NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS & STATISTICS

SEMESTER 1 2010/2011

MAS8303

Modern Bayesian Inference: SPECIMEN

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 EIGHT questions on this paper.

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

- Beta $Beta(\alpha, \beta)$ distribution. It has density function

$$f(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)},\ 0 < x < 1,\ \alpha > 0, \beta > 0$$

where $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$.
Also, $\mathrm{E}(X) = \alpha/(\alpha + \beta)$ and $\mathrm{Var}(X) = \alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)]$.

- Exponential $Exp(\lambda)$ distribution. It has density

$$f(x|\lambda) = \lambda e^{-\lambda x},\ x > 0,\ \lambda > 0.$$

Also, $\mathrm{E}(X) = 1/\lambda$ and $\mathrm{Var}(X) = 1/\lambda^2$.

- Gamma $Ga(\alpha, \lambda)$ distribution. It has density

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)},\ x > 0,\ \alpha > 0,\ \lambda > 0.$$

Also, $\mathrm{E}(X) = \alpha/\lambda$ and $\mathrm{Var}(X) = \alpha/\lambda^2$.

- Normal $N(\mu, \sigma^2)$ distribution. It has density

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},\ -\infty < x < \infty,$$
$$-\infty < \mu < \infty,\ \sigma > 0.$$

Also, $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

- Lognormal $LN(\mu, \sigma^2)$ distribution. It has density

$$f(x|\mu, \sigma) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\},\ x > 0,\ -\infty < \mu < \infty,\ \sigma > 0.$$

Also, $\mathrm{E}(X) = \exp(\mu + \sigma^2/2)$ and $\mathrm{Var}(X) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$.

- Rayleigh $R(\theta)$ distribution. It has density

$$f(x|\theta) = 2x\theta e^{-\theta x^2},\quad x > 0,\quad \theta > 0.$$

Also, $\mathrm{E}(X) = \sqrt{\pi/(4\theta)}$ and $\mathrm{Var}(X) = (4 - \pi)/(4\theta)$.

- Poisson $Po(\lambda)$ distribution. It has probability function

$$Pr(X = j \mid \lambda) = \frac{e^{-\lambda}\lambda^j}{j!},\quad j = 0, 1, \ldots,\quad \lambda > 0.$$

Also,

$$\mathrm{E}(X) = \lambda,\qquad \mathrm{Var}(X) = \lambda.$$

- Dirichlet $D_d(a_1, \ldots, a_d)$ distribution. It has density

$$f(x_1, \ldots, x_n \mid a_1, \ldots, a_n) = \frac{\Gamma(A)}{\prod_{i=1}^{d} \Gamma(a_i)} \prod_{i-1}^{d} x_i^{a_i - 1}, \quad A = \sum_{i=1}^{d} a_i,$$

$$0 < a_i,\ 0 < x_i < 1,\ \sum_{i=1}^{d} x_i = 1.$$

Also

$$\mathrm{E}(X_i) = \frac{a_i}{A},\ \mathrm{Var}(X_i) = \frac{a_i(A - a_i)}{A^2(A+1)},\ \mathrm{Covar}(X_i, X_j) = -\frac{a_i a_j}{A^2(A+1)},\ (i \neq j).$$

