Queen Mary
University of London

Main Examination period 2018

MTH6116 / MTH6116P: Design of Experiments

Duration: 2 hours

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 available are shown next to the questions.

Only non-programmable calculators that have been approved from the college list of non-programmable calculators are permitted in this examination. Please state on your answer book the name and type of machine used.

The New Cambridge Statistical Tables are provided.

Complete all rough work 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.

Examiners: D. S. Coad, H. Maruri-Aguilar

 Turn Over

Question 1. [24 marks]

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

(ii) A study requires a completely randomised design for comparing two treatments. There is interest in detecting a difference of five units between treatment means with a significance level of 5%. The analysts have two candidate designs.

D_1 This is a design with equal replication $r = 6$.

D_2 Due to budget constraints and the fact that treatment 1 is considerably more expensive than treatment 2, the analysts are considering whether they can use this design with replication $r_1 = 4$ and $r_2 = 10$.

Help the analysts decide which design to use by computing the power of each one. It is known from a previous study that the variance of individual units is estimated to be $\sigma^2 = 8$. Comment on your results. **[12]**

(b) (i) Describe the correct construction and randomisation of a completely randomised design. **[3]**

(ii) A technician intends to produce a plan for a completely randomised design with seven runs and three treatments with unequal replication. She applied a permutation scheme to her systematic layout and obtained a design plan. The random permutation that she used, her systematic design and the design plan that she obtained are given below.

Systematic Design

Plot	1	2	3	4	5	6	7
Treatment	B	A	C	A	B	C	B

Random Permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 7 & 3 & 4 & 2 & 6 \end{pmatrix}$$

Design Plan

Plot	1	2	3	4	5	6	7
Treatment	B	A	B	C	B	A	C

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

(iii) Using the description of the correct method and the same random permutation, produce the plan for the randomised design. As part of your solution, explain how you carried out the randomisation. **[5]**

Question 2. [30 marks] An engineer is studying the effects of five different formulations, A, B, C, D and E, of an explosive mixture used in the manufacture of dynamite on the observed explosive force. Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. Furthermore, the formulations are prepared by five operators, and there may be substantial differences in the skills and experience of the operators. The row-column design used is given below, together with the data.

		Operator				
		1	2	3	4	5
Batch	1	A	B	C	D	E
	2	B	C	D	E	A
	3	C	D	E	A	B
	4	D	E	A	B	C
	5	E	A	B	C	D

		Operator				
		1	2	3	4	5
Batch	1	24	20	19	24	24
	2	17	24	30	27	36
	3	18	38	26	27	21
	4	26	31	26	23	22
	5	22	30	20	29	31

(a) Explain why this is a Latin square and describe how it was constructed. **[5]**

(b) Write down the orthogonal decomposition of the vector space V for this design in terms of suitable subspaces. **[4]**

(c) Given that $\sum_{\omega\in\Omega} y_\omega^2 = 16,805$, complete the analysis of variance table and test factor formulation. **[15]**

(d) Find the standard error of a difference for comparing two treatments. **[2]**

(e) To analyse the data with `GenStat`, what factors must be declared? How would you use the `GenStat` Dialogue Box to specify the structures that `GenStat` needs to know for carrying out the analysis of variance? **[4]**

Question 3. [25 marks] A study was undertaken to investigate the effects of three chocolate cake recipes and three baking temperatures, in degrees Fahrenheit, on the quality of cakes. There were three replications of the experiment with replications serving as blocks. At each replication, a recipe was selected at random and enough cake batter was prepared for three cakes. After making a particular batch, it was split into three equal parts and each part was assigned at random to one of the three oven temperatures. There were three ovens available. The data are provided below.

		Recipe								
		1			2			3		
	Temperature	175	195	215	175	195	215	175	195	215
Block	1	28	31	41	31	29	40	21	31	33
	2	24	27	30	21	24	37	26	27	35
	3	26	32	37	21	28	27	21	25	31

(a) Explain what is meant by a **split-plot design**. **[5]**

(b) Complete the analysis of variance table for the experiment. **[15]**

Stratum	Source	SS	df	MS	VR
Mean	Mean	22,765.04	1	22,765.04	
Blocks	Blocks	93.85			
Batches	Recipe				
	Residual	96.37			
	Total	135.78			
Parts	Temperature	479.19			
	Recipe $\wedge$ Temperature				
	Residual	110.44			
	Total	595.33			
Total		23,590.00	27		

(c) By performing appropriate tests, what conclusions can you draw about the effects of the two factors and their interaction? **[5]**

Question 4. [21 marks] A food scientist wants to study whether quality differences exist between yoghurt made from skimmed milk with and without the pre-culture of psychrotropic (PC) bacteria. Samples of skimmed milk are procured from seven dairy farms. One half of the milk sampled from each farm is inoculated with PC and the other half is not. After yoghurt is made with these milk samples, the firmness of the curd is measured. These measurements are given below.

	Dairy Farm						
	A	B	C	D	E	F	G
With PC	68	75	62	86	52	46	72
Without PC	61	69	64	76	52	38	68

(a) State the condition for a block design to be orthogonal. **[2]**

(b) Show that the design above is orthogonal and explain how it should have been randomised. **[5]**

(c) Write down orthogonal basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ for the treatment subspace V_T. **[2]**

(d) Show that $\mathbf{u}_1 - \mathbf{u}_2$ is a basis for W_T and hence that $\dim(W_T) = 1$. **[3]**

(e) Using the basic model equation $Y_\omega = \tau_{T(\omega)} + Z_\omega$, write down $E(\mathbf{Y})$ and then compute the expected mean square $\mathrm{EMS}(W_T) = E(||P_{W_T}\mathbf{Y}||^2)$. **[9]**

End of Paper.