Experimental Design

(Fall 2005)

(Dr. M. Indralingam)

Aids and facilities allowed: All usual (books, notes, solutions to exercises etc.)

The questions were answered by

________________________ ________________________ ________

(name) (signature) (table no.)

Introduction : The examination is a multiple choice. Only one answer should be given for each question. If more than one answer is given, the question is considered "not answered". If, by mistake, a wrong answer is given, it can be corrected by "blackening" out the wrong answer and adding the correct answer and "Corrected by ......" . If there is doubt about the meaning of a correction, the question is considered "not answered".

There are 30 questions distributed over 7 problems with numbering I through VII written in the text. The numbering of the individual questions are given as (1),(2),...,(30) in the text.

Problem	I							II	III		IV			V	
Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Answer															

Problem				VI				VII							
Question	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Answer															

Only this page must be delivered back.

5 points are obtained for a correct answer and -1 point is given for a wrong answer. If a question is unanswered or if the answer is '6' (for 'do not know') 0 points is given. The number of points corresponding to speci_c marks or needed to pass the examination is ultimately determined during censoring.

Problem 1:

It is studied whether different types of powder-mix are influential with respect to the surface quality of certain items for which the powder-mix is used. The powder in question occurs in principally 4 different types (type 0, type 1, type 2 and type 3) which differ by the amount of stearate added. Stearate acts as binder and it is expected to reduce friction.

Experiments were conducted in which the powder in the form of pressed tablets is rubbing a surface film. The amount of loss of material is taken as expression of friction. Since each single experiment is lengthy (several days), 3 sets of measuring equipments (that the laboratory owned) were used. Thus 3 single experiments were carried out in parallel. The design used and the results obtained are shown in the following table. As it appears 3 experiments were carried out within one week. The experiments in one week were all based on powder material from the same batch of raw material. The measurement values are the amount of loss of material (mg) during the experiment. Data and some calculations were:

Powder type	type 0	type 1	type 2	type 3	sum
Week 19 (batch A19)	84	96		105	285
Week 22 (batch L23)	70		72	76	218
Week 24 (batch K10)		100	98	84	282
Week 29 (batch M12)	84	94	93		271
sum	238	290	263	265	1056

$$238 - (285 + 218 + 271)/3 = -20.00$$
$$290 - (285 + 282 + 271)/3 = 10.67$$
$$263 - (218 + 282 + 271)/3 = 6.00$$
$$265 - (285 + 218 + 282)/3 = 3.33$$
$$(285^2 + 218^2 + 282^2 + 271^2)/3 = 93904.67$$
$$84^2 + 96^2 + ... + 94^2 + 93^2 = 94358.00$$
$$1056^2/12 = 92928.00$$

In order to analyse the results for the experiment the usual model was adopted:

$$Y_{ij} = \mu + W_i + p_j + \epsilon_{ij} \quad ; \quad \sum_j p_j = 0$$

where W_i corresponds to week no. i, while p_j corresponds to the effect from powder type j. Finally $\in_{ij}$ Iis the experimental error of the (ij)'th measurement, and it is assumed to have a constant variance denoted by $\sigma_\in^2$

An analysis of variance is carried out and so far the following sums of squares have been calculated:

Source	SSQ
Powder type	210.33
Weeks	976.67
Residual	243.00
Total	1430.00

Q1: First, the variance σ_ϵ^2 should be estimated. The usual estimate is :

1 ☐ $\widehat{\sigma}_\epsilon^2 = 3.19^2$
2 ☐ $\widehat{\sigma}_\epsilon^2 = 6.97^2$
3 ☐ $\widehat{\sigma}_\epsilon^2 = 8.37^2$
4 ☐ $\widehat{\sigma}_\epsilon^2 = 11.40^2$
5 ☐ $\widehat{\sigma}_\epsilon^2 = 15.59^2$
6 ☐ Do not know

Q2: Likewise the parameters p_0, p_1, p_2, p_3 should be estimated. The usual estimate is :

