Chapter 19 Split-Plot Designs

Split-plot designs are needed when the levels of some treatment factors are more difficult to change during the experiment than those of others. The designs have a nested blocking structure: split plots are nested within whole plots, which may be nested within blocks.

Example. An experiment is to compare the yield of three varieties of oats (***factor A with a=3 levels***) and four different levels of manure (***factor B with b=4 levels***). Suppose 6 farmers agree to participate in the experiment and each will designate a farm field for the experiment (***blocking factor with s=6 levels***). Since it is easier to plant a variety of oat in a large field, the experimenter uses a split-plot design as follows:

1. To divide each block into three equal sized plots (***whole plots***), and each plot is assigned a variety of oat according to a randomized block design.
2. Each whole plot is divided into 4 plots (***split-plots***) and the four levels of manure are randomly assigned to the 4 split-plots.

A model for such a split-plot design is the following:

$$
\begin{aligned}
Y_{hij} = \mu + \theta_h + \alpha_i + \epsilon^W_{i(h)} \\
+ \beta_j + (\alpha\beta)_{ij} + \epsilon^S_{j(hi)},
\end{aligned}
$$

where $\epsilon^W_{i(h)}$ $i.i.d. \sim N(0, \sigma^2_W)$, $\epsilon^S_{j(hi)}$ $i.i.d. \sim N(0, \sigma^2_S)$ and are mutually independent, h=1, 2, ..., s, i=1, 2, ..., a, j=1, 2, ..., b.

Note the nested blocking structure: whole plots are nested within the blocks, and split-plots are nested within the whole plots.

Two kinds of errors: $\epsilon^{W}_{i(h)}$ representing the random effects of the whole plots, and $\epsilon^{S}_{j(hi)}$ representing the random effects of split-plots and random noises.

ANOVA Table for split-plot designs:

Source of Variation	D.F.	SS	MS	Ratio
Block	s-1	SS θ	—	—
A	a-1	SSA	MSA	MSA/MSE^w
Whole-plot error	(s-1)(a-1)	SSE^w	MSE^w	
B	b-1	SSB	MSB	MSB/MSE^s
AB	(a-1)(b-1)	SSAB	MSAB	MSAB/MSE^s
Split-plot error	a(b-1)(s-1)	SSE^s	MSE^s	
Total	n-1	Sstotal		

For calculations of the sums of squares, please see table 19.2 on page 679.

Note for testing of equal effects of factor A, the whole-plot mean square error is used. It should also be used for main effect contrasts of factor A. For tests or main effect contrasts of factor B, or AB interaction contrasts, the split-plot mean square error is used.

For main effects and interaction contrasts, the methods of multiple comparison of Bonferroni, Scheffe, Tukey, Dunnett, and Hsu can be used as usual.

Remark: If either levels of factor are assigned to whole plots as an incomplete block design, or the levels of factor B are assigned to split-plots as an incomplete design, the formulas of the sum of squares should be adjusted. But the degrees of freedom will remain the same. Estimates for main effects and interaction contrasts should be adjusted also.

In general, within-whole-plot comparisons will generally be more precise than between-whole-plot comparisons. If the levels of all factors are easy to change, split-plot designs are recommended only when there is considerably less interest in one or more of the treatment factors.

SAS Programs

1. Complete block designs

```
*** analysis of variance; * method 1;
PROC GLM;
  CLASSES BLOCK A B WP;
  MODEL Y = BLOCK A WP(BLOCK) B A*B/E1;
  RANDOM BLOCK WP(BLOCK) /TEST;
  MEANS   A / DUNNETT('0')   ALPHA=0.01   CLDIFF  E=WP(BLOCK);
  MEANS   B / DUNNETT('0')   ALPHA=0.01   CLDIFF;
RUN;
```

2. Complete block designs or incomplete block designs

```
*** analysis of variance; * method 2;
PROC GLM;
  CLASSES  BLOCK A B;
  MODEL  Y  =  BLOCK  A  BLOCK*A  B  A*B;
  RANDOM BLOCK A*BLOCK/TEST;
  MEANS   A / DUNNETT('0')   ALPHA=0.01   CLDIFF  E=BLOCK*A;
  MEANS   B / DUNNETT('0')   ALPHA=0.01   CLDIFF;
Run;
```

Note the second method does not use the whole-plot as a random factor as in methods one. It makes use of the fact that the whole-plot error sum of squares uses the same degrees of freedom as the interactions between the block factor and the whole-plot factor.

Example of Split-Plot Design and Analysis: The Oats Experiment

An experiment on the yield of three varieties (factor A) and four different levels of manure (factor B) was described by Yates (*Complex Experiments*, 1935). The experiment area was divided into s=6 blocks. Each of these was then subdivided into a=3 whole plots. The varieties of oats were sown on the whole plots according to a randomized complete block design. Each whole plot was then divided into b=4 split-plots and the levels of manure were applied to the split plots according to a randomized complete block design. The design and data were shown in Table 19.3, page 682.

1. Write down an appropriate model for this experiment.

2.Do the varieties of oats and the levels of manure have significant interaction effects?

3. Do the varieties of oats have significantly different effects?

4. Do the levels of manure have significantly different effects?

5. Find simultaneous 95% confidence intervals for all treatment-versus-control comparisons for the varieties(Variety 0 is the control).

