Bayesian Inference and Computational Methods 2009

Homework 2

1. *Inference for a binomial parameter* p. Recall that the class of *beta* distributions form the natural *conjugate class* for the *binomial/Bernoulli* family, and that if y *successes* are observed in n *trials*, a $\beta(a, b)$ *prior distribution* is transformed to a $\beta(a + y, b + n - y)$ *posterior distribution.*

   Suppose now that a coin is to be tossed so as to make inference for the probability p of a head.

   (a) It is natural to choose a symmetric prior distribution $\beta(a, a)$ for some $a > 0$. Suppose it is believed that the coin is so badly made that there is a 20% chance that the value of p lies outside the range $[0.4, 0.6]$. Find, to the nearest integer, the value of a appropriate to the belief.

   (b) Suppose now that the coin is tossed 10 times, and that 1 head is observed. Using the prior distribution identified above, find the posterior distribution for p given these observations. Identify its mean and standard deviation, and use **R** to sketch it. Suppose further that the coin is to be tossed a further 5 times. Compute the predictive distribution of the number of heads to be obtained, the predictive probability that at most 2 heads will be obtained, and the predictive mean of the number of heads to be obtained.

   (c) Suppose instead that, in the original 10 tosses of the coin, it is only observed that *at most* 1 head is obtained. Find the new posterior distribution for p, and compare it graphically with that obtained previously.

2. *Inference for a geometric parameter* p. Let $\boldsymbol{y} = (y_1, \dots, y_n)$ be a vector of independent observations from a Geo(p) distribution (i.e. with probability function $p(y) = p(1-p)^{y-1}$, $y = 1, 2, \dots$). Write down the associated likelihood function. Deduce that Bayesian inference for the parameter p is exactly the same as if p were a binomial parameter and that in $\sum_{i=1}^{n} y_i$ trials a total of n successes were observed. (*Understand why this is!*) Deduce that the *beta* distributions again form the natural *conjugate class* for the *geometric* family. For the $\beta(a, b)$ prior distribution, what is the posterior distribution induced by the above data.

3. *Inference for a Poisson parameter* λ. Recall that the class of *gamma* distributions form the natural *conjugate class* for the *Poisson* family, and that for Poisson observations $\boldsymbol{y} = (y_1, \dots, y_n)$, a $\Gamma(a, b)$ *prior distribution* is transformed to a $\Gamma(a + \sum_{i=1}^{n} y_i, b + n)$ *posterior distribution.*

   Suppose now that, within a given portfolio, motorcycle insurance claims are assumed to arise as a Poisson process with rate λ per week. The insurance company intends a Bayesian estimation of λ.

   (a) Previous experience suggests the use of a prior distribution whose 0.1-quantile and 0.9-quantile are approximately 2.9 and 5.2 respectively (e.g., the probability that λ should be less than 2.9 is approximately 0.1). Show that if the prior distribution is taken to be $\Gamma(a, b)$ where, for simplicity, a and b are taken to be integer valued, then the best choices of these are $a = 20$ and $b = 5$.

   (b) Over a 45-week period the total number of claims is observed to be 280. Calculate the posterior distribution of λ and sketch its density along with that of the prior distribution. Comment.

   (c) Use this posterior distribution to estimate the predictive distribution of the number n of claims occurring in the next two weeks. Find the value of n which is exceeded with probability 0.05.

(d) Suppose instead that the company had had little previous experience of the claim rate, and had therefore decided to use an Exp(0.1) prior distribution. (*Why might this have been sensible?*) Given the above data, what would the posterior distribution of λ then have been, and how different would it have been from that found previously?

4. *Inference for a Poisson parameter λ—continued.* Show that the Jeffreys' prior distribution for a Poisson parameter λ is given by the *improper* prior distribution with density

$$\pi(\lambda) \propto \lambda^{-1/2}.$$

What is the corresponding posterior distribution, given Poisson observations $\boldsymbol{y} = (y_1, \ldots, y_n)$? Interpret this.

