

Worksheet on “Decision Theory (incl. Bayes Classifiers)”

CS5691 PRML Jul–Nov 2025

August 9, 2025

- 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 $(X|Y = -1) \sim \text{Unif}(-2, 1)$.
 - What is the marginal distribution of X ? Specifically, plot the pdf of X denoted $f_X(x)$.
 - Write down the pdf $f_X(x)$ of X .
 - Write down and plot the posterior $Y|X$.
 - What is the optimal classifier for predicting Y from X , given the above assumptions?
- Derive the Bayes classifier for binary classification ($Y = \pm 1$) under the below assumptions:

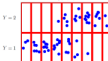
$$\begin{aligned}
 P(Y = 1) &= 0.7 \text{ and } P(Y = -1) = 0.3 \\
 X|Y = 1 &\sim \text{Unif}(-1, 3) \\
 X|Y = -1 &\sim \text{Unif}(-2, 0)
 \end{aligned}$$

- For a binary classifier 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.
- Consider the four examples of two jointly distributed rvs (X, Y) from Slide 19 of “M0a. Background on Probability”, 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.

Conditional distbn./expectation: Example rvs

What is $p(X|Y = 2)$ and $E[X|Y = 2]$ for each of these cases?

- X discr., Y discr.: Already seen example \rightarrow



- X cont., Y discr.: Let $p(X, Y) = p(Y)p(X|Y) = 0.5 \times \mathcal{N}(X | 2Y, \sigma^2)$

pmf pdf assuming uniform prior for any fn. of Y (here it is 2Y)

- X cont., Y cont.: Let W_1, W_2 be two indept. Gaussian rvs, i.e.,

$$p(W_1) = \mathcal{N}(W_1 | \mu_1, \sigma^2), p(W_2) = \mathcal{N}(W_2 | \mu_2, \sigma^2), \text{ \& } p(W_1, W_2) = p(W_1)p(W_2).$$

- Independent rvs: Let $X = W_1$, and $Y = W_2$
- Dependent rvs: Let $X = W_1 + W_2$, and $Y = W_2$