UNIVERSIDADE DE SÃO PAULO
Campus "Luiz de Queiroz"
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"

DEPARTMENTO DE CIÊNCIAS EXATAS

Estatística e Experimentação Agronômica

LCE802 Planejamento de Experimentos e Análise de Dados

FINAL PAPER

PART A — Practical

General Instructions to Candidates　　　　　　　**Time allowed: 4 hours**

1.　This examination is an open book examination. You may need a calculator and statistical tables.

2.　Attempt **ALL** questions. The marks for each question are shown with the question and in the following table.

PART	Question	Max. Mark
A — Practical	1	14
	2	34
	3	15
B — Theory	4	7
	5	6
	6	7
	7	12
	Total	95

3.　In carrying out statistical procedures, use a 5% level of significance and a 95% confidence level, unless otherwise specified. Also, you should carry out the complete procedure even if the assumptions underlying it are not met.

Q1 An experiment was conducted to investigate the effects of four diets on the growth of young chicks. There were 16 cages of 15 chicks with the cages arranged in four rows (layers) by four columns (stacks). The four diets were assigned to the cages using a Latin square design. The four diets were:

C: Control diet that includes lysine;

L: Lysine enriched diet that is C plus 0.2% artificial lysine;

W: Wheat-added diet that is C, except that some starch and sugar of C is replaced with sufficient wheat to add an extra 0.2% lysine — note that, because wheat naturally contains amino acids, this diet also has additional amino acids to C and L;

A: Amino enriched diet that is L, plus the amino acids added naturally by the wheat to diet W.

These 4 treatments form natural groupings as follows:

1. (C) versus (L, W, A) compares no lysine with lysine;
2. within the lysine treatments, (L) versus (W, A) compares no amino acid to amino acid;
3. within the amino acid treatments, (W) versus (A) compares the two sources of amino acid.

The mean weights of the 15 chicks in each cage are given in the following table:

				Stack				
		1		2		3		4
Layer	Diet	Weight	Diet	Weight	Diet	Weight	Diet	Weight
1	C	161	L	228	A	224	W	190
2	L	246	W	199	C	169	A	210
3	W	186	A	251	L	241	C	175
4	A	229	C	170	W	183	L	247

a) Enter the data into Genstat and obtain a printout out the factors and variates that have been entered.

[6 marks]

b) Perform the analysis of variance for a Latin square that includes a source for Diets with 3 degrees of freedom.

[3 marks]

c) Using either orthogonal contrasts or nested factors, perform a second analysis of variance in which the Diets sums of squares has been partitioned into 3 single-degree-of-freedom sources corresponding to the three natural groupings of the four diets.

[5 marks]

[Total: 14 marks = 6 + 3 + 5 marks]

Q2 An experiment was conducted to compare the effects of 5 row spacings on the yields of 2 varieties of soya beans. The 2 varieties were assigned to plots using a randomized complete block design of 6 blocks. Each plot was subdivided into 5 subplots and the 5 row spacings randomized to the subplots within each plot. The yields, in unrandomized order, are given in the table below.

Data for a soya bean experiment

Variety	Spacing	Block 1	2	3	4	5	6
OM	18	33.6	37.1	34.1	34.6	35.4	36.1
	24	31.1	34.5	30.5	32.7	30.7	30.3
	30	33.0	29.5	29.2	30.7	30.7	27.9
	36	28.4	29.9	31.6	32.3	28.1	26.9
	42	31.4	28.3	28.9	28.6	*	33.4
B	18	28.0	25.5	28.3	29.4	27.3	28.3
	24	23.7	26.2	27.0	25.8	26.8	23.8
	30	23.5	26.8	24.9	23.3	21.4	22.0
	36	25.0	25.3	25.6	26.4	24.6	24.5
	42	25.7	23.2	23.4	25.6	24.5	22.9

An initial analysis of variance for the data is appended to this exam paper. The data is in the Genstat spreadsheet file *SplSoyMV.gsh* in *G:\Disciplina\Genstat*.

a) Assuming that the unrandomized factors are random and the randomized factors are fixed, what are the maximal expectation and variation models, the variance components and their multipliers and the expected mean squares for the appended analysis?

[9 marks]

b) Based on the appended analysis, what conclusions do you draw about the effects of the factors Variety and Spacing on the yield of soya beans? Make sure you indicate how you reached your conclusions.

[5 marks]

c) Perform diagnostic checking and write a summary of your conclusions from each graph or analysis.

[7 marks]

d) Given your conclusions in b), obtain plots of appropriate tables of means. Using either multiple comparisons or polynomial submodels, examine the differences between the treatments. Give reasons for your choice. If you use multiple comparisons, summarize what you conclude about the differences. If you use polynomial submodels, summarize what you can say about the fitted trend and give the fitted equation(s).

