IMM - DTU Statistical Design and Analysis of Experiments
H. Spliid, October 20, 2010

Solution to examination set 1994, part 1.

Problem I

1): The design is a symmetrically balanced incomplete block design. Compare e.g. with exercise 39 about the twin-cows or exercise 47 about the noise experiment with litters of pigs. **Answer = 2**.

2) The experiment becomes more precise with two single experiments pr plate. See exercise 68 about the plates with wells containing penicillin in which we saw that comparisons of treatments had smaller uncertainty in an incomplete balanced block design. **Answer = 4**.

3) Use orthogonal polynomials on the Q-values (see text book or slides):

$$SSQ_{Contrast} = \frac{k \cdot [C]^2}{\lambda \cdot a(c_1^2 + c_2^2 + \ldots + c_a^2)}$$

where $\lambda = 1$, $k = 2$, $a = 3$. See slides for orthogonal polynomial coefficients (or textbook table X)

$C_{linear} = -1 \cdot Q_1 + 0 \cdot Q_2 + 1 \cdot Q_3 = 3.85$ and $C_{squa} = 1 \cdot Q_1 - 2 \cdot Q_2 + 1 \cdot Q_3 = -1.35$.

$$SSQ_{linear} = \frac{2(3.85)^2}{1 \cdot 3 \cdot (1^2 + 0^2 + 1^2)} = 4.94$$

$$SSQ_{squa} = \frac{2(-1.35)^2}{1 \cdot 3 \cdot (1^2 + 2^2 + 1^2)} = 0.20$$

Answer = 1

4) The sequence should be arranged in such a way that balance with respect to both rows and columns is attained. Compare e.g. with questions 3a and 3b in exercise 39. **Answer = 3**.

Problem II

5) There are three factors of which “Locality” is random, while the two other factors are deterministic. The design is completely randomised. **Answer = 1**

6) $\widehat{t}_j = \overline{y}_{.j..} - \overline{y}_{tot}$, that is

$$\widehat{t}_2 = (463 + 313 + 576)/18 - 2479/36 = 6.25 \text{ and } \widehat{t}_1 = -\widehat{t}_2$$

Answer = 3

7) From the EMS-table it is seen that the quotient:

$$\frac{6\phi_{to} + 2\sigma^2_{LTO} + \sigma^2_E}{2\sigma^2_{LTO} + \sigma^2_E}$$

points out that we shall use $F_{to} = s^2_{to}/s^2_{LTO} = 383.25/49.88 \in F(2, 4)$:
. **Answer = 2**

8)From the EMS-table it is seen that the estimate is:

$$\widehat{\sigma}^2 = (s^2_{LO} - s^2_{LTO})/4 = (71.07 - 49.88)/4$$

Answer = 5

9) The pooled estimate for σ^2_E is:

$$(SAK_{LTO} + SSQ_E)/(f_{LTO} + f_E) = (199.5 + 665.49)/(4 + 18)$$

Answer = 2

10) Since Locality (L) is a random factor it is the component of variance σ^2_L which is of interest. In the model $\mathrm{E}\{\mathrm{S}^2_\mathrm{L}\} = 12\sigma^2_\mathrm{L} + \sigma^2_\mathrm{E}$, which is seen from the original EMS-table by putting $\sigma^2_{LT} = 0$, $\sigma^2_{LO} = 0$ and $\sigma^2_{LTO} = 0$. The residual variation give has been computed as $28.5 + 284.28 + 119.50 + 665.49 = 1177.77$ with $2 + 4 + 4 + 18 = 28$ degrees of freedom. In this way we get $\widehat{\sigma}^2_L = (s^2_L - s^2_E)/12$:
. **Answer = 5**

Problem III

11) The whole design is replicated on two different days. One day in this way becomes a block:
. **Answer = 3**

12) It is seen that one day is a 'superblock' while one temperature level essentially also represents an 'underblock'. There is complete randomisation within temperature levels, but not accross temperature levels. The design is a typical split plot design in which the temperature is the 'whole plot' factor: **Answer = 4**

13) By carrying out the design twice (that is on two different days) it is accomplished that it is possible to construct a test for the temperature effect. Answer no. 2: Rubbish. Answer possibility 3: It is possible to test the treatment time in all circumstances since it is easy to randomise. Answer possibility 4: Rubbish, there is allready balance if all temperatures are used om both days. Answer possibility 5: It is not a prime purpose, the temperature and the treament time are the factors of interest: **Answer = 1**.

14) The temperature is to be tested against the"whole plot" error, which is computed as the DT-interaction: **Answer = 2**.

