Queen Mary
University of London

MTH6116 / MTH6116P: Design of Experiments

Duration: 2 hours

Date and time: 12th May 2016, 14:30–16:30

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 L. Pettit

Question 1 (24 marks). A rugby coach studied different types of run with the aim of later establishing game strategies. The coach used eight players, making each player do four 50 metre runs. Each player ran once with the ball in the left hand, once in the right hand and on one occasion with the ball in both hands. For comparison he also made the players run without a ball, thus totalling four treatments which he termed as L, R, B and N, respectively. In each run, the coach took the time elapsed between the runner passing the ten metre and forty metre lines. The response time (in seconds) was multiplied by a factor of 100 to simplify computations.
The coach used a crossover design for which he created two 4×4 latin squares which were properly randomized. He now wants to analyze his data.

(a) Working on the set of treatments $\mathcal{T} = \{B, L, N, R\}$, together with the usual factors U and $E \equiv \texttt{tmnt}$, create a factor `Ball` that identifies whether the treatment used (or not) a ball; also create a factor `Hand` which counts the number of hands the player used in each treatment. Build equivalence classes for each factor, determine which factors are finer than others and complete the Hasse diagram over $\mathcal{T}$.

After this step, factor `Ball` will test the difference between using a ball or not; factor `Hand` will compare running with the ball in one hand versus in two hands; while factor `tmnt` will compare running with ball in left hand versus doing so with the ball in the right hand. **[6]**

(b) Are the factors in your Hasse diagram over $\mathcal{T}$ orthogonal? Briefly explain why or why not. **[2]**

(c) The following table contains totals per treatment. Build crude sums of squares and sums of squares for all the treatment factors you have.

Treatment	N	L	R	B
Total	4491.6	4489.7	4485.4	4415.2
Replications	8	8	8	8

[6]

(d) For plot factors, the following relations were identified: $E \prec \texttt{player} \prec U$ and $E \prec \texttt{run} \prec U$. Ignoring treatment structure, crude sums of squares and sums of squares were computed:

Factor	E	`run`	`player`	U
CSS	10006527	10001658	9996888	9992573
SS	554.36	9084.64	4314.64	9992573

Using all the information you have, complete the anova table, test the effect of treatment factors you have previously identified and make conclusions. **[10]**

Question 2 (28 marks).

(a) Define

(i) a matched pairs design; [4]

(ii) an orthogonal block design. [4]

(b) Consider a matched pairs design with three pairs.

(i) Determine number of treatments t, number of replications per treatment r, number of blocks b, block size k and number of units N. [2]

(ii) Show that this design is orthogonal. [5]

(iii) Write explicitely the vector u_0 and the orthogonal basis vectors $u_1, \ldots, u_t$ for treatment subspace V_T. Also write orthogonal vectors $v_1, \ldots, v_b$ for block subspace V_B. Determine $\dim(V_T)$ and $\dim(V_B)$. [2]

(iv) Show that $u_1 - u_2$ is a basis for W_T and thus $\dim(W_T) = 1$. [3]

(v) Under the fixed effects model for this design $Y_\omega = \tau_{T(\omega)} + \zeta_{B(\omega)} + Z_\omega$, use Theorem 2.2 given below to compute EMS(W_T). [8]
Hint: Write the explicit vector version of $E(Y_\omega) = \tau_{T(\omega)} + \zeta_{B(\omega)}$ and then compute $E(||P_{W_T}Y||^2)$.

Appendix for Question 2.

