Queen Mary
University of London

B. Sc. Examination by course unit 2015

MTH6134: Statistical Modelling II

Duration: 2 hours

Date and time: 29th May 2015, 10:00–12:00

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.

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 that is not to be assessed**.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately. It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms it shall 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. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

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

 TURN OVER

Question 1 (23 marks). A study was performed to determine the preference of customers for three different types of energy drinks, codenamed as A, B, C. Ten specimens of drink B were given to customers, while five specimens of each of A and C were used.

The following facts using the notation seen in lectures will be required. The model under consideration is the standard model $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ with independent errors $\epsilon_{ij} \sim N(0, \sigma^2)$. For this model and a general contrast $L_k = \sum_{i=1}^{t} \lambda_{ki}\alpha_i$, its estimator is $\hat{L}_k = \sum_{i=1}^{t} \lambda_{ki} y_{i\cdot}$; and the sum of squares for contrast L_k is

$$S_k = \frac{\hat{L}_k^2}{\sum_{i=1}^{t} \frac{\lambda_{ki}^2}{r_i}}$$

where r_i and $y_{i\cdot}$ are the replication and mean for treatment i, respectively.

(a) There is interest in performing orthogonal comparisons. One comparison of interest is to test equality of effects of A and C. Write this comparison as a contrast and call it L_1. **[2]**

(b) Propose an orthogonal contrast L_2. Show that L_1 and L_2 are orthogonal. **[4]**

(c) Write down the estimator $\hat{L}_2$ of your proposed contrast and show that $\hat{L}_2$ is an unbiased estimator of L_2. **[4]**

(d) Find the variance of the estimator $\hat{L}_2$. **[6]**

(e) Combine your results to determine the expected mean square of contrast L_2, that is, compute $E(M_2) = E(S_2)$. Explain how this expectation relates to the hypothesis test of contrast L_2. **[7]**

Question 2 (27 marks). The loss of product in the manufacture of an organic chemical was studied in an experiment involving three blocks. Fifteen batches were used and five blends of acetanilde were considered. The data and sums of squares are given below.

Loss	Block	Blend
18.2	I	B
17.2	I	A
17	I	C
18.5	I	E
15.1	I	D
16.9	II	A
18.3	II	E
19.2	II	B
18.1	II	C
16	II	D
17.1	III	B
16	III	D
17.3	III	C
19.8	III	E
17.3	III	A

Source of variation	s.s.
Blend	17.047
Block	0.633
Residual	4.173
Total	21.853

(a) Complete the analysis of variance table and test the relevant hypothesis for the effect of factors acetanilde blend and block. [8]

(b) Construct a table of means for acetanilde blend and give the standard error of a difference between means in that table. [5]

(c) Using the method of least significant difference, perform multiple comparisons between all pairs of acetanilde blends. [9]

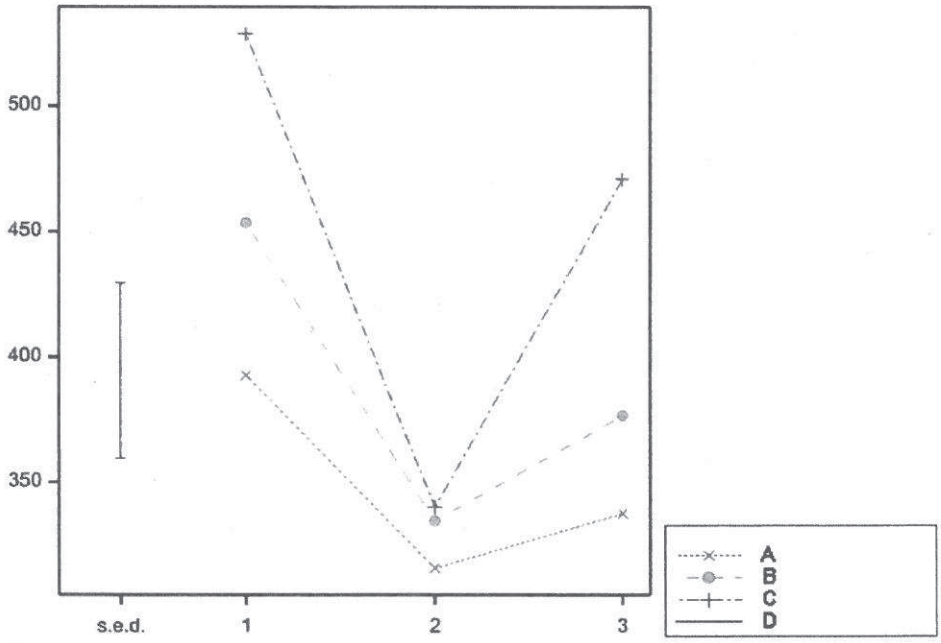
(d) Give the explicit version of the matrices X and β involved in writing the matrix form version of the model $Y = X\beta + \epsilon$ for these data. [5]

Question 3 (28 marks). A factorial experiment was carried out to investigate the effects of the physical properties of vulcanised rubber (wear) by varying two factors. The first factor was type of `filler` and the second factor was `method` of pretreatment. Both factors were given coded levels $1, 2, 3$. The data are given below; in addition we know that $404^2 + 392^2 + 348^2 + \cdots + 249^2 + 482^2 = 2932360$ and that $404 + 392 + 348 + \cdots + 249 + 482 = 7100$.

		Replications					
		1			2		
		Method			Method		
		1	2	3	1	2	3
Filler	1	404	478	530	381	429	528
	2	392	418	431	239	251	249
	3	348	381	460	327	372	482

(a) Compute a table of means per treatment combination. [4]

(b) The following is a plot of means per treatment combination.



Identify which factor belongs to the horizontal axis and which to the cathegories indicated by line types; also state which levels correspond to labels A, B, C in plot and what the bar represents (full line, label D). Interpret the plot. [9]

(c) Complete computations of analysis of variance and test the effect for factors filler and method and their interaction. [13]

(d) Give the standard error for the difference between means per treatment combination. [2]

Question 4 (22 marks). Batches of plastic-like material were sampled, with each batch being sampled four times; two manufacturers were considered, each providing three batches of material to total 24 measurements.

(a) The following two panels correspond to analyses performed in `GenStat`.

Panel 1

Variate: Strength

Source of variation	d.f.	s.s.	m.s.	v.r.
Manufacturer	1	3.38	3.38	0.09
Manufacturer.Batch	4	398.83	99.71	2.59
Residual	18	692.75	38.49	
Total	23	1094.96		

Panel 2

Variate: Strength

Source of variation	d.f.	s.s.	m.s.	v.r.
Manufacturer.Batch stratum				
Manufacturer	1	3.38	3.38	0.03
Residual	4	398.83	99.71	2.59
Manufacturer.Batch.*Units* stratum				
	18	692.75	38.49	
Total	23	1094.96		

For each panel above, and assuming that data have been captured in factors "Manufacturer" and "Batch", give the correct instruction of what to put in the boxes

Treatment Structure: ☐
Block Structure: ☐

[6]

(b) Now assume that the data come from an experiment in which manufacturers were selected randomly from a large number of suppliers and that for each manufacturer, batches were also selected at random. Write down the appropriate model and assumptions. [4]

(c) Test the hypothesis for the factors manufacturer and batches and conclude. [12]

End of Paper.