

## 14.0 Other ANOVA Designs

- Answer Questions; Review
- Latin Squares
- Greco-Latin Squares

## 14.1 Latin Squares

A Latin Square extends the Randomized Complete Block Design to the case in which there are two blocking factors and one treatment. The assumption is that there is not interaction between the blocking factors or between either or both blocking factor and the treatment factor.

These designs are called “Latin Squares” because they use the Roman alphabet to denote the assignment of the experimental units to the treatment and block levels.

Note that you have to have the same number of levels for each block factor as there are for the levels of treatment. So these are a bit delicate.

## Latin-Square Design with 4 Levels

	Column Blocking Variable			
Row BV	1	2	3	4
1	A	B	C	D
2	C	D	A	B
3	B	A	D	C
4	D	C	B	A

Note that each treatment (Roman letter) appears exactly once in each row and each column. This enables one to estimate all three effects.

## Latin Square ANOVA Table

Source	df	SS	MS	F
treatment	$t - 1$	$SS_{\text{trt}}$	$SS_{\text{trt}}/(t - 1)$	$MS_{\text{trt}}/MS_{\text{err}}$
Block 1	$t - 1$	$SS_{\text{B1}}$	$SS_{\text{B1}}/(t - 1)$	$MS_{\text{B1}}/MS_{\text{err}}$
Block 2	$t - 1$	$SS_{\text{B2}}$	$SS_{\text{B2}}/(t - 1)$	$MS_{\text{B2}}/MS_{\text{err}}$
error	$(t - 1)(t - 2)$	$SS_{\text{err}}$	$SS_{\text{err}}/(t - 1)(t - 2)$	
total	$t^2 - 1$	$SS_{\text{tot}}$		

Note that the degrees of freedom add up as they should. And the formulae for the calculation of the SS terms are similar to what we have seen before.

## 14.2 Graeco-Latin Squares

The thirty-six officers problem is a mathematical puzzle proposed by Leonhard Euler in 1782.

The problem asks if it is possible to arrange 6 regiments consisting of 6 officers each of different ranks in a  $6 \times 6$  square so that no rank or regiment will be repeated in any row or column. Such an arrangement would form a Graeco-Latin square, since it finds two Latin squares, one for rank and one for regiment. Euler correctly predicted there was no solution to this problem, and Gaston Tarry proved this in 1901; but the problem has led to important work in combinatorics.

It is possible to solve the equivalent problem for all other sizes of square other than a  $2 \times 2$  squares.



He did everything: analysis, number theory, combinatorics, graph theory, physics, etc. In particular, he showed that  $\exp(i\theta) = \cos \theta + i \sin \theta$ , implying  $e^{i\pi} + 1 = 0$  (which Feynman liked because it connects all five of math's special numbers).

		Column Blocking Variable			
Row BV		1	2	3	4
	1	$\alpha:A$	$\beta:B$	$\gamma:C$	$\delta:D$
	2	$\beta:C$	$\alpha:D$	$\delta:A$	$\gamma:B$
	3	$\delta:B$	$\gamma:A$	$\beta:D$	$\alpha:C$
	4	$\gamma:D$	$\delta:C$	$\alpha:B$	$\beta:A$

One can think of this as having three blocking variables. Clearly, the row and column blocking variables give us a Latin Square. Each treatment appears only once in each row and once in each column. But the Greek letters indicate a third block; note that each treatment is paired with the same Greek letter only once, and that each Greek letter appears exactly once in each row and once in each column.

Source	df	SS	MS	F
treatment	$t - 1$	$SS_{\text{trt}}$	$SS_{\text{trt}}/(t - 1)$	$MS_{\text{trt}}/MS_{\text{err}}$
Block 1	$t - 1$	$SS_{\text{B1}}$	$SS_{\text{B1}}/(t - 1)$	$MS_{\text{B1}}/MS_{\text{err}}$
Block 2	$t - 1$	$SS_{\text{B2}}$	$SS_{\text{B2}}/(t - 1)$	$MS_{\text{B2}}/MS_{\text{err}}$
Block 3	$t - 1$	$SS_{\text{B3}}$	$SS_{\text{B3}}/(t - 1)$	$MS_{\text{B3}}/MS_{\text{err}}$
error	$(t - 3)(t - 1)$	$SS_{\text{err}}$	$SS_{\text{err}}/(t - 3)(t - 1)$	
total	$t^2 - 1$	$SS_{\text{tot}}$		

For  $t < 3$ , there are no degrees of freedom for the error term. Thus at least one additional observation is needed. The general SS formulae are similar to those we had before.



