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

MAS364

The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS — **Autumn Semester 2010–11**

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 In an investigation into the size of the errors produced by a new measurement instrument, n measurements are taken of the mass of a standard sample of mass 1000 grammes. The values obtained (in grammes), $x_1, \ldots, x_n$, can be modelled as exchangeable normal random variables with known mean $\mu = 1000$ and unknown variance σ^2.

(i) Show that the Inverse Gamma family of distributions is a conjugate family of priors for σ^2 in this situation, and that the parameters are updated as follows:

$$\begin{aligned} d &\to d + n/2, \\ a &\to a + \sum (x_i - \mu)^2/2. \end{aligned}$$

(8 marks)

Recall that σ^2 has an inverse gamma distribution with parameters d and a, written $IG(d, a)$, if it has density

$$f(\sigma^2) = \frac{a^d (\sigma^2)^{-(d+1)}}{\Gamma(d)} \exp\left(-\frac{a}{\sigma^2}\right),$$

for $\sigma^2 > 0$, and that provided $d > 2$, then

$$E(\sigma^2) = \frac{a}{d-1}$$

and

$$Var(\sigma^2) = \frac{a^2}{(d-1)^2(d-2)}.$$

(ii) Before taking any measurements, a scientist has priors beliefs summarised by $E(\sigma^2) = 0.01, Var(\sigma^2) = (0.005)^2$. Obtain a suitable conjugate prior representing these beliefs. ***(6 marks)***

(iii) An initial measurement is taken with the new instrument, and gives $x_1 = 1000.01$. What is the posterior distribution for σ^2, for the scientist in part (ii)? ***(3 marks)***

(iv) A further 9 measurements are taken, and the whole series, including the measurement from part (iii), can be summarised by $\sum_{i=1}^{10} (x_i - 1000)^2 = 0.025$. Given these measurements, calculate the posterior mean and variance for σ^2, for the scientist in part (ii). ***(6 marks)***

(v) The new instrument is one of a collection, all manufactured in a similar way. The properties of these instruments are considered to be exchangeable, and similar kinds of test data are available for each of them. Explain briefly how a hierarchical model could be used to describe the relationships between the properties of the instruments and the data. ***(5 marks)***

2 (i) 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 (a) the posterior distribution for θ given X, and (b) the predictive distribution for Y given X. (You do not need to *derive* these results.) ***(3 marks)***

(ii) Two gamblers are interested in the long-run probability of getting heads, θ, when a particular coin is tossed repeatedly. They agree that their beliefs are symmetric around $\theta = 1/2$. Gambler 1 has prior variance $1/20$ for θ, and Gambler 2 has prior variance $1/100$. In each case, obtain suitable Beta distributions to represent these prior beliefs. ***(5 marks)***

(iii) A series of 8 tosses of the coin is observed (by both gamblers), and produces 4 heads and 4 tails. For each gambler, obtain the posterior distribution for θ and the predictive probabilities that (a) the next toss of the coin gives a head, and (b) the next 4 tosses of the coin all give heads. Comment briefly on how the gamblers' predictions are influenced by their priors. ***(16 marks)***

(iv) Without further calculation, explain what would happen to these predictive probabilities after the gamblers had seen a large number of tosses of the coin, of which a proportion h were heads. ***(4 marks)***

3 Prior uncertainty about a population mean μ is described by a $N(50, 49)$ distribution. Four observations, $x_1, \ldots, x_4$, conditionally independent given μ, are available with $x_i|\mu \sim N(\mu, 25)$, with $x_1 = 55$, $x_2 = 41$, $x_3 = 46$ and $x_4 = 42$.

(i) Suppose that μ is to be estimated by d under the zero-one loss function $L(d, \mu) = 0$ if $|d-\mu| < 2$, and $L(d, \mu) = 1$ otherwise, using prior information only. If d is chosen to be the prior mode, calculate the expected loss using this estimate. ***(5 marks)***

(ii) State the posterior distribution of μ given $x = \{x_1, \ldots, x_4\}$, calculate the posterior mean and variance, and give a 95% highest posterior density interval for μ. Given this posterior distribution, state the optimum estimate of μ under quadratic loss, and give the expected loss using this estimate. ***(10 marks)***

(iii) A fifth observation x_5 is available with $x_5|\mu, \tau^2 \sim N(\mu, \tau^2)$. Derive the largest value of τ^2 such that

$$Var(\mu|x_1, \ldots, x_5) \leq 0.5 \times Var(\mu|x_1, \ldots, x_4).$$

(5 marks)

(iv) Comment briefly on how your answers to (i), (ii) and (iii) would be affected if the prior distribution given were replaced by a weak prior for μ. ***(8 marks)***

4 An investigation into the numbers of cases of a certain rare disease in 2009, in various towns and cities in the UK, gave the following figures.

City/town	Population (thousands)	Number of cases
Wolverhampton	251	20
Derby	229	13
Norwich	174	11
Oxford	143	7
St Helens	103	6
Crawley	101	5

The Winbugs code below defines a possible model for the occurrence of the disease.

```
model
{
for (j in 1:6)
{
lambda[j] <- alpha[j] * theta[j]
x[j] ~ dpois(lambda[j])
theta[j] ~ dgamma(psi,rho)
}
psi ~ dgamma(0.001,0.001)
rho ~ dgamma(0.001,0.001)
}

list(alpha=c(251,229,174,143,103,101),x=c(20,13,11,7,6,5))
```

(i) Draw a Directed Acyclic Graph (DAG) to illustrate the model represented by the above code. ***(7 marks)***

(ii) Explain the structure and assumptions of the above model, and the meaning of the variables, in a form suitable for a Bayesian statistician who is not familiar with Winbugs. ***(8 marks)***

(iii) The quantities

```
mu <- psi/rho
cv <- 1/sqrt(psi)
```

(in Winbugs notation) represent the mean and the coefficient of variation of a particular distribution. Explain their interpretations. ***(5 marks)***

(iv) For the city of Stoke-on-Trent, population 259,000, the number of cases in 2009 is unknown. Write down the additional Winbugs code necessary to sample from the posterior distribution for this quantity, explaining any additional assumptions you are making. ***(8 marks)***

End of Question Paper