NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS & STATISTICS

SEMESTER 1 2011/2012

MAS8303

Modern Bayesian Inference: SOLUTIONS

Time allowed: 2 hours

Candidates should attempt all questions. Marks for each question are indicated. However you are advised that marks indicate the relative weight of individual questions, they do not correspond directly to marks on the University scale.

There are SIX questions on this paper.

Calculators may be used. Extracts from the WinBUGS manual are included in this paper.

1. JQS

[15 marks]

2. JQS

[10 marks]

3. (a) Given the values of $\alpha, \beta_1, \beta_2, \beta_3$, the lifetime T_i of belt i, $i = 1, \ldots, N$, has a gamma distribution,

$$T_i \sim \text{Ga}(\alpha, \beta_{m_i}),$$

where belt i is made of material m_i, and T_i is independent of T_k for $k \neq i$.

The prior distributions are as follows:

$$\alpha \sim \text{Ga}(1.5, 0.1)$$

and, for $j = 1, 2, 3$,

$$\beta_j = \beta_0 + \delta_j$$

where $\beta_0 \sim \text{Ga}(3, 500)$ and $\delta_j \sim \text{Ga}(2, 500)$ with $\alpha, \beta_0, \delta_1, \delta_2, \delta_3$ independent.

[5 marks]

(b) We have $\beta_1 = \beta_0 + \delta_1$.

$$\begin{aligned}
\text{E}(\beta_0) &= \frac{3}{500} = 0.006 \\
\text{Var}(\beta_0) &= \frac{3}{500^2} = 1.2 \times 10^{-5} \\
\text{E}(\delta_1) &= \frac{2}{500} = 0.004 \\
\text{Var}(\delta_1) &= \frac{2}{500^2} = 0.8 \times 10^{-5}
\end{aligned}$$

Hence

$$\begin{aligned}
\text{E}(\beta_1) &= 0.006 + 0.004 = 0.01 \\
\text{Var}(\beta_1) &= (1.2 + 0.8) \times 10^{-5} = 2 \times 10^{-5}
\end{aligned}$$

[4 marks]

(c)

$$\text{Covar}(\beta_1, \beta_2) = \text{Var}(\beta_0) = 1.2 \times 10^{-5}$$

[3 marks]

[Total: 12 marks]

4. (a) The linear predictor is

$$\eta_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2}.$$

The link function is log:

$$\eta_i = \ln\{\mathrm{E}(Y_i \mid \lambda_i)\}.$$

[4 marks]

(b) Let $\mathbf{x} = (1,\ 2.5,\ -0.7)$ and $\beta = (\beta_0, \beta_1, \beta_2)'$.

$$\begin{aligned}
\eta_i &= \mathbf{x}\beta \\
\mathrm{E}(\eta_i) &= 1 \times (-1) + 2.5 \times 0.1 + (-0.7) \times 0.2 = -0.89 \\
\mathrm{Var}(\eta_i) &= \mathbf{x}V\mathbf{x}' \\
&= (0.5,\ 0.043,\ -0.004)\mathbf{x}' \\
&= 0.6047
\end{aligned}$$

95% interval for η_i is $-0.89 \pm 1.96\sqrt{0.6047}$. That is

$$-2.4141 < \eta_i < 0.6341.$$

This converts to

$$0.0894 < \lambda_i < 1.8854.$$

[5 marks]

(c) Let $I_i = 1$ if $x_{i,2}$ is observed and $I_i = 0$ if $x_{i,2}$ is missing. Then the values of $x_{i,2}$ are missing at random if I_i and $X_{i,2}$ are conditionally independent given $y_1, \ldots, y_n, x_{1,1}, \ldots, x_{n,1}$ and the observed values among $x_{1,2}, \ldots, x_{n,2}$.

I would accept:

...conditionally independent given $y_i, x_{i,1}$.

[4 marks]

[Total: 13 marks]

5. Answer *EITHER* **PART (a)** *OR* **PART (b)** below.

(a) JQS

(b) JQS

[Total: 25 marks]

6. Answer *EITHER* **PART (a)** *OR* **PART (b)** below.

(a) (i) Distribution function:

$$\begin{aligned} F_i(t) &= 1 - \exp\{-\lambda_i t^\alpha\} \\ &= 1 - \exp\{-(\rho_i t)^\alpha\} \\ F_i(m_i) = \frac{1}{2} &= 1 - \exp\{-(\rho_i m_i)^\alpha\} \\ &= \exp\{-(\rho_i m_i)^\alpha\} \\ \text{So } \ln 2 &= (\rho_i m_i)^\alpha \\ \text{and } m_i &= (\ln 2)^{1/\alpha} \rho_i^{-1} \end{aligned}$$

[4 marks]

