22S:138, Bayesian Statistics
Fall 2007, Homework 4

Due: Mon. 10/01

1. (problem from Johnson and Albert) (From Antleman, 1997) Suppose that a trucking company owns a large fleet of well-maintained trucks and assume that breakdowns appear to occur at random times. The president of the company is interested in learning about the daily rate R at which breakdowns occur. (Realistically, each truck would have a breakdown rate that depends possibly on its type, age, condition, driver, usage, etc. The breakdown rate for the whole company can be viewed as the sum of the breakdown rates of the individual trucks.) For a given value of the rate parameter R, it is known that the number of breakdowns y on a particular day has a Poisson distribution with mean R:

$$p(y|R) \;=\; \frac{e^{-R}\,R^y}{y!}, \quad y = 0, 1, 2, \ldots$$

   (a) Suppose that one observes the number of truck breakdowns for n consecutive days — denote these numbers by $y_1, \ldots, y_n$. If one assumes that these are exchangeable measurements (conditionally independent given R), find the joint probability distribution of $y_1, \ldots, y_n$.

   (b) The number of breakdowns for 5 days are recorded to be 2, 5, 1, 0, and 3. Find the likelihood function $L(R)$ of the rate parameter R for these observations. Graph this function. (You may either use Splus or other computer software, or you may do it "by hand" by calculating the likelihood for the values $R = .1, .5, 1, 2, 4, 8$ and 16, and connecting the points with a smooth curve.

   (c) Find the MLE of R and its associated asymptotic standard error.

   (d) Use the results in part (c) to find an approximate 95% frequentist confidence interval for R.

2. The president of the company has some knowledge about the location of the Poisson rate parameter R based on the observed number of breakdowns from previous years. His prior beliefs about R are represented in the following:

$$p(R) \;\propto\; R^3\; exp(-2R), \quad R > 0$$

   (a) Is this prior a member of a particular parametric family? If so, what family, and what are the prior parameters?

   (b) Plot this prior density, either using software or by picking a few values at which to evaluate it as in the previous problem. Based on the plot, describe the president's prior beliefs about the rate parameter R.

   (c) Write out the mathematical form of the unnormalized posterior density. Identify its parametric family and parameters.

   (d) Find the posterior mean and 95% central credible set for R based on this posterior.

   (e) Was the president's prior from a conjugate family for the Poisson likelihood? How could you tell?

3. A new employee of the trucking firm wishes to learn about the breakdown rate. She does not have the previous information available to the president, so she wishes to assign a noninformative prior density.

   (a) Derive the Jeffrey's prior that goes with the Poisson likelihood.

   (b) Compute the resulting posterior distribution, using this prior and the data given earlier for the 5 days.

   (c) Find the posterior mean and 95% central credible set for this posterior.

   (d) Compare the Bayesian point and interval estimates with the classical estimates found in the first exercise.

   (e) Compare the new employee's estimates with the president's estimates.

   (f) Suppose the employee has a personal rule that she would not drive for a company whose fleet of trucks had a daily breakdown rate > 2. What is her posterior probability that $R > 2$ for this company?

   (g) Contrast this with the president's posterior probability that $R > 2$.

4. We used the following identity in deriving the likelihood for the mean μ of a normal distribution. Verify that it is true.

$$\sum_{i=1}^{n}(y_i - \mu)^2 \;=\; \sum_{i=1}^{n}(y_i - \bar{y})^2 + n\;(\bar{y} - \mu)^2$$