1 ☐ $\{\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{-7.50, 4.00, 2.25, 1.25\}$
2 ☐ $\{\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{-20.00, 10.67, 6.00, 3.33\}$
3 ☐ $\{\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{-8.67, 8.67, -0.33, 0.33\}$
4 ☐ $\{\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{7.00, -15.33, 6.00, 2.33\}$
5 ☐ $\{\widehat{p}_0, \widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{-15.00, 8.00, 4.50, 2.50\}$
6 ☐ Do not know

Q3: The usual test statistics for the difference between the 4 powder types and the corresponding test distribution are found as:

1 ☐ F = 1.96 ~ F(3,6)
2 ☐ F = 0.86 ~ F(3,5)
3 ☐ F = 4.33 ~ F(3,9)
4 ☐ F = 1.44 ~ F(3,5)
5 ☐ F = 2.88 ~ F(3,11)
6 ☐ Do not know

Powder type 0 (p_0) is a standard type, which actually does not contain any stearate, whereas for the three other types stearate is added in different amounts. It could therefore be of interest to split the variation between the powder types in one SSQ from the difference between p_0 and the three other types and one SSQ from the variation between p_1, p_2 and p_3.

These two SSQ's are called $SSQ_{p_0-others}$ and $SSQ_{between\ others}$, respectively.

Q4: We find that

1 ☐ $SSQ_{p_0-others} = 100.00$ with df=2 and $SSQ_{between\ others} = 110.33$ with df=1
2 ☐ $SSQ_{p_0-others} = 200.00$ with df=2 and $SSQ_{between\ others} = 10.33$ with df=1
3 ☐ $SSQ_{p_0-others} = 150.00$ with df=1 and $SSQ_{between\ others} = 60.33$ with df=3
4 ☐ $SSQ_{p_0-others} = 130.00$ with df=1 and $SSQ_{between\ others} = 60.33$ with df=2
5 ☐ $SSQ_{p_0-others} = 200.00$ with df=1 and $SSQ_{between\ others} = 10.33$ with df=2
6 ☐ Do not know

It turns out, that the results obtained from the described experiment are not sufficiently convincing. Therefore it is considered to do the experiment once more. This time, however, the possible variation between the three equipments used should be eliminated.

Q5: If the three equipments are called a, b and c, the following alternatives have been suggested for distributing the single measurements on the equipments a, b or c, in that the rows are still the weeks and the columns are the powder types:

Powder type	0	1	2	3																
1. Week	a	b		c	a	b		c	b	a		c	a	c		b	b	b		b
2. Week	b		c	a	a		c	b	c		b	a	b		c	a	a		a	a
3. Week		b	a	b		b	c	a		b	a	c		b	a	c		a	b	c
4. Week	c	a	b		a	b	c		a	c	b		c	a	b		c	c	c	
	Suggestion 1				Suggestion 2				Suggestion 3				Suggestion 4				Suggestion 5			

1 ☐ Suggestion 1 can be used

2 ☐ Suggestion 2 can be used

3 ☐ Suggestion 3 can be used

4 ☐ Suggestion 4 can be used

5 ☐ Both suggestions 2 and 5 can be used

6 ☐ Do not know

We now suppose that the experiment was in fact carried out using the design shown below where also the allocation of the equipments is shown as indicated in the parentheses as (a), (b) or (c):

Powder type	type 0	type 1	type 1	type 3	sum
Week 19 (batch A19)	84 (c)	96 (a)		105 (b)	285
Week 22 (batch L23)	70 (a)		72 (b)	76 (c)	218
Week 24 (batch K10)		100 (b)	98 (c)	84 (a)	282
Week 29 (batch M12)	84 (b)	94 (c)	93 (a)		271
sum	238	290	263	265	1056

Source	SSQ
Powder types	210.33
Weeks	976.67
Equipments	SSQ_{Equip}
Residual	$SSQ_{Residual}$
Total	1430.00

Q6: Now determine in the above table SSQ_{Equip} and $SSQ_{Residual}$ and the corresponding degrees of freedom (df). The result is:

