H. Spliid, October 20, 2010

Solution to exercise 50

The design is a split plot design, where one block is defined by the amount of paint produced for one type on one day (for example 'A' on day 1). This block is called a "whole plot" since the factor 'type' is fixed in that block. There are 4 whole plots per day. The block size is 3, which corresponds to the fact that 3 methods of application are used for each block (that is per paint type per day) in a completely randomized fashion.

The factors paint type (t) and application method (m) are deterministic factors, while day (D) is a random factor. These 3 factors are crossed in relation to each other in the design.

The (restricted) randomisation is shown in the text where the numbering indicates a possible sequence of the 36 single measurements:

Randomisation					
Day	Application method	Type of paint A	B	C	D
	1	3	11	5	7
1	2	1	12	4	9
	3	2	10	6	8
	1	20	13	17	24
2	2	21	14	16	22
	3	19	15	18	23
	1	28	31	36	27
3	2	29	32	34	26
	3	30	33	35	25

For example, the 3 measurements with numbers 13, 14 and 15 are all carried out on day nr II in the same block $\sim$ whole plot.

The factor "type" is called t_j and the factor "method" m_k, while the random factor "Day" is called D_i. Finally the residual error term is denoted $E_{l(ijk)}$.

We may therefore write the mathematical model for the experimental results as:

$$Y_{ijkl} = \mu + D_i + t_j + DT_{ij} + m_k + DM_{ik} + tm_{jk} + DTM_{ijk} + E_{l(ijk)}$$

in that indices $i = 1, 3$, $j = 1, 4$, $k = 1, 3$ and $l = 1$ (since there is only one observation per (ijk)–combination. The term DT_{ij} corresponds to the whole plot error (the variation between the above discussed blocks).

The structure $m_k + DM_{ik} + tm_{jk} + DTM_{ijk}$ defines the "split plot", and the "split plot error" is the DTM_{ijk} term. One split plot is precisely defined by one (i, j, k) combination, which essentially is a small block nested within the whole plot (allthough it only consists of one measurement in the present case).

Term in model	3 i	4 j	3 k	1 l	σ_D^2	ϕ_t	σ_{DT}^2	ϕ_m	σ_{DM}^2	ϕ_{tm}	σ_{DTM}^2	σ_E^2
					EMS contributions							
D_i	1	4	3	1	12		3		4		1	1
t_j	3	0	3	1		9	3				1	1
DT_{ij} (whole plot error)	1	1	3	1			3				1	1
m_k	3	4	0	1				12	4	0	1	1
DM_{ik}	1	4	1	1					4		1	1
tm_{jk}	3	0	0	1						3	1	1
DTM_{ijk} (split plot error)	1	1	1	1							1	1
$E_{l(ijk)}$	1	1	1	1	(no degrees of freedom)							1

There cannot be computed a sum of squares for the term $E_{l(ijk)}$ (0 degrees of freedom).

Normally you do not test factorial effects "across the line", such that the whole–plot factor (t_j) is generally tested by means of the whole–plot error while the split–plot factor (m_k) is generally tested by means of the split–plot error, (if possible, that is when the DM-term is small).

So we start with the term DM_{ik} in the model (which is difficult to interprete anyway) and we hope that it is not significant.

ANOVA			
Source of variation	SSQ	Degr. fr.	s^2
Days $= D_i$	2.167	$3-1=2$	1.084
Types $= t_j$	308.000	$4-1=3$	102.667
DT_{ij}	5.833	$(3-1)\cdot(4-1)=6$	0.972
Methods $= m_k$	228.667	$3-1=2$	114.333
DM_{ik}	1.667	$(3-1)\cdot(3-1)=4$	0.417
tm_{jk}	12.667	$(4-1)\cdot(3-1)=6$	2.111
DTM_{ijk}	11.000	$(3-1)\cdot(4-1)\cdot(3-1)=12$	0.917
$E_{l(ijk)}$	—	0	—

Test of the DM–effect is

$$\mathrm{F_{DM}} = \frac{1.667/4}{11.000/12} = 0.45 < \mathrm{F}(4,12)_{0.95} = 3.26$$

and, hence, it is reasonable to exclude the DM effect from the model.

From the EMS table it is seen, that the sum of squares for DM can be pooled with the sum of squares for the DTM effect (but not with the residual variation E). The resulting new DTM-SSQ is then used to test tm_{jk} and m_k.

SSQ_{DTM}(pooled) $= (11.000 + 1.667)$ with $(12 + 4)$ degrees of freedom.

$$\mathrm{F_{tm}} = \frac{12.667/6}{(11.000 + 1.667)/(12 + 4)} = 2.66 < \mathrm{F}(6, 16)_{0.95} = 2.74$$

$$\mathrm{F_{m}} = \frac{228.667/2}{(11.000 + 1.667)/(12 + 4)} = 144.4 >> \mathrm{F}(2, 16)_{0.99} = 6.23$$

$$\mathrm{F_{t}} = \frac{308.000/3}{5.833/6} = 105.6 >> \mathrm{F}(3, 6)_{0.99} = 9.78$$

$$\mathrm{F_{D}} = \frac{2.167/2}{5.833/6} = 1.11 < \mathrm{F}(2, 6)_{0.95} = 5.14$$

From this it can be concluded that there is not a significant interaction between types of paint and method of application (besides the F-value is much smaller than the F-values for the main effects). Also, the variation beween days is not significant.

The factors type of paint and method of application are both highly significant.

We may assess the whole plot variance in relation to the split plot variance by comparing the DT term with the DTM term:

$$\mathrm{F_{DT/DTM}} = \frac{5.833/6}{(11.000 + 1.667)/(12 + 4)} = 1.23 < \mathrm{F}(6, 16)_{0.75} = 1.48$$

and it is seen that the whole plot error is not very pronounced.

It can be concluded that the following model is a good description of the data obtained in the experiment:

$$Y_{ijkl} = \mu + DT_{ij} + t_j + m_k + (DTM_{ijk} + E_{l(ijk)})$$

where DT_{ij} represents the whole plot error. Generally the whole plot error is not removed from the model, even if it is small because there inevitably must be some (small or large) variation between the whole plots and, furthermore, it may turn out larger in a later experiment.

Estimates for the significant effects are:

$$\hat{m}_1 = \frac{270 + 267 + 271}{12} - \frac{2532}{36} = -3.00, \;\; \hat{m}_2 = -0.17, \;\; \hat{m}_3 = 3.17$$

$$\hat{t}_1 = -2.33, \;\; \hat{t}_2 = -1.00, \;\; \hat{t}_3 = 5.00, \;\; \hat{t}_4 = -1.67$$

$$\hat{\mu} = 2532/36 = 70.33$$

By means of the data and the EMS table one can also estimate the split plot variance from:

$$\hat{\sigma}^2_{DTM+E} = \frac{1.667 + 12.667 + 11.000}{4 + 6 + 12} = 1.15 = 1.07^2$$

This estimate includes both the split plot error DTM variance and the measurement error E variance.

Generally one would not include the variance between days or the whole plot variance in this estimate. They could be given as separate estimates (how?).

. o O o .