Set 2. Due February 13, 2017

Problem 5 Consider a binary classification problem in which the observation X is real valued, $\mathbf{P}\{Y=0\} = \mathbf{P}\{Y=1\} = 1/2$, and the class-conditional cumulative distribution functions are

$$\mathbf{P}\{X \le x | Y=0\} = \begin{cases} 0 & \text{if } x \le 0 \\ x/2 & \text{if } 0 < x \le 2 \\ 1 & \text{if } x > 2 \end{cases} \quad \text{and} \quad \mathbf{P}\{X \le x | Y=1\} = \begin{cases} 0 & \text{if } x \le 1 \\ (x-1)/3 & \text{if } 1 < x \le 4 \\ 1 & \text{if } x > 4 \end{cases}$$

Determine $\eta(x) = \mathbf{P}\{Y=1|X=x\}$. Compute the Bayes classifier and the Bayes risk R^*. Compute the asymptotic risk R_{1-NN} of the nearest neighbor classifier.

Problem 6 Consider a binary classification problem in which both class-conditional densities are multivariate normal of the form

$$f_i(x) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma_i}} e^{-\frac{1}{2}(x-m_i)^T \Sigma_i^{-1}(x-m_i)}, \qquad i = 0, 1,$$

where $m_i = \mathbf{E}[X|Y=i]$ and Σ_i is the covariance matrix for class i. Let $q_0 = \mathbf{P}\{Y=0\}$ and $q_1 = \mathbf{P}\{Y=1\}$ be the a priori probabilities.

Determine the Bayes classifier. Characterize the cases when the Bayes decision is linear (i.e., it is obtained by thresholding a linear function of x).

Problem 7 Let the joint distribution of (X, Y) be such that X is uniform on the interval $[0,1]$, and for all $x \in [0,1]$, $\eta(x) = x$. Determine the prior probabilities $\mathbf{P}\{Y=0\}, \mathbf{P}\{Y=1\}$ and the class-conditional densities $f(x|Y=0)$ and $f(x|Y=1)$.

Calculate R^*,R_{1-NN}, and R_{3-NN} (i.e., the Bayes risk and the asymptotic risk of the 1-, and 3-nearest neighbor rules).

Problem 8 Write a program that generates training data of n i.i.d. pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$ of random variables distributed such that X takes values in $\mathbb{R}^d$ and $Y \in \{0,1\}$. The joint distribution is such that X is uniformly distributed in $[0,1]^d$ and $\mathbf{P}\{Y=1|X=x\} = x^{(1)}$ (where $x^{(1)}$ is the first component of $x = (x^{(1)}, \ldots, x^{(d)})$.

Classify X using the $1,3,5,7,9$-nearest neighbor rules. Re-draw (X,Y) many times so that you can estimate the risk of these rules. Try this for various values of n and d and plot the estimated risk. Explain what you observe.