1 ☐ $SSQ_{Equip} = 40.5$, df=2 and $SSQ_{Residual} = 202.5$, df=3

2 ☐ $SSQ_{Equip} = 40.5$, df=3 and $SSQ_{Residual} = 202.5$, df=2

3 ☐ $SSQ_{Equip} = 54.0$, df=2 and $SSQ_{Residual} = 189.00$, df=3

4 ☐ $SSQ_{Equip} = 54.0$, df=3 and $SSQ_{Residual} = 189.00$, df=2

5 ☐ $SSQ_{Equip} = 40.5$, df=1 and $SSQ_{Residual} = 283.50$, df=4

6 ☐ Do not know

Q7: The variable "Week (batch)" can (shortly) be characterized in one of the following ways:

1 ☐ A deterministic factor
2 ☐ An incomplete block
3 ☐ A complete block
4 ☐ A whole plot
5 ☐ A hierarchial factor
6 ☐ Do not know

Problem 2:

We want to assess whether a number of factors influence the yield of an extraction process. The raw material which is used consists of leaves from a certain plant. The plants studied were grown at 3 randomly selected localities where they grow naturally. From each locality 12 samples were taken, as shown in the data table below.

In the laboratory 3 potential extraction compounds were tested, namely ether, ethanol and acetone, and extraction was performed at 2 different temperatures.

All other sources of variation, including the sequence in which the single measurements were carried out, are completely randomized. The following results were obtained (yield in mg/kg):

Locality		I		II		III		sum
Eter	15^0C	62	47	67	60	46	49	331
	20^0C	76	84	80	89	69	65	463
Ethanol	15^0C	47	39	57	65	57	48	313
	20^0C	46	55	58	61	51	42	313
Acetone	15^0C	81	92	88	84	76	62	483
	20^0C	89	99	112	105	83	88	576
Sum		817		926		736		2479

The three factors are denoted by L_i , t_j and o_k for locality, temperature and extraction compound, respectively, and distinction between fixed and random factors is made by using small or capital letters. The following mathematical model is considered to be adequate for describing the results of the experiment in that the pure experimental error is called *E*:

$$Y_{ijk\nu} = \mu + L_i + t_j + LT_{ij} + o_k + LO_{ik} + to_{jk} + LTO_{ijk} + E_{\nu(ijk)}$$

A computer program for doing analysis of variance has produced the following printout:

Source of variation	Sum of square	Degrees of freedom	Mean square	Expected mean squares
Locality(L)	1515.06	2	757.53	$12\sigma_L^2 + 6\sigma_{LT}^2 + 4\sigma_{LO}^2 + 2\sigma_{LTO}^2 + \sigma_E^2$
Temperature (T)	1406.25	1	1406.25	$18\phi_t + 6\sigma_{LT}^2 + 2\sigma_{LTO}^2 + \sigma_E^2$
L×T	28.50	2	14.25	$6\sigma_{LT}^2 + 2\sigma_{LTO}^2 + \sigma_E^2$
Extraction compound (O)	7942.72	2	3971.36	$12\phi_o + 4\sigma_{LO}^2 + 2\sigma_{LTO}^2 + \sigma_E^2$
L×O	284.28	4	71.07	$4\sigma_{LO}^2 + 2\sigma_{LTO}^2 + \sigma_E^2$
T×O	766.50	2	383.25	$6\phi_{to} + 2\sigma_{LTO}^2 + \sigma_E^2$
L×T×O	199.50	4	49.88	$2\sigma_{LTO}^2 + \sigma_E^2$
Residual	665.49	18	36.97	σ_E^2
Total	12808.30	35		

Q1: The experimental design shown could be called

1 ☐ A mixed fixed and random effect factorial design

2 ☐ An incomplete balanced block design

3 ☐ An unbalanced block design

4 ☐ A split plot design

5 ☐ A nested factorial design

6 ☐ Don't know

Q2: The influence from the temperature, that is the effect t_j, can be estimated directly. In that "+" , "-" correspond to 20°C and 15°C, we find:

1 ☐ $\pm$ (1406.25/18–14.25/6)

2 ☐ $\pm$ (1406.25–14.25)16

3 ☐ $\pm$ 6.25

4 ☐ $\pm$ 1406.25/18

5 ☐ $\pm$ 1352/18 = $\pm$ 75.11

6 ☐ Don't know

Q3: Testing of the interaction between temperature and extraction compound can be done directly by computing one of the following F{test quantities. Which one?

