

CSE 555 Spring 2010 Homework 1: Bayesian Decision Theory

Jason J. Corso

Computer Science and Engineering

SUNY at Buffalo SUNY

jcorso@buffalo.edu

Date Assigned 13 Jan 2010

Date Due 1 Feb 2010

Homework must be submitted in class. No late work will be accepted.

Problem 1: Bayesian Decision Rule (30%)

Suppose the task is to classify the input signal x into one of K classes $\omega \in \{1, 2, \dots, K\}$ such that the action $\alpha(x) = i$ means classifying x into class i . The Bayesian decision rule is to maximize the posterior probability

$$\alpha_{\text{Bayes}}(x) = \omega^* = \arg \max_{\omega} p(\omega|x) .$$

Suppose we replace it by a *randomized decision rule*, which classifies x to class i following the posterior probability $p(\omega = i|x)$, i.e.,

$$\alpha_{\text{rand}}(x) = \omega \sim p(\omega|x) .$$

1. What is the overall risk R_{rand} for this decision rule? Derive it in terms of the posterior probability using the zero-one loss function.
2. Show that this risk R_{rand} is always no smaller than the Bayes risk R_{Bayes} . Thus, we cannot benefit from the randomized decision.
3. Under what conditions on the posterior are the two decision rules the same?

Problem 2: Bayesian Classification Boundaries for the Normal Distribution (30%)

Suppose we have a two-class recognition problem with salmon ($\omega = 1$) and sea bass ($\omega = 2$).

1. First, assume we have one feature, the pdfs are the Gaussians $\mathcal{N}(0, \sigma^2)$ and $\mathcal{N}(1, \sigma^2)$ for the two classes, respectively.

Show that the threshold τ minimizing the average risk is equal to

$$\tau = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{21}P(\omega_2)}{\lambda_{12}P(\omega_1)} \quad (1)$$

where we have assumed $\lambda_{11} = \lambda_{22} = 0$.

TYP0 — That equation should be

$$\tau = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)} \quad (2)$$

(Tried a question from a different book that abbreviated the λ loss differently!)

2. Next, suppose we have two features $\mathbf{x} = (x_1, x_2)$ and the two class-conditional densities, $p(x|\omega = 1)$ and $p(x|\omega = 2)$, are 2D Gaussian distributions centered at points (3, 9) and (9, 3) respectively with the same covariance matrix $\Sigma = 3\mathbf{I}$ (with \mathbf{I} is the identity matrix). Suppose the priors are $P(\omega = 1) = 0.6$ and $P(\omega = 2) = 0.4$.
 - (a) Suppose we use a Bayes decision rule, write the two discriminant functions $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$.
 - (b) Derive the equation for the decision boundary $g_1(\mathbf{x}) = g_2(\mathbf{x})$. Draw the boundary on the feature space (the 2D plane).

Problem 4: Bayesian Reasoning (40%)

Monty Hall Formulate and solve this classical problem using the Bayes rule. Imagine you are on a gameshow and you're given the choice of three doors: behind one door is a car and behind the other two doors are goats. You have the opportunity to select a door (say No. 1). Then the host, who knows exactly what is behind each door and will not reveal the car, will open a different door (i.e., one that has a goat). The host then asks you if you want to switch your selection to the last remaining door.

1. (3pts) Formulate the problem using the Bayes rule, i.e., what are the random variables and the input data. What are the meaning of the prior and the posterior probabilities in this problem (one sentence each).
2. (3pts) What are the probability values for the prior?
3. (3pts) What are the probability values for the likelihood?
4. (3pts) Derive the posterior probability (include intermediate steps).
5. (3pts) Is it in the contestant's advantage to switch his/her selection? Why?
6. (10pts) Now, consider the following twist. The host is having trouble remembering what is behind each of the doors. So, we cannot guarantee that he will not accidentally open the door for the car. We only know that he will not open the door the contestant has opened. Indeed, if he accidentally opens the door for the car, the contestant wins. How does this change the situation? Is it now in the contestant's advantage to switch? Rederive your probabilities to justify your answer.