## 14.3 Split Plot Designs

Split plot designs are used when practical constraints on a problem require the that crossed factors occur within each level of a nested factor.

The term derives from agricultural research, where one must experiment on plots of land. Levels of one factor are assigned to geographical areas called “whole plots” (e.g., farms), while levels of the other factor are applied to smaller plots within the whole plots (or “split plots”). The implication is that the whole plot factor cannot be applied to many small areas, but only to large areas.

In this class we shall consider only two-factor split plot designs, but in principle the whole plot can have complex structure (e.g., many more factors, RCBD, Latin Square).

For example, the whole-plot factor might be whether or not one crop dusts. The split plot factor would be different strains of corn. You can only crop dust large regions, but within those regions one can have small plots with each variety of corn.

Split plots are also often used to handle repeated measurements on the same individual. For example, you might give someone a memory test before coffee, after a nap, and after listening to music. The response variable is the score on the test, the person is the whole plot, and the different experimental conditions are the split plot factor.

A single whole plot is a block effect. The whole plot treatment effect estimated by the differences among all whole plots that receive the same level of the whole plot factor. So there is an automatic block effect that is included.

3 A	5 B	1 B	4 B	2 A	1 A	4 A	3 B	2 B	5 A	Block 1
2 A	5 B	4 B	2 B	4 A	3 A	1 A	1 B	3 B	5 A	Block 2
1 A	3 B	4 B	5 B	3 A	4 A	2 A	2 B	1 B	5 A	Block 3

Factorial Arrangement of Treatments in a Randomized Complete Block Design

5 A	2 A	1 A	4 A	3 A	1 B	3 B	5 B	4 B	2 B	Block 1
5 B	3 B	1 B	2 B	4 B	4 A	3 A	2 A	1 A	5 A	Block 2
4 A	3 A	5 A	1 A	2 A	2 B	1 B	3 B	5 B	4 B	Block 3

Factorial Arrangement of Treatments in a Split-Plot Design

Once the whole plots are chosen, they are randomly assigned to the whole plot treatment levels (say. crop dusting or no crop dusting).

Within the whole plots, split plots are identified. Each split plot is randomly assigned to one of the treatment levels of the split plot factors (say, corn variety).

Suppose we have a split plot experiment in which the whole plot factor is Treatment A and the split plot factor is treatment B. The model is:

$$Y_{ijk} = \mu + \rho_i + \alpha_j + \beta_k + (\rho\alpha)_{ij} + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

where  $\rho_i$  denotes the block effect and various constraints are put on the effects to ensure a unique solution (e.g., requiring that  $\sum_i \rho_i = 0$ , etc.).

There are  $n$  whole plots,  $a$  levels of factor A, and  $b$  levels of factor B. This keeps the formulae simple.

$$\begin{aligned}
\mathbf{SS}_{\text{tot}} &= \left( \sum_{i=1}^n \sum_{j=1}^a \sum_{k=1}^b X_{ijk}^2 \right) - \frac{X_{\dots}^2}{nab} \\
\mathbf{SS}_{\text{Blk}} &= ab \sum_{i=1}^n (\bar{X}_{i..} - \bar{X}_{\dots})^2 \\
\mathbf{SS}_{\text{A}} &= bn \sum_{j=1}^a (\bar{X}_{.j.} - \bar{X}_{\dots})^2 \\
\mathbf{SS}_{\text{Blk}^* \text{A}} &= b \sum_{i=1}^n \sum_{j=1}^a (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{\dots})^2 \\
\mathbf{SS}_{\text{B}} &= an \sum_{k=1}^b \sum_{i=1}^n (X_{..k} - \bar{X}_{\dots})^2 \\
\mathbf{SS}_{\text{AB}} &= n \sum_{j=1}^a \sum_{k=1}^b (\bar{X}_{.jk} - \bar{X}_{.j.} - \bar{X}_{..k} + \bar{X}_{\dots})^2 \\
\mathbf{SS}_{\text{err}} &= \mathbf{SS}_{\text{tot}} - \mathbf{SS}_{\text{Blk}} - \mathbf{SS}_{\text{A}} - \mathbf{SS}_{\text{Blk}^* \text{A}} - \mathbf{SS}_{\text{B}} - \mathbf{SS}_{\text{AB}}
\end{aligned}$$