1 ☐ F(6,2) = 766.50/199.50

2 ☐ F(2,4) = 383.25/49.88

3 ☐ F(2,18) = 383.25/36.97

4 ☐ F(2,4) = (766.50/2)/(284.25/4)

5 ☐ $F(f_1,f_2)$ = (383.25–36.97)/(49.88–36.97), f_1 and f_2 are computed in a special way (approximate F–test)

6 ☐ Don't know

Q4: One can (based on the given model) estimate a component of variance corresponding to the interaction between locality and extraction compound. The result is

1 ☐ σ^2_{LO} = (284.28–199.50)/(4–2)

2 ☐ σ^2_{LO} = (284.28–199.50)/(4+2)

3 ☐ σ^2_{LO} = (71.07–49.88)/(4–2)

4 ☐ σ^2_{LO} = 71.07/4

5 ☐ σ^2_{LO} = (71.07–49.88)/4

6 ☐ Don't know

The *LTO*–effect is tested firstly at a 5% level of significance in order, subsequently, to test the two–factor interactions LT_{ij}, LO_{ik} and to_{jk}.

Q5: After the initial test of the *LTO* -effect an improved estimate for the pure experimental error variance can be computed. Which of the following proposals should be preferred?

1 ☐ σ^2_E = (49.88+36.97)/(2)

2 ☐ σ^2_E = (199.50+665.49)/(4+18)

3 ☐ σ^2_E = 36.97+(49.88–36.97)/(18–4)

4 ☐ σ^2_E = 49.88/4

5 ☐ σ^2_E = 48.05 = 6.93^2

6 ☐ Don't know

After testing *LTO* and afterwards testing the terms LT_{ij}, LO_{ik} and to_{jk} it is concluded that the following model describes the data well:

Q6:Further we are interested in assessing how different localities influence the yield. Which estimate do you think is the relevant one.

1 ☐ $(\hat{L}_1, \hat{L}_2, \hat{L}_3) = (\frac{817}{12} - \frac{2479}{36} \;\; , \;\; \frac{926}{12} - \frac{2479}{36} \;\; , \;\; \frac{736}{12} - \frac{2479}{36})$

2 ☐ $(\hat{\mu} + \hat{L}_1, \hat{\mu} + \hat{L}_2, \hat{\mu} + \hat{L}_3) = (\frac{817}{12}, \frac{926}{12}, \frac{736}{12})$

3 ☐ $\hat{\sigma}_L^2/12 = 1515.06$

4 ☐ $\hat{\sigma}_L^2 = 12 \cdot 1515.05 + 42.06$

5 ☐ $\hat{\sigma}_L^2 = (1515.06/2 - 42.06)/12$

6 ☐ Don't know

Problem 3:

An experiment is to be carried out. The purpose is to evaluate how the friction (loss of energy) in a certain type of transmission (a bearing in combination with a gearing) depends on the following factors, which at the same time are organized in accordance with their expected importance:

A : Type of grease (grafite/oil)
B : Radial load on bearing (low/high)
C : Horizontal load of bearing (low/high)
D : Speed of rotation in transmission (low/high)
E : Temperature during operation (60_0C,90_0C)
F : Moment transmitted in transmission (low/high)

As can be seen the factors are to be assessed on only two levels.
It is assumed that the factors in question essentially act additively in relation to the measured friction and in the first experiment and using this assumption it is decided to take as few single Measurements as possible .

Q1: Which of the following designs would you recommend under the describe

1 ☐ A $2^{-1} \times 2^6$ faktorial in 4 blocks constructed from I_1=ABC and I_2=DEF.

2 ☐ A fractional $2^{-1} \times 2^6$ factorial design.

