STOR 565 Homework

1. Consider a classification problem in which the predictor X is uniformly distributed on the unit interval $[0,1]$ and the response $Y \in \{0,1\}$ as usual. For $x \in [0,1]$ let $\eta(x) = \mathbb{P}(Y = 1 \,|\, X = x)$. Specify the Bayes rule ϕ^* and the Bayes risk R^* in each of the following cases.

   a. $\eta(x) = 1/2$ for all x

   b. $\eta(x) = 1/3$ for all x

   c. $\eta(x) = x$

   d. $\eta(x) \in \{0,1\}$ for all x

In each of the cases above, find the prior probability $\pi_1 = \mathbb{P}(Y = 1)$, or indicate why this is not possible without more information.

2. Let $(X,Y) \in \mathbb{R} \times \{-1,+1\}$ be a random predictor-response pair. Suppose that Y has prior probabilities $\pi_1 = \mathbb{P}(Y = 1)$ and $\pi_0 = \mathbb{P}(Y = 0)$, and that X is continuous with marginal density f and class conditional densities f_0 and f_1.

   a. Derive an expression for the Bayes rule $\phi^*(x)$ in terms of the logarithm of the ratio $\pi_1 f_1(x)/\pi_0 f_0(x)$.

Suppose that f_1 is $\mathcal{N}(\mu_1, \sigma^2)$ and that f_0 is $\mathcal{N}(\mu_0, \sigma^2)$ where $\mu_1 > \mu_0$.

   b. Using the result of part (a), find an expression for the Bayes rule $\phi^*(x)$ in terms of the parameters π_0, π_1, μ_0, μ_1, and σ^2.

   c. What is the form of the rule in part (b) when $\pi_1 = 1/2$? Explain why this makes intuitive sense.

   d. Suppose for simplicity that $\mu_1 = u$ and $\mu_0 = -u$ for some $u > 0$. What form does the Bayes rule take when u increases (tends to infinity), and in particular, how does the rule depend on π_1 versus π_0? A informal but clear answer is fine.

3. Find the gradient and Hessian of the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = x_1^2 x_2 + 3x_1 - 5x_2 + 1$$

4. Use Jensen's inequality to find relations among the following. Explain your reasoning.

   a. $\mathbb{E}(1/X)$ and $1/\mathbb{E}X$ for $X > 0$.

   b. $\mathbb{E}(X^{1/3})$ and $(\mathbb{E}X)^{1/3}$ for $X \geq 0$.

5. Consider the labeled data set $(-2, 1), (-1, 1), (0, 0), (1, 1), (2, 0) \in \mathbb{R} \times \{0, 1\}$.

   a. Sketch the 1-nearest neighbor rule for this dataset by drawing a line and indicating which points are assigned to zero and which are assigned to one.

   b. Sketch the 3-nearest neighbor rule for this dataset by drawing a line and indicating which points are assigned to zero and which are assigned to one.

6. Let X have a $\mathcal{N}(\mu, \sigma^2)$ distribution. Find the expected value of X.

7. Let $Z \sim \mathcal{N}(0, 1)$. Use the CDF method to find the density $X = aZ + b$.

8. Let $X \in \mathbb{R}^k$ be a random vector and $A \in \mathbb{R}^{r \times k}$. Establish the following.

   a. $\mathbb{E}(AX) = A\,\mathbb{E}X$

   b. $\text{Var}(X)_{ij} = \text{Cov}(X_i, X_j)$

   c. $\text{Var}(X)$ is symmetric and non-negative definite

   d. $\text{Var}(AX) = A\,\text{Var}(X)A^t$

9. Let X and Y be two jointly distributed random variables. Suppose that we wish to predict the value of Y based on the value of X via a function $g(\cdot)$. Suppose that we judge the quality of the prediction $g(X)$ by the expected squared error $\mathbb{E}(Y - g(X))^2$. Then it turns out that the best estimate of Y given X is the conditional expectation $E(Y \mid X)$.

   a. Let $g(\cdot)$ be a function. Show that $\mathbb{E}[(Y - \mathbb{E}[Y|X])\,(\mathbb{E}[Y|X] - g(X)) \mid X] = 0$. (Hint: expand the product and use basic properties of conditional expectations.)

   b. Note that $(Y - g(X))^2 = (Y - E(Y \mid X) + E(Y \mid X) - g(X))^2$. Show by expanding the square and using the result from (a) that

$$\mathbb{E}[(Y - g(X))^2|X] \;=\; \mathbb{E}[(Y - E[Y|X])^2 \mid X] \,+\, \mathbb{E}[(\mathbb{E}[Y|X] - g(X))^2 \mid X]$$

c. Deduce from (b) that $\mathbb{E}[(Y - g(X))^2] \geq \mathbb{E}[(Y - \mathbb{E}[Y \mid X])^2]$

10. Let $(X, Y) \in \mathbb{R}^2 \times \{-1, +1\}$ be a random predictor-response pair. Suppose that the predictor X is a pair (X_1, X_2) where $X_1, X_2 \in [0, 1]$ are independent, X_1 is uniform on $[0, 1]$, and X_2 has density $g(x_2) = 3x_2^2$ for $0 \leq x_2 \leq 1$. Suppose that $\eta(x_1, x_2) = (x_1 + x_2)/2$.

a. Find the Bayes rule ϕ^* for this problem and identify its decision boundary.

b. Find the unconditional density of X

c. Find the Bayes risk associated with (X, Y)

d. Find the prior probability that $Y = +1$.

e. Find the class-conditional density of X given $Y = 1$.