RESTRICTED OPEN BOOK EXAMINATION
Data provided: Neave's Statistical Tables

The University Of Sheffield.

MAS364

SCHOOL OF MATHEMATICS AND STATISTICS **Autumn Semester 2012–13**

Bayesian Statistics **2 hours**

Restricted Open Book Examination.
Candidates may bring to the examination lecture notes and associated lecture material (but no textbooks) plus a calculator which conforms to University regulations.
Marks will be awarded for your best ***three*** *answers. Total marks 84.*

Please leave this exam paper on your desk
Do not remove it from the hall

Registration number from U-Card (9 digits)
to be completed by student

Blank

1 The true length (in millimetres, mm) of an important engineering component is denoted by θ. It is possible to take measurements x_j which are conditionally independent, given θ, such that $x_j \sim N(\theta, \sigma^2)$ with known standard deviation (in mm) $\sigma = 0.02$.

(i) Given a prior distribution $N(m, v)$ for θ, state carefully the posterior distribution for θ after n measurements $x_1, \ldots x_n$ as above. (You do not need to derive the result). ***(4 marks)***

(ii) If measurements $x_1, \ldots x_5$ are taken, giving values

$$62.607, 62.594, 62.582, 62.610, 62.602,$$

calculate the posterior distribution for θ:

(a) if the prior distribution for θ is $N(62.8, 0.04^2)$, based on past information for similar components, and

(b) if a limiting 'flat' prior is used for θ. ***(4 marks)***

(iii) Calculate the posterior probability that $\theta < 62.6$ based on each posterior distribution from (ii). ***(4 marks)***

(iv) An engineer who believes that the information in (ii)(a) is not directly relevant wants to formulate a prior distribution based on her own experience. Before seeing the data, she believes that θ is likely to be close to 62.5, and has probability 0.9 of being between 62.3 and 62.7. Formulate a suitable prior to represent her views, and give one example of a probability that could be 'fed back' to her to check the appropriateness of the prior. ***(6 marks)***

(v) The following Winbugs model represents a generalisation of the situation described above. (Variables not defined in the model can be assumed to be given fixed numerical values.) Explain statistically what the model does, and draw a Directed Acyclic Graph to represent it.

```
model
{
p <- 1/v
theta ~ dnorm(m,p)
sigma2 <- 1/tau
tau ~ dgamma(a,b)
for (j in 1:n)
{
  x[j] ~ dnorm(theta,tau)
}
i <- step(theta-c)
}
```

(10 marks)

2 A scientist is interested in estimating the decay rate θ of a rare isotope, relative to a known rate; she designs a series of experiments giving observations $X_1, X_2, \ldots$ with $X_i \sim Poisson(\theta)$, so that

$$P(X_i = x|\theta) = \frac{\theta^x \exp(-\theta)}{x!}$$

with X_i and X_j conditionally independent given θ, for $i \neq j$. Before carrying out the experiments, she considers her prior beliefs about θ and decides that they can represented by an $Exponential(0.4)$ distribution. She wants to update her beliefs using data x and give a point estimate for θ, but is unsure of the appropriate form for her loss function.

Recall that if θ has the $Gamma(a, b)$ distribution then its probability density function is

$$f(\theta) = \frac{b^a \theta^{a-1} \exp(-b\theta)}{\Gamma(a)}$$

for $\theta > 0$, and it has mean a/b and variance a/b^2; and that the $Exponential(b)$ distribution is the special case with a=1, and has cumulative distribution function

$$F(\theta) = 1 - \exp(-b\theta)$$

for $\theta > 0$.