[9 marks]

e) From the appended analysis, do you think that the precision of the Variety main effects is less than it might have been because of the use of a split-plot design? Provide evidence for your conclusion.

[4 marks]

[Total: 34 marks = 9 + 5 + 7 + 9 + 4 marks]

Q3 An experiment is to be conducted to investigate suspected cobalt and copper deficiency of pasture. A flock of sheep is to be run on a paddock suspected to be deficient in these trace elements. The experiment will be a 2^2 factorial experiment made up of the two factors cobalt (administered or not administered) and copper (administered or not administered). The treatments will be randomized to the sheep using a completely randomized design and all sheep will be grazed on the suspect pasture. It is believed that the variance $\left(\sigma^2\right)$ of sheep within a treatment will be about 30 kg. It is desired to be able to detect a change of 4 kg in the cobalt difference between the two copper treatments with power of at least 80% and with significance 0.05.

a) How many replicates should there be of each treatment combination to achieve the desired power?

[9 marks]

b) Using Genstat to generate a layout for the experiment with the number of replicates you have computed in a). Use a seed of 443322. If you have not got a satisfactory answer to a), generate a design for 15 replicates. Obtain a printout of the layout.

[6 marks]

[Total: 15 marks = 9 + 6 marks]

```
Genstat 5  Release 4.1  (PC/Windows NT)              10 May 2000 13:37:30
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)
          ──────────────────────────────────────────────────

                    Genstat 5 Fourth Edition - (for Windows)
                    Genstat 5 Procedure Library Release PL11
          ──────────────────────────────────────────────────

   3   "Data taken from File: D:/ANALYSES/LM/MULTIFAC/SPLSOYMV.GSH"
   4   DELETE [redefine=yes] Blocks,Plots,SubPlots,Variety,Spacing,Yields
   5   FACTOR [modify=yes;nvalues=60;levels=6] Blocks
   6   READ Blocks; frepresentation=ordinal

     Identifier    Values    Missing    Levels
         Blocks        60          0         6

   9   FACTOR [modify=yes;nvalues=60;levels=2] Plots
  10   READ Plots; frepresentation=ordinal

     Identifier    Values    Missing    Levels
          Plots       60          0         2

  13   FACTOR [modify=yes;nvalues=60;levels=5] SubPlots
  14   READ SubPlots; frepresentation=ordinal

     Identifier    Values    Missing    Levels
       SubPlots       60          0         5

  17   FACTOR [modify=yes;nvalues=60;levels=2;labels=!t('OM','B')] Variety
  18   READ Variety; frepresentation=ordinal

     Identifier    Values    Missing    Levels
        Variety       60          0         2

  21   FACTOR [modify=yes;nvalues=60;levels=!(18,24,30,36,42)] Spacing
  22   READ Spacing; frepresentation=ordinal

     Identifier    Values    Missing    Levels
        Spacing       60          0         5

  25   VARIATE [nvalues=60] Yields
  26   READ Yields

     Identifier    Minimum      Mean    Maximum    Values    Missing
         Yields      21.40     28.27      37.10        60          1

  31
  32   BLOCK Blocks/Plots/SubPlots
  33   TREAT Variety*Spacing
  34   ANOVA [FPROB=Y] Yields


34........................................................................

***** Analysis of variance *****

Variate: Yields

Source of variation      d.f.(m.v.)       s.s.        m.s.     v.r.   F pr.

Blocks stratum                5          11.473      2.295     2.61

Blocks.Plots stratum
Variety                       1         542.673    542.673   616.67   <.001
Residual                      5           4.400      0.880     0.31

Blocks.Plots.SubPlots stratum
Spacing                       4         171.774     42.944    14.92   <.001
Variety.Spacing               4          15.049      3.762     1.31   0.284
Residual                     39(1)      112.230      2.878

Total                        58(1)      855.752
```

```
* MESSAGE: the following units have large residuals.

Blocks 6    Plots 1    SubPlots 5          3.74   s.e. 1.37


***** Tables of means *****

Variate: Yields

Grand mean  28.30

   Variety       OM       B
             31.30    25.29

   Spacing    18.00    24.00    30.00    36.00    42.00
             31.47    28.59    26.91    27.38    27.13

   Variety  Spacing    18.00    24.00    30.00    36.00    42.00
        OM              35.15    31.63    30.17    29.53    30.04
         B              27.80    25.55    23.65    25.23    24.22


*** Standard errors of differences of means ***

Table               Variety     Spacing    Variety
                                            Spacing
rep.                     30          12          6
s.e.d.               0.242       0.693      0.909
d.f.                     5          39      43.22
Except when comparing means with the same level(s) of
 Variety                                     0.979
 d.f.                                           39

(Not adjusted for missing values)


***** Missing values *****

Variate: Yields

 Unit   estimate
   29      29.64

Max. no. iterations 2
```

