

Solution to exercise 33

Question 1

The design is an incomplete balanced block design. We use the following model:

$$Y_{ij} = \mu + T_i + f_j + E_{ij}$$

and it is assumed that there is no interaction between twin pair and diet (T_i og f_j respectively).

By using twins pairs (small homogeneous blocks) the best possible precision is obtained (small experimental error variance).

Question 2

$$\begin{aligned} b &= 10 \\ t &= 5 \\ k &= 2 \\ r &= 4 \\ N &= 20 \\ \lambda &= 1 (= r(k-1)/(t-1) = 4(2-1)/(5-1)) \end{aligned}$$

$$SSQ_{tot} = 210.17 - 3058.09/20 = 57.27, \quad f = 19$$

$$SSQ_{blocks,unadjusted} = 361.03/2 - 3058.09/20 = 27.61, \quad f = 9$$

$$\begin{aligned} Q_1 &= 2.50 - 4.4/2 - 3.7/2 - 6.1/2 - 2.1/2 &= -5.65 \\ Q_2 &= 7.50 - 2.1/2 - 9.2/2 - 3.2/2 - 6.7/2 &= -3.10 \\ Q_3 &= 11.50 - 3.7/2 - 9.3/2 - 3.9/2 - 3.2/2 &= 1.45 \\ Q_4 &= 14.60 - 6.1/2 - 6.7/2 - 3.9/2 - 6.7/2 &= 2.90 \\ Q_5 &= 19.20 - 4.4/2 - 9.3/2 - 6.7/2 - 9.2/2 &= 4.40 \end{aligned}$$

Sum of square for diets based on the Q's is now:

$$SSQ_{diet,adjusted} = \frac{2(5.65^2 + 3.10^2 + 1.45^2 + 2.90^2 + 4.40^2)}{(\lambda t - 1)^{1.5}} = 28.56, \quad f = 4$$

$$SSQ_{residual} = 57.27 - 27.61 - 28.56 = 1.10; \quad f = 19 - 9 - 4 = 6$$

ANOVA

Source of variation	SSQ	f	F -value
Twin pairs (blocks)	27.61	9	
Diets (the factor)	28.56	4	$\frac{28.56/4}{1.10/6} = 38.95$
Residual	1.10	6	is strongly significant
Total	57.27	19	

Conklusion: the diet has a significant influence on the intake of the animals.

Question 3

A contrast corresponding to a linear polynomium (1st order):

$$\begin{aligned}
 C_L &= -2 \cdot (-5.65) - 1 \cdot (-3.10) + 0 \cdot 1.45 + 1 \cdot 2.90 + 2 \cdot 4.40 \\
 &= 26.1 \\
 SSQ_L &= k \cdot C_L^2 / [(2^2 + 1^2 + 0^2 + 1^2 + 2^2)\lambda t] \\
 &= 2 \cdot 26.1^2 / (10 \cdot 1 \cdot 5) = 27.25 \\
 SSQ_{higher\ order} &= 28.56 - 27.25 = 1.31, \quad f = 3
 \end{aligned}$$

This splitting of the variation from diets gives the following ANOVA

Source of variation	SSQ	f	F -value
Twin pairs (blocks)	27.61	9	
Linear polynomial diets	27.25	1	$\frac{27.25/1}{1.10/6} = 148.64$
2nd, 3rd and 4th order	1.31	3	$\frac{1.31/3}{1.10/6} = 2.38$
Residual	1.10	6	
Total	57.27	19	

Test of the linear term gives: $F(1, 6) = \frac{27.25/1}{1.10/6} = 148.64$, which is highly significant.

Test of all terms of order higher than linear gives: $F(3, 6) = \frac{1.31/3}{1.10/6} = 2.38$, which is hardly significant.

Conclusion: predominantly linear dependence (1st order polynomial is chosen).

Question 4

The regression model is:

$$Y(z) = \alpha_0 + \alpha_1 \cdot P_1(z) + E$$

The linear polynomium is:

with x -values $\{0.5, 1.0, 1.5, 2.0, 2.5\}$ then $z = (x - 1.5)/0.5 = 2x - 3$ and for a 5-point 1st order polynomial:

$$P_1(z) = \lambda_1 \cdot z = 1 \cdot \left(\frac{x - 1.5}{0.5}\right) = 2x - 3$$

If the Q 's are computed as shown above and a polynomial contrast is computed as

$$C_{degree} = c_1 \cdot Q_1 + c_2 \cdot Q_2 + \dots + c_t \cdot Q_t$$

then the estimate for the corresponding coefficient is

$$\hat{\alpha}_{degree} = \frac{C_{degree}}{[c_1^2 + \dots + c_t^2]} \cdot \frac{k}{\lambda t}$$

This gives the estimate for the linear coefficient:

$$\hat{\alpha}_1 = \frac{C_l \cdot 2}{[2^2 + 1^2 + 0 + 1^2 + 2^2]\lambda t} = \frac{52.2}{50} = 1.044$$

and for the constant

$$\hat{\alpha}_0 = \bar{y} = 55.3/20 = 2.765$$

The final polynomial then becomes

$$\hat{Y}(z) = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot z = 2.765 + 1.044 \cdot z$$

or by inserting $z = 2x - 3$

$$\hat{Y}(x) = 2.765 + 1.044 \cdot (2x - 3) = 2.088 \cdot x - 0.367$$

. o O o .

To repeat: the general formulas for sum of squares and coefficient for the *degree* order orthogonal polynomial are simply

$$C_{degree} = [c_1 \cdot Q_1 + c_2 \cdot Q_2 + \dots + c_t \cdot Q_t]$$

$$SSQ_{degree} = \frac{C_{degree}^2}{c_1^2 + c_2^2 + \dots + c_t^2} \cdot \frac{k}{\lambda t} \text{ with d.f.} = 1$$

$$\hat{\alpha}_{degree} = \frac{C_{degree}}{c_1^2 + c_2^2 + \dots + c_t^2} \cdot \frac{k}{\lambda t} \text{ , which also is } SSQ_{degree}/C_{degree}$$

. o O o .