MAS8303 Modern Bayesian Inference
Solutions to Problems 4

M. Farrow
Newcastle University

Semester 1, 2012-3

1. Full conditional distribution

   Certain components are manufactured in batches. Each batch contains n components. The components in N batches are then tested and some are found to be defective. Let the number of defective components in batch i be x_i. We suppose that, given the value of π_i, where $0 < \pi_i < 1$, the value of x_i is an observation from a binomial distribution $X_i \sim \text{Bin}(n, \pi_i)$ and X_i and X_j are independent given π_i and π_j when $i \neq j$. Let $\eta_i = \log_e\{\pi_i/(1-\pi_i)\}$. We suppose that, given the values of μ and τ, η_i is an observation from the normal $N(\mu, \tau^{-1})$ distribution and η_i and η_j are independent, when $i \neq j$, given the values of μ and τ. Finally we have independent prior distributions for μ and τ with μ having a normal prior, $\mu \sim N(m, v)$, and τ having a gamma prior, $\tau \sim \text{Ga}(a, b)$.

   Find a function proportional to the density of the full conditional distribution (fcd) of η_i, that is the distribution of η_i given x_i and values for μ and τ.

   Solution

   The conditional prior density of η_i given μ and τ is proportional to

   $$\exp\left\{-\frac{\tau}{2}(\eta_i - \mu)^2\right\}$$

   since the distribution is $N(\mu, \tau^{-1})$.

   The relevant likelihood, that is the probability of observing $X_i = x_i$ given η_i, is proportional to $\pi_i^{x_i}(1-\pi_i)^{n-x_i}$ since the distribution is $\text{Bin}(n, \pi_i)$. However $\eta_i = \log\{\pi_i/(1-\pi_i)\}$ so

   $$\pi_i = \frac{e^{\eta_i}}{1 + e^{\eta_i}}$$

   and the likelihood is proportional to

   $$\begin{aligned} \pi_i^{x_i}(1-\pi_i)^{n-x_i} &= \left(\frac{e^{\eta_i}}{1+e^{\eta_i}}\right)^{x_i}\left(1 - \frac{e^{\eta_i}}{1+e^{\eta_i}}\right)^{n-x_i} \\ &= \left(\frac{e^{\eta_i}}{1+e^{\eta_i}}\right)^{x_i}\left(\frac{1}{1+e^{\eta_i}}\right)^{n-x_i} = \frac{e^{x_i\eta_i}}{(1+e^{\eta_i})^n}. \end{aligned}$$

   Therefore the fcd is proportional to

   $$\exp\left\{-\frac{\tau}{2}(\eta_i - \mu)^2\right\}\frac{e^{x_i\eta_i}}{(1+e^{\eta_i})^n}.$$

   (10 marks)

2. Piston rings

Four compressors are located in the same building. Each has three "legs". The compressors are of the same design and are oriented the same way. The three legs of each are labelled "North", "Centre" and "South." Over a certain period of time the number of failures of piston rings in each leg of each compressor is counted. These numbers are your data.

The model is as follows. Let the number of failures in leg i of compressor j be $y_{i,j}$ (where $i = 1$ for North, $i = 2$ for Centre and $i = 3$ for South). Given the value of a quantity $\lambda_{i,j} > 0$, we assume that $y_{i,j}$ is an observation from a Poisson distribution $Y_{i,j} \sim \text{Po}(\lambda_{i,j})$, with $Y_{i,j}$ independent of $Y_{i',j'}$ unless $(i, j) = (i', j')$, given the values of $\lambda_{i,j}$ and $\lambda_{i',j'}$.

The prior distribution is as follows. Let $\eta_{i,j} = \log_e(\lambda_{i,j})$. Then

$$\eta_{i,j} = \mu + \alpha_i + \beta_j + \gamma_{i,j}$$

where, $\alpha_1, \ldots, \alpha_3, \beta_1, \ldots, \beta_4, \gamma_{1,1}, \ldots, \gamma_{3,4}$ and μ are mutually independent and

$$\begin{aligned}
\mu &\sim N(3, 4), && \\
\alpha_i &\sim N(0, 1), && i = 1, \ldots, 3, \\
\beta_j &\sim N(0, 1), && j = 1, \ldots, 4, \\
\gamma_{i,j} &\sim N(0, 0.25), && i = 1, \ldots, 3,\ j = 1, \ldots, 4.
\end{aligned}$$

- Use MCMC to take samples from the posterior distribution of the unknowns in the model.