(i) If she observes $X_1, \ldots, X_{24}$ with $\sum_{i=1}^{20} x_i = 44$, show that her posterior distribution for θ is $Gamma(a^\star, b^\star)$ where $a^\star = 45, b^\star = 24.4$. ***(6 marks)***

(ii) Under the assumption of a quadratic loss function

$$L_Q(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$$

what would be the scientist's point estimate, and associated expected loss,

(a) given her prior distribution and

(b) given her posterior distribution? ***(6 marks)***

(iii) Under the assumption of an absolute loss function

$$L_A(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$$

what would be the appropriate point estimate given the scientist's prior distribution? If her posterior distribution can be approximated by a normal distribution, what is her posterior point estimate? ***(6 marks)***

(iv) Under the assumption of a zero-one loss function

$$L_Z(\theta, \hat{\theta}) = \begin{cases} 0 & |\theta - \hat{\theta}| \leq c \\ 1 & |\theta - \hat{\theta}| > c, \end{cases}$$

what would be the scientist's prior point estimate (as a function of the constant c) and the associated expected loss? If c is small, derive her posterior point estimate (**without** assuming normality) and, by taking the posterior density to be approximately constant close to the estimate, obtain an approximate expression for the expected loss. ***(10 marks)***

3 (i) (a) Define $X \sim Binomial(n, \theta)$ and $Y \sim Binomial(m, \theta)$ to be conditionally independent, conditional on the value of θ, and let θ have a $Beta(a, b)$ prior distribution. Write down the posterior distribution for θ given X, and the predictive distribution for Y given X. (You do not need to derive these results.) ***(4 marks)***

(b) A geneticist is unsure about the proportion θ of individuals in a (large) population who carry a particular gene. His prior distribution for θ, based on experience of other similar genes, is $Beta(1/2, 1/2)$. He tests three randomly sampled individuals and finds that none of them carry the gene. Calculate his predictive probability that the next individual sampled carries the gene, and his predictive probability that the next two individuals sampled both carry the gene. ***(6 marks)***

(ii) Observations X and Y each have the distribution $N(0, \theta)$ and are conditionally independent given θ. The parameter θ has prior distribution given by the inverse gamma distribution with parameters d and a, written $IG(d, a)$, and so has density

$$f(\theta) = \frac{a^d \theta^{-(d+1)}}{\Gamma(d)} \exp\left(-\frac{a}{\theta}\right),$$

for $\theta > 0$.

(a) Show that the posterior distribution for θ given X is also of the form $IG(D, A)$ and give expressions for D and A. ***(6 marks)***

(b) Derive the predictive distribution for Y given X and show that, for large y,

$$f(y|x) \propto y^{-(2D+1)}.$$

(10 marks)

Comment briefly on the shape of the predictive distribution, compared with the distribution of $Y|\theta$. ***(2 marks)***

4 (i) If θ has a t distribution with location μ, precision λ and ν degrees of freedom, then

$$f(\theta) \propto \left(1 + \frac{\lambda(\theta - \mu)^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

A sample from the above distribution can be generated using the Metropolis-Hastings algorithm, with proposal distribution $N(\theta_t, \sigma^2)$ given current value θ_t.

(a) Derive the form of the acceptance probability for this algorithm. ***(7 marks)***

(b) Explain how to obtain a realisation of the Markov chain $\theta_1, \theta_2, \ldots$ given series of random numbers from the standard normal distribution, $Z_1, Z_2, \ldots$ and the standard uniform distribution, $U_1, U_2, \ldots$. ***(5 marks)***

(c) In practice, how might σ^2 be chosen? ***(2 marks)***

(ii) The Winbugs code below represents a hierarchical model for the weights of different adult voles in n separate populations on n isolated islands.

```
model
{
k~dpois(kappa)
df<-k+1
for (i in 1:n)
{
mean[i] ~ dt(overall.mean,overall.prec,df)
for (j in 1:m[i])
{
x[i,j] ~ dnorm(mean[i],prec)
}
}
}
```

Fixed numerical values are available for all variables used but not defined in the model, and for the observations `x[,]`.

Explain in statistical terms the structure of the model, in particular the interpretation of the variables `df`, `mean[i]` and `prec`. ***(11 marks)***

The following commands are added to the above model.

```
mnew ~ dt(overall.mean,overall.prec,df)
y ~ dnorm(mnew,prec)
```

What is the interpretation of the variable `y`? ***(3 marks)***

End of Question Paper