BUGS functions

Function	Usage	Definition
Complementary	`cloglog(p)<-a+b*x`	$\log[-\log(1-p)] = a + bx$
log log	`y<-cloglog(p)`	$y = \log[-\log(1-p)]$
Logical equals	`y<-equals(x,z)`	$y = 1$ if $x = z$
		$y = 0$ if $x \neq z$
Exponential	`y<-exp(x)`	$y = e^x$
Inner product	`y<-inprod(a[],b[])`	$y = \sum_i a_i b_i$
Matrix inverse	`y[,]<-inverse(x[,])`	$y = x^{-1}$
		$y,\ x$ both $n \times n$ matrices
Natural logarithm	`log(lambda)<-a+b*x`	$\log(\lambda) = a + bx$
	`y<-log(x)`	$y = \log x$
Log determinant	`y<-logdet(x[,])`	$y = \log \lvert x \rvert$
		x is a $n \times n$ matrix
Log factorial	`y<-logfact(x)`	$y = \log(x!)$
Log(gamma function)	`y<-loggam(x)`	$y = \log[\Gamma(x)]$
Logit	`y<-logit(p)`	$y = \log[p/(1-p)]$
	`logit(p)<-a+b*x`	$\log[p/(1-p)] = a + bx$
Maximum	`c<-max(a,b)`	$c = \max(a, b)$
Mean	`x.bar<-mean(x[])`	$\bar{x} = \sum_i x_i/n$
Minimum	`c<-min(a,b)`	$c = \min(a, b)$
Standard Gaussian	`p<-phi(x)`	$p = \int_{-\infty}^{x} (2\pi)^{-1/2} e^{-t^2/2}\ dt$
distribution function		i.e. $p = \Phi(x)$
Power	`z<-pow(x,y)`	$z = x^y$
Probit	`y<-probit(p)`	$y = \Phi^{-1}(p)$
	`probit(p)<-a+b*x`	$\Phi^{-1}(p) = a + bx$
Standard deviation	`s<-sd(x[])`	$s = \sqrt{\sum_i (x_i - \bar{x})^2/n}$
Square root	`sigma<-sqrt(tau)`	$\sigma = \sqrt{\tau}$
Unit step	`y<-step(x)`	$y = 0$ if $x < 0$
		$y = 1$ if $x \geq 0$
Sum	`x.sum<-sum(x[])`	$x_{\text{sum}} = \sum_i x_i$

BUGS distributions

Distribution	Usage	Definition
Bernoulli	`r~dbern(p)`	$f(r \mid p) = p^r(1-p)^{1-r}$; $r = 0, 1$
beta	`p~dbeta(a,b)`	$f(p \mid a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}$; $0 < p < 1$
binomial	`r~dbin(p,n)`	$f(r \mid p, n) = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$; $r = 0, \ldots, n$
categorical	`r~dcat(p[])`	$f(r \mid p_1, \ldots, p_R) = p_r$; $r = 1, 2, \ldots, R$ where $R = \dim(\mathbf{p})$
chi-squared	`x~dchisq(k)`	$f(x \mid k) = 2^{-k/2} x^{k/2-1} e^{-x/2} / \Gamma(\frac{k}{2})$; $x > 0$
double exponential	`x~ddexp(mu,tau)`	$f(x \mid \mu, \tau) = \frac{\tau}{2} e^{-\tau \lvert x-\mu \rvert}$; $-\infty < x < \infty$
Dirichlet	`p[]~ddirch(alpha[])`	$f(\mathbf{p} \mid \alpha) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \prod_i p_i^{\alpha_i - 1}$; $0 < p_i < 1, \ \sum_i p_i = 1$
exponential	`x~dexp(lambda)`	$f(x \mid \lambda) = \lambda e^{-\lambda x}$; $x > 0$
gamma	`x~dgamma(r,mu)`	$f(x \mid r, \mu) = \mu^r x^{r-1} e^{-\mu x} / \Gamma(r)$; $x > 0$
lognormal	`x~dlnorm(mu,tau)`	$f(x \mid \mu, \tau) = \sqrt{\frac{\tau}{2\pi}} x^{-1} \exp[-\frac{\tau}{2}(\log x - \mu)^2]; \quad x > 0$
logistic	`x~dlogis(mu,tau)`	$f(x \mid \mu, \tau) = \tau e^{\tau(x-\mu)} / (1 + e^{\tau(x-\mu)})^2; \quad -\infty < x < \infty$
multivariate normal	`x[]~dmnorm(mu[],T[,])`	$f(\mathbf{x} \mid \mu, \mathbf{T}) = (2\pi)^{-N/2} \lvert \mathbf{T} \rvert^{1/2} \exp[-\frac{1}{2}(\mathbf{x} - \mu)' \mathbf{T} (\mathbf{x} - \mu)]$; $-\infty < x_i < \infty$
multinomial	`x[]~dmulti(p[],N)` $0 < p_i < 1, \ \sum_i p_i = 1$	$f(\mathbf{x} \mid \mathbf{p}, N) = \frac{(\sum_i x_i)!}{\prod_i x_i!} \prod_i p_i^{x_i}$; $\sum_i x_i = N$
negative binomial	`x~dnegbin(p,r)`	$f(x \mid p, r) = \frac{(x+r-1)!}{x!(r-1)!} p^r (1-p)^x$; $x = 0, 1, 2, \ldots$
normal	`x~dnorm(mu,tau)`	$f(x \mid \mu, \tau) = \sqrt{\frac{\tau}{2\pi}} \exp[-\frac{\tau}{2}(x - \mu)^2]; \quad -\infty < x < \infty$
Pareto	`x~dpar(alpha,c)`	$f(x \mid \alpha, c) = \alpha c^\alpha x^{-(\alpha+1)}$; $x > c$

