Stat 8003, HW7

Due: Monday, Dec 8th, 2014

1. For a given model with the density function $f(y|\theta)$, the Jefferey's prior is defined as

$$J(\theta) \propto (J(\theta))^{1/2},$$

where $J(\theta)$ is the Fisher information for θ

$$J(\theta) = -E(\frac{d^2 \log p(y|\theta)}{d\theta^2}|\theta).$$

(a) Suppose that $y|\theta \sim Poisson(\theta)$. Find Jeffrey's prior density for θ;

(b) Calculate the posterior distribution;

(c) Consider Example 5.3.1 regarding the Traffic lights. Assume the Poisson model and Jeffrey's prior, estimate the rate parameter and construct 95% credible interval.

2. Statistical Decision Theory: a decision-theoretical approach to the estimation of an unknown parameter θ introduces the loss function $L(\theta, a)$ which, loosely speaking, gives the cost of deciding that the parameter has the value a, when it is in fact equal to θ. The estimate a can be chosen to minimize the posterior expected loss,

$$E(L(a|y)) = \int L(\theta, a)p(\theta|y)d\theta.$$

This optimal choice of a is called a Bayes estimate for the loss function L. Show that:

(a) If $L(\theta, a) = (\theta - a)^2$, (squared error loss), then the posterior mean, $E(\theta|y)$, if it exists , is the unique Bayes estimate of θ;

(b) If $L(\theta, a) = |\theta - a|$, then any posterior median of θi s a Bayes estimate of θ;

(c) If k_0 and k_1 re nonnegative numbers, not both zero, and

$$L(\theta, a) = \begin{cases} k_0(\theta - a), & if \theta \geq a \\ k_1(a - \theta), & if \theta < a, \end{cases}$$

then any $\frac{k_0}{k_0+k_1}$ quantile of the posterior distribution $p(\theta|y)$ is a Bayes estimate of θ.

3. Overdispersed Poisson. It is known that the number of faults in a fabric production relates to the length. Let Y_i and l_i be the number of faults and the logarithm of the length of the i-th fabric. Let λ be the rate parameter. Assume the Poisson model

$$Y_i \sim Poisson(l_i \lambda).$$

The dataset is available on the course website. The two columns are the length and the count of faults.

`http://astro.temple.edu/~zhaozhg/Stat8003/data/fabric.txt`

(a) Use the frequentist approach, estimate λ and construct 95% confidence interval;

(b) There exists an issue of overdispersion in this dataset. To deal with that, assume that $\lambda \sim Gamma(\alpha, \beta)$. Calculate the posterior distribution of λ;

(c) Suggest an informative prior. Set the test quantities as the sample variance of Y, use the posterior predictive p-values to check your model;

(d) Based on your suggested model, construct 95% credible interval λ;

(e) Estimate α, β based on the method of moments.

(f) Similar to part (c), check the model and construct 95% credible interval for λ.