Homework 1: Very basic probability theory

CSC 2515, Fall 2004

Due: Friday, September 24, 3pm. Page numbers refer to the first printing of MacKay's textbook.

1. **Problem 2.30, MacKay** (p. 38). Does the result make sense intuitively?
2. **Problem 2.37, MacKay** (p. 39)
3. **Monty Hall with n doors**. The Monty Hall problem is a fascinating case where probability theory disagrees with common sense, but common sense is utterly wrong. The basic problem is described in Exercise 3.8 (p. 57), and the solution is provided on p. 60. (If you are not convinced, you might want to write a short computer program that tests the two strategies over thousands of trials.)

   Now, consider the case where the rules are the same as in the basic problem, but now there are more doors. Specifically, there are n doors, only one of which hides a prize (with uniform prior probability). You select door 1, and the host opens all other doors, except for door k. What is the probability that the prize is behind door k? What if $n = 1,000,000$?
4. **Monty Hall with unknown host**. MacKay's solution to the Monty Hall problem (three doors) makes some assumptions that are not present in the problem statement, and that we might not be able to make in real life. In this problem, we consider the case that we do not know the host's strategy. In this case we assume that

$$P(D = 2|\mathcal{H}_1) = \theta \quad (1)$$

$$P(D = 3|\mathcal{H}_1) = 1 - \theta \quad (2)$$

   for some unknown variable θ. For a given value of θ, what is $P(\mathcal{H}_2|D = 1, \theta)$? Is there any θ for which it is better *not* to switch doors?

   Ideally, we would assume a prior $p(\theta)$, and then make our choice by marginalizing over θ. Without going through these calculations, do you think that this will give a different strategy from before?
5. **Gaussian learning.** Suppose we have a set of N scalar data values $x_{1:N}$, drawn from a Gaussian with unknown mean μ and known variance σ^2. Assume that the prior over the mean is a Gaussian with specified mean $\bar{\mu}$ and variance σ_μ^2. What is the MAP estimate of μ? How does the prior influence the result as N gets larger? Suppose that the mean is known, but not the variance; what is the maximum likelihood estimate of the variance?