Data provided: Neave's Tables

MAS6062

The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2010–2011**

MAS6062 Bayesian Methods 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 In an investigation into the accuracy of 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:

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

(6 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. ***(4 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)? ***(2 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). ***(4 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. ***(4 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.) ***(2 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 is $1/20$ for θ, and Gambler 2 has prior variance $1/100$. In each case, obtain suitable Beta distributions to represent these prior beliefs. ***(4 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. ***(11 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. ***(3 marks)***

3 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. ***(5 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. ***(6 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. ***(4 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. ***(5 marks)***

4. Eight randomised parallel group trials have been undertaken to compare a new treatment against placebo in patients who have suffered a myocardial infarction. The primary efficacy response is binary and is the proportion of patients who died within a year. The results from eight trials are given below

	Test			Placebo		
Study	**Deaths**	**Total**	**Proportion**	**Deaths**	**Total**	**Proportion**
1	20	80	0.250	15	85	0.176
2	70	375	0.187	100	355	0.282
3	25	220	0.114	25	210	0.119
4	20	165	0.121	15	160	0.094
5	15	100	0.150	30	105	0.286
6	20	265	0.075	25	255	0.098
7	45	300	0.150	45	295	0.153
8	20	155	0.129	30	160	0.188

The estimate of the treatment effect is a probability difference and the initial calculations for a fixed effects meta analysis are given below

	Proportion		$\hat{\theta}_i$	w_i	w_i^2	$\hat{\theta}_i w_i$	$\hat{\theta}_i^2 w_i$
Study	**Test**	**Placebo**					
1	0.250	0.176	0.074	247	60861	18.1	1.33
2	0.187	0.282	-0.095	1026	1052299	-97.5	9.26
3	0.114	0.119	-0.005	1045	1091339	-5.7	0.03
4	0.121	0.094	0.027	850	722366	23.3	0.64
5	0.150	0.286	-0.136	311	96529	-42.2	5.72
6	0.075	0.098	-0.023	1639	2686752	-37.0	0.83
7	0.150	0.153	-0.003	1158	1342032	-2.9	0.01
8	0.129	0.188	-0.058	596	355493	-34.9	2.04
Total				**6872**	**7407671**	**-178.8**	**19.86**

a. Using the probability difference as the measure of treatment difference calculate the fixed effects estimate of the treatment difference and its associated 95% confidence interval. Comment on the results. **(3 marks)**

b. Undertake a test for evidence of a treatment difference to assess statistical significance. Comment on the results **(2 marks)**

c. Calculate the inconsistency statistic, I^2, and undertake a test of heterogeneity between the trials. Comment on the results **(3 marks)**

d. Calculate the between study variance, τ^2, for these data **(1 mark)**

The initial calculations for a random effects analysis are given below

	Proportion				
Study	**Treated**	**Placebo**	$\hat{\theta}_i$	w_i^*	$\hat{\theta}_i w_i^*$
1	0.250	0.176	0.074	183	13.4
2	0.187	0.282	-0.095	418	-39.7
3	0.114	0.119	-0.005	421	-2.3
4	0.121	0.094	0.027	385	10.6
5	0.150	0.286	-0.136	216	-29.3
6	0.075	0.098	-0.023	493	-11.1
7	0.150	0.153	-0.003	438	-1.1
8	0.129	0.188	-0.058	323	-18.9
Total				2877	-78.4

e. How do the relative weights for the random and fixed effects analysis, taken from w_i and w_i^* in the estimate of overall treatment difference compare? (**1 mark)**

f. Using the probability difference as the measure of treatment difference calculate the random effects estimate of the treatment difference and its associated 95% confidence interval. Comment on the results (**3 marks)**

