Queen Mary
University of London

B. Sc. Examination by course unit 2014

MTH6116: Design of Experiments

Duration: 2 hours

Date and time: 13th May 2014, 14:30h–16:30h

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 unauthorised 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 unauthorised 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 B. Bogacka

TURN OVER

Question 1 (24 marks)

(a) (i) Define what is meant by power in a statistical test. [2]

(ii) A study requires a completely randomized design for comparing two treatments. There is interest in detecting a difference of six units between treatment means with a significance level $\alpha = 0.05$. The analysts have two candidate designs.

D_1 This is a design with equal replication $r = 5$.

D_2 Due to budget constraints and the fact that treatment two is considerably more expensive than treatment one, the analysts are considering if they can use design D_2 with replication pattern $r_1 = 8, r_2 = 4$.

Help the analysts decide which design to use by computing the power of each proposal. It is known from a previous study that the variance of individual units is estimated to be $\sigma^2 = 9$. Comment on your results. [12]

(b) (i) Describe the correct construction and randomization of a completely randomized design. [3]

(ii) A technician intended to produce a plan for a completely randomized design with six runs and three treatments with equal replication. He applied a permutation scheme to his systematic layout and obtained a design plan. The random permutation he used, his systematic design and design plan he obtained are given below.

Systematic design

Plot	1	2	3	4	5	6
Tmnt	A	A	B	C	B	C

Random permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 4 & 2 & 3 & 1 \end{pmatrix}$$

Design plan

Plot	1	2	3	4	5	6
Tmnt	C	C	B	A	B	A

You are asked to revise if this method was correct by identifying which elements of the technician's report are not consistent with the randomization method. [2]

(iii) Using the description of the correct method and the same random permutation, produce the plan for the randomized design. As part of your solution, explain how you did the randomization. [5]

Question 2 (25 marks)

In a greenhouse experiment to study the effects of three different fertilisers on three different varieties of tomato plant, eighteen plants were used, each in its own pot, and all of the same age. There were two pots for each combination of variety and fertiliser. The pots were arranged in random order on the greenhouse bench, and the allocation of treatment combinations to plots was random. The growth of each plant over a certain period was measured in logarithmic units.

		Fertiliser					
		A		B		C	
Variety	1	1.1	1.7	2.9	2.1	1.4	2.2
	2	1.4	1.8	2.6	2.2	3.3	2.7
	3	2.3	2.7	1.8	1.2	3.5	2.3

(a) Give the analysis of variance table for these data (the total uncorrected sum of squares is 93.5 and $\text{SUM}^2/N = 85.369$). [13]

(b) Report briefly on the results of the table. Use a significance level $\alpha = 0.05$. [3]

(c) Give the two-way table of treatment means; calculate standard errors for use in interpreting this table. [6]

(d) Give a 95% confidence interval for the best mean growth coming from this experiment. [3]

Question 3 (29 marks)

(a) Define what is a row-column design. [3]

(b) Define what is a Latin square. [2]

(c) Using Latin squares $S_1 = \begin{array}{|cc|}\hline A & B \\ B & A \\ \hline\end{array}$ and $S_2 = \begin{array}{|cc|}\hline B & A \\ A & B \\ \hline\end{array}$, construct the Latin square $S_2 \otimes S_1$. [3]

(d) The purpose of an investigation was to find the best method of preparing an insecticide. Three factors were involved: (a) diluent used, (b) method of mixing and (c) ingredient. It was decided to use seven variants of each factor in a Latin Square design shown below. Letters A-G represent the seven diluents, and numbers in the cells are measurements of the persistence of the insecticide, i.e. its ability to remain effective after spreading. Totals are given in the table.

		ingredient							
		1	2	3	4	5	6	7	Total
mixing	1	A 98	B 117	C 89	D 64	E 63	F 132	G 244	807
	2	B 69	E 67	A 70	G 70	F 111	D 60	C 218	665
	3	C 37	F 83	G 83	B 74	D 70	A 75	E 169	591
	4	D 65	G 60	E 91	F 56	C 61	B 59	A 150	542
	5	E 56	D 44	B 70	C 68	A 88	G 111	F 220	657
	6	F 113	C 105	D 65	A 51	G 83	E 57	B 233	707
	7	G 64	A 62	F 65	E 86	B 45	C 108	D 187	617
	Total	502	538	533	469	521	602	1421	4586

diluent	A	B	C	D	E	F	G
Total	594	667	686	555	589	780	715

(i) For each of factors ingredient, method and diluent, calculate a table of means and describe the main features you observe. [6]

(ii) Complete the analysis of variance table for this data and perform the relevant hypothesis tests for the three factors using a significance level $\alpha = 0.05$. Use the facts that $\sum_{\omega\in\Omega} y_\omega^2 = 553714$ and SS(mean) $= 429212.1633$. [15]

Question 4 (22 marks)

A study was conducted to find out the effect of temperature and amount of coloring (dye) in the amount of material lost (grams) of certain fabric. Two amounts of dye were used together with two temperatures to create a full factorial structure to which control treatment was added. The following table summarizes the treatment structure used.

Treatment	1	2	3	4	5
Control	Y	N	N	N	N
Dye	0	0.1	0.1	0.3	0.3
Temperature	40	60	80	60	80

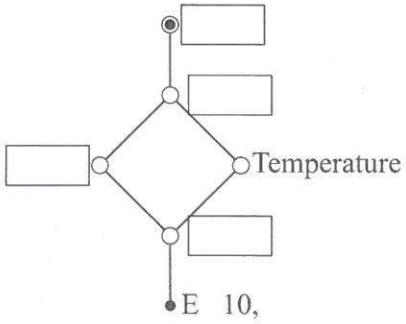
Two replications were allowed for each treatment and the allocation of treatments to plots was completely at random. The following was computed: $\sum_{\omega\in\Omega} y_\omega^2 = 2512.45$. The following table summarizes the results of the experiment.

Treatment	1	2	3	4	5
Replications	2	2	2	2	2
Total	31.1	29.5	18.8	40.9	34
Squared total	967.21	870.25	353.44	1672.81	1156

(a) List the equivalence classes of treatment factors control, dye and temperature. [2]

(b) Show that dye$\vee$temperature$\equiv$control and that dye$\wedge$temperature$\equiv$E, where E is the equivalent factor in the set of treatments. [4]

(c) Copy the combined Hasse diagram shown below left. Then complete it by filling the missing factor labels, levels and degrees of freedom. [4]



Source	SS	d.f.	MS	VR
		1		
Control			0.04	0.07
Dye	88.45	1		
Temp	38.72			
Dye.Temp				3.48
Residual	2.59		0.52	
Total	2512.45	10		

(d) Complete the missing entries in the analysis of variance table above, test the relevant hypotheses using $\alpha = 0.05$ and conclude. [12]

End of Paper