, it may be a desirable action in some cases. Let 

\[ \lambda(a_i \mid \omega_j) = \begin{cases} 
0, & i = j \ (i, j = 1, \ldots, c) \\
\lambda_r, & i = c + 1 \\
\lambda_s, & i \neq j \ (i, j = 1, \ldots, c),
\end{cases} \]

where \( \lambda_r \) is the loss incurred for rejecting an input pattern and \( \lambda_s \) is the loss incurred for misclassifying the input pattern (known as substitution error).

(a) Show that the minimum risk rule results in the following decision policy.

Assign pattern \( x \) to class \( \omega_i \) if \( P(\omega_i \mid x) \geq P(\omega_j \mid x) \) \( \forall j \) \( \text{AND} \ P(\omega_i \mid x) \geq 1 - \lambda_r / \lambda_s \), else reject it.
(b) Explain what happens if \( \lambda_r = 0 \)? Also, explain what happens if \( \lambda_r > \lambda_s \)?

3. [20 points] Consider a 2-category classification problem involving a single feature \( x \). Assume equal class priors and a 0-1 loss function. The class-conditional densities are as follows:

\[
p(x|\omega_1) = \begin{cases} 
\frac{3}{2} & \text{for } 0 \leq x \leq \frac{2}{3} \\
0 & \text{otherwise}, 
\end{cases}
\]

\[
p(x|\omega_2) = \begin{cases} 
\frac{3}{2} & \text{for } \frac{1}{3} \leq x \leq 1 \\
0 & \text{otherwise}.
\end{cases}
\]

(a) Sketch the two class-conditional densities.
(b) Compute the Bayes decision boundary and show the decision regions in the sketch of (a).
(c) Compute the Bayes classification error.
(d) What will be the Bayes decision boundary if the prior for class \( \omega_1 \) is increased to 0.7? Show the new decision regions in the sketch of (a).

4. Consider a two-class problem with the following class-conditional pdfs:

\[ p(x | \omega_1) \sim N(50, 5) \]

and

\[ p(x | \omega_2) \sim N(40, 10). \]

(a) [5 points] Plot the two class-conditional pdfs in the interval \( x \in [10, 75] \) on the same graph.
(b) [5 points] In another graph, plot the likelihood ratio, \( \frac{p(x|\omega_1)}{p(x|\omega_2)} \), in the range \( x \in [10, 75] \).
(c) [5 points] Recall that the likelihood ratio can be compared against a threshold, say \( \theta \), in order to assign a feature value \( x \) to one of the two classes, \( \omega_1 \) or \( \omega_2 \). Assuming that \( \lambda_{11} = \lambda_{22} = 0, \lambda_{12} = 2, \lambda_{21} = 1, P(\omega_1) = 0.5 \) and \( P(\omega_2) = 0.5 \), what is the value of \( \theta \)? Show this threshold in graph 4(b) and mark the decision regions corresponding to \( \omega_1 \) and \( \omega_2 \).
(d) [5 points] Now suppose that \( \lambda_{11} = \lambda_{22} = 0, \lambda_{12} = 1, \lambda_{21} = 2, P(\omega_1) = 0.5 \) and \( P(\omega_2) = 0.5 \), what is the value of \( \theta \)? Show this threshold also in graph 4(b) and mark the decision regions corresponding to \( \omega_1 \) and \( \omega_2 \). Explain in words, the underlying reason for the change in decision regions in 4(c) and 4(d).

5. [10 points] Consider the following class-conditional densities for a three-class problem involving two-dimensional features:

\[ p(x|\omega_1) \sim N((0,0)^t,I) ; \]

\[ p(x|\omega_2) \sim N((1,1)^t,I) ; \]

\[ p(x|\omega_3) \sim \frac{1}{2} N((0.5,0.5)^t,I) + \frac{1}{2} N((-0.5,0.5)^t,I). \]

(Here, class \( \omega_3 \) conforms to a Gaussian Mixture Model (GMM) with two components - one component is \( N((0.5,0.5)^t,I) \) and the other component is \( N((-0.5,0.5)^t,I) \) - whose weights are equal (i.e., \( \frac{1}{2} \)).)
(a) In a 2D graph, mark the mean of $\omega_1$, $\omega_2$, and the two components of $\omega_3$. In the same graph, mark the point $x = (0.3, 0.3)^t$.