(5 marks)

- Display your results appropriately.

(4 marks)

- Explain your method and show your BUGS model specification and the commands which you have used.

(4 marks)

- Show how you have checked convergence.

(3 marks)

- Give summaries of the posterior distributions of the model unknowns. In particular, compare the failure rates in the twelve legs using the posterior distribution. What can you conclude?

(4 marks)

Solution

The solution here is for data set 10.

The posterior distribution of the model unknowns was computed using MCMC by using BRugs. The model specification is shown in Figure 1.

Convergence and mixing properties were checked using the following commands.

```
> modelCheck("pistonbug.txt")
> modelData("mypistondata.txt")
> modelCompile(2)
> modelGenInits()
> samplesSet(c("mu","alpha","beta","gamma","eta"))
```

```
model piston

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

 mu~dnorm(3,0.25)


 for (i in 1:3)
     {alpha[i]~dnorm(0,1)
      for (j in 1:4)
          {gamma[i,j]~dnorm(0,4)
           }
      }

 for (j in 1:4)
     {beta[j]~dnorm(0,1)
      }

 }
```

Figure 1: Model specification for the piston-rings problem.

```
> modelUpdate(1000)
> samplesHistory("mu")
> samplesHistory("alpha")
> samplesHistory("beta")
> samplesHistory("gamma")
> samplesHistory("eta")
> modelUpdate(2000)
> samplesHistory("mu")
```

While convergence and mixing for the individual linear predictors η_i seemed satisfactory, mixing was less good for the underlying random effects and, particularly, for the overall mean μ. These quantities are less well identified than the individual linear predictors. Illustrative history plots are shown in Figure 2. Because of the poor mixing, a long burn-in (5000 iterations) and a large number of samples (10000), with two parallel chains, were used.

The following commands were used.

```
> modelCheck("pistonbug.txt")
> modelData("mypistondata.txt")
> modelCompile(2)
> modelGenInits()
> modelUpdate(5000)
> samplesSet(c("mu","alpha","beta","gamma","eta"))
> modelUpdate(10000)
> samplesStats("mu")
> samplesStats("alpha")
> samplesStats("beta")
> samplesStats("gamma")
> samplesStats("eta")
```

