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

MAS6062

The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2011–2012**

MAS6062 Bayesian Statistics and Clinical Trials **3 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 **five** answers. Total marks 100.*

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. ***(2 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). ***(6 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.) ***(4 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

(4 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. ***(3 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? ***(5 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)? ***(6 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 $P(\theta \in H|x)$.) ***(6 marks)***

3 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? ***(4 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. ***(6 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. ***(6 marks)***

3 (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[j] ~ dbinom(survival.adult[j], adults[j])
young.surviving[j] ~ dbinom(survival.young[j], young[j])
adults[j+1] <- adults.surviving[j] + young.surviving[j]
young[j+1] ~ dbinom(reproduction[j], adults.surviving[j])
}
}
```

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. ***(4 marks)***

4. A cluster randomised controlled trial was undertaken in a primary care setting designed to evaluate the effectiveness of new nurse-led intervention compared to usual care in treating depression. Patients were followed up for 12 months.

There were 100 clusters (GP practices) with 1745 (in 63 clusters) patients allocated to intervention and 914 patients to (in 37 clusters) usual care.

The primary outcome was the SF-36 Mental Component Summary (MCS) scale measured at six-months post-randomisation. Table 4.1 gives the regression coefficients from fitting a marginal generalised linear model, with coefficients estimated using generalized estimating equations with robust standard errors and an exchangeable autocorrelation matrix to analyse the 6-month MCS outcome and allow for the clustered nature of the data.

The estimated Intra cluster correlation (ICC) from the model was 0.016.

Table 4.1: Estimated regression coefficients from a marginal model to show the effect of group on outcome, 6-month SF-36 Mental Component Summary (MCS) score, after adjustment for baseline

Outcome: 6-month MCS score	**B**	**Semi-robust SE(B)**	**95% CI Lower**	**Upper**
Baseline MCS	0.4	0.1		
Group (0 = Control, 1 = Intervention)	1.4	0.5		
Constant	29.6	1.4		

Note The SF-36 Mental Component Summary Scale (MCS) is scored with a higher score indicating better mental health.

a) What is meant by an exchangeable autocorrelation matrix? **(1mark)**

b) Define the ICC in words and explain how it affects the design and analysis of cluster randomised trials **(2 marks).**

c) Comment on the estimated ICC. Does it suggest that within a cluster individual MCS scores are strongly correlated? **(2 marks)**

4. (continued)

d) Calculate an approximate 95% confidence interval for the treatment group coefficient for the model in Table 4.1. Comment on your results **(3 Marks)**

e) Assume the Minimum Clinically Important Difference (MCID) for the SF-36 MCS is 5 points. Comment on the size of the difference in mean MCS scores between the two groups. **(1 mark)**

Table 4.2 shows the mean Quality adjusted life years (QALYs) and mean costs of treatment over 12 months for depressed patients in the nurse-led intervention and control groups respectively.

f) Calculate the mean difference in QALYs between the Intervention and control groups and its associated 95% confidence interval. Comment on the results. **(2 marks)**

g) Calculate the incremental cost-effectiveness ratio (ICER) ,cost per QALY,to compare the cost-effectiveness of the nurse led intervention with usual care for treatment of patients with depression. Comment on the results **(2 mark).**

h) Assume that service providers would be willing to pay no more than £20,000 per QALY.. Do you think the nurse led intervention for treating depression is cost effective compared to usual treatment? **(1 mark)**

Table 4.2. Mean QALYs and costs for depressed patients over the 12 month follow-up period by treatment group

	Control (Usual care)			**Nurse led Intervention**		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**
QALY	914	0.62	0.46	1745	0.50	0.47
Cost (£)	914	11,345	5,673	1745	9,023	6,665

4. (continued)

A pre-specified analysis was to look at young women in the study, aged 25 years or less, as it was believed a larger effect may be observed in this population. A simple summary table is given in Table 4.3 of the number of young women (aged 25 or less) whose 6-month MCS score improved by 10 or more points compared to their baseline value.

Table 4.3 Numbers of young women whose 6-month SF-36 Mental Component Summary (MCS) score improves by 10 or more (0 = Control, 1 = Intervention) compared to baseline

	Treatment Group	
Change from baseline in 6-month MCS score	1	0
<10	82	51
≥10	104	43
Total	186	94
$\tilde{m}$	3.6	1.8

i) Find the chi-squared value unadjusted for clustering. Comment on the results **(3 marks)**

j) Find the chi-squared value adjusted for clustering . Note the estimated ICC is 0.010. Comment on the results. **(3 marks)**