(ii)

$$k = \frac{m_A}{m_b} = \frac{(\ln 2)^{1/\alpha}}{(\ln 2)^{1/\alpha}} \frac{\rho_A^{-1}}{\rho_B^{-1}} = \frac{\rho_B}{\rho_A}$$

$$\begin{aligned} \ln k &= \ln \rho_B - \ln \rho_A \\ &= (\beta_0 + \beta_1 x_{A,1} + \beta_2 x_{A,2} - \beta_3) - (\beta_0 + \beta_1 x_{A,1} + \beta_2 x_{A,2} + \beta_3) \\ &= -2\beta_3 \end{aligned}$$

So

$$\begin{aligned} 1/2 < k < 2 &\Leftrightarrow -\ln 2 < \ln k < \ln 2 \\ &\Leftrightarrow -\ln 2 < -2\beta_3 < \ln 2 \\ &\Leftrightarrow -(\ln 2)/2 < \beta_3 < (\ln 2)/2 \end{aligned}$$

So

$$\begin{aligned} \mathrm{E}(\beta_3) &= 0 \\ \text{Std.dev.}(\beta_3) &= \frac{\ln 2}{2 \times 1.96} \\ \mathrm{Var}(\beta_3) &= \left(\frac{\ln 2}{2 \times 1.96}\right)^2 = 0.03127 \end{aligned}$$

[6 marks]

Question 6 continued on next page

(iii)
```
model Q7a

{for ( i in 1:n)
     {t[i]~dweib(alpha,lambda[i])
      lambda[i]<-exp(alpha*eta[i])
      eta[i]<-beta0+beta[1]*x[i,1]+beta[2]*x[i,2] \
                                   +beta[3]*x[i,3]
      }

 alpha~dgamma(1.5,1.5)
 beta0~dnorm(-3,10)
 beta[1]~dnorm(0,6.25)
 beta[2]~dnorm(0,0.625)
 beta[3]~dnorm(0,31.98)
 }
```

[**8 marks**]

(iv) Simply replace the line

```
t[i]~dweib(alpha,lambda[i])
```

with

```
t[i]~dweib(alpha,lambda[i]) I(c[i],)
```

In a case where we observe t_i, we set `c[i]=0` and `t[i]` is the observed value of t_i. Then the `I(c[i],)` construction has no effect.
In a case where the observation is censored, we set `c[i]` equal to the censoring time `c[i]` and enter `NA` for `t[i]` Then `t[i]` will be treated as missing and, at each iteration, it will be sampled from the correct truncated distribution.

Note: The method in Part (iv) is UNSEEN.

[**7 marks**]

Question 6 continued on next page

(b) (i)

$$\mathrm{E}(Y_i) = \frac{p_{j(i)}\alpha}{\alpha/\rho} = p_{j(i)}\rho$$

so the mean amount per pig is ρ.

[2 marks]

(ii)

$$\frac{a}{b} = 0.04 \quad \text{and} \quad \frac{\sqrt{a}}{b} = 0.025$$

so $\sqrt{a} = 0.04/0.025 = 1.6$ and

$$a = 2.56.$$

Then

$$b = \frac{a}{0.04} = 64.$$

[2 marks]

(iii)

```
model Q7b

{for (i in 1:N)
     {y[i]~dgamma(a[i],b)
      a[i]<-p[j[i]]*alpha
      }
 b<-alpha/rho

 alpha~dgamma(1.5,0.5)
 rho~dgamma(2.56,64)
 }
```

[7 marks]

(iv)

```
model Q7biv

{for (i in 1:N)
     {c[i]~dcat(pi[])
      y[i]~dgamma(a[i],b[i])
      a[i]<-p[j[i]]*alpha
      b[i]<-alpha/rho[c[i]]
      }
```

Question 6 continued on next page

```
alpha~dgamma(1.5,0.5)
rho[1]~dgamma(2.56,64)
rho[2]~dgamma(4,400)
pi[1]~dbeta(2,1)
pi[2]<-1-pi[1]
}
```

We might experience "label switching". While we would expect $\rho_1 > \rho_2$, the model does not force this to be the case and we might choose samples where $\rho_1 < \rho_2$ and cases previously allocated to component 1 are allocated to component 2 and *vice versa.* This will lead to a bimodal posterior.

We could avoid this, for example, by imposing the constraint $\rho_1 < \rho_2$ using, eg,

```
rho[1]~dgamma(2.56,64) I(rho[2],)
rho[2]~dgamma(4,400)   I(,rho[1])
```

[**8 marks**]

(v) The BUGS model specification stays the same. However `c[i]` is now observed.

For a case where $C_i = 1$, ie $Y_i > 0$, we enter the observed value of Y_i for `y[i]`.

For a case where $C_i = 2$, ie $Y_i = 0$, we enter `NA` for `y[i]`. These observations then make no contribution to the likelihood for ρ but they do, of course, contribute to the likelihood for π.

Note: The method in Part (v) is UNSEEN.

[**6 marks**]

[**Total: 25 marks**]

THE END