## ANOVA Table for Split Plot: First Part

Source	df	SS	MS
Blk	$n-1$	$SS_{\text{Blk}}$	$SS_{\text{Blk}}/(n-1)$
A	$a-1$	$SS_A$	$SS_A/(a-1)$
Blk*A	$(a-1)(n-1)$	$SS_{\text{Blk}*A}$	$SS_{\text{Blk}*A}/(a-1)(n-1)$
B	$b-1$	$SS_B$	$SS_B/(b-1)$
A*B	$(a-1)(b-1)$	$SS_{AB}$	$SS_{AB}/(a-1)(b-1)$
err	$a(b-1)(n-1)$	$SS_{\text{err}}$	$SS_{\text{err}}/a(b-1)(n-1)$
tot	$abn-1$	$SS_{\text{tot}}$	

As usual, one can check the degrees of freedom. As always the mean squared terms are the SS terms divided by the corresponding degrees of freedom.

## ANOVA Table for Split Plot: Second Part

Source	EMS: Fixed Blocks	F statistic
Blk	$\sigma^2 + \frac{ab}{n-1} \sum_{i=1}^n \rho_i^2$	$MS_{\text{Blk}}/MS_{\text{err}}$
A	$\sigma^2 + \frac{nb}{a-1} \sum_{j=1}^a \alpha_j^2$	$MS_A/MS_{\text{err}}$
Blk*A	$\sigma^2 + \frac{b}{(n-1)(a-1)} \sum_{i=1}^n \sum_{j=1}^a (\rho\alpha)_{ij}^2$	$MS_{\text{Blk*A}}/MS_{\text{err}}$
B	$\sigma^2 + \frac{na}{(b-1)} \sum_{k=1}^b \beta_k^2$	$MS_B/MS_{\text{err}}$
A*B	$\sigma^2 + \frac{n}{(a-1)(b-1)} \sum_{j=1}^a \sum_{k=1}^b (\alpha\beta)_{jk}^2$	$MS_{AB}/MS_{\text{err}}$
error	$\sigma^2$	
tot		

This is the easy case. Every MS term of interest is divided by the mean squared error.

## ANOVA Table for Split Plot: Third Part

Source	EMS: Random Blocks	F
Blk	$\sigma^2 + ab\sigma_{Blk}^2$	$MS_{Blk}/MS_{err}$
A	$\sigma^2 + b\sigma_{Blk*A} + \frac{nb}{a-1} \sum_{j=1}^a \alpha_j^2$	$MS_A/MS_{Blk*A}$
Blk*A	$\sigma^2 + b\sigma_{Blk*A}^2$	$MS_{Blk*A}/MS_{err}$
B	$\sigma^2 + \frac{na}{(b-1)} \sum_{k=1}^b \beta_k^2$	$MS_B/MS_{err}$
A*B	$\sigma^2 + \frac{n}{(a-1)(b-1)} \sum_{j=1}^a \sum_{k=1}^b (\alpha\beta)_{jk}^2$	$MS_{AB}/MS_{err}$
error	$\sigma^2$	
tot		

This is the tricky case. The divisor mean squares have to be chosen so that under the null, the expected mean square terms in the numerator and denominator are equal.

Things get more complicated when the levels of Factors A and/or B are random, but that goes beyond our purview.



Consider the following experiment: there are 8 men and 8 women. Each subject undertakes two different kinds of exercise warm-up routine (say I and II) several days apart. The response variable is the unresisted ankle reaction time.

- What is the whole factor treatment?
- What is the block?
- What is the split plot treatment?

Additionally, identify whether each of these is a random effect or a fixed effect.

The advantages of split plot designs include:

- The split plot treatments usually has smaller variance than the whole plot treatments, so their test is reasonably powerful.
- Can combine experiments in which some factors require large amounts of experimental material and other factors require very little.
- If there is a pre-planned whole plot experiment, the split plot treatment can usually be added at little cost.
- It is a natural way to handle repeated measurements.

The disadvantages of split plot designs include:

- The analysis is fairly complicated.
- There is often high variance and small sample numbers for the whole plots, so the test may not be powerful.

Example of SAS code for split plot analyses:

```
data EXAMPLE;
input Block A B Y @@;
datalines;
1 1 1 56 1 1 2 41
1 2 1 50 1 2 2 36
1 3 1 39 1 3 2 35
{etc.}
4 3 1 17 4 3 2 18
;
proc mixed;
class A B Block;
model Y = A B A*B;
random block A*Block;
run;
```