3 ☐ A fractional $2^{-3} \times 2^6$ factorial design.

4 ☐ A partially confounded $2^{-2} \times 2^6$ factorial design.

5 ☐ A $1/4 \times 2^6$ factorial design.

6 ☐ Don't know

Q2: It is decided to carry out an experiment based on the complete factorial for the factors A, B and C and on introducing the factors D, E and F into this factorial, resulting in a design for 6 factors in 8 measurements. Which of the following generator equations do you think is best suited as basis for the design under construction?

1 ☐ D=AC, E=ABD and F=BCD

2 ☐ I_1=ABC, I_2=DEF and I_3=ABCDEF

3 ☐ D=ABC, E=ABCE and F=CDE

4 ☐ D=ABE, E=ABD, F=DE

5 ☐ None of the suggestions are adequate

6 ☐ Don't know

Q3: A design based on the generator equations D=AB, E=AC and F=BC is considered. In this design the factor F will be confounded as follows:

1 ☐ F=BC=ACEF=ABDF

2 ☐ F=ABDF=ACEF=BC=BCDEF=ABE=ACD=DE

3 ☐ F=DF=ABF=EF=ACF=BC

4 ☐ $F=ABDF=ABDF^2=ACEF=ACEF^2=ABDF=ABDF^2=BCF$

5 ☐ None of the suggestions are adequate

6 ☐ Don't know

Q4: The above generator equations (D=AB, E=AC and F=BC) are used to construct an experiment. Which of the following possibilities correspond to these equations?

1 ☐ (1) ade bdf abef cef acdf bcde abc

2 ☐ (1) adf bde abef cef acde bcdf abc

3 ☐ (1) ad be abdf cf acde bdcf abce

4 ☐ (1) adef be ab ce acde bcef abcd

5 ☐ (1) abdf bded abef cdef acde bcdf abc

6 ☐ Don't know

Q5: The same generator equations (D=AB, E=AC and F=BC) are used to construct the following design:

de a abdf bef acef cdf bc abcde,

as one of the 8 possibilities that can be chosen. For reasons of randomization the experiment has to be carried out in 2 blocks each containing 4 single measurements by confounding blocks with the interaction term ABC.

Which one of the following designs corresponds to this choice?

1 ☐ a de abdf acef and bc bef cdf abcde

2 ☐ de abdf cdf bc and a bef acef abcde

3 ☐ de abdf acef bef and a bc cdf abcde

4 ☐ de abdf acef bc and a bef cdf abcde

5 ☐ a abdf acef de and de bef cdf abcde

6 ☐ Don't know

Problem 4 :

A $\frac{1}{8} x\ 2^6$ factorial has been carried out with factors A, B, C, D, E and F. The design is constructed by introducing the factors D, E and F into a complete factorial defined by the factors A, B and C by means of the relations D = AB, E = AC and F = BC. The following experiment was chosen in that standard notation is used:
1

The following experiment was chosen in that standard notation is used:

a	de	abdf	bef	acef	cdf	abcde	bc

Q1: The principal fraction corresponding to the fraction shown is:

1 ☐ (1) ade abdf bef acef df abcde bc

2 ☐ (1) ade bdf abef cef acdf bcde abc

3 ☐ (1) a b ab c ac bc abc

4 ☐ (1) ad bd abd ce ace bcf abcf

5 ☐ (1) ade bde abdf cef acde bcdef abcf

6 ☐ Don't know

For the 8 measurements actually carried out the following results were obtained

a=19	de=23	abdf=9	bef=25	acef=16	cdf=15	abcde=8	bc=31

Q2: The sum of squares corresponding to the main effect for the factor 'F' is found to be:

1 ☐ $[(19+9+16+8)-(23+25+15+31)]^2/4$

2 ☐ $[-19-23+9+25+16+15-8-31]^2/8$

3 ☐ $[-19-23+9+25+16+15-8-31]^2/4$

4 ☐ $[(-19-23+9+25)-(16+15-8-31)]^2/8$

5 ☐ $[(-19-23+9+25)^2+(16+15-8-31)^2]/8$

6 ☐ Don't know

Problem 5:

An investigation concerning a fermentation method has been carried out. In the experiment interest was concerned about which factors might influence the purity of the product produced and to which degree. With the method in question unwanted by-products closely related to the primary product are often experienced. 5 factors were considered in the experiment, and the measurements were carried out on two parallel sets of equipment (I and II) in order to carry out the whole experiment within a reasonably short period of time. The two sets of equipments are thought to act as blocks.