5. A trial of a new drug for the treatment of stroke against placebo is being planned. The primary efficacy outcome is the improvement in patients' quality of life (QoL) at 6 months post-randomisation as measured by the StrokeQoL. The StrokeQoL is a patient reported outcome measure scored on a 0 (poor) to 100 (good) quality of life scale.

a) If a mean improvement of 5 units, on the StrokeQoL, is required for the trial with an anticipated population standard deviation of 24 units what is the required sample size per arm (for 90% power and two sided significance levels of 5%)? **(3 marks)**

b) If 12.5% of patients are expected to not have evaluable information at 6 months what sample size would be required to ensure sufficient numbers of patients with data for analysis? **(1 mark)**

c) If the trial is planned as a cluster randomised trial with GP practices randomised rather than individual stroke patients: with an intra-class correlation of 0.03 and a cluster size of 25 patients what would be the evaluable sample size? Comment on your results **(2 marks).**

From now on for the calculations the assumption is the study is an individually randomised study.

d) The investigator reports that due to budget constraints they can only have 80% power. What impact does this constraint have on the required sample size again assuming 12.5% of patients are not evaluable at 6 months? **(3 mark)**

e) There is to be safety monitoring of the trial. Efficacy will be also assessed with O'Brien-Fleming rules for the efficacy assessments at the interim analyses. There will be 2 interim analyses after a third and two-thirds of patients have reached the primary endpoint. Comment on the effect these efficacy analyses will have on the evaluable sample size **(2 marks)**

f) The trial is expected to recruit 100 patients per month. How would this impact on the planning of the interim analyses? **(1 mark)**

g) The randomisation is 1-to-1 and a sample size re-estimation is undertaken at [or after] the first interim analysis. The total variance observed at the interim analysis is 650 (units) estimated from the two groups combined. What is the re-estimate of the evaluable sample size (assume 80% power)? **(3 marks)**

h) A futility assessment was being planned for the second interim analysis using the observed results and the anticipated response under the alternative hypothesis. What is the implication for the trial if the predictive power is low (say less than 20%)? **(1 mark)**

5. (continued)

i) A key secondary outcome is the proportion of patients who have responded well to treatment as assessed by a physician. It is anticipated that 55% of patients on the placebo arm will respond while 65% on the drug are anticipated to respond. For a two sided level of statistical significance at 5% and 80% power what is the sample size? If you have the evaluable sample size estimated from part (d); would the study have sufficient power for this key secondary? **(4 marks)**

6. The compartmental pharmacokinetic model for a single person to assess the concentration c(t) at a given time t is given as

$$c(t) = 6\left(e^{-0.05t} - e^{-1.2t}\right)$$

a. Estimate AUC0-∞ (units) from this model. **[2 marks]**

b. What is the estimate of the elimination rate half-life in hours? **(1 mark)**

c. In designing a cross-over study with a pharmacokinetic primary endpoint what would be the minimum time for washout between periods? **(1 mark)**

d. What is AUC0-12 from the model? **(2 marks)**

e. The dosing interval is to be 12 hours for repeat dosing. What is the predicted accumulation ratio? **(1 mark)**

f. What is the anticipated area under the curve within a dosing interval (AUC0-τ) at steady state? **(1 mark)**

g. If Tmax is at 3 hours from the model what is Cmax estimated to be after a single dose and at steady state? **(2 marks).**

A dose proportionality study was undertaken to assess the dose proportionality of the drug being developed. The results of the analysis of AUC (units) and logDose are shown in Table 6.1. Both AUC and dose were log transformed for the analysis.

Table 6.1. Analysis of logAUC against logDose for the assessment of dose proportionality using the Power Method

	b	**SE(b)**	**95% Confidence Interval**
logDose	1.05	0.04	(0.97 to 1.13)

h. Estimate the AUC ratio, per doubling per doubling of dose. Comment on the results. (**2 marks**)

i. Estimate the AUC dose normalised ratio, per doubling of dose. Comment on the results. (**2 marks**)

j. Is there evidence to suggest from AUC that the pharmacokinetics are dose proportional? (**1 mark**)

6. (continued)

The study where the results came from was designed to assess three doses (25mg, 50mg and 100mg) with placebo in four treatment periods.

k. Construct a Williams square of sequences that the subjects could be randomised to. **(2 marks)**

l. Twenty subjects are expected to complete the trial. For logAUC the within subject standard deviation on the log scale is expected to be 0.35. On the log scale what would be the precision (the half width of a 95% confidence interval) for estimates of a given dose compared to placebo **(3 marks)**

Note for this calculation Z statistics or t-statistics could be used.

End of Question Paper