Homework 4
No due Date

Do the following exercises without turning them in. There is no credit for this assignment, but doing these problems will help you prepare for the mid sem and end sem exams.

Problem 1. Do Exercise 4.1 from Murphy's book. We did this in the lecture but try to do it on your own, before you check your notes.

Problem 2. A die is rolled N times, where face c comes up N_c times on top, where $c \in \{1, \ldots, 6\}$. Find the maximum likelihood estimate for the probablility of the c-th face $\theta_c = p(Y = c)$. *Hint*: The probablily of getting the c-th face N_c times is

$$\mathcal{L}(\theta) = \prod_{c=1}^{6} \theta_c^{N_c}$$

Then proceed in a similar manner as with the coin-toss example that we discussed in the lecture (Chap 4, pg 1). You have to take into account that $\sum_{c=1}^{6} \theta_c = 1$. You can check your work by reading Section 4.2.4 in Murphy's book.

Problem 3. In the lecture we derived a quadratic formula for the decision boundaries in Gaussian discriminant analysis (see the lecture notes Chap 9, pg. 2). This was done by setting the posteriors of the classes equal

$$p(X = x|Y = 1)\pi_1 = p(X = x|Y = 2)\pi_2.$$

Re-do this derivation for the special case of one input variable and two categories $D = 1$, $C = 2$. In this case we get a quadratic equation of the form

$$\frac{1}{2}\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right) x^2 + bx + c = 0.$$

This equation can have either zero, one, or two real solutions. Illustrate the three different situations by drawing the graphs of two Gaussians that intersect zero, one, or two times.

Problem 4 Suppose we collect data for a group of students in a statistics class with variables X_1 =, hours studied, X_2 = current grade average, and Y = receive an A or better. We fit a logistic regression and produce estimated coefficient, $b = -6$, $w_1 = 0.05$, $w_2 = 1$.

a. Estimate the probability that a student who studies for 40 hours and has a grade average of 3.5 gets an A in the class.

b. How many hours would the student in part (a) need to study to have a 50 percent chance of getting at least an A in the class?

Problem 5 Suppose that we wish to predict whether a given stock will issue a dividend this year (Yes or No) based on X=last year's percent profit. We examine a large number of companies and discover that the mean value of X for companies that issued a dividend was $\mu = 10$, while the mean for those that didn't was $\mu = 0$. In addition, the variance of X for these two sets of companies was $\sigma^2 = 36$. Finally, 80 percent of companies issued dividends. Assuming that X follows a normal distribution, predict the probability that a company will issue a dividend this year given that its percentage profit was $X = 4$ last year.