Bayesian Inference and Computational Methods 2009

Homework 1

1. Player Y has a pack of 4 cards (Ace and Queen of clubs, Ace and Queen of Hearts) from which he deals a random selection of 2 to player X.
   (a) What is the probability that X receives both Aces conditional on receiving at least 1 Ace?
   (b) Suppose now that Y deals X two cards from the pack of 4, after which X says *I have an Ace.*
      i. Discuss whether the above information is sufficient to calculate the conditional probability
      $$\mathbf{P}(X \text{ has 2 Aces} \,|\, X \text{ says } \textit{I have an Ace}).$$
      ii. If it is not, what other information would be required in order to calculate this conditional probability?

2. An urn is known to contain n differently coloured balls where n can be any integer in the set $\{1, 2, 3\}$. Your prior information tells you that n is equally likely to be any of these values.
   (a) A ball is drawn randomly from the urn and is found to be red.
      i. Alice argues that, since the probability of the red ball being drawn conditional on there being n balls in the urn is $1/n$,
      $$\mathbf{P}(n = 1 \,|\, \text{red ball drawn}) = \frac{\frac{1}{3} \times \frac{1}{1}}{\frac{1}{3} \times \frac{1}{1} + \frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{3}}$$
      and calculates the posterior probabilities of n being 1, 2, and 3 as 6/11, 3/11 and 2/11 respectively. She then expresses her surprise that the her beliefs regarding n have changed having observed only the colour of a single draw from the urn. Explain the fallacy in her argument and why the above the information alone does not define a posterior probability for n.
      ii. Bertie assumes that the n balls placed in the urn are drawn uniformly at random from a large stock of differently coloured balls. Calculate Bertie's posterior probabilities for $n = 1, 2, 3$.
      iii. Under what circumstances would Alice's posterior probabilities be correct?
   (b) Suppose now that two balls are drawn from the urn *with replacement* and the event that both are the same colour is observed. Calculate the posterior probabilities for $n = 1, 2, 3$ in this case.

3. (*More balls and urns*) Five balls are drawn uniformly randomly from a very large population of black and white balls where the proportion of black balls is 1/3. You do not know the colours of the balls selected. The balls are then placed in an urn.
   (a) Give suitable prior probabilities for the number of black balls in the urn.
   (b) You now select two balls uniformly at random from the urn with replacement. They are both white. Calculate the posterior probabilities for the number of black balls in the urn.
   (c) Suppose that the two balls were selected from the urn without replacement and were both white. Calculate your posterior probabilities for the number of black balls in the urn for this case.

4. (a) A fair coin is tossed n times where n can take the values $1, 2, 3, 4, 5$ with equal probability. Suppose that 2 heads result from the n tosses. Determine the posterior distribution (i.e. work out the probability function) of n and identify the value of n that is *a posteriori* most likely.

   (b) Suppose now the coin is instead tossed repeatedly until m tails are obtained where the value m is first selected from a $Geometric(1/3)$ distribution, i.e. the probability function of m is $(2/3)^{m-1}(1/3)$. Suppose that 2 heads are obtained in the sequence. What is the posterior distribution of m given this information? (It is sufficient to write an expression involving infinite sums!)

5. In a sequence of n independent trials each has a success with unknown probability $p \in (0, 1)$.

   (a) Write down the distribution of the total number X of successes.

   (b) A Bayesian approach is used to estimate p. A $\beta(a, b)$ prior distribution is used. (Note that, in particular, $a = b = 1$ defines a $U(0, 1)$ distribution.) A total of k successes are observed. Determine the posterior distribution of p, together with its mode, its mean and its standard deviation. Assuming that n and k are large, show that these quantities are relatively insensitive to the choice of a and b (provided that these quantities are not large).

   (c) Suppose $a = b = 2$ and that $n = 200$, $k = 70$. Plot the prior and posterior distributions for p and determine an equal-tailed 95% Bayesian credible interval for p.

6. The lifetime T days of a component is known to have an $\text{Exp}(\lambda)$ distribution, where λ is unknown, but modelled as having an $\text{Exp}(2)$ prior distribution. In a random sample of 5 such components, their *total* lifetime is observed to be 3.0 days.

   (a) Find the posterior distribution of λ. Sketch it. Determine also an equal-tailed 95% Bayesian credible interval for λ.

   (b) Write down, as a function of λ, the probability that a further component will have a lifetime of at least 0.25 days. Using simulation, or otherwise, estimate the expected value of this probability under the posterior distribution found above.

   (c) Suppose now that a further 2 components are observed and found to have lifetimes in excess of 1 day. Find the new posterior distribution of λ and again determine also an equal-tailed 95% Bayesian credible interval for λ.