UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering

ECE 598NA Pattern Recognition
Fall 2006

Mid-Term Exam

Thursday, October 5, 2006

- This is an OPEN BOOK exam.
- You must SHOW YOUR WORK to get full credit.

Problem	Score
1	
2	
3	
4	
5	
Total	

Name: ____

Problem 1 (25 points)

Consider a univariate symmetric Gaussian dichotomizer, that is, a classifier whose likelihood functions $p(x|\omega_1) \sim \mathcal{N}(\mu_1, \sigma)$ and $p(x|\omega_2) \sim \mathcal{N}(\mu_2, \sigma)$ have the same variance σ^2. As the decision boundary x^* is varied, P_M and P_{FA} also vary, where

$$P_{FA} = P(\text{error}|\omega_1)$$
$$P_M = P(\text{error}|\omega_2)$$

A plot of P_{FA} vs. P_M is called a "receiver operating characteristic" (ROC). In order to compare the performance of different systems, we often wish to summarize the ROC using a single summary score. One common summary score, that you have already learned, is d' (d prime). Two other commonly used summary scores are the "equal error rate" (EER) and the "figure of merit" (F-score).

(a) The EER is defined to be the value of P_{FA} for which $P_M = P_{FA}$. Given the EER for a univariate symmetric Gaussian dichotomizer, is it possible to compute d'? Prove your answer.

(b) The F-score is defined to be the harmonic mean of $1 - P_{FA}$ and $1 - P_M$, i.e.

$$\frac{1}{F} = \frac{1}{2}\left(\frac{1}{1 - P_{FA}} + \frac{1}{1 - P_M}\right)$$

Prove that F is not the same for all values of the decision boundary x^*. Find a singular point of the function $F(x^*)$ (you need not determine whether the singular point is a minimum or maximum).

Problem 2 (20 points)

Consider a univariate symmetric Gaussian dichotomizer, with $\mu_2 > \mu_1$. Suppose that the two classes have equal priors, $P(\omega_1) = P(\omega_2) = 1/2$, but unequal costs. In particular, suppose that the costs are given by

$$\begin{aligned} \lambda(\alpha_2|\omega_1) &= 1 \\ \lambda(\alpha_1|\omega_2) &= A \end{aligned}$$

and $\lambda(\alpha_1|\omega_1) = \lambda(\alpha_2|\omega_2) = 0$. Write the minimum Bayes risk classification rule, $\alpha^*(x) = \arg\min \mathcal{R}(\alpha_i|x)$, in terms of the cost parameter A, and in terms of the Gaussian means μ_1 and μ_2 and the Gaussian variance σ^2. There are values of A for which $\alpha^*(x)$ is independent of x. What are they?

Problem 3 (15 points)

Consider a multivariate symmetric Gaussian dichotomizer, i.e., a classifier whose likelihood functions $p(x|\omega_i) \sim \mathcal{N}(\mu_i, \Sigma)$ have the same covariance matrix Σ. Let w be the Fisher linear discriminant vector, and let $g(x) = w^T x$ be the Fisher discriminant. Define the "discriminability" of the Fisher discriminant to be

$$d'_g = \frac{|w^T \mu_1 - w^T \mu_2|}{\sigma_g} \tag{1}$$

where σ_g is the class-conditional standard deviation of the random variable $g(x)$, i.e., $p(g|\omega_1) \sim \mathcal{N}(w^T \mu_1, \sigma_g)$. Write d'_g in terms of μ_1, μ_2, and Σ (i.e., eliminate w and σ_g from the equation). Simplify as far as you can.

Problem 4 (20 points)

Richard Duda's hardware store carries a wide variety of nails. The length of a nail, x, is a random variable described by a PDF of the following form:

$$p(x) = \begin{cases} \theta x^{-(\theta+1)} & x > 1 \\ 0 & \text{otherwise} \end{cases}$$

where the parameter θ is unknown, but we know that $\theta > 0$. Given a large number of observed nail sizes, $\mathcal{D} = \{x_1, \ldots, x_n\}$, you wish to construct a maximum likelihood estimate θ_{ML} of the parameter θ. Find a sufficient statistic s that can be used to estimate θ_{ML}, and determine the value of θ_{ML} in terms of s.

Problem 5 (20 points)

Suppose that you are given n samples from an exponential distribution, $\mathcal{D} = \{x_1, \ldots, x_n\}$, where

$$p(x|\theta) = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Suppose that we are given an *a priori* distribution for the parameter θ: suppose that θ is itself exponentially distributed with hyperparameter a, thus

$$p(\theta) = \begin{cases} a e^{-a\theta} & \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $p(\theta|\mathcal{D})$, the posterior probability of θ given the observations. You may find the following identities useful.

$$\int_a^b u dv = [uv]_a^b - \int_a^b v du$$

$$\sum_{m=0}^{n-1} \frac{a^m}{m!} = \left(1 - \frac{a^n}{n!}\right) e^a$$