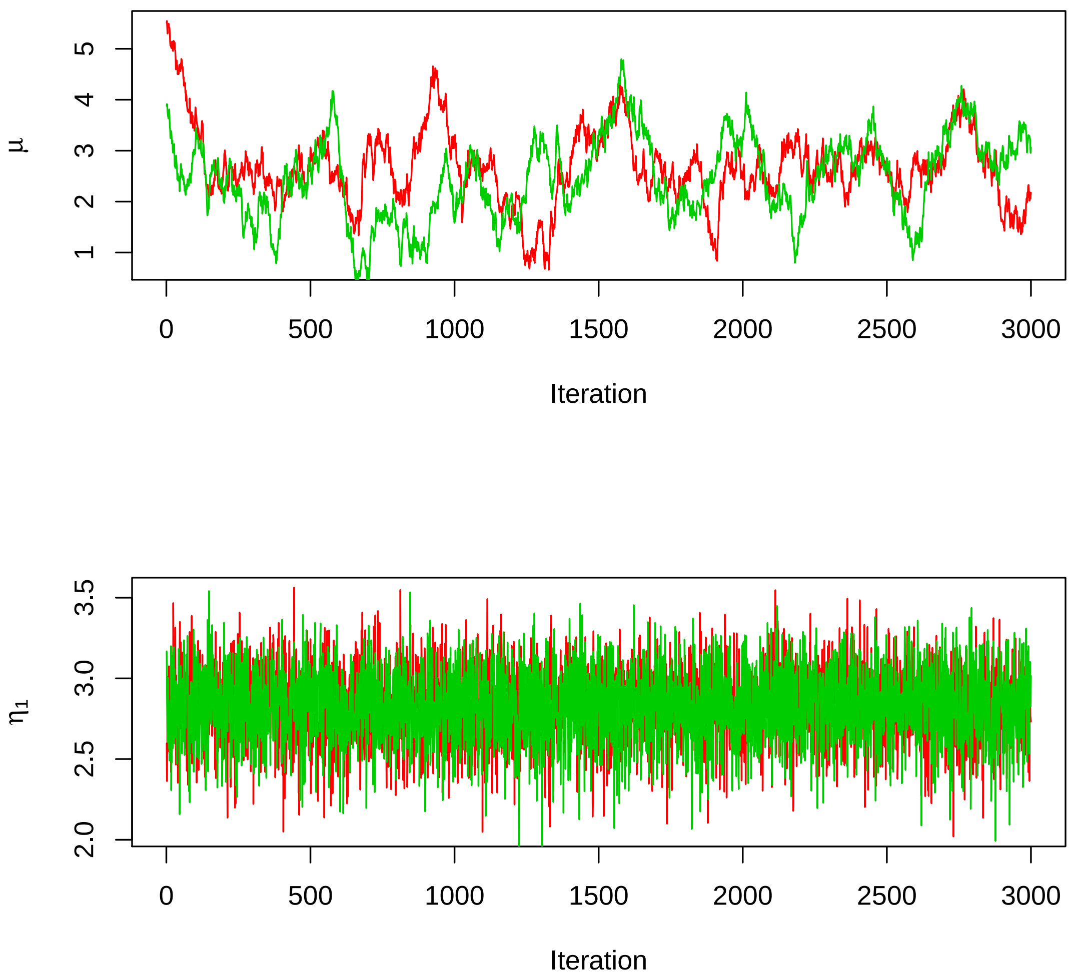

Figure 2: History plots for μ and $\eta_{1,1}$ (Compressor 1, North leg), piston-rings problem.

Compressor	Leg	Quantity	Prior		Posterior			
			Mean	S.D.	Mean	S.D.	95% Interval	
1	North	$\eta_{1,1}$	3	2.5	2.825	0.2311	2.3470	3.253
2	North	$\eta_{1,2}$	3	2.5	3.006	0.2123	2.5750	3.404
3	North	$\eta_{1,3}$	3	2.5	1.935	0.3365	1.2350	2.546
4	North	$\eta_{1,4}$	3	2.5	2.662	0.2456	2.1590	3.126
1	Centre	$\eta_{2,1}$	3	2.5	2.211	0.2994	1.5910	2.766
2	Centre	$\eta_{2,2}$	3	2.5	2.909	0.2216	2.4510	3.321
3	Centre	$\eta_{2,3}$	3	2.5	1.713	0.3659	0.9542	2.386
4	Centre	$\eta_{2,4}$	3	2.5	2.265	0.2950	1.6490	2.811
1	South	$\eta_{3,1}$	3	2.5	3.569	0.1630	3.2370	3.873
2	South	$\eta_{3,2}$	3	2.5	2.760	0.2346	2.2790	3.195
3	South	$\eta_{3,3}$	3	2.5	2.071	0.3185	1.4130	2.652
4	South	$\eta_{3,4}$	3	2.5	2.694	0.2417	2.1920	3.144
		μ	3	2	2.572	0.7437	1.109	4.119
	North	α_1	0	1	0.05074	0.5919	-1.1330	1.2150
	Centre	α_2	0	1	-0.26280	0.5926	-1.4550	0.8765
	South	α_3	0	1	0.19910	0.5996	-0.9876	1.3620
1		β_1	0	1	0.27920	0.5305	-0.7905	1.2810
2		β_2	0	1	0.30370	0.5472	-0.7843	1.3560
3		β_3	0	1	-0.61220	0.5448	-1.7080	0.4295
4		β_4	0	1	-0.02611	0.5430	-1.1190	0.9916
1	North	$\gamma_{1,1}$	0	0.5	-0.07681	0.3767	-0.8122	0.6588
2	North	$\gamma_{1,2}$	0	0.5	0.08032	0.3740	-0.6513	0.8217
3	North	$\gamma_{1,3}$	0	0.5	-0.07489	0.3983	-0.8601	0.7110
4	North	$\gamma_{1,4}$	0	0.5	0.06575	0.3786	-0.6652	0.8139
1	Centre	$\gamma_{2,1}$	0	0.5	-0.37720	0.3914	-1.1540	0.3776
2	Centre	$\gamma_{2,2}$	0	0.5	0.29620	0.3825	-0.4581	1.0410
3	Centre	$\gamma_{2,3}$	0	0.5	0.01669	0.4058	-0.7817	0.8094
4	Centre	$\gamma_{2,4}$	0	0.5	-0.01761	0.3924	-0.7906	0.7484
1	South	$\gamma_{3,1}$	0	0.5	0.51920	0.3677	-0.2074	1.2380
2	South	$\gamma_{3,2}$	0	0.5	-0.31430	0.3838	-1.0690	0.4253
3	South	$\gamma_{3,3}$	0	0.5	-0.08795	0.3962	-0.8656	0.6824
4	South	$\gamma_{3,4}$	0	0.5	-0.05094	0.3802	-0.8074	0.6915

