MAS4303 Modern Bayesian Inference
Part 2

TEST: SOLUTION

Semester 1, 2008-9

Davies and Goldsmith (1972) give the following data on piston ring failures in steam-driven compressors. There were four identical compressors in the same compressor house, each oriented the same way, and each had three legs. The data give the number of failures in each leg of each compressor over a period of some years.

Compressor Number	North Leg	Centre Leg	South Leg
1	17	17	12
2	11	9	13
3	11	8	19
4	14	7	28

Let the number of failures in leg j (North: $j = 1$, Centre: $j = 2$, South: $j = 3$) of compressor i be $Y_{i,j}$. Suppose that we believe that, given parameters $\lambda_{i,j} > 0$, each $Y_{i,j}$ has a Poisson($\lambda_{i,j}$) distribution, independently of the others. That is, given $\lambda_{i,j}$,

$$Y_{i,j} \sim \text{Poisson}(\lambda_{i,j}).$$

We also believe that

$$\log(\lambda_{i,j}) = \alpha_i + \beta_j + \gamma_{i,j}.$$

- The compressor effects α_i are regarded as random. Given the value of τ_α, we have $\alpha_1, \ldots, \alpha_4$ independent with
$$\alpha_i \sim N(0,\ \tau_\alpha^{-1}).$$
The prior distribution for τ_α is
$$\tau_\alpha \sim \text{gamma}(1,\ 1).$$
- The interaction effects $\gamma_{i,j}$ are regarded as random. Given the value of τ_γ, we have $\gamma_{1,1}, \ldots, \gamma_{4,3}$ independent with
$$\gamma_{i,j} \sim N(0,\ \tau_\gamma^{-1}).$$
The prior distribution for τ_γ is
$$\tau_\gamma \sim \text{gamma}(1,\ 1).$$
- In the prior distribution τ_α is independent of τ_γ and of $\beta_1, \beta_2, \beta_3$ and τ_γ is independent of $\beta_1, \beta_2, \beta_3$.
- The leg means $\beta_1, \beta_2, \beta_3$ are regarded as fixed but unknown. The prior distribution for $\beta_1, \beta_2, \beta_3$ is multivariate normal with

$$\begin{aligned} \text{E}(\beta_1) = \text{E}(\beta_1) = \text{E}(\beta_3) &= 12.0, \\ \text{var}(\beta_1) = \text{var}(\beta_2) = \text{var}(\beta_3) &= 25.0, \\ \text{covar}(\beta_1, \beta_2) = \text{covar}(\beta_1, \beta_3) = \text{covar}(\beta_2, \beta_3) &= 16.0. \end{aligned}$$

1. A hierarchical prior structure is proposed for $\beta_1, \beta_2, \beta_3$ with

$$\beta_j \mid \mu \sim N(\mu,\ v_\beta)$$

and

$$\mu \sim N(m,\ v_\mu).$$

Find the values of $m,\ v_\mu$ and v_β.

We obtain

$$\begin{aligned} m &= \underline{12.0}, \\ v_\mu &= \underline{16.0}, \\ v_\beta = 25.0 - 16.0 &= \underline{9.0}. \end{aligned}$$

2. Use BRugs to evaluate the posterior distribution.

   (a) Write down your BRugs model specification.

   A suitable BRugs model specification is as follows.

```
model compressor

{for (i in 1:12)
   {y[i]~dpois(lambda[i])
    log(lambda[i])<-alpha[comp[i]]+beta[leg[i]]+gamma[i]
    gamma[i]~dnorm(0,tau.gamma)
    }

 for (k in 1:4)
   {alpha[k]~dnorm(0,tau.alpha)
    }

 for (j in 1:3)
   {beta[j]~dnorm(mu,tau.beta)
    }

 mu~dnorm(m,tau.mu)

 m<-12.0
 tau.mu<-1/16.0
 tau.beta<-1/9.0

 tau.alpha~dgamma(1,1)
 tau.gamma~dgamma(1,1)
 }
```

   (b) Write down the commands which you use.

```
> modelCheck("test09bug.txt")
> modelData("test09data.txt")
> modelCompile(2)
> modelGenInits()
> samplesSet(c("tau.alpha","tau.gamma","alpha","beta","gamma","lambda"))
> modelUpdate(1000)
> samplesHistory("tau.alpha")
> modelUpdate(10000)
> samplesHistory("tau.alpha")
> samplesHistory("tau.gamma")
> samplesHistory("alpha")
> samplesHistory("beta")
> samplesHistory("gamma")
> samplesHistory("lambda")

        (convergence checked)

> modelCheck("test09bug.txt")
> modelData("test09data.txt")
> modelCompile()
> modelGenInits()
> modelUpdate(10000)
> samplesSet(c("beta","lambda","tau.alpha"))
> modelUpdate(30000)
> samplesStats(c("beta","lambda","tau.alpha"))
              mean     sd MC_error val2.5pc median val97.5pc start sample
beta[1]      2.725 0.7717  0.04562   1.5530  2.645     4.696 10001  30000
beta[2]      2.445 0.7945  0.04662   1.1920  2.368     4.477 10001  30000
beta[3]      3.005 0.7821  0.04640   1.7670  2.935     4.979 10001  30000
lambda[1]   16.700 3.9200  0.02205   9.9250 16.400    25.270 10001  30000
lambda[2]   16.030 3.8450  0.02107   9.5000 15.700    24.390 10001  30000
lambda[3]   13.200 3.4810  0.02277   7.2720 12.890    20.870 10001  30000
lambda[4]   11.110 3.1300  0.01843   5.9080 10.810    18.060 10001  30000
lambda[5]    8.948 2.7910  0.01676   4.3930  8.654    15.270 10001  30000
lambda[6]   13.370 3.4840  0.01982   7.4570 13.070    20.970 10001  30000
lambda[7]   11.250 3.1710  0.01981   5.9590 10.940    18.310 10001  30000
lambda[8]    8.330 2.6690  0.01682   3.9630  8.041    14.310 10001  30000
lambda[9]   18.600 4.1500  0.02238  11.4100 18.280    27.550 10001  30000
lambda[10]  14.100 3.5540  0.01809   8.0230 13.810    21.840 10001  30000
lambda[11]   7.924 2.6080  0.01931   3.6890  7.627    13.790 10001  30000
lambda[12]  26.970 5.0850  0.02630  17.9500 26.610    37.840 10001  30000
tau.alpha    2.021 1.3800  0.03177   0.1766  1.738     5.451 10001  30000
```

(c) Find the posterior mean of τ_α.

The posterior mean of τ_α is $\underline{2.021}$.

(d) Find a symmetric 95% posterior interval for β_1.

The interval is $\underline{1.5530 < \beta_1 < 4.696}$.

(e) Find a symmetric 95% posterior interval for $\lambda_{1,1}$.

The interval is $\underline{9.9250 < \lambda_{1,1} < 25.270}$.

(f) Comment on the behaviour of the MCMC sampler (e.g. convergence, mixing).

Convergence appears to take up to about 8000 iterations for some quantities and mixing is poor so a large burn-in and a large number of samples are required for a good approximation to the posterior distribution.