The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2013–2014**

Medical Statistics **2 hours**

RESTRICTED OPEN BOOK EXAMINATION

Candidates may bring to the examination lecture notes and associated lectures material (but no textbooks) plus a calculator that conforms to University regulations. All questions will be marked, but credit will be given for only the best ***THREE*** *answers. All questions carry equal marks. Total marks 60.*

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) A doctor intends to conduct a trial of a non-steroidal anti-inflammatory drug (NSAID) for rheumatoid arthritis against a placebo. She plans to measure effectiveness by use of a standard grip test. This will produce responses on a continuous scale. From a pilot study of untreated patients (whom she assumes will be similar to placebo-treated patients), she expects responses will be acceptably normally distributed, with a typical value of around 4.8 units (standard deviation 1.847 units). She hopes the NSAID will increase grip strength by at least 1 unit, but be no more or less variable. She will use a standard 5% significance level and wants to know the number of patients she should enrol in order to achieve a power of 80% when testing for that level of effectiveness. Specify what test she should use and why and perform the appropriate sample size calculation for her (note that selected standard normal percentage points are given below). ***(5 marks)***

p	qnorm(p)
0.025	-1.96
0.05	-1.64
0.1	-1.28
0.2	-0.84

(ii) It is anticipated that 20% of prospective particpants might fail the trial's inclusion/exclusion criteria.

(a) Explain what 'inclusion/exclusion criteria' are. ***(2 marks)***

(b) Amend your sample size calculation to take this additional information into account. ***(3 marks)***

(iii) As the patients are enrolled, they must be assigned to receive either NSAID or placebo. Identify three factors that should be considered when deciding how to make the allocation and explain briefly how they can be addressed by a suitable allocation scheme. You need not actually carry out the allocation you suggest. ***(3 marks)***

1 (continued)

(iv) As part of her background research into what is already known about the drug's effects, she discovers a journal article which presents the following summary results from a trial on a closely-related NSAID. The trial appears to have measured the drug's effectiveness in 8 different ways, giving rise to 8 'outcomes', some discrete, some continuous, some binary. For each, the result given appears to be the legitimate p-value from an appropriate test assessing whether or not there is a difference in the specified outcome for those patients under drug and placebo treatments.

Outcome	p-value
grip test	0.017
pain on movement	0.009
joint swelling	0.048
use of analgesics	0.047
time to fatigue	0.066
duration of morning stiffness	0.087
joint mobility index	0.022
number of tender joints	0.034

(a) What can the doctor safely conclude about the effects of the drug on rheumatoid arthritis? Explain your reasoning. ***(4 marks)***

(b) She reads the article more carefully and notes that the trial was powered and conducted with the grip test as its 'primary endpoint'. Should her conclusion be different? Explain your answer.
(3 marks)

2 A 2-treatment, 2-period crossover trial has been appropriately powered and conducted to compare a new drug A with a placebo B. Low values of the response are good. 8 patients receive the drugs in each order and the trial data are organized in the following manner:

Variable name	V1	V2	V3	V4
Group	1	1	2	2
Period	1	2	1	2
Treatment	A	B	B	A

The following table shows the data (adapted from Altman(1995)).

Variable name	V1	V2	V3	V4
	1.6	1.2	1.8	1.2
	2.6	1.9	1.2	0.4
	0.8	2.0	4.6	3.7
	3.7	4.4	5.1	5.8
	0.9	2.5	2.8	0.2
	4.1	3.6	2.9	1.8
	5.2	3.6	5.1	4.4
	1.0	1.1	4.6	1.4
	1.1	2.0	1.8	3.0
	3.0	2.7	4.4	0.4

Additional variables are constructed in R as follows:

```
> V5<-0.5*(V1+V2)
> V6<-0.5*(V3+V4)
> V7<-V1-V2
> V8<-V3-V4
> V9<--V8
```

(i) Write down the standard model for such a trial, explaining your notation clearly. *(3 marks)*

2 (continued)

(ii) Use three of the four tests reported in the following R output to provide an appropriate analysis of the trial. Explain your reasoning at each stage by reference to the model in (i). Ensure you check for:

