BCCS 2008/09: Graphical models and complex stochastic systems: Exercises 2

1. Two rational complexity science students, Amy and Ben, play the following game. An unknown amount of money, £w, is placed in an envelope, £$2w$ is placed in another envelope, and both are sealed. The envelopes are then shuffled, and one given to Amy and one to Ben. Before they each open their envelopes, Amy is given the option of switching envelopes with Ben, if Ben agrees. In either case, each student keeps the money in the envelope he opens.

   Amy argues that "There is an unknown sum x in my envelope. The other envelope is equally likely to contain $2x$ or $x/2$. On average, therefore, I will get $(2x + x/2)/2 = 1.25x$ if I switch, which is more than x. Therefore I should switch." Ben says the expected gain from switching is zero, so he is happy to accept Amy's suggestion. They switch, and both are happy.

   But who is right? Use probability to explain.

2. On the "Monty Hall" TV game show (in the US in the 70's), a contestant is shown 3 doors. Behind one is a valuable prize (say, a car), behind the other two something worthless (traditionally, a goat). The host of the show knows where the car is. The contestant, who does not know, chooses a door to be opened. Before that door is opened, the host opens one of the *other* doors, and exposes a goat. The contestant is given an option to reconsider her choice (among the two still closed). She thinks that this option gives her an even chance of getting the car, instead of the 2 to 1 odds against that she started with. Should she switch her choice? Or is it true that the host is giving her no information, since whatever door she chooses initially, there remains one that he can open to show a goat?

   There is a lot about this problem on the Internet; *some* of what you read there is correct. One fun page is `http://math.ucsd.edu/~crypto/Monty/monty.html`

3. Find the maximum likelihood estimator of λ given a random sample of size n from the Poisson distribution with parameter λ, $p(x|\lambda) = \exp(-\lambda)\lambda^x/x!$. [The likelihood for λ given data $x_1, x_2, \ldots, x_n$ is $p(x_1|\lambda) \times p(x_2|\lambda) \times \cdots p(x_n|\lambda)$.]

4. Suppose that X_i has a Poisson distribution with expectation λt_i, for $i = 1, 2, \ldots, n$, where t_i are known numbers. All the X_i are independent. What is the maximum likelihood estimate of λ, given observed values $x_1, x_2, \ldots, x_n$? [This is a generalisation of the previous question. This time the data are not i.i.d. But don't be alarmed – just write down the joint distribution of all the random variables.]

5. There are 100 raffle tickets in a hat. Three of them are numbered the same – with the value θ, say – the others are numbered $\theta + 1000, \theta + 1001, \ldots, \theta + 1096$. One ticket is drawn, and its number X, say, is recorded. What is the maximum likelihood estimator of θ? Do you think this is a good estimator of θ? Or would $X - 1048$ be a better one?

6. Find the maximum likelihood estimates of (θ_1, θ_2) given independent observations $x_1, x_2, \ldots, x_n$ from the Uniform distribution on the interval $[\theta_1, \theta_2]$.

7. Three examples of judgement evaluation:

   - A woman tea-drinker claims she can detect from a cup of tea whether the milk was added before or after the tea. She does so correctly for ten cups.
   - A music expert claims that she can distinguish between a page of Haydn's work and a page of Mozart. She correctly categorises 10 pieces.
   - A drunk friend claims she can predict the outcome of tossing a fair coin, and does so correctly for 10 tosses.

A possible strategy here would be to assume the number of correct guesses $X \sim \text{Binomial}(10, p)$. Then the probability of observing the given data is *very* small if $p = 0.5$; we would naturally conclude that p is *not* 0.5. (More formally, the hypothesis $H_0 : p = 0.5$ would be rejected in favour of $H_1 : p > 0.5$.) But does it make sense to take the same decision in three cases?

8. From past experience, 1 in 1000 patients arriving at a clinic with a particular symptom have the serious form of a disease, the remainder have only a harmless form. A blood test is available, which gives a positive result with probability 99% when the serious form is present, but also 5% given the harmless form. If a patient has a positive test result, what is the probability that he has the serious form of the disease? Suppose the test is administered twice, with results that can be assumed independent: what do you conclude if both tests show positive?

9. You have two coins, with probabilities of getting a Head equal to 0.6 and 0.4 respectively. You pick a coin at random, and toss that coin repeatedly. If the first two tosses show Tail, what is the probability that the 3rd is a Tail too?

10. The 'Tramcar example' (Jeffreys). A man travelling in a foreign country has to change trains at a junction, and goes into the town, the existence of which he has only just heard. He has no idea of its size. The first thing that he sees is a tramcar numbered 100. What can he infer about the number of tramcars in the town? It may be assumed that they are numbered consecutively from 1 upwards.

    Suppose your prior distribution for the number θ of tramcars in the town is $\pi(\theta) = c/\theta^a, \theta = b, b+1, b+2, \ldots$, where $a > 1$, b is a positive integer, and $c = c(a, b)$ is a constant depending on a and b that ensures the distribution sums to 1 (there is no need to work it out). Find the posterior distribution for θ given the single observation of a tramcar numbered 100.

11. Independent observations $X_1, X_2, \ldots, X_n$ are available from a Geometric distribution, with $P\{X_i = x\} = \theta(1-\theta)^{x-1}, x = 1, 2, \ldots$. Suppose that your prior information about θ can be represented by the Beta(α, β) distribution. What is the resulting posterior distribution of θ? What is the posterior mean? Compare these with the corresponding results for Binomially distributed observations, seen in Lecture 2 or 3.

12. For n successive days you try to complete the crossword in your newspaper. Suppose that you are successful with probability θ, independently from day to day. Your initial belief about θ is expressed by a Beta(α, β) prior distribution. What is your posterior distribution if you are successful X times out of n? Now suppose that after each attempt you update your distribution for θ, and use the resulting posterior as the prior for the next day, and so on. Show that your posterior for θ after n days is the same as if you considered all the data together, so that, in particular, the order of the successes and failures is irrelevant. [Hint: let X_i be 1 or 0 for success or failure on day i; note that $X_i \sim \text{Binomial}(1, \theta)$.]