UNIVERSIDADE DE SÃO PAULO
Campus "Luiz de Queiroz"
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"

DEPARTMENTO DE CIÊNCIAS EXATAS

Estatística e Experimentação Agronômica

LCE802 Planejamento de Experimentos e Análise de Dados

FINAL PAPER

PART B — Theory

General Instructions to Candidates   **Time allowed: 2 hours**

1.   This examination is an open book examination. You may need a calculator and statistical tables.

2.   Attempt **ALL** questions. The marks for each question are shown with the question and in the following table.

PART	Question	Max. Mark
A — Practical	1	14
	2	34
	3	15
B — Theory	4	7
	5	6
	6	7
	7	12
	Total	95

3.   In carrying out statistical procedures, use a 5% level of significance and a 95% confidence level, unless otherwise specified. Also, you should carry out the complete procedure even if the assumptions underlying it are not met.

Q4 An experiment is to be conducted on the effects of air and soil heating on the yield of glasshouse tomatoes. There are 8 compartments of a glasshouse to be used in the experiment with these compartments arranged in four rows each containing two compartments. Two air temperatures (55 and 60 C) were randomized to the two compartments in each row. In each compartment there were two long troughs divided in half with the soil in one half of each trough to be heated and the other half not heated. For each trough the half to be unheated was randomly selected. Also, the two temperatures (65 and 75 C) for the heated half was randomly assigned to the two troughs in a compartment. In each half trough 4 tomato plants were planted and the total yield of the four plants was measured.

What are the components of this study?

1.   the observational unit
2.   response variable
3.   unrandomized factors
4.   randomized factors
5.   type of study

[Total: 7 marks]

Q5 An experiment is conducted to investigate the differences between eight cultivars of corn. The experimental area is divided into four rows by four columns. Each row-column area is divided into eight plot, each plot consisting of 3 lines of 20 plants. The eight cultivars are randomized to the eight plots in each row-column area. The centre line of each plot is harvested and the total yield of the 20 plants is recorded.

The components of this experiment are:

1.   the observational unit   a plot
2.   response variable   Yield
3.   unrandomized factor   Rows, Columns, Plots
4.   randomized factors   Cultivar
5.   type of study   ?

a)   What is the experimental structure, including the numbers of levels of each of the factors, for the experiment?

[3 marks]

b)   What type of study is this? Give reasons for your answer.

[1 marks]

c)   Which of the factors are likely to be random and which fixed? Give reasons for your answer.

[2 marks]

[Total: 6 marks = 3 + 1 + 2 marks]

Q6 An experiment that employs a split-plot design with five main-plot treatments assigned using a Latin square and three subplot treatments assigned completely at random would have the following unrandomized structure:

(*5* Rows**5* Columns)/*3* Subplots
= Rows + Columns + Rows.Columns + Rows.Columns.Subplots

What is the Hasse diagram of term marginalities, including degrees of freedom, for these unrandomized terms?

 [Total: 7 marks]

Q7 Suppose the data from a Latin square experiment with t treatments is to be analysed. Let $\psi = E[\mathbf{Y}] = \mathbf{X}_T\tau$, $\mathbf{V} = \sigma^2\mathbf{I}_n$, and $R(\tau\,|\,\mu) = \mathbf{Y}'\mathbf{P}_T\mathbf{R}_G\mathbf{Y}$ where $\mathbf{R}_G = \mathbf{I} - \mathbf{P}_G$,

$\mathbf{P}_G = t^{-2}\mathbf{J}_t \otimes \mathbf{J}_t$ and $\mathbf{P}_T = \mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^{-1}\mathbf{X}_T'$. Also, it is known that $\mathbf{P}_T\mathbf{P}_G = \mathbf{P}_G\mathbf{P}_T = \mathbf{P}_G$ and $trace(\mathbf{P}_T\mathbf{R}_G) = t - 1$.

a)   Prove that $\mathbf{P}_T\mathbf{R}_G$ is symmetric and idempotent, assuming that $\mathbf{P}_T$ and $\mathbf{P}_G$ are symmetric and idempotent.

[4 marks]

b)   Prove that

$$E\left[R(\tau\,|\,\mu)/(t-1)\right] = \sigma^2 + f_T(\psi)$$

where $f_T(\psi) = \sum_{j=1}^{t} t(\tau_j - \bar{\tau}_.)^2 \Big/ (t-1)$, $\bar{\tau}_. = \sum_{j=1}^{t}\tau_j / t$, τ_j is the jth element of the t-vector τ.

[8 marks]

Note that in answering this question you can use the results of any theorems given in the lecture notes, other than the theorem stating the results you are asked to prove, but you must cite the theorem when it is used.

[Total: 12 marks = 4 + 8 marks]