Lab 7

Questions with a (*) next to them form part of homework 2. Due in 2/11/2015. [Q 4]

1. Suppose $X_1, \dots X_n \sim \text{Poisson}(\lambda)$ iidrv so $\mathbb{P}(X_i = k) = e^{-\lambda}\frac{\lambda^k}{k!}$.
   (a) What, roughly speaking, is the distribution of $\bar{X} = \sum_{i=1}^n X_i/n$? [i.e. use the CLT]
   (b) Given λ, find an interval such that $\bar{X}$ belongs to that interval with probability 95%.
   (c) Re-arrange your answer to part (b) to produce an approximate 95% confidence interval for λ.
   (d) Let $\lambda = 2$ and $n = 10$. Using R, do the following a large number of times: Sample $X_1, \dots X_n \sim Poisson(\lambda)$ and check if your confidence interval includes λ. What fraction of the time does your confidence interval include λ.
   (e) How about when $n = 100$?
   (f) Let $n = 1$ and assume λ is large. Argue that $\sqrt{X} \approx N(\sqrt{\lambda}, 1)$. Use that approximation to form a confidence interval for $\sqrt{\lambda}$ and therefore for λ. [This is an example of a variance stabilising transform!]

2. Suppose $X_1, \dots X_n \sim N(\mu, \sigma^2)$ i.i.d.r.v. with unknown μ, σ^2.
   (a) Produce 95% confidence intervals for μ and σ^2 in terms of $\bar{X}$ and S^2.
   (b) Use R to sample some data from the $N(1, 10^2)$ distribution and apply your answer from part (a).

3. Let $X_1,, \dots, X_n \sim \text{Uniform}(0, 1/\theta)$ where $\theta > 0$ is some unknown parameter; $X = (X_1, \dots, X_n)$. Taking the Bayesian route, take (θ, X) to have a joint distribution f. Take your prior distribution to be the distribution to be

$$f(\theta) = 1/\theta, \qquad f(\theta) = \int f(\theta, x)dx.$$

   (a) Is this a proper prior? i.e. is $\int f(\theta)d\theta = 1$? In words, what does our choice of prior say about our beliefs about θ?
   (b) Use the formula

$$f(\theta \mid x) \propto f(\theta) \prod_{i=1}^n f(X_i \mid \theta)$$

   to find the posterior distribution? Is it a proper distribution? i.e. is $\int f(\theta \mid x)d\theta = 1$?
   (c) What is the posterior mean? Does that seem reasonable?

4. (*) Let $X_1, \dots, X_n \sim N(\theta, 1)$ and take your prior beliefs about θ be $N(a, b^2)$. You will observe $X_1, ..., X_n$ from the distribution $N(\theta, \sigma^2)$ where σ is known.
   (a) Find the posterior distribution. Try to express your answer as simply as possible.
   (b) Calculate the limit of the posterior distribution as $n \to \infty$ with a, b, σ fixed.
   (c) Let $a = 0$, $b = 10$. Setting $\theta = 100$, $\sigma = 1$ and $n = 8$, sample $X_1, ..., X_n$.Calculate the posterior distribution. Find a 98% credible interval for θ by calculating the 1st and 99th quantiles of the posterior distribution using `qnorm`.
   (d) Repeat part (c) but $b = 1/10$. Comment on the difference.

5. An experiment will produce a result X with distribution Binomial$(100, p)$. A scientist tells you that their prior beliefs is that p is roughly between 0.7 and 0.8. Formulate a suitable prior distribution using the Beta family of distributions. You observe a sample $x = 65$.
   (a) Plot the prior distribution on the interval $[0, 1]$.
   (b) Calculate the posterior distribution and plot it. Compare with part (a).
   (c) Calculate a 95% posterior credible interval for p. [Hint: use `qbeta`] Does this seem to be compatible with your observation?
   (d) Repeat parts (b) and (c) but based on an observation of $x = 10$.

6. Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$ where λ is unknown.

(a) Calculate the Jeffreys prior for the $\text{Poisson}(\lambda)$ distribution

$$\text{prior} \propto \sqrt{I(\theta)} \qquad I(\theta) = \text{Var}_{X\sim\text{Poisson}(\theta)}\left(\frac{\partial}{\partial\theta}\log f(X \mid \theta)\right) \quad f(x \mid \theta) = e^{-\theta}\frac{\theta^x}{x!}$$

Calculate the posterior distribution with respect to the Jeffreys prior.

(b) After taking expert advice you decide instead that your prior beliefs about λ are given by a Gamma distribution with mean 100 and variance 200. Calculate the posterior distribution.