Table 1: Prior and posterior summaries, piston-rings problem.

Table 1 shows a summary of the prior and posterior distributions. Figure 3 shows posterior means and 95% equitailed posterior credible intervals for the log failure rates $\eta_{i,j}$. The dashed line represents the posterior mean for μ. We can see that the rate for the South leg of Compressor 1 seems to be unusually great. It seems that Compressor 3 may have a generally lower failure rate.

Note that we can easily convert the 95% posterior intervals for the log failure rates $\eta_{i,j}$ into 95% posterior intervals for the actual failure rates. These are shown in Table 2.

3. Fraud

Banks and credit card companies attempt to detect fraud by looking for unusual observations in the withdrawal data for customers. This potentially involves quite complicated models. The model in this question is a somewhat simplified version but the principal is the same.

You will each be supplied with data for five customers. For each of these customers you will be given the total withdrawals from the customer's account for each of twenty weeks. The value for customer j in week i is $y_{i,j}$.

For each customer in each week there is a small probability π that a fraud takes place. We therefore use a mixture model with two components. The component indicator for customer

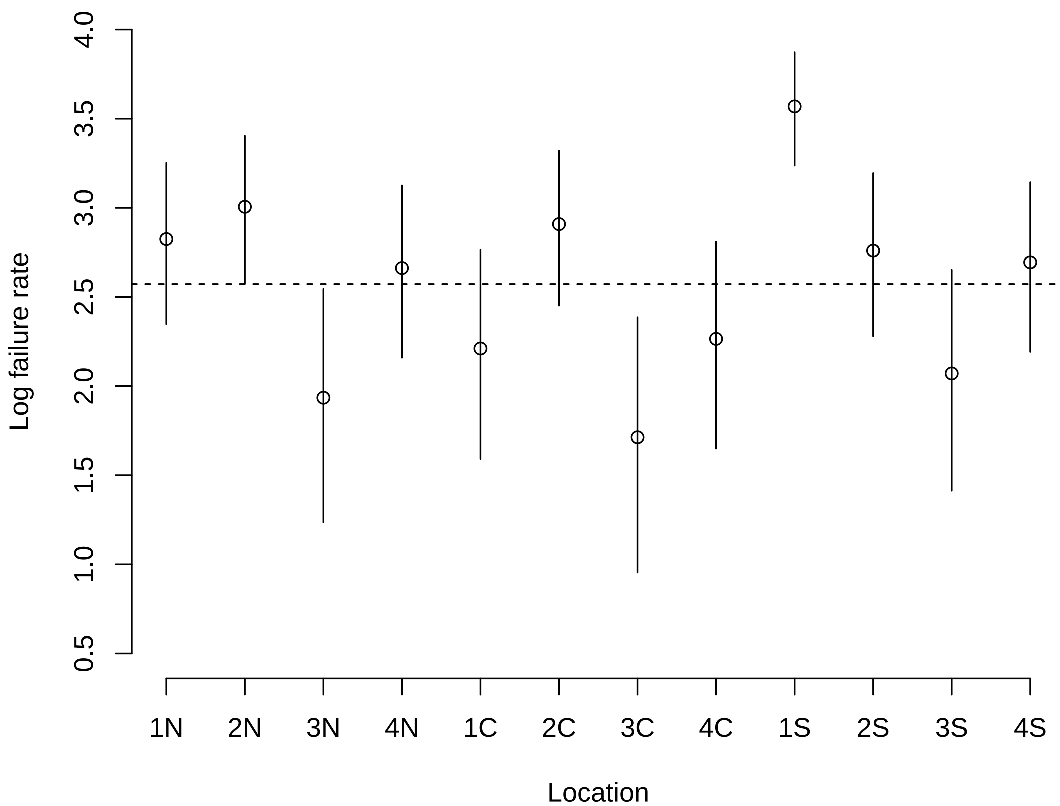

Figure 3: Posterior means and 95% equitailed posterior credible intervals for the log failure rates $\eta_{i,j}$, piston-rings problem. "1N" denotes "Compressor 1, North Leg" and so on.