(a) a possible carryover effect from Period 1 to Period 2 ***(4 marks)***

(b) a possible treatment effect ***(4 marks)***

(c) a possible period effect. ***(4 marks)***

(d) For the test given in the R output but which is NOT used above, explain the circumstances in which it might be a useful test to conduct. ***(1 mark)***

```
> t.test(V1,V3)

        Welch Two Sample t-test

data:  V1 and V3
t = -1.5052, df = 17.971, p-value = 0.1496
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -2.4677776  0.4077776
sample estimates:
mean of x mean of y
     2.40      3.43

> t.test(V5,V6)

        Welch Two Sample t-test

data:  V5 and V6
t = -0.6125, df = 17.384, p-value = 0.5481
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -1.6867372  0.9267372
sample estimates:
mean of x mean of y
     2.45      2.83
```

2 (continued)

```
> t.test(V7,V8)

        Welch Two Sample t-test

data:  V7 and V8
t = -2.1534, df = 14.79, p-value = 0.04821
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -2.5883311 -0.0116689
sample estimates:
mean of x mean of y
     -0.1       1.2
```

```
> t.test(V7,V9)

        Welch Two Sample t-test

data:  V7 and V9
t = 1.8221, df = 14.79, p-value = 0.08872
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -0.1883311  2.3883311
sample estimates:
mean of x mean of y
     -0.1      -1.2
```

(iii) Do your findings suggest that the new drug A should
a) definitely be introduced
b) definitely not be introduced
c) be introduced if it passes further checks?
Explain your reasoning, including, if you select option c), what further checks are needed. *(4 marks)*

3 A car manufacturer makes two different types of engine components (A and B) and wants to see if there is any difference between their lifetimes. The components were fitted to an engine that was run until failure. The length of time (in hrs) that the engines functioned with the components were recorded and can be seen below. Here the status variable records whether the component was still operating when something else on the engine broke (status = 0) or whether the engine broke due to the component itself failing (status = 1).

	Type A		Type B	
	Time (hrs)	Status	Time (hrs)	Status
	5.10	1	5.20	1
	16.30	1	14.20	1
	0.30	0	7.20	0
	4.30	1	4.90	0
	10.00	1	18.40	1
	0.10	0	7.80	1
Total	36.1	4	57.7	4

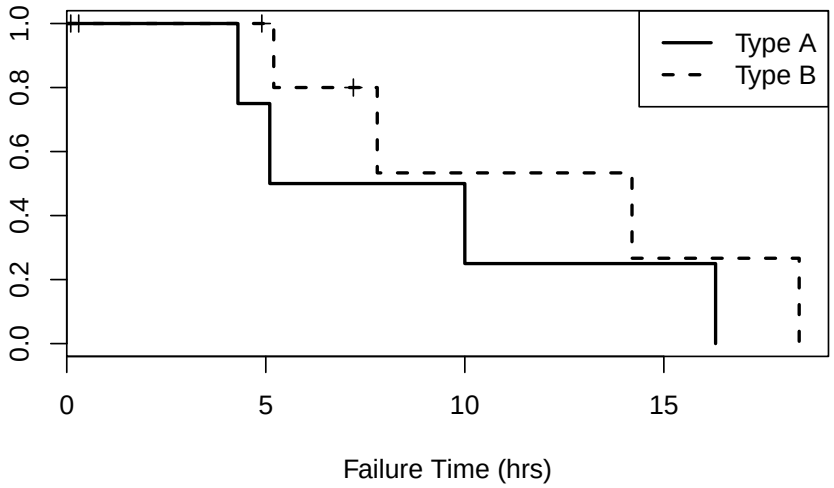
Given below are the results of a Kaplan-Meier preliminary graphical analysis of the data

```
> Comp.sv <- Surv(time, status, type = "right")
> summary(survfit(Comp.sv ~ type))
Call: survfit(formula = Comp.sv ~ type)

                type=A
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
  4.3      4       1     0.75   0.217       0.4259            1
  5.1      3       1     0.50   0.250       0.1877            1
 10.0      2       1     0.25   0.217       0.0458            1
 16.3      1       1     0.00     NaN           NA           NA

                type=B
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
  5.2      5       1    0.800   0.179       0.5161            1
  7.8      3       1    0.533   0.248       0.2142            1
 14.2      2       1    0.267   0.226       0.0507            1
 18.4      1       1    0.000     NaN           NA           NA
```