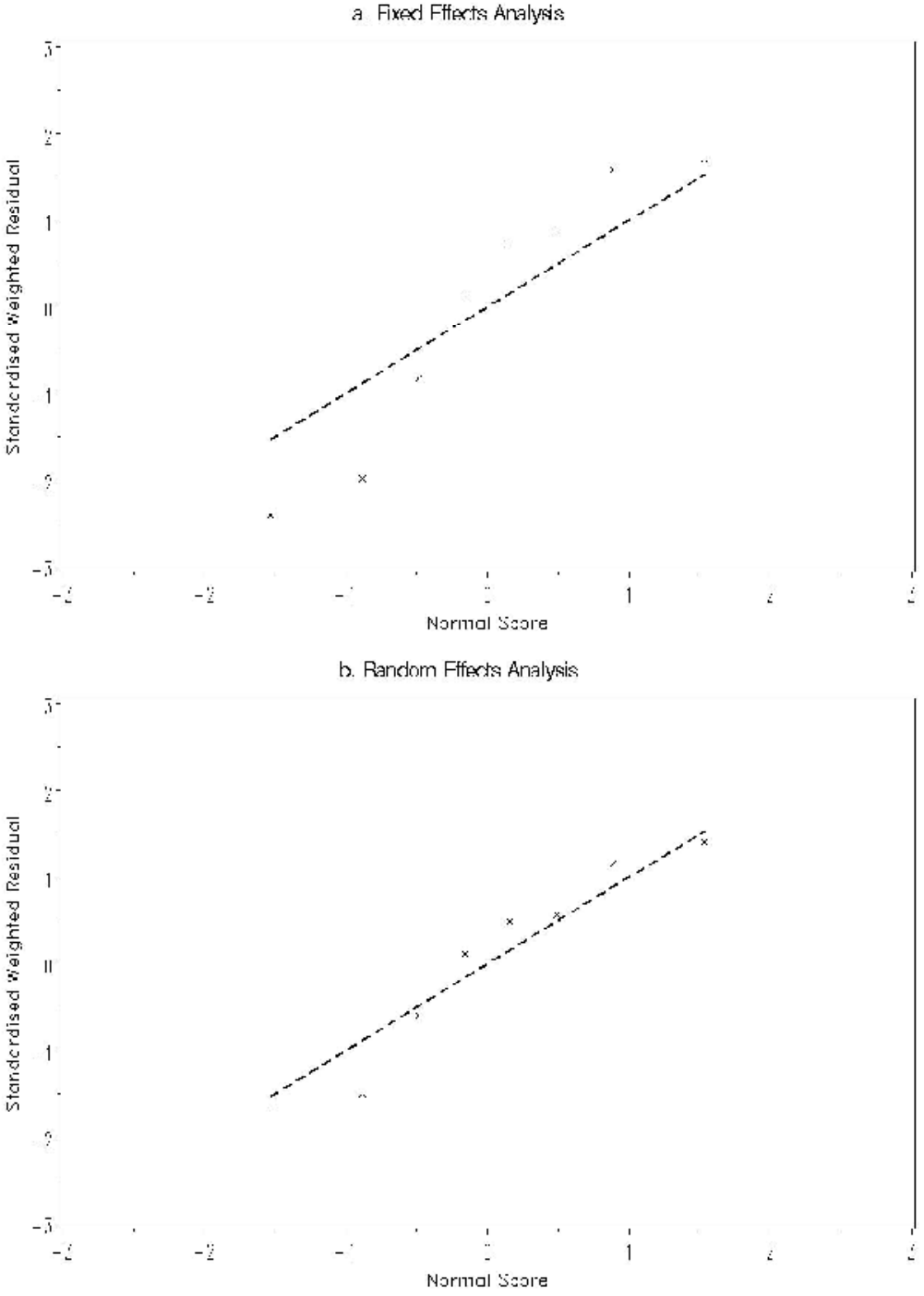
g. Undertake a test for evidence of a treatment difference to assess statistical significance. Comment on the results **(2 marks)**

Figure 4.1 gives the standardised weighted residuals for the fixed effects and random effects meta analyses respectively

h. Comment on the Normal probability plots of the standardised weighted residuals for the fixed and random effects analysis given in Figure 4.1 **(2 marks).**

i. Compare and comment on the fixed and random effects estimates and their associated confidence intervals **(3 marks)**

Figure 4.1. Normal probability plots of the standardised weighted residuals for the fixed effects and random effects meta analyses



5. A bioavailability study is being designed to investigate three new formulations against the standard formulation. The study will be a 4 period cross-over study in 16 subjects. After completion and analysis the best new formulation will be selected for further development. The sample size is based on feasibility with no formal power considerations. The pharmacokinetic parameters area under the curve to infinity, AUC0-inf (units), and maximum concentration, Cmax (units), will be used for the primary assessment (the units here are arbitrary)

a. Construct a Williams square of treatment sequences for subjects to be randomised to for the study. **(1 mark)**

b. The within subject standard deviation on the log scale for both AUC0-inf and Cmax is assumed to be 0.35. What is the precision on the log scale (the half width of a 95% confidence interval) for the comparison of a given new formulation to standard? **(3 marks)**

c. Suppose the design is now a balanced incomplete block design of 4 regimens in 3 periods. Construct the treatment sequences for subjects to be randomised to **(1 mark)**

d. If the sample size is increased to 24 subjects what would be the precision for AUC0-inf and Cmax on the log scale for the same standard deviation as from b. but with the design from c? **(3 marks)**

The results for the best performing formulation across all subjects is given below. AUC0-24 is given as the dosing schedule for the treatment as it is to be given once a day (every 24 hours).