Compressor	Leg	95% Posterior Interval	
1	North	10.5	25.9
2	North	13.1	30.1
3	North	3.4	12.8
4	North	8.7	22.8
1	Centre	4.9	15.9
2	Centre	11.6	27.7
3	Centre	2.6	10.9
4	Centre	5.2	16.6
1	South	25.5	48.1
2	South	9.8	24.4
3	South	4.1	14.2
4	South	9.0	23.2

Table 2: Equitailed 95% posterior intervals for failure rates, piston-rings problem.

j in week i is $c_{i,j}$.

If $c_{i,j} = 1$ then a fraud against customer j takes place in week i.

If $c_{i,j} = 2$ then no fraud takes place against customer j in week i.

We assume that, given the model parameters, $c_{i,j}$ is independent of $c_{i',j'}$ for $(i,j) \neq (i',j')$. Given π, we have $\Pr(c_{i,j} = 1) = \pi$. Our prior distribution for π is Beta$(1, 99)$.

If $c_{i,j} = 1$ then $y_{i,j} \sim \text{Ga}(2,\ 0.0002)$. If $c_{i,j} = 2$ then, given $\alpha,\ \beta_j$, we have $y_{i,j} \sim \text{Ga}(\alpha, \beta_j)$. We assume that $y_{i,j}$ is independent of $y_{i',j'}$ for $(i,j) \neq (i',j')$, given $\alpha,\ \beta_j$ and $\beta_{j'}$. Our prior distribution for α is Ga$(2,\ 0.5)$.

Let $\beta_j = \alpha/\lambda_j$ and $\lambda_j = \exp(\mu_j)$. Given $\mu_0,\ \tau$, we have $\mu_j \sim N(\mu_0, \tau^{-1})$ with μ_j independent of $\mu_{j'}$ for $j \neq j'$. Our prior distribution for μ_0 is $\mu_0 \sim N(5.3, 1.4)$. Our prior distribution for τ is Ga$(3, 4)$.

Unless otherwise stated, the prior distributions are independent.

- Use MCMC to find the posterior means for $p_{i,j} = 2 - c_{i,j}$ and hence find the posterior probabilities of fraud for each customer in each week and identify any cases where fraud is likely to have occurred.

(5 marks)

- Display your results appropriately.

(4 marks)

- Explain your method and show your BUGS model specification and the commands which you have used.

(4 marks)

- Show how you have checked convergence.

(3 marks)

- Give summaries of the posterior distributions of the model parameters.

(4 marks)

<u>Solution</u>

The solution below refers to data set 30.

The model (as described above) was specified in the BUGS model specification given in Figure 4. Note that the variable `prob[i,j]` is included as an indicator of fraud in week j for customer i. Its posterior expectation is the posterior probability of fraud in that case.

The posterior distribution was evaluated using a Gibbs sampler, implemented using the BRugs software. In order to check the convergence and mixing of the MCMC sampler, an initial run of 5000 iterations was made with two parallel chains started with different initial values. The two initial value files were as follows.

```
list(mu=c(5,5,5,5,5),pi=0.01)

list(mu=c(4,4,4,4,4),pi=0.05)
```

Note: We can sometimes choose initial values for convergence checking by trying an initial run of the sampler and then setting the starting values outside the main range of the sampled values in each direction.