3 (continued)



(i) Without making any model assumptions, estimate the median failure times for the two component types. ***(3 marks)***

(ii) It is suggested that the survival times for components A and B are Exponentially distributed with rates λ_A and λ_B respectively. Under this assumption:

(a) Estimate λ_A and λ_B and hence the mean failure times with approximate 95% confidence intervals (note that selected standard normal percentage points are given below). ***(4 marks)***

p	qnorm(p)
0.025	-1.96
0.05	-1.64

(b) Use the likelihood ratio test to assess whether there is a difference in the failure time distribution of the two components (note that selected chi-squared percentage points are given below). Why is the LRT more suitable than an MLE test in this case? ***(4 marks)***

p	qchisq(p,1)
0.9	2.71
0.95	3.84

(c) Do the assumptions of Exponential survival distributions seem plausible? Explain your answer. ***(2 marks)***

3 (continued)

(iii) By copying and completing the partially filled table below, perform a non-parametric comparison of the two survival distributions (note that selected chi-squared percentage points are given below). Compare your findings with the parametric test in part (ii)(b).

p	qchisq(p,1)
0.9	2.71
0.95	3.84

		Number at risk			Number of deaths			Expected number of deaths	
i	t_i	r_{Ai}	r_{Bi}	r_i	d_{Ai}	d_{Bi}	d_i	e_{Ai}	e_{Bi}
1	4.3								
2	5.1								
3	5.2								
4	7.8								
5	10.0								
6	14.2								
7	16.3								
8	18.4								
Total					$O_A = 4$	$O_B = 4$		$E_A = 2.8$	$E_B = 5.2$

(7 marks)

4 (i) A cohort of live born babies was followed up for one year to explore the effect of birth weight and sex on survival. The results from fitting a proportional hazards regression model are shown in the Table below. The model relates survival to birth weight (grouped into four categories) and the sex of the baby, and the age of the mother (grouped into two categories).

Variable	Coefficient	Standard Error
Sex of Baby		
Male	Reference	—
Female	-0.39	0.07
Birthweight		
$\geq$ 4000 g	0.12	0.08
3500–3999g	Reference	—
3000–3499g	0.27	0.15
$<$ 3000g	1.61	0.23
Age of mother		
$>$ 40 yrs	0.65	0.33
$\leq$ 40 yrs	Reference	—

Table: Log hazard ratio of infant mortality with standard errors

(a) Specify the form of the hazard model used for this analysis, carefully defining each predictor variable. ***(3 marks)***

(b) Describe in detail the effects of these variables on infant survival (note that selected standard normal and chi-squared percentage points are given below). ***(6 marks)***

p	qnorm(p)	p	qchisq(p,3)
0.001	-3.09	0.95	7.81
0.005	-2.58	0.99	11.34
0.025	-1.96		
0.05	-1.64		

(c) How would you assess whether the proportional hazards model was appropriate for these data? ***(2 marks)***

(d) Using the model, calulate the estimate of the hazard ratio comparing the following two babies:

- A female baby with birth weight 3000–3499g and a 35 yr old mother,
- A male baby with birth weight 3500–3999g and a 41 yr old mother.

(3 marks)

4 (continued)

(ii) A key task for radiologists is to interpret mammograms to see if they provide evidence of breast cancer. As part of his final examination, a student radiologist is asked to assess 85 mammograms and assign each to one of the 4 categories: 'Normal', 'Benign Disease', 'Suspected Cancer', 'Cancer'. A senior consultant radiologist has already categorized the mammograms. The cross-classified categorizations (adapted from Altman (1995)) are shown below. Do you think the student should pass the examination? Justify your answer. ***(6 marks)***

		Consultant				
		Normal	Benign	Suspected Cancer	Cancer	Total
Student	Normal	21	12	0	0	33
	Benign	4	17	1	0	22
	Suspected Cancer	3	9	15	2	29
	Cancer	0	0	0	1	1
	Total	28	38	16	3	85

End of Question Paper