National Institute of Technology Calicut

Department of Computer Science & Engineering

Pattern Recognition (CS4044)

Assignment 1

Due Date: Feb 15 2016 - Monday, (In Class)

---

1. (Example 1.10, Ross, Intro to Probability Models) (**Pairwise Independent Events That are Not Independent**). Pairwise independence is an important concept in probability.

   Let a ball be drawn from a box containing four balls, numbered 1,2,3,4. Let $E = \{1, 2\}, F = \{1, 3\}, G = \{1, 4\}$. If all four outcomes are equally likely, what can you say about the independence of pairwise events, E and F, E and G, F and G? What can you say about the independence of E, F and G?

2. (Chapter 1, Ex-29, Ross, Intro to Probability Models) Suppose that $P(E) = 0.6$, what can say about $P(E|F)$, when

   (a) E and F are mutually exclusive?

   (b) $E \subset F$?

   (c) $F \subset E$?

3. (Chapter 1, Ex-40, Ross, Intro to Probability Models)

   (a) A gambler has in his pocket a fair coin and a two-headed coin. He selects one of the coins at random, and when he flips it, it show heads. What is the probability that it is the fair coin?

   (b) Suppose that he flips the same coin a second time and again it show heads. Now what is the probability that it is the fair coin?

   (c) Suppose that he flips the same coin a third time and it shows tails. Now what is the probability that it is the fair coin?

   (See how probabilities vary depending on the observations).

4. Derive $E[X]$, when X is

   1. Bernoulli, Binomial Geometric, Poisson.
   2. Uniformly distributed over (a,b); exponential with parameter λ.

5. Find $Var[X]$, when X is

   1. Bernoulli, Binomial.
   2. Normal

6. At a party, N men throw their hats into the center of a room. The hats are mixed up and each man randomly selects one. Find the expected number of men who select their own hats.

7. Consider a 2 - class problem with feature space $\mathbb{R}$ and equal prior probabilities.
   Let the class conditional densities be given by the pdfs

   $$p_i(x/c_i) = \frac{1}{\pi b} \frac{1}{1 + (\frac{x - a_i}{b})^2}, i = 1, 2$$

   where a_1, a_2 and b are parameters of the class conditional density functions. Assume $a_1 < a_2$

(a) Show that the class conditional densities given by the above equations are indeed pdfs. (ie. they integrate to 1)

$$Hint : \int \frac{1}{1+x^2} = tan^{-1}(x)$$

(b) Find the Bayes/Optimal classifier. (ie. mark the decision regions on the real line)

(c) Show that the minimum probability of error is given by

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

8. Consider a 2-class PR problem with feature vectors in $\mathbb{R}^2$. The class conditional density for class-I is uniform over [1, 3][1, 3] and that for class-II is uniform over [2, 4] [2, 4]. Suppose the prior probabilities are equal. Which is Bayes classifier? (For this problem, assume that we are using 0-1 loss function and hence Bayes classifier minimizes proba- bility of misclassification). Is Bayes Classifier unique for this problem? What would be the probability of misclassification by Bayes Classifier? Consider a hyperplane given by x + y = 5 in $\mathbb{R}^2$. Is this a Bayes (optimal) clasifier? Suppose the prior probabilities are changed to p 1 = 0.4 and p 2 = 0.6. What is a Bayes classifier now? How good would the earlier hyperplane be now?

9. Suppose for a 2-class classification problem, $L(0,1) = 4L(1,0)$ and $L(0,0) = L(1,1) = 0$. The Bayes classifier is then given by

$$h_B(X) = \begin{cases} 0 & \text{if } \frac{p(c_0/x)}{p(c_1/x} > 4 \\ 1 & \text{otherwise} \end{cases}$$

Prove that this Bayes classifier minimizes risk.

10. The error function for regulalrized least square regression is given by

$$\tilde{E} = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, w) - t_n\}^2 + \frac{\lambda}{2} \|w\|^2$$

where $y(x_i, w) = w^T x_i$. Show that the optimal parameter $\hat{w}$ giving minimum error is given by

$$\hat{w} = (X^T X + \lambda I_d)^{-1} X^T \mathbf{t}$$

where X is the design matrix and $\mathbf{t}$ is the label vector (as discussed in class).