MAST20005/MAST90058: Week 10 Problems

1. A food hygiene inspection of Melbourne restaurants is carried out routinely. The outcome of each inspection is either a pass or a fail. The inspectors are not always reliable at picking up all food hygiene issues. From past experience, it is known that the probability of a pass if food hygiene is satisfactory is 0.9, and if it is unsatisfactory is 0.3.

   (a) Fast food restaurants are known to have unsatisfactory standards about 20% of the time. If an inspection of such a restaurant results in a fail, what is the probability it has unsatisfactory food hygiene standards?

   (b) Fine dining restaurants are known to have unsatisfactory standards only about 2% of the time. If an inspection of such a restaurant results in a fail, what is the probability it has unsatisfactory food hygiene standards?

   (c) Draw tree diagrams to represent this problem.

2. You are running a clinical trial of a new treatment for depression. Of the first 40 patients in your trial, 4 have resulted in a successful outcome.

   (a) Using a uniform prior distribution for the probability of a successful outcome, what is the posterior mean?

   (b) Your colleagues overseas, who are running similar trials, have observed successful outcomes in about 4% of their patients. You know their results are relevant and their trials are larger than yours, but there are some differences that make them only partly comparable. You decide to use their results as a prior to get a sense of the likely outcome of your own trial. You deem their information to be equivalent to about 60 pseudo-observations.

      i. Construct a prior that encapsulates this information.

      ii. Using this prior, what is your new posterior distribution?

      iii. What is your new posterior mean?

3. Let $X \sim \mathrm{N}(\theta, \sigma^2)$ where σ is a known value. Use a normal distribution as a prior for θ. Given a single observation on X, derive the posterior distribution. Show full working.

4. A standard mathematics test is given to all students about to start high school in Australia. The test scores are known to be normally distributed with a variance of 25. You need to review a random selection of 16 of the test papers. The mean score for this subset of students was 70. Using an appropriate uninformative prior distribution, give a 95% credible interval for the overall mean score for all students.

5. Let Y be the sum of n observations from a Poisson distribution with mean θ. Let the prior for θ be a gamma distribution with parameters α and β, parameterised as follows:

$$f(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\theta\beta}, \quad 0 \leqslant \theta < \infty \quad (\text{Note: } \mathbb{E}(\theta) = \alpha/\beta,\ \mathrm{var}(\theta) = \alpha/\beta^2)$$

   (a) Find the posterior pdf of θ given $Y = y$. (Hint: You only need to consider the terms that involve θ. What distribution do those terms correspond to?)

   (b) What is the posterior mean?

   (c) Show that the posterior mean is a weighted average of the maximum likelihood estimate and the prior mean.

6. Consider a random sample $X_1, \ldots, X_n$ from a distribution with pdf

$$f(x \mid \theta) = 3\theta x^2 e^{-\theta x^3}, \quad 0 < x < \infty.$$

Let the prior for θ be a gamma distribution with $\alpha = 4$ and $\beta = 4$. Find the posterior mean of θ.

7. Suppose that the time in minutes required to serve a customer at a certain shop has an exponential distribution with mean $1/\theta$. As a prior for θ use a gamma distribution with mean 0.2 and standard deviation 0.1. If the average time to serve 20 customers is observed to be 3.8 minutes, what is the posterior distribution of θ?

8. Take a random sample $X_1, \ldots, X_n \sim \text{Unif}(0, \theta)$. Use an improper uniform prior for θ.

   (a) Derive the posterior distribution of θ.

   (b) Show that the posterior mode is $x_{(n)}$.

   (c) Derive a 95% credible interval of the form $(x_{(n)}, c \cdot x_{(n)})$.
   (This will require some knowledge from the Module 10 lectures.)

   (d) Derive a 95% confidence interval of the same form.
   (Hint: you will get a different value of c.)

   (e) What prior distribution would result in these two interval estimates being identical (i.e. the same value of c)?