(b) Assuming a 0-1 loss function and equal priors, determine the class to which you will assign the two-dimensional point $x = (0.3, 0.3)^t$ based on the Bayes decision rule.

6. [15 points] Consider the following bivariate density function for the random variable $x$:

$$p(x) \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 25 & 10 \\ 10 & 15 \end{bmatrix} \right).$$

(a) What is the whitening transform of $x$? (You can use matlab (or any other programming language) to compute the eigenvectors/values).

(b) Generate 10,000 bivariate random patterns from this density function (if you are using matlab, then the \textit{mvnrnd} function can be used to generate these patterns). Plot these patterns in a graph.

(c) When the whitening transformation is applied to $x$, what is the density function of the resulting random variable?

(d) Apply the whitening transformation to the 10,000 bivariate patterns generated above. Plot the transformed patterns in a graph. Compare these patterns with those in (b) above. \textbf{What do you observe?}

7. Consider a two-category classification problem involving one feature, $x$. Let both class-conditional densities conform to a Cauchy distribution as follows:

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

Assume a 0-1 loss function and equal priors.

(a) [5 points] Compute the Bayes decision boundary in terms of the parameters $a_1$, $a_2$ and $b$. What is the Bayes decision rule?

(b) [10 points] Show that the probability of error (i.e., misclassification) according to the Bayes decision rule is

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

8. Consider a two-category ($\omega_1$ and $\omega_2$) classification problem with equal priors. Each feature is a two-dimensional vector $x = (x_1, x_2)^t$. The class-conditional densities are:

$$p(x \mid \omega_1) \sim N(\mu_1 = (0, 0)^t, \Sigma_1 = 2I),$$

$$p(x \mid \omega_2) \sim N(\mu_2 = (1, 2)^t, \Sigma_2 = I).$$

(a) [10 points] Compute the Bayes decision boundary.

(b) [5 points] Generate 10,000 bivariate random patterns from each of the two densities (if you are using matlab, then the \textit{mvnrnd} function can be used to generate these patterns). Plot these patterns in a graph using different markers to distinguish the two classes. On the same graph, plot the Bayes decision boundary.
9. Consider a two-category classification problem involving two-dimensional feature vectors of the form \( x = (x_1, x_2)^t \). The two categories are \( \omega_1 \) and \( \omega_2 \), and

\[
\begin{align*}
    p(x \mid \omega_1) &\sim N \left((0, 0)^t, I\right), \\
    p(x \mid \omega_2) &\sim N \left((1, 1)^t, I\right), \\
    P(\omega_1) &= P(\omega_2) = \frac{1}{2}.
\end{align*}
\]

(a) [10 points] Calculate the Bayes decision boundary and write down the Bayes decision rule assuming a 0-1 loss function. The Bayes decision rule has to be written in terms of the Bayes decision boundary.

(b) [5 points] What are the Bhattacharyya and Chernoff theoretical bounds on the probability of misclassification, \( P(\text{error}) \)?

(c) [5 points] Generate \( n = 25 \) test patterns from each of the two class-conditional densities and plot them in a two-dimensional feature space using different markers for the two categories (if you are using matlab, then the \textit{mvnrnd} function can be used to generate these patterns). Draw the Bayes decision boundary on this plot for visualization purposes.

(d) [10 points] What is the confusion matrix and empirical error rate when classifying the generated patterns using the Bayes decision rule? Can the empirical error rate exceed the theoretical bounds on the probability of misclassification? Why or why not?