5. *Inference for an exponential parameter λ.*

(a) Show that the class of *gamma* distributions also form the natural *conjugate class* for the *exponential* family. Given a $\Gamma(a, b)$ *prior distribution*, and for exponential observations $\boldsymbol{y} = (y_1, \ldots, y_n)$, what is the corresponding *posterior distribution*? Compare this situation with that for inference for a Poisson parameter λ and comment intelligently. (*Hint:* recall the close connection between the Poisson and exponential distributions via the Poisson process.)

(b) Suppose that, in a certain slightly strange country (where the sole significant cause of death is accidents), insured lives have durations in years which are independently Exp(λ)-distributed. As usual a Bayesian estimation of λ is required. The prior distribution of λ is taken to be Exp(10) (to what sort of belief about likely lifetimes does this correspond?). A total of 40 new policyholders are followed for a 10 year period. At the end of this time it is observed that 18 of these have died, with the following insured lifetimes (years).

0.54	0.63	0.93	1.92	2.10	2.13	2.73	2.81	2.82
2.86	2.97	3.30	3.77	5.47	7.22	7.54	8.22	9.68

The remaining policyholders are still alive.

i. Use (e.g.) a Q-Q plot to investigate the assumption that lifetimes are exponentially distributed, and to make an informal estimate of λ.

ii. Obtain the posterior distribution of λ, and compare it with the prior.

iii. Given this posterior distribution, find the predictive distribution of the insured lifetime of a further policyholder. In particular identify the predictive probability that a further policyholder survives at least 10 years.

iv. Similarly, given this posterior distribution, find the predictive probability that two further policyholders both survive at least 10 years. Why is this not the square of the probability found for a single policyholder?

6. *[1997 Statistical Inference Exam, Q4 (modified)]* In a raid on a coffee shop, Bayesian trading inspectors take a random sample of n packets of coffee, each of nominal weight 125 g. They model these data as independent observations $Y_1, \ldots, Y_n$ from a Normal $\mathrm{N}\left(\mu, \sigma^2\right)$ distribution. They take σ^2 to be known, while for μ they assume a prior distribution of $\mathrm{N}\left(\mu_0, \sigma_0^2\right)$, where μ_0 and σ_0^2 are specified values.

   The data they obtain are (weights in grams):

   ```
   105.3 113.3 114.5 121.2 122.9 123.7 124.0 124.6 124.9 124.9
   124.9 125.1 125.5 125.9 126.8 127.7 128.2 128.3 128.5 130.2
   ```

   ($n = 20$, $\sum y_i = 2470.4$, $\sum y_i^2 = 305828.98$).

   The parameter values they assume are $\mu_0 = 126$, $\sigma_0^2 = 1$, $\sigma^2 = 4$.

   The inspectors can impose a fine if their 95% credible interval for μ falls wholly below the claimed value of $\mu = 125$ g.

   (a) Recall the result of the lecture notes for the posterior distribution for the $\mathrm{N}\left(\mu, \sigma^2\right)$ model (with σ^2 known) with $\mathrm{N}\left(\mu_0, \sigma_0^2\right)$ prior distribution for μ. Show that the inspectors' 95% credible interval for μ for these data does lie wholly below 125 g; they therefore impose a fine on the owners of the coffee shop.

   (b) Use a normal Q-Q plot (or other appropriate graphical technique) to investigate the validity of the assumed normal distribution for the data.

   (c) Comment briefly as to whether the inspectors are justified in imposing a fine on the basis of this sample.

   (d) Using the posterior distribution found above, calculate the predictive probabilities that

      i. a single further randomly chosen packet of coffee weighs at least 125 g;

      ii. two further randomly chosen packets of coffee have a mean weight of at least 125 g.

7. Suppose that $y_1, y_2, \ldots, y_n$ form a random sample of observations from an $\mathrm{Exp}(\lambda)$ distribution where λ is unknown.

   (a) Show that the Jeffreys' prior for λ in this case is given by $\pi(\lambda) \propto \lambda^{-1}$.

   (b) For $n = 5$ and $\sum_{i=1}^{5} y_i = 10$ calculate a 95% equal-tailed credible interval for λ using the Jeffreys prior.