15) The lack of randomisation with respect to temperature makes the experiment unsuited for assessing (testing) the temperature effect: **Answer = 3**

Problem IV

16) When computing $SSQ_{P(TDB)}$ the term 'outside' the parentheses (here "P") is to be combined with all possible combinations of factors inside the parentheses, that is, for example:

$$P(TDB) = P + PT + PD + PTD + PB + PTB + PDB + PTDB$$

giving $SSQ_{P(TDB)} = 129.89$ degrees of freedom $f_{P(TDB)} = 36$. **Answer = 2**

17)

.

	i	j	k	l	m
Model	3	2	3	3	2
t_i	0	2	3	3	2
d_j	3	0	3	3	2
td_{ij}	0	0	3	3	2
B_k	3	2	1	3	2
TB_{ik}	1	2	1	3	2
DB_{jk}	3	1	1	3	2
TDB_{ijk}	1	1	1	3	2
$P(TDB)_{l(ijk)}$	1	1	1	1	2
$G(PTDB)_{m(ijkl)}$	1	1	1	1	1

covering the i-column and multiplying for all terms containing the index 'i' we find

$$\mathrm{E}\{S^2_t\} = 36\phi_t + 0\phi_{td} + 12\sigma^2_{TB} + 6\sigma^2_{TDB} + 2\sigma^2_{P(TDB)} + \sigma^2_G$$

. **Answer = 1**

Solution to examination set 1994, part 2.

Problem V

18) Provided the factor effects are additive one can as starting point use a complete 2^3 factorial design for the factors A, B and C (the underlying complete factorial) and introduce the other 3 factors D, E and F into it. The resulting design is fractional $2^{-3} \times 2^6$ factorial design. **Answer = 3**

19) None of the suggestions are correct which can be seen by writing the standard sequence for model terms (I,A,B,AB,C,AC,BC,ABC) in the 2^3 design, and subsequently introducing D, E and F by confounding with 3 different effects among (I,A,B,AB,C,AC,BC,ABC) (not I though!): in practice only AB, AC, BC or ABC. **Answer = 5**

20) The cosen generator equations yield:

I = ABD = ACE = BCDE = BCF = ACDF = ABEF = DEF

(that is $I = I_1 = I_2 = I_1I_2 = I_3 = I_1I_3 = I_2I_3 = I_1I_2I_3$) by which we for F find:

F = ABDF = ACEF = BCDEF = BC = ACD = ABE = DE

. **Answer = 2**

21) The chosen contrasts with index equations corresponding to the principal fraction are

$$I_1 = ABD \Rightarrow (i+j+l)_2 = 0$$
$$I_2 = ACE \Rightarrow (i+k+m)_2 = 0$$
$$I_3 = BCF \Rightarrow (j+k+n)_2 = 0$$

The principal fraction can be determined by deriving 3 linearly independent solutions to the index equations. If the solutions are called x, y and z the principal fraction becomes (1),x,y,xy,z,xz,yz,xyz.

A	B	C	D	E	F	treatment
i	j	k	l	m	n	
1	0	0	1	1	0	x = adc
0	1	0	1	0	1	y = bdf
0	0	1	1	1	0	z = cef

. **Answer = 1**

22) With the confounding chosen (I = ABC) the two blocks have index-values $(i+j+k)_2$ (0 or 1). By finding $(i+j+k)_2$ for the individual experiments the two blocks are constructed. For the experiment 'de' $(i+j+k)_2 = 0$, for 'a' it is $(i+j+k)_2 = 1$. Thus 'de' is going in the block '0', while 'a' is going into block '1'. Alternatively one can use the 'even-uneven' rule or the tabular method for distributing the treatments to the blocks. **Answer = 4**

Problem VI

23)The only correct answer is D = ABC. The remaining suggestions are more or less rubbish. **Answer = 4** .

24) We have to determine 3 lienarly independent solutions to the equation $(i+j+2k)_3$=l.

i	j	k	l	treatment
1	0	0	1	x=ad
0	1	0	1	y=bd
0	0	1	2	z=cd^2

This corresponds to possibility 1. The other possibilities do not fulfill $(i+j+2k)_3$=l.

. **Answer = 1**

25) For the block confounfounding I=AB^2C the index equation is $(i+2j+k)_3$=0. Furthermore we have D=ABC^2, which then gives rise to $(i+j+2k+2l)_3$=0, corresponding to I=ABC^2D^2. In the same way as above we find (only two now) solutions:

i	j	k	l	treatment
1	0	2	2	x=ac^2d^2
0	1	1	0	y=bc

It is (coincidently) possibility 2. The other possibilities do <u>not</u> fulfill $(i+2j+k)_3=0$ and $(i+j+2k+2l)_3=0$. **Answer = 2**

26) With the confounding (the generator equation) $D=ABC^2$ we get the defining relation $I=ABC^2D^2$ by which the alis relation for the factor A becomes $A=A^2BC^2D^2=A^3BC^2D^2$, which is reduced to $A=AB^2CD=BC^2D^2$. **Answer = 5**

27) Batch is a (typical) block effect. **Answer = 3**