BUGS distributions continued

Distribution	Usage	Definition
Poisson	`r~dpois(lambda)`	$f(r \mid \lambda) = e^{-\lambda}\frac{\lambda^r}{r!}$; $r = 0, 1, \ldots$
Student's t	`x~dt(mu,tau,k)`	$f(x \mid \mu, \tau, k) = \frac{\Gamma([k+1]/2)}{\Gamma(k/2)}\sqrt{\frac{\tau}{k\pi}}[1 + \frac{\tau}{k}(x-\mu)^2]^{-(k+1)/2}$; $-\infty < x < \infty$
uniform	`x~dunif(a,b)`	$f(x \mid a, b) = 1/(b-a)$; $a < x < b$
Weibull	`x~dweib(v,lambda)`	$f(x \mid v, \lambda) = v\lambda x^{v-1}\exp(-\lambda x^v)$; $x > 0$
Wishart	`x[,]~dwish(R[,],k)`	$f(\mathbf{x} \mid \mathbf{R}, k) \propto \lvert\mathbf{R}\rvert^{k/2}\lvert\mathbf{x}\rvert^{(k-p-1)/2}\exp(-\frac{1}{2}\mathrm{tr}[\mathbf{R}\mathbf{x}])$; $\mathbf{x}$ symmetric and positive definite

1. Dr Shi's question.

[7 marks]

2. Dr Shi's question.

[8 marks]

3. Dr Shi's question.

[10 marks]

4. In an experiment people are tested to see how they learn to perform a difficult task. Different groups of subjects (i.e. people) are given different lengths of time practising the task and each person is then tested once to see whether they complete the task successfully. The number y_i of subjects who successfully complete the task in Group i is counted ($i = 1, \ldots, N$). The number of subjects tested in Group i is n_i. Members of Group i were given x_i hours of practice before the test.

The following BUGS code is used in the analysis of the data. In the BUGS code `y[i]` stands for y_i, `n[i]` stands for n_i and `x[i]` stands for x_i. We will use β_0, β_1 and p_i to refer to the quantities denoted by `beta0`, `beta1` and `p[i]` in the code.

```
{
 for (i in 1:N)
   { y[i] ~ dbin(p[i],n[i])
     logit(p[i])<-beta0+beta1*x[i]
    }

 beta0~dnorm(-1.0, 1.0)
 beta1~dnorm(0.0, 5.0)
 }
```

(a) Write down a description of the model expressed by this code using standard mathematical notation.

(b) Find the prior lower and upper quartiles of p_i if $x_i = 5.0$.
Note that, if $Z \sim N(0,1)$, then the quartiles of Z are ± 0.6745.

[9 marks]

5. In a medical study, patients are given a simple diagnostic test T_1. The result Y_1 is either "positive" or "negative." If Y_1 is positive then the patient is given a second test T_2 which gives a result Y_2. If the result of T_1 is negative then the patient is either given T_2, with probability 0.2, or not given any further test, with probability 0.8. Determine whether the missing Y_2 values are "missing at random."

[5 marks]

6. (a) Explain briefly what is meant by "data augmentation."

(b) An ecologist counts the number of bumble bees seen in a series of two-hour visits, on different days, to a site. Let the number seen on visit i be Y_i. The ecologist wishes to use a negative binomial distribution with probability function

$$\Pr(Y_i = j) = \frac{\Gamma(a+j)}{j!\Gamma(a)} p^a (1-p)^j \quad (j = 0, 1, 2, \ldots)$$

to model these data.

Show that, by introducing random quantities $\lambda_i > 0$ and supposing that

$$\begin{aligned} Y_i \mid \lambda_i &\sim \mathrm{Po}(\lambda_i) \\ \text{and} \quad \lambda_i &\sim \mathrm{Ga}(a,\ b) \end{aligned}$$

where $p = b/(b+1)$, we obtain the same distribution for Y_i and comment on how this fact might be used in a MCMC algorithm

[11 marks]

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

(a) Dr Shi's question.