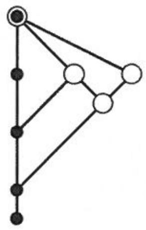
The following are definitions of subspaces and Theorem 2.2 from the course.
Definition (Null subspace) The null subspace V_0 is the span of the vector u_0, that is $V_0 = \{cu_0 \text{ for } c \in \mathbb{R}\}$.
Definition (Treatment subspace) The treatment subspace V_T is the span of basis vectors $u_1, \ldots, u_t$: $V_T = \left\{\sum_{i=1}^{t} c_i u_i \text{ for constants } c_1, c_2, \ldots, c_t \in \mathbb{R}\right\}$.
Definition (Subspace W_T) The subspace W_T consists of those vectors in V_T that are orthogonal to vectors in V_0: $W_T = \{v \in V_T \text{ such that } \langle v, x\rangle = 0 \text{ for } x \in V_0\}$.
Theorem 2.2 Let $U \subseteq \mathbb{R}^N$ be a vector space and let W be a d-dimensional subspace of U. Assume that $E(Y) \in U$ and $\text{Cov}(Y) = \sigma^2 I$ where I is the identity matrix. Then

i) $E(P_W Y) = P_W E(Y)$, and

ii) $E(||P_W Y||^2) = ||P_W E(Y)||^2 + d\sigma^2$.

Question 3 (25 marks). A study about the impact of fertilizer and pruning method on the weight of apples of a certain variety was performed. The study involved two farms, each farm with four orchards of apple trees. Each orchard in turn consisted of six trees, thus totalling 48 trees in the study. Two types of fertilizer were involved in the study, each type being applied in bulk to an orchard; two pruning methods were applied directly to trees inside orchards, and each pruning method applied to three trees in each orchard. During harvest, two apples were taken at random from each tree and the weight in grams of each apple was recorded. The combined weight of all the apples was 9513.1 grams.

(a) Copy the Hasse diagram below and complete it by adding factor labels as well as number of levels and degrees of freedom. **[10]**



(b) Complete the analysis of variance table for this experiment by carefully studying the partial `GenStat` results labelled Output 1 and 2 in the appendix below and extracting information from them. You are **not** supposed to analyze Outputs 1 and 2 directly, but to use the figures to build your anova table with the help of your Hasse diagram. **[9]**

(c) At the 90% significance level (i.e. $\alpha = 0.1$), test the effect of factors fertilizer, pruning and their interaction. **[6]**

The next page contains an appendix for part (b) of this question

Appendix for Question 3.

Output 1

```
63 "General Analysis of Variance."
64 BLOCK Farm/Orchard/Tree/Apple
65 TREATMENTS
Analysis of variance
Variate:  Weight
```

Source of variation	d.f.	s.s.	m.s.	v.r.
Farm stratum	1	0.143	0.143	0.01
Farm.Orchard stratum	6	115.871	19.312	2.11
Farm.Orchard.Tree stratum	40	366.729	9.168	1.35
Farm.Orchard.Tree.Apple stratum	48	325.585	6.783	
Total	95	808.327		

Output 2

```
69 "General Analysis of Variance."
70 BLOCK "No Blocking"
71 TREATMENTS Fertilizer*Pruning
Analysis of variance
Variate:  Weight
```

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Fertilizer	1	37.375	37.375	4.66	0.033
Pruning	1	31.855	31.855	3.97	0.049
Fertilizer.Pruning	1	1.105	1.105	0.14	0.711
Residual	92	737.992	8.022		
Total	95	808.327			

Question 4 (23 marks). A factorial experiment was performed in an industrial baking process. Two recipes (italian, brioche) were used to produce bread loaves. Three different amounts of flour together with two different amounts of yeast were applied to loaves from each recipe. The height of the loaf in centimetres was recorded. The combined total height of all loaves was 193.2 centimetres.
The following is partial `GenStat` output.

```
Variate:  Height
 Source of variation          s.s.

 Recipe stratum             2.8033

 Recipe.*Units* stratum
 Flour                      9.7800
 Yeast                      7.6800
 Flour.Yeast                1.4600
 Residual                   2.7367

 Total                     24.4600
```

(a) Give the correct instruction of what to put in the `GenStat` boxes below in order to analyze this data. [4]

Treatment Structure:	
Block Structure:	

(b) Complete the analysis of variance table and perform the relevant tests for factors flour and yeast. [13]

(c) Give the standard error for comparing means per level of flour and for level of yeast. [6]

End of Paper.