

**Assignment #9 — Due Monday, April 13, 2009, by 5:00 P.M.**

Turn in homework in lecture, discussion, or your TA's mailbox. Indicate the discussion section in which you expect to attend to pick up this assignment on the assignment.

**311:** Monday 1:20–2:10

**312:** Monday 12:05–12:55

This short assignment involves writing a short MCMC program in R. See the function `hw09.R` which contains most of what you need.

1. A sample  $s = (753.0, 1396.9, 1528.2, 1646.9, 717.6, 446.5, 848.0, 1222.0, 748.5, 610.1)$  is modeled as an i.i.d sample from a  $\text{Gamma}(\alpha, \lambda)$  distribution. You may think of this data as a sample of lifetimes measured in hours.

We wish to analyze this data with a Bayesian model with the joint prior distribution for  $\theta = (\alpha, \lambda)$  having density

$$\pi(\alpha, \lambda) = \frac{3\alpha^3(4-\alpha)}{32(400)^2\lambda^3} e^{-\alpha/(400\lambda)}, \quad 0 < \alpha < 4, \lambda > 0$$

using MCMC. (This prior density arises from  $\alpha/4 \sim \text{Beta}(2, 2)$ ,  $\lambda^{-1} | \alpha \sim \text{Gamma}(2, \alpha/800)$ .)

- (a) For a fixed  $\alpha_0$ , find an expression for  $\lambda$  that maximizes the likelihood and evaluate this for  $\alpha_0 = 2$ .
- (b) Plot  $h(\alpha, \lambda_0)$  versus  $\alpha$  for  $\lambda_0$  equal to the value from (a) where  $h(\alpha, \lambda) = \pi(\alpha, \lambda) f_{\theta}(s)$ . By eye, what is an approximate good initial value to choose for  $\alpha$  where  $h$  is high? For this  $\alpha$ , which value of  $\lambda$  will maximize the likelihood? This pair is an excellent candidate to start your MCMC.
- (c) Run an initial run of MCMC of 1000 steps with the starting value from the previous part and some choices for tuning parameters using `mcmc()` in `hw09.R`. Find the mean and standard deviation of each column of the output. Ideally, the mean acceptance probability will be between 0.1 and 0.4 (say close to 0.25). Experiment with selecting values of the tuning parameters until you achieve a mean acceptance probability in this range. Which parameter values do you find work? *Hint: using a window whose width is a few posterior sd's is a good idea.*

Example:

```
> out.1 = mcmc(1000, 2, 0.001, s, w.alpha=0.1, w.lambda=0.0005)
> apply(out.1, 2, mean)
[1] 2.050840e+00 2.079436e-03 7.734877e-32 7.610560e-01
> apply(out.1, 2, sd)
[1] 4.899213e-01 6.703536e-04 5.292465e-32 2.884865e-01
> xyplot(out.1[,1] ~ out.1[,2])
```

Here, the average acceptance probability is 0.76. The window sizes of 0.1 and 0.0005 are about 0.2 of the sd for  $\alpha$  and 0.8 of the sd for  $\lambda$ , both of which are too small.

- (d) Take a large MCMC sample using good starting values and proposal densities. On the basis of this sample, find approximate 95% credible regions for  $\alpha$  and for  $\lambda$ . *Hint: Use a command like the following: `apply(out.2[, 1:2], 2, quantile, probs=c(0.025, 0.975))` to find the empirical quantiles of columns 1 and 2 from the matrix `out.2`.*
- (e) Summarize the sample to estimate the posterior mean and sd for both  $\alpha$  and  $\lambda$ . Are the intervals you found in the previous part close to the the mean plus and minus 1.96 times the sd?

- (f) Use `xyp1ot()` in the `lattice` package to graph the sample  $\lambda$  versus  $\alpha$ , (`xyp1ot(out.2[,2] ~ out.2[,1])`). Are  $\alpha$  and  $\lambda$  nearly independent in the posterior distribution, or does information about one parameter say something about the other?

**Work to do, but not turn in.**

- Read Chapter 9.
-