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

The University Of Sheffield.

MAS364

SCHOOL OF MATHEMATICS AND STATISTICS **Autumn Semester 2011–12**

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 (i) The Geometric distribution is defined by

$$P(x|\theta) = (1-\theta)^x\theta, \; x = 0, 1, 2, \ldots, \;\; 0 < \theta < 1.$$

Show that the Beta family of distributions is conjugate for θ. ***(4 marks)***

(ii) Given a Beta(a, b) prior for θ, obtain the posterior probability density for θ given $x_1, \ldots, x_n$, if $x_j \sim$ Geometric (θ) and the x_js are independent conditional on θ. ***(4 marks)***

(iii) An engineer carries out several inspections of a production line, which produces items that may be either working or faulty. In each case, he observes a series of items until he sees one that is faulty. In three inspections, the numbers of working items he sees (before the first faulty one) are 4, 8, 9 respectively. Using the results of (ii), and explaining your notation and assumptions, obtain his posterior mean and variance for the proportion of items that are faulty, given that his prior beliefs are uniform on (0,1). ***(8 marks)***

(iv) Given the model, prior and observations in (ii), derive the predictive distribution for a future observation Y of the same form as x_j. (You may leave your answer in terms of Beta functions.) ***(6 marks)***

(v) The following table shows some probabilities from the predictive distribution in (iv), based on the data from (iii), and also from a simpler approach to prediction involving 'plugging in' $\hat{\theta}$, the maximum likelihood estimate of θ. Describe the main similarities and differences between these distributions, and explain why they arise.

y	0	1	2	3	4	5	6	7	8
$p(y\|\mathbf{x})$	0.154	0.125	0.103	0.085	0.071	0.06	0.05	0.043	0.036
$p(y\|\hat{\theta})$	0.143	0.122	0.105	0.09	0.077	0.066	0.057	0.049	0.042
y	9	10	11	12	13	14	15	16	17
$p(y\|\mathbf{x})$	0.031	0.027	0.023	0.02	0.018	0.015	0.014	0.012	0.011
$p(y\|\hat{\theta})$	0.036	0.031	0.026	0.022	0.019	0.017	0.014	0.012	0.010
y	18	19	20	21	22	23	24	25	26
$p(y\|\mathbf{x})$	0.009	0.008	0.007	0.007	0.006	0.005	0.005	0.004	0.004
$p(y\|\hat{\theta})$	0.009	0.008	0.007	0.006	0.005	0.004	0.004	0.003	0.003

(6 marks)

2 To assess a patient's health, the level of an important hormone is measured on several occasions. On a logarithmic scale, the measurements x_i can be regarded as normally distributed, with known variance σ^2, and are separated in time enough to be conditionally independent given the patient's true hormone level θ. The measurements are also adjusted for the patient's age, sex, mass etc, so that zero represents a nominal healthy level.

(i) A doctor wants to formalise her prior beliefs before considering measurements from a routine health check for a new patient. She regards adjusted hormone levels of above 0.3 or below -0.3 as unhealthy; in her experience, new patients have an average adjusted level of zero, and 25% of them have unhealthy levels. Find a suitable normal distribution to represent her prior beliefs about the true level for the new patient, θ. ***(4 marks)***

(ii) The doctor takes a first measurement on her new patient, with standard deviation $\sigma = 0.15$, and obtains a value $x = 0.2$. Calculate the doctor's posterior distribution for θ. What are her posterior probabilities that (a) the patient's true hormone level is negative; (b) it is in the healthy range? ***(7 marks)***

(iii) The doctor is considering a 'screening' approach, which would involve making a decision on a patient after their first measurement. The possible actions would be d_1 : declare the patient to be healthy, and d_2 : arrange for follow-up measurements. Let H represent the interval of 'healthy' parameter values $[-0.3, 0.3]$. Her loss function for the decision is then, for some constant $c \geq 1$,

$$L(d_1, \theta \in H) = 0;$$
$$L(d_1, \theta \notin H) = c;$$
$$L(d_2, \theta \in H) = 1;$$
$$L(d_2, \theta \notin H) = 0.$$

If the doctor assesses $P(\theta \in H|x) = p$, say, for a patient, derive the expected loss for each possible decision. If $c = 3$, what is the optimal decision for the patient in (ii)? ***(8 marks)***