Name	Meaning	Levels 0	Levels 1
A :	Temperature	18°C	24°C
B :	pH in solution	6.80	7.40
C :	Stabiliser	0%	5%
D :	Concentration of fertiliser	2%	4%
E :	Concentration of raw material in sample	30%	50%

The measured response was the relative content of by-product in the product given as %. The design, the data and some preliminary computations are given in the following table:

Design : '−1' corresponds to level 0, and '+1' to level 1							Data			
A	B	C	D=−AC	E=−BC	Equip=ABC	Code	%	Yates Algorithm		
−1	−1	−1	−1	−1	−1 ~ I	(1)	3.7	14.9	24.8	57.6
+1	−1	−1	+1	−1	+1 ~ II	a d	11.2	9.9	32.8	15.8
−1	+1	−1	−1	+1	+1 ~ II	b e	4.6	17.9	8.2	-8.0
+1	+1	−1	+1	+1	−1 ~ I	ab de	5.3	14.9	7.6	-1.4
−1	−1	+1	+1	+1	+1 ~ II	c de	8.4	7.5	-5.0	8.0
+1	−1	+1	−1	+1	−1 ~ I	a ce	9.5	0.7	-3.0	-0.6
−1	+1	+1	+1	−1	−1 ~ I	bc d	4.2	1.1	-6.8	2.0
+1	+1	+1	−1	−1	+1 ~ II	abc	10.7	6.5	5.4	12.2

That is

Equipment I			
(1)	abde	ace	bcd

Equipment II			
ad	be	cde	abc

It is suggested that the following model can be used:

$$Y_{ijklmn} = \mu + A_i + B_j + AB_{ij} + C_k + D_l + E_m + Equip_n + \epsilon_{ijklmn}$$

where all indices, i - o, in the usual way may take one of the values {0,1}. All interactions except AB_{ij} are assumed to be unimportant (zero). Standard assumptions such{ as, for example, $A_0 = - A_1$, $B_0 = - B_1$ are used.

Q1: : In this design the defining relation and the alias-relation for the factor A (with proper signs) are respectively:

1 ☐ $I = -AC = -BC = AB$ and $A = -C = -ABC = B$

2 ☐ $I = -ACD = -BCE = -ABD$ and $A = -CD = -ABCE = -BD$

3 ☐ $I = -ACD = -BCE = ABDE$ and $A = +CD = +ABCE = -BDE$

4 ☐ $I = -ACD = -BCE$ and $A = -CD = -ABCE$

5 ☐ $I = -ACD = -BCE = ABDE$ and $A = -CD = -ABCE = BDE$

6 ☐ Don't know

Q2: We are interested in estimating the effects of the model. The Effect is defined as $E = E_{-} - E_0$. For this quantity we find

1 ☐ $\hat{E} = 2/8 = 0.25$

2 ☐ $\hat{E} = -2/8 = -0.25$

3 ☐ $\hat{E} = -2/4 = -0.5$

4 ☐ $\hat{E} = 4.2/4 = 1.05$

5 ☐ $\hat{E} = -12.2/4 = -3.05$

6 ☐ Don't know

Q3: The sum of squares corresponding to the blocks of the experiment is found to be

1 ☐ $SSQ_{blocks} = 37.21$

2 ☐ $SSQ_{blocks} = 18.605$

3 ☐ $SSQ_{blocks} = 9.303$

4 ☐ $SSQ_{blocks} = 414.72$

5 ☐ $SSQ_{blocks} = 18.467$

6 ☐ Don't know