The commands used were as follows.

```
model fraud

{for (j in 1:5)
   {for (i in 1:20)
      {c[i,j]~dcat(q[])
       prob[i,j]<-2-c[i,j]
       y[i,j]~dgamma(alpha[c[i,j]],beta[c[i,j],j])
    }
 beta[2,j]<-alpha[2]/lambda[j]
 beta[1,j]<-0.0002
 lambda[j]<-exp(mu[j])
 mu[j]~dnorm(mumean,tau)
 }

 mumean~dnorm(5.3,p.mu)
 p.mu<-1/1.4
 tau~dgamma(3,4)
 alpha[1]<-2
 alpha[2]~dgamma(2,0.5)

 pi~dbeta(1,99)
 q[1]<-pi
 q[2]<-1-pi
 }
```

Figure 4: BUGS model specification, fraud problem.

```
> modelCheck("fraudbug.txt")
> modelData("myfrauddata.txt")
> modelCompile(2)
> modelInits("fraudinits1.txt")
> modelInits("fraudinits2.txt")
> modelGenInits()
>  samplesSet(c("pi","alpha","tau","mumean"))
> modelUpdate(5000)
> samplesHistory("pi")
> samplesHistory("alpha")
> samplesHistory("tau")
> samplesHistory("mumean")
```

Note that, even though some unknowns were initialised, it was necessary to generate initial values for others.

Figure 5 shows plots of the sampled values of $\pi,\ \tau,\ \mu_0$ and α against interation number. The graphs show little indication of problems, although it seems that in the cases of π and τ the posterior distributions may have long right-hand tails. It seems that we can safely assume that convergence has been achived after 5000 iterations and that a further 10000 iterations should be sufficient for inference purposes.

To evaluate the posterior distribution, a further run of the sampler was used, with a burn-in of 50000 iterations and values collected over 10000 iterations, with two parallel chains. The commands used were as follows.

```
> modelCheck("fraudbug.txt")
> modelData("myfrauddata.txt")
> modelCompile(2)
```

