Worksheet on "Decision Theory (incl. Bayes Classifiers)"

CS5691 PRML Jul–Nov 2025

August 9, 2025

1. Consider a continuous random variable X and a discrete random variable Y. Let
   - $P_Y(Y = 1) = 0.5$ and $P_Y(Y = -1) = 0.5$, and
   - $(X|Y = 1) \sim \text{Unif}(-1, 2)$ and $(X|Y = -1) \sim \text{Unif}(-2, 1)$.

   a. What is the marginal distribution of X? Specifically, plot the pdf of X denoted $f_X(x)$.

   b. Write down the pdf $f_X(x)$ of X.

   c. Write down and plot the posterior $Y|X$.

   d. What is the optimal classifier for predicting Y from X, given the above assumptions?

2. Derive the Bayes classifier for binary classification ($Y = \pm 1$) under the below assumptions:

$$P(Y = 1) = 0.7 \text{ and } P(Y = -1) = 0.3$$
$$X|Y = 1 \sim Unif(-1, 3)$$
$$X|Y = -1 \sim Unif(-2, 0)$$

3. For a binary classifer h, let $L = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ be the loss matrix; and $C_{\text{train}} = \begin{bmatrix} 100 & 10 \\ 20 & 120 \end{bmatrix}$, and $C_{\text{test}} = \begin{bmatrix} 90 & 45 \\ 30 & 85 \end{bmatrix}$ be the confusion matrix when h is applied on the training and test data respectively. All three matrices have ground-truth classes t along the rows and predictions h along the columns in the same order for the two classes. Express your estimate of the expected loss of h in terms of p to s above.

4. Consider the four examples of two jointly distributed rvs (X, Y) from Slide 19 of "M0a. Background on Probabillity", a screenshot of which is shown below. For each of these examples, write down the optimal (Bayes) classifier for predicting Y given X (in case of discrete Y) and optimal regressor for predicting Y given X (in case of continuous Y). Assume that standard loss functions (0-1 loss function for classification and squared loss function for regression) need to be optimized.