	AUC0-inf	AUC0-24	Cmax
N	16	16	16
Geometric Mean	37.16	34.76	1.74
Standard Deviation of the Logs	0.72	0.70	0.80

e. What is the predicted accumulation ratio of the repeat dose pharmacokinetics at steady state to single dose pharmacokinetics based on these results? **(1 mark)**

f. What is the prediction for the area under the curve at steady state, AUC0-τ , and Cmax at steady state based on these data? **(1 mark)**

The new formulation may have a within subject standard deviation on the log scale as high as 0.50 for both AUC0-inf and Cmax. A bioequivalence study is being planned to demonstrate bioequivalence of the new formulation to the standard formulation. The bioequivalence criteria of (0.80, 1.25) are being used i.e. the 90% Confidence Intervals for both AUC and Cmax must be wholly contained within the interval (0.80, 1.25) to be able to declare bioequivalence. It is also assumed that there is no difference between the means for the test and reference formulations ie $\mu_T/\mu_R = 1$

g. Write down a null and alternative hypothesis for this study **(1 mark)**

h. What sample size would be required to demonstrate bioequivalence assuming: the study is designed with 90% power; within subject standard deviation on the log scale of 0.5 and a one tailed Type I error rate of 5% **(2 marks)**

i. If this is a replicate ABBA/BAAB cross-over study what would the sample size be? **(1 mark)**

A group sequential trial is being considered. The first cohort of subjects will be recruited based on a within subject standard deviation of 0.35. Depending of the results of the interim analysis a second cohort of patients will be recruited where the total sample size across both cohorts will be based on the within subject standard deviation of 0.50

j. If the Pocock approach was used to calculate the nominal levels of significance, what would the new nominal level of significance be for each analysis (interim and final), assuming a maximum of two analyses can be undertaken? **(2 marks)**

k. Assuming the design will be a two period cross-over study and using the nominal level of significance from j. what would the sample size be in cohort 1 (based on a within subject standard deviation of 0.35) and the maximum sample size across both cohorts (based on a within subject standard deviation of 0.50)? (**3 marks)**

l. What considerations would need to made with respect to the choice of endpoint(s) for the interim analysis? (**1 mark)**

6

Suppose we are planning a randomised controlled trial in new mothers identified as at high risk of postnatal depression (PND) to see if a new psychological intervention will lead to an improved quality of life (QoL) compared with usual care offered by health visitors (HV). Health visitors trained in detecting depressive symptoms weeks postnatally, using the Edinburgh Postnatal Depression Scale (EPDS) and clinical assessment, will offer 'at-risk women' a psychological intervention of one hour per week for 8 weeks.

The primary outcome will be the mean score on the EPDS at six months postnatally. The EPDS is a QoL questionnaire with a higher score indicting more symptoms. From a previous study of 220 women, the EPDS score at 6 months postnatally was 7.0 with a standard deviation of 6.0. Approximately 25% (55/220) scored 12 or more on the EPDS and were regard as greater risk of PND.

a. How many women would need to be individually randomised per group to have an 80% chance of detecting a difference of two or more points in mean 6-month EPDS scores between the intervention and control groups as statistically significant at the 5% (two-sided) level? **(3 marks)**

b. If 20% of patients drop out by 6-month follow-up, how many patients will we need to recruit and randomise? **(1 mark)**

c. If the power was increased to 90% for the same effect size as in a) what would the sample size be? **(3 marks)**

d. Comment on the sizes from a) and c) **(1 mark)**

e. Suppose the primary outcome is binary i.e. a 6-month EPDS score of 12 or more vs. less than 12. Let us assume that approximately 25% of new mothers in the control group will have a 6-month EPDS score of 12 or more and are at higher risk of PND. We anticipate that the new psychological intervention, offered by health visitors will reduce (improve) this proportion to 15% having a score of 12 or more in the intervention group. The primary analysis will be a chi-squared test. How many women would be needed to detect this difference in proportions with 80% power and 5% (two-sided) significance? **(3 marks)**

f. If the power was increased to 90% for the effect size as in e) what would the sample size be? **(3 marks)**

Suppose we believe that since the health visitors deliver the intervention they may behave differently with women under their care individually randomised to the control group. Therefore we believe a cluster randomised design is more appropriate, that is health visitors rather than individual women will be randomised to deliver the psychological intervention or usual care. Let us assume that health visitors will recruit about 10 women from their caseload during the trial period and the intra cluster correlation (ICC) for the EPDS outcome is 0.05.

g. Re-estimate the sample size from part a) to take into account the clustered nature of the design. How many HV (and women) would need to be randomised per group to have an 80% chance of detecting a difference of two or more points in mean 6-month EPDS scores between the intervention and control groups as statistically significant at the 5% (two-sided) level? **(2 mark)**

h. Re-estimate the sample size from part e), treating the EPDS as a binary outcome, to take into account the clustered design. How many HV (and women) would be needed to detect this difference in proportions (of 25% vs. 15%) with 80% power and 5% (two-sided) significance? **(2 marks)**

For the next two questions use the results from a) and e)

i. The new psychological intervention costs £1000 per patient. How does the study design impact on costs? (**1 mark)**

j. Suppose only 50% of eligible patients enter the individual randomised trial but all eligible patients enter the cluster randomised trial. It is anticipated that there will be 40 eligible patients per month who could enter the trial. How does the study design impact on study duration? **(1 mark)**

End of Question Paper