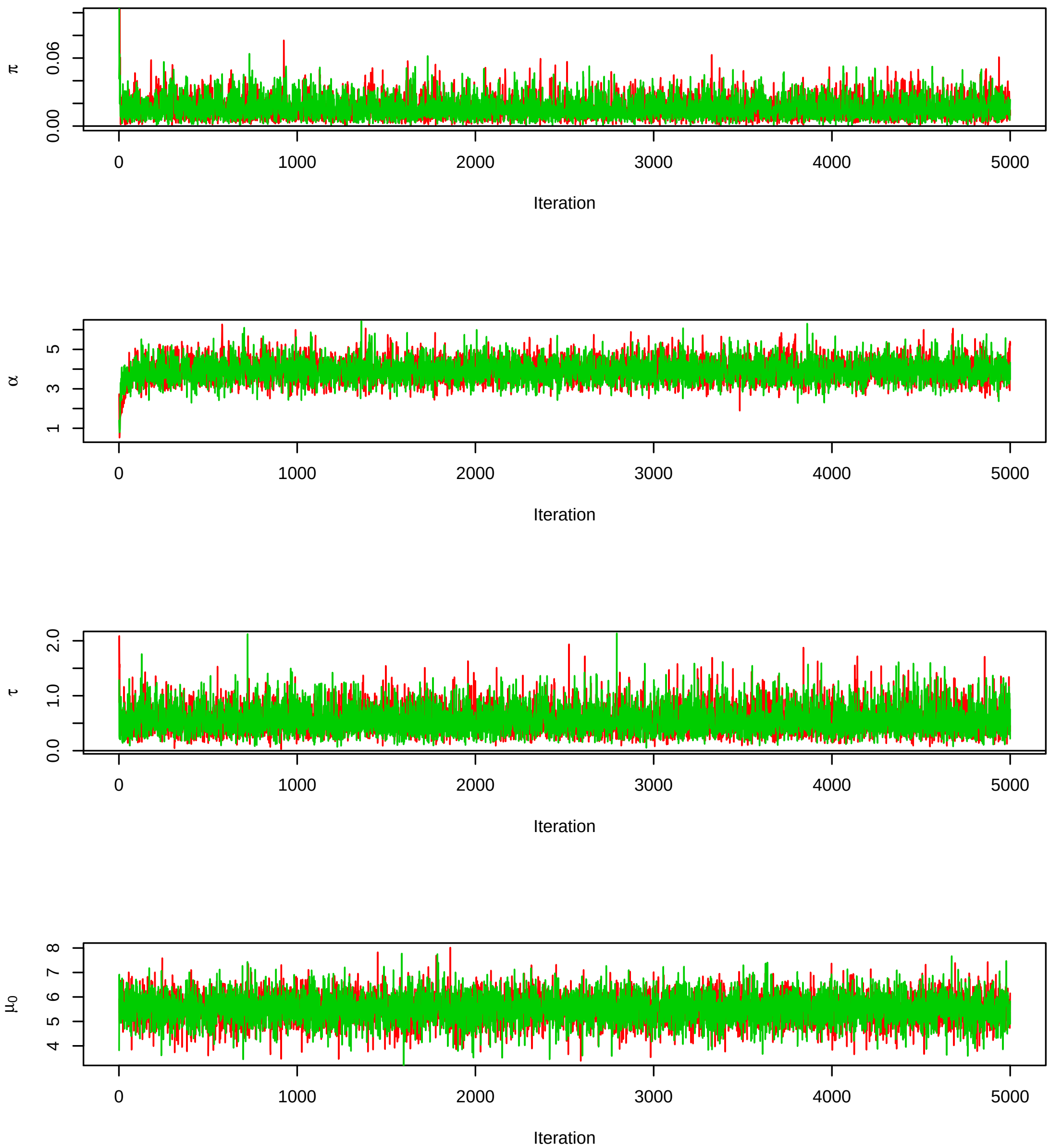

Figure 5: Convergence check, fraud problem.

Parameter	Mean	Std.Dev.	MC error	2.5% point	Median	97.5% point
π	0.01602	0.009263	0.000089	0.003341	0.01431	0.03852
α	3.941	0.5556	0.004477	2.913	3.921	5.11
τ	0.5631	0.2481	0.001971	0.1846	0.5276	1.143
μ_0	5.521	0.5656	0.004312	4.374	5.528	6.622
λ_1	379.50	43.470	0.66700	302.00	376.80	473.30
λ_2	41.66	4.812	0.07125	33.19	41.37	52.04
λ_3	802.70	101.900	1.68300	631.10	794.10	1029.00
λ_4	1684.00	200.900	3.26300	1334.00	1670.00	2119.00
λ_5	65.81	7.754	0.11400	52.46	65.17	82.59

Table 3: Summaries of posterior distributions of model parameters

```
> modelGenInits()
> modelUpdate(5000)
> samplesSet(c("pi","alpha","tau","mumean","lambda","prob"))
> modelUpdate(10000)
```

Summary values of the posterior distributions of the main model parameters and the mean weekly withdrawals for the five customers are shown in Table 3. It is clearly seen that there is wide variation between the mean withdrawal amounts for the five customers.

Figure 6 shows marginal prior and posterior densities for the four main model parameters. It can be seen that the data have had little effect on the distributions in some cases, particularly of τ. The posterior distribution is similar to the prior distribution. On the other hand there is a clear difference between the prior and posterior densities in the cases of μ_0 and α.

Table 4 shows the posterior probability of fraud for each customer in each week. The results for Customer 3 in Week 2 and Customer 5 in Week 8 show that these were almost certainly cases of fraud (or, at least, very unusual behaviour). Customer 4 in Week 12 gets a probability of almost 13% which merits investigation. There is only one other probability, Customer 4 in Week 5, which is greater than 1%.