(b) Dr Shi's question.

[**25 marks**]

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

(a) (i) We wish to construct a model for the one-year survival of patients with a particular disease. Let $p(\underline{x}, \underline{\theta})$ be the conditional probability that a patient survives one year given a vector of covariate values $\underline{x}$ and the model parameters $\underline{\theta}$. Let

$$\eta(\underline{x}, \underline{\theta}) = \log_e \left(\frac{p(\underline{x}, \underline{\theta})}{1 - p(\underline{x}, \underline{\theta})} \right).$$

A. For a particular covariate vector $\underline{x}$, we wish to give $\eta(\underline{x}, \underline{\theta})$ a normal prior distribution with mean m and variance v. Find values for m and v such that

$$\Pr[p(\underline{x}, \underline{\theta}) < 0.5] = \Pr[p(\underline{x}, \underline{\theta}) > 0.75] = 0.05.$$

Note that, if $Z \sim N(0, 1)$, then $\Pr(Z < 1.645) = \Pr(Z > 1.645) = 0.05$.

B. Suppose that

$$\eta(\underline{x}, \underline{\theta}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

where x_1 and x_2 are scalar covariates. Suppose that we make the following prior judgements.
When $x_1 = 2$ and $x_2 = 0$

$$\eta(\underline{x}, \underline{\theta}) = \eta_1 \sim N(1.8,\ 0.04).$$

When $x_1 = 0$ and $x_2 = 1$

$$\eta(\underline{x}, \underline{\theta}) = \eta_2 \sim N(0.0,\ 0.10).$$

When $x_1 = 1$ and $x_2 = 1$

$$\eta(\underline{x}, \underline{\theta}) = \eta_3 \sim N(1.0,\ 0.07).$$

These three distributions are independent.
Find the joint prior distribution of $\theta_0, \theta_1, \theta_2$.

(ii) Patients suffering from a certain long-term chronic illness suffer events which require medical attention from time to time. After a patient is put on long-term treatment, the times of these events are recorded. Let the time till the first event for patient i be $T_{i,1}$, the time between the first and second events be $T_{i,2}$ and time between the second and third events be $T_{i,3}$. These three times are recorded for n patients, $i = 1, \ldots, n$.
Our model says that, given the values of α_j and λ_i, for $j = 1, 2, 3$, we have

$$T_{i,j} \sim \text{Ga}(\alpha_j, \lambda_i)$$

with $T_{i,j}$ independent of $T_{i',j'}$ unless $i = i'$ and $j = j'$.
Given the values of $\pi,\ \alpha_\lambda,\ \beta_1, \beta_2$, where $0 < \pi < 1$, the other parameters are all positive and $\beta_2 < \beta_1$, we have a two-component mixture distribution for λ_i with density

$$f_\lambda(\lambda_i) = \pi f_{\lambda,1}(\lambda_i;\ \alpha_\lambda, \beta_1) + (1 - \pi) f_{\lambda,2}(\lambda_i;\ \alpha_\lambda, \beta_2)$$

where $f_{\lambda,k}$ is a gamma density

$$f_{\lambda,k} = \frac{\beta_k^{\alpha_\lambda} \lambda_i^{\alpha_\lambda - 1} \exp(-\beta_k \lambda_i)}{\Gamma(\alpha_\lambda)}.$$

A. Verify that this model can be reformulated by writing

$$f_\lambda(\lambda_i \mid Z = z) = z f_{\lambda,1}(\lambda_i;\ \alpha_\lambda, \beta_1) + (1 - z) f_{\lambda,2}(\lambda_i;\ \alpha_\lambda, \beta_2)$$

where Z is a binary random variable with $\Pr(Z = 1) = \pi$ and $\Pr(Z = 0) = 1 - \pi$.