6. Find simultaneous 95% confidence intervals for all treatment-versus-control comparisons for the levels of manure (Level 0 is the control).

SAS Program:

```
*** analysis of variance; * method 1;
PROC GLM;
  CLASSES BLOCK A B WP;
  MODEL Y = BLOCK A WP(BLOCK) B A*B / E1;
  RANDOM BLOCK WP(BLOCK) / TEST;
  MEANS   A / DUNNETT('O')   ALPHA=0.01   CLDIFF  E=WP(BLOCK);
  MEANS   B / DUNNETT('O')   ALPHA=0.01   CLDIFF;
title 'method 1';
;
*** analysis of variance; * method 2;
DATA; SET OAT;
PROC GLM;
  CLASSES  BLOCK A B;
  MODEL  Y  =  BLOCK  A  BLOCK*A  B  A*B;
  RANDOM BLOCK A*BLOCK/TEST;
  LSMEANS   A / PDIFF=CONTROL CL ADJUST=DUNNETT  ALPHA=0.01 E=BLOCK*A;
  LSMEANS   B / PDIFF=CONTROL CL ADJUST=DUNNETT  ALPHA=0.01;
title 'Method 2';
run;
```

The two methods give identical results. Result from method 1 is given in the book (p689). Provided below are results from method 2.

```
                    General Linear Models Procedure

Dependent Variable: Y
                                     Sum of          Mean
Source                  DF          Squares        Square  F Value    Pr > F
Model                   26        44017.194      1692.969     9.56    0.0001
Error                   45         7968.750       177.083
Corrected Total         71        51985.944

                  R-Square             C.V.      Root MSE             Y Mean
                  0.846713         12.79887        13.307             103.97

Source                  DF        Type I SS   Mean Square  F Value    Pr > F
BLOCK                    5        15875.278      3175.056    17.93    0.0001
A                        2         1786.361       893.181     5.04    0.0106
BLOCK*A                 10         6013.306       601.331     3.40    0.0023
B                        3        20020.500      6673.500    37.69    0.0001
A*B                      6          321.750        53.625     0.30    0.9322

Source                  DF      Type III SS   Mean Square  F Value    Pr > F
BLOCK                    5        15875.278      3175.056    17.93    0.0001
A                        2         1786.361       893.181     5.04    0.0106
BLOCK*A                 10         6013.306       601.331     3.40    0.0023
B                        3        20020.500      6673.500    37.69    0.0001
A*B                      6          321.750        53.625     0.30    0.9322

                    General Linear Models Procedure
Source      Type III Expected Mean Square
BLOCK       Var(Error) + 4 Var(BLOCK*A) + 12 Var(BLOCK)
A           Var(Error) + 4 Var(BLOCK*A) + Q(A,A*B)
BLOCK*A     Var(Error) + 4 Var(BLOCK*A)
B           Var(Error) + Q(B,A*B)
A*B         Var(Error) + Q(A*B)

        Tests of Hypotheses for Mixed Model Analysis of Variance
Dependent Variable: Y
Source: BLOCK
Error: MS(BLOCK*A)
                          Denominator    Denominator
     DF    Type III MS             DF             MS     F Value   Pr > F
      5   3175.0555556             10   601.33055556      5.2801   0.0124
Source: A *
Error: MS(BLOCK*A)
                         Denominator    Denominator
     DF    Type III MS            DF             MS     F Value   Pr > F
      2   893.18055556            10   601.33055556      1.4853   0.2724
* - This test assumes one or more other fixed effects are zero.

Source: BLOCK*A
Error: MS(Error)
                         Denominator    Denominator
     DF    Type III MS            DF             MS     F Value   Pr > F
     10   601.33055556            45   177.08333333      3.3957   0.0023

Source: B *
Error: MS(Error)
                         Denominator    Denominator
     DF    Type III MS            DF             MS     F Value   Pr > F
      3         6673.5            45   177.08333333     37.6856   0.0001
* - This test assumes one or more other fixed effects are zero.
Source: A*B
Error: MS(Error)
                         Denominator    Denominator
     DF    Type III MS            DF             MS     F Value   Pr > F
      6         53.625            45   177.08333333      0.3028   0.9322
```

A	99% Lower Confidence Limit	Y LSMEAN	99% Upper Confidence Limit
0	81.761076	97.625000	113.488924
1	88.636076	104.500000	120.363924
2	93.927743	109.791667	125.655591

Least Squares Means
Adjustment for multiple comparisons: Dunnett
Least Squares Means for effect A
99% Confidence Limits for LSMEAN(i)-LSMEAN(j)

i	j	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit
2	1	-18.122987	6.875000	31.872987
3	1	-12.831320	12.166667	37.164653

Least Squares Means

B	99% Lower Confidence Limit	Y LSMEAN	99% Upper Confidence Limit
0	70.952864	79.388889	87.824914
1	90.452864	98.888889	107.324914
2	105.786197	114.222222	122.658247
3	114.952864	123.388889	131.824914

Method 2

Least Squares Means
Adjustment for multiple comparisons: Dunnett
Least Squares Means for effect B
99% Confidence Limits for LSMEAN(i)-LSMEAN(j)

i	j	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit
2	1	5.879616	19.500000	33.120384
3	1	21.212949	34.833333	48.453718
4	1	30.379616	44.000000	57.620384