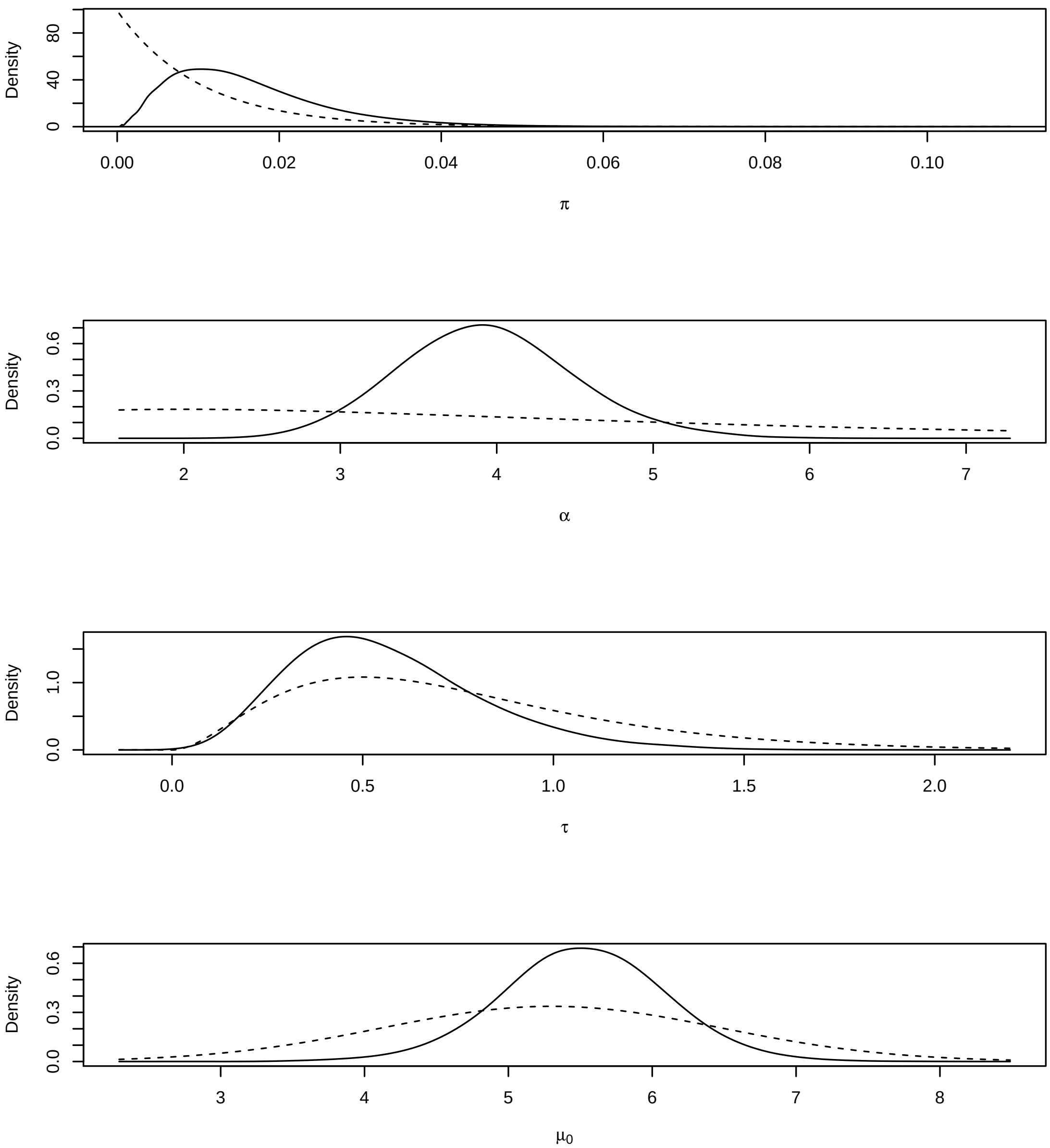

Figure 6: Prior (dashed) and posterior (solid) marginal densities for the four main model parameters

Week	Customer				
	1	2	3	4	5
1	0.00000	0.00000	0.00020	0.00090	0.00005
2	0.00000	0.00000	0.97750	0.00410	0.00000
3	0.00020	0.00000	0.00000	0.00110	0.00000
4	0.00000	0.00000	0.00010	0.00445	0.00000
5	0.00000	0.00000	0.00015	0.01010	0.00000
6	0.00035	0.00000	0.00085	0.00290	0.00000
7	0.00015	0.00000	0.00415	0.00620	0.00000
8	0.00090	0.00000	0.00060	0.00110	1.00000
9	0.00070	0.00000	0.00000	0.00405	0.00000
10	0.00000	0.00000	0.00020	0.00140	0.00045
11	0.00070	0.00000	0.00025	0.00160	0.00000
12	0.00120	0.00000	0.04075	0.12830	0.00000
13	0.00000	0.00000	0.00075	0.00125	0.00000
14	0.00020	0.00000	0.00030	0.00120	0.00000
15	0.00000	0.00000	0.00485	0.00185	0.00000
16	0.00000	0.00000	0.00045	0.00455	0.00000
17	0.00010	0.00000	0.00035	0.00075	0.00005
18	0.00010	0.00000	0.00130	0.00135	0.00000
19	0.00045	0.00000	0.00045	0.00145	0.00000
20	0.00000	0.00000	0.00170	0.00260	0.00000

Table 4: Posterior probabilities of fraud