B. Write down the full conditional distribution of Z_i, the value of Z for patient i.

C. Write suitable BUGS (or BRugs or WinBUGS or OpenBUGS) code for the model. The model is completed with the following prior distribution for the parameters which are independent apart from the constraint on β_1, β_2.

$$\begin{aligned}
\alpha_j &\sim \text{Ga}(2,\ 0.5) \quad (j = 1, 2, 3), \\
\pi &\sim \text{Beta}(1,\ 2), \\
\alpha_\lambda &\sim \text{Ga}(2,\ 1), \\
\beta_1 &\sim \text{Ga}(2,\ 0.2), \\
\beta_2 &\sim \text{Ga}(2,\ 0.5).
\end{aligned}$$

Question 8 continued on next page

D. Explain briefly why it is advisable to impose a constraint such as $\beta_2 < \beta_1$.

(b) In a medical experiment, five measurements of the same variable are to be made on each of n patients. The measurements are to be made at times $t = 1, 2, 3, 4, 5$. Let the measurement on patient i at time t be $Y_{i,t}$.

We have a model as follows. Given the values of the unknown quantities μ_i and τ_y the conditional distribution of $Y_{i,t}$ is

$$Y_{i,t} \mid \mu_i, \tau_y \sim N(\mu_i,\ \tau_y^{-1})$$

independently for $i = 1, \ldots, n$ and $t = 1, \ldots, 5$. Given the values of the unknown quantities μ and τ_μ the distribution of μ_i is

$$\mu_i \mid \mu, \tau_\mu \sim N(\mu,\ \tau_\mu^{-1})$$

independently for $i = 1, \ldots, n$. The prior distribution for μ is

$$\mu \sim N(m, \tau_0^{-1}).$$

(i) Find the following conditional variance and covariances, given the values of $\tau_0,\ \tau_\mu$ and τ_y.

A. The conditional variance of $Y_{i,t}$.

B. The conditional covariance of $Y_{i,t}$ and $Y_{i,s}$ where $s \neq t$.

C. The conditional covariance of $Y_{i,t}$ and $Y_{j,s}$ where $i \neq j$.

(ii) the values $m = 45.0,\ \tau_0 = 0.002$ are used but τ_μ and τ_y are given independent gamma prior distributions with

$$\begin{aligned} \tau_\mu &\sim \text{Ga}(1,\ 0.01), \\ \tau_y &\sim \text{Ga}(1,\ 0.01). \end{aligned}$$

Write suitable BUGS (or BRugs or WinBUGS or OpenBUGS) code to specify the model.

(iii) We now wish to modify the model to allow the value of the variable to change over time within a patient. For $t = 1, \ldots, 5$ we replace μ_i with $\mu_{i,t}$ and, given $\mu_{i,t},\ \tau_y$, we have

$$Y_{i,t} \mid \mu_{i,t},\ \tau_y \sim N(\mu_{i,t},\ \tau_y^{-1}).$$

Question 8 continued on next page

Then

$$\mu_{i,1} = \mu_i$$

and, for $t = 2, \ldots, 5$, given $\mu_{i,t-1},\ \delta_i,\ \tau_x$,

$$\mu_{i,t} \mid \mu_{i,t-1}, \delta_i, \tau_x \sim N(\mu_{i,t-1} + \delta_i,\ \tau_x^{-1}).$$

Given δ and τ_δ,

$$\delta_i \mid \delta, \tau_\delta \sim N(\delta,\ \tau_\delta^{-1})$$

independently for $i = 1, \ldots, n$. The prior distribution for δ is

$$\delta \sim N(d,\ \tau_1^{-1}).$$

The values of $d = 0$ and $\tau_1 = 0.01$ are used but τ_δ and τ_x are given independent gamma prior distributions with

$$\begin{aligned} \tau_\delta &\sim \text{Ga}(1,\ 0.002), \\ \tau_x &\sim \text{Ga}(1,\ 0.002). \end{aligned}$$

A. Write modified BUGS (or BRugs or WinBUGS or OpenBUGS) code to specify this new model.

B. Find the conditional variance of $Y_{i,5}$ given the values of $\tau_0, \tau_\mu, \tau_y, \tau_1, \tau_\delta, \tau_x$.

C. Find the conditional covariance of $Y_{i,2}$ and $Y_{i,5}$ given the values of $\tau_0, \tau_\mu, \tau_y, \tau_1, \tau_\delta, \tau_x$.

D. If these are measurements taken over time on patients with a chronic condition and expected to be in a steady state, do you think this model is reasonable? Comment.

[**Total: 25 marks**]

THE END