(iv) Given the prior distribution in (i), $\sigma = 0.15$ as in (ii), and the loss function in (iii), how large would c have to be to make d_2 the optimal action for *all* patients, regardless of their initial measurement? (Hint: consider which x would lead to the maximum value of $P(\theta \in H|x)$.) ***(9 marks)***

3 A hierarchical model for observations $x_i, i = 1, \ldots, n$ on the log-concentration of a particular pollutant, on n different farms around the UK, can be represented as

$$\theta_i = \phi + \gamma_i$$
$$x_i = \theta_i + \delta_i$$

where

$$\phi \sim N(m, v)$$
$$\gamma_i \sim N(0, \rho^2)$$
$$\delta_i \sim N(0, \tau^2).$$

The quantities m and v define the prior distribution for ϕ, the mean level nationally, with $i = 1, \ldots, n$ indexing the different farms on which observations can be made, and θ_i representing the true level on farm i. The parameter τ^2 represents the variance of the measurement error, and ϕ, γ_i and δ_j are independent for all i, j.

(i) Explain why in general x_i and x_j are exchangeable but not independent, and give an example of an event A such that x_i and x_j are conditionally independent given A (for $i \neq j$). ***(6 marks)***

(ii) By writing $\epsilon_i = \gamma_i + \delta_i$, say, obtain the posterior distribution for ϕ given $x_i, i = 1, \ldots, n$. ***(6 marks)***

(iii) By considering $\theta_1|\phi$ and $x_1|\theta_1$, derive the full conditional distribution $p(\theta_1|\theta_2, \ldots, \theta_n, \phi, x_1, \ldots, x_n)$ for θ_1. ***(8 marks)***

(iv) In each of the following cases, comment briefly on the interpretation of the given limits and on their implications for learning about ϕ and about $\theta_i, i = 1, \ldots, n$ from $x_i, i = 1, \ldots, n$.

(a) $v \to \infty$.

(b) $\rho^2 \to 0$.

(c) $\rho^2 \to \infty$.

(d) $\tau^2 \to 0$. ***(8 marks)***

4 An ecologist is interested in how a population of animals changes over time. The species reproduces at most once a year, with adult females (those who are over a year old) producing at most one offspring. (All population counts below relate to females; males are ignored in this model!)

(i) As part of his initial data collection, the ecologist monitors six adult females; only one of them is 'successful' i.e. produces a female offspring. If the ecologist is initially completely uncertain about the rate of reproduction, give a suitable prior for θ, the probability that an adult female is successful in a given year, and obtain the corresponding posterior distribution. If the ecologist is to monitor *one* further adult female, what would be his predictive probability of her being successful? ***(6 marks)***

(ii) The following WinBUGS code represents a model for changes over a single year which can be used to learn about survival and reproduction rates, given values for the current numbers of adults and young, and the corresponding numbers for the previous year.

```
model
{
as <- 1
bs <- 1
survival.adult ~ dbeta(as,bs)
survival.young ~ dbeta(as,bs)
ar <- 2
br <- 6
reproduction ~ dbeta(ar,br)
adults.surviving ~ dbinom(survival.adult, adults.last)
young.surviving ~ dbinom(survival.young, young.last)
adults.now <- adults.surviving + young.surviving
young.now ~ dbinom(reproduction, adults.surviving)
}
```

Draw a Directed Acyclic Graph to represent this model. ***(8 marks)***

Explain briefly the structure of the model, in a form suitable for a Bayesian statistician who is not familiar with WinBUGS, and comment on the priors for the survival and reproduction rates. ***(8 marks)***

 Question 4 continued on next page

4 (continued)

(iii) An extension of the above model to multiple years is defined by the following WinBUGS code.

```
model
{
as ~ dgamma(1,1)
bs ~ dgamma(2,1)
ar ~ dgamma(2,1)
br ~ dgamma(2,1/3)
for (j in 1:n)
{
survival.adult[j] ~ dbeta(as,bs)
survival.young[j] ~ dbeta(as,bs)
reproduction[j] ~ dbeta(ar,br)
adults.surviving ~ dbinom(survival.adult[j], adults[j])
young.surviving ~ dbinom(survival.young[j], young[j])
adults[j+1] <- adults.surviving + young.surviving
young[j+1] ~ dbinom(reproduction[j], adults.surviving)
}
}
```

Describe briefly the main similarities and differences from the model in part (ii). Explain what assumptions are being made about the structure of survival and reproduction rates across ages and years. ***(6 marks)***

End of Question Paper