Queen Mary
University of London

B. Sc. Examination by course unit 2014

MTH6134P Statistical Modelling II

Duration: 2 hours

Date and time: 23 May 2014, 10:00h–12:00h

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

You should attempt all questions. Marks awarded are shown next to the questions.

Calculators ARE permitted in this examination. The unauthorized use of material stored in pre-programmable memory constitutes an examination offence. Please state on your answer book the name and type of machine used.

Statistical functions provided by the calculator may be used provided that you state clearly where you have used them.

The New Cambridge Statistical Tables are provided.

Complete all rough workings in the answer book and cross through any work which is not to be assessed.

Important note: the Academic Regulations state that possession of unauthorized material at any time by a student who is under examination conditions is an assessment offence and can lead to expulsion from QMUL.

Please check now to ensure you do not have any notes, mobile phones or unauthorised electronic devices on your person. If you have any, then please raise your hand and give them to an invigilator immediately. Please be aware that if you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms, it will be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. Disruption caused by mobile phones is also an examination offence.

Exam papers must not be removed from the examination room.

Examiner(s): H. Maruri-Aguilar and D. S. Coad

 TURN OVER

Question 1 (11 marks)
Acidity measurements of soil were made at different distances A_1 and A_2 from the surface at three different locations. The coded values are given in the following table.

A_1	A_2
2	4
10	12
2	6

(a) Define the null subspace V_0 and the subspace $V_0^{\perp}$. [2]

(b) Define the treatment subspace V_T; also define subspaces $V_0^{\perp} \cap V_T$ and $V_T^{\perp}$. [3]

(c) Compute the following projections of the data: $P_{V_T} y$ and $P_{V_T^{\perp}} y$. [6]

Question 2 (22 marks)
Four different catapult designs are being tested for their ability to launch water balloons. Four students (Peter, Shawn, Mark and Jake) were involved in the experiment. Because the person firing the catapult could represent a source of variability, the experimenter decides to use a block design and block by person firing. Below are the experimental data, with distances given in feet. Catapult designs "Mark IV V2", "Balloon Ballista", "Hydrolaunch" and "Waterworks II" were coded as MIV, BB, H and W2.

	MIV	BB	H	W2
Peter	92	94	97	100
Shawn	94	93	98	99
Mark	92	94	95	98
Jake	97	96	101	103

`GenStat` output is also given.

```
  Variate: distance

Source of variation  d.f.      s.s.      m.s.
catapult                3  105.6875  35.2292
person                  3   45.6875  15.2292
Residual                9    8.5625   0.9514
Total                  15  159.9375
```

(a) Describe an appropriate model for the data and assumptions required. [4]

(b) Write down the hypotheses for the factors `catapult` and for `person`. Then perform the hypothesis tests using a significance level of 5%. [7]

(c) Compute the standard error of a difference between treatment means. [2]

(d) Compare all pairs of means using the least significant difference method and a significance level of 5%. [6]

(e) Describe how you would analyze these data using `GenStat`. [3]

Question 3 (23 marks)

A factorial experiment with cathode ray tubes was performed with two types of glass (I, II) and three kinds of phosphorus (A,B,C). The table below shows the values of the current (in milliamperes) necessary to produce a given amount of brightness in a cathode ray tube.

	Phosphorus		
Glass	A	B	C
I	285	302	282
	280	301	284
II	235	245	225
	228	215	226

(a) Draw the interaction means plot for the data and interpret the plot. [10]

(b) For the fixed effects model for both factors `glass` and `phosphorus`, complete the analysis of variance table. For your computations use the result $285^2 + 302^2 + \cdots + 226^2 = 816770$. [9]

(c) Perform the hypothesis tests concerning the factors glass and phosphorus and their interaction using a significance level of 5%. Relate your results to what you interpreted in your plot. [4]

Question 4 (20 marks)

Samples of water are to be tested for their quality. The study involved only two lakes. From each lake, among many candidate locations the scientist selected two and then took two samples from each location, making a total of eight water samples. The scheme is represented in the following table:

Lake	1				2			
Location	1		2		1		2	
Measurement	y_{111}	y_{112}	y_{121}	y_{122}	y_{211}	y_{212}	y_{221}	y_{222}
	5.8	3	2.8	5.5	6.7	6.6	5.4	6.9

(a) Describe an appropriate model for these data. [7]

(b) Given that $5.8^2 + 3^3 + \cdots + 6.9^2 = 245.95$, that $5.8 + 3 + \cdots + 6.9 = 42.7$ and that $(5.8 + 3)^2 + (2.8 + 5.5)^2 + (6.7 + 6.6)^2 + (5.4 + 6.9)^2 = 474.51$, complete the analysis of variance table. [6]

(c) Perform the hypothesis tests on factors lake and location using a significance level of 5%. [4]

(d) Estimate the variance components. [3]

Question 5 (24 marks)

Consider the one-way anova model $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ for $i = 1, \ldots, t$, equal replication $r_1 = \ldots = r_t = r$ so that $j = 1, \ldots, r$ and independent errors $\epsilon_{ij} \sim N(0, \sigma^2)$.

(a) Define what a contrast is and explain what contrasts are used for. [6]

(b) Given two contrasts L_i and L_j, define the condition for them to be orthogonal in the above setting. [4]

(c) How many orthogonal contrasts can we build and test in the above setting? [3]

(d) A survey about customer preference for four types of crisps, named A, B, C, D, is to be performed. For each type of crisps, ten customers will be interviewed. The analyst proposes three contrasts of interest: $L_1 = \alpha_A - \alpha_B$, $L_2 = \alpha_C - \alpha_D$ and $L_3 = \alpha_A + \alpha_B - \alpha_C - \alpha_D$.

(i) Show that the contrasts L_1, L_2 and L_3 are orthogonal. [4]

(ii) A technician suggested instead the contrasts $M_1 = L_1 + L_2$, $M_2 = L_1 - L_2$ and $M_3 = L_3$. Help him to determine if his set of contrasts is orthogonal. [7]

End of Paper