

Chapter 14. Supplemental Text Material

S14-1. The Staggered, Nested Design

In Section 14-1.4 we introduced the staggered, nested design as a useful way to prevent the number of degrees of freedom from “building up” so rapidly at lower levels of the design. In general, these designs are just unbalanced nested designs, and many computer software packages that have the capability to analyze general unbalanced designs can successfully analyze the staggered, nested design. The general linear model routine in Minitab is one of these packages.

To illustrate a staggered, nested design, suppose that a pharmaceutical manufacturer is interested in testing the absorption of a drug two hours after the tablet is ingested. The product is manufactured in lots, and specific interest focuses on determining whether there is any significant lot-to-lot variability. Excessive lot-to-lot variability probably indicates problems with the manufacturing process, perhaps at the stage where the coating material that controls tablet absorption is applied. It could also indicate a problem with either the coating formulation, or with other formulation aspects of the tablet itself.

The experimenters select $a = 10$ lots at random from the production process, and decide to use a staggered, nested design to sample from the lots. Two samples are taken at random from each lot. The first sample contains two tablets, and the second sample contains only one tablet. Each tablet is test for the percentage of active drug absorbed after two hours. The data from this experiment is shown in Table 1 below.

Table 1. The Drug Absorption Experiment

Lot	Sample	
	1	2
1	24.5, 25.9	23.9
2	23.6, 26.1	25.2
3	27.3, 28.1	27.0
4	28.3, 27.5	27.4
5	24.3, 24.1	25.1
6	25.3, 26.0	24.7
7	27.3, 26.8	28.0
8	23.3, 23.9	23.0
9	24.6, 25.1	24.9
10	24.3, 24.9	25.3

The following output is from the Minitab general linear model analysis procedure.

General Linear Model

Factor	Type	Levels	Values
Lot	random	10	1 2 3 4 5 6 7 8 9 10
Sample(Lot)	random	20	1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2

Analysis of Variance for Absorp., using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Lot	9	58.3203	52.3593	5.8177	14.50	0.000
Sample(Lot)	10	4.0133	4.0133	0.4013	0.71	0.698
Error	10	5.6200	5.6200	0.5620		
Total	29	67.9537				

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 Lot	(3) + 1.3333(2) + 2.6667(1)
2 Sample(Lot)	(3) + 1.3333(2)
3 Error	(3)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 Lot	10.00	0.4013	(2)
2 Sample(Lot)	10.00	0.5620	(3)

Variance Components, using Adjusted SS

Source	Estimated Value
Lot	2.0311
Sample(Lot)	-0.1205
Error	0.5620

As noted in the textbook, this design results in $a - 1 = 9$ degrees of freedom for lots, and $a = 10$ degrees of freedom for samples within lots and error. The ANOVA indicates that there is a significant difference between lots, and the estimate of the variance component for lots is $\hat{\sigma}_{Lots}^2 = 2.03$. The ANOVA indicates that the sample within lots is not a significant source of variability. This is an indication of lot homogeneity. There is a small negative estimate of the sample-within-lots variance component. The experimental error variance is estimated as $\hat{\sigma}^2 = 0.526$. Notice that the constants in the expected mean squares are not integers; this is a consequence of the unbalanced nature of the design.

S14-2. Inadvertent Split-Plots

In recent years experimenters from many different industrial settings have become exposed to the concepts of designed experiments, either from university-level DOX courses or from industrial short courses and seminars. As a result, factorial and fractional factorial designs have enjoyed expanded use. Sometimes the principle of randomization is not sufficiently stressed in these courses, and as a result experimenters may fail to understand its importance. This can lead to **inadvertent split-plotting** of a factorial design.

For example, suppose that an experimenter wishes to conduct a 2^4 factorial using the factors A = temperature, B = feed rate, C = concentration, and D = reaction time. A 2^4 with the runs arranged in random order is shown in Table 2.

Table 2. A 2^4 Design in Random Order

Std	Run	Block	Factor A: Temperature DegC	Factor B: Feed rate gal/h	Factor C: Concentration gm/l	Factor D: Reaction time h	Response Yield
16	1	Block 1	150	8	30	1.2	
9	2	Block 1	100	5	25	1.2	
7	3	Block 1	100	8	30	1	
12	4	Block 1	150	8	25	1.2	
2	5	Block 1	150	5	25	1	
13	6	Block 1	100	5	30	1.2	
1	7	Block 1	100	5	25	1	
10	8	Block 1	150	5	25	1.2	
3	9	Block 1	100	8	25	1	
14	10	Block 1	150	5	30	1.2	
6	11	Block 1	150	5	30	1	
4	12	Block 1	150	8	25	1	
5	13	Block 1	100	5	30	1	
15	14	Block 1	100	8	30	1.2	
11	15	Block 1	100	8	25	1.2	
8	16	Block 1	150	8	30	1	

When the experimenter examines this run order, he notices that the level of temperature is going to start at 150 degrees and then be changed eight times over the course of the 16 trials. Now temperature is a hard-to-change-variable, and following every adjustment to temperature several hours are needed for the process to reach the new temperature level and for the process to stabilize at the new operating conditions.

The experimenter may feel that this is an intolerable situation. Consequently, he may decide that fewer changes in temperature are required, and rearrange the temperature levels in the experiment so that the new design appears as in Table 3. Notice that only three changes in the level of temperature are required in this new design. In effect, the experimenter will set the temperature at 150 degrees and perform four runs with the other three factors tested in random order. Then he will change the temperature to 100 degrees

and repeat the process, and so on. The experimenter has inadvertently introduced a split-plot structure into the experiment.

Table 3. The Modified 2⁴ Factorial

Std	Run	Block	Factor A: Temperature DegC	Factor B: Feed rate gal/h	Factor C: Concentration gm/l	Factor D: Reaction time h	Response Yield
16	1	Block 1	150	8	30	1.2	
9	2	Block 1	150	5	25	1.2	
7	3	Block 1	150	8	30	1	
12	4	Block 1	150	8	25	1.2	
2	5	Block 1	100	5	25	1	
13	6	Block 1	100	5	30	1.2	
1	7	Block 1	100	5	25	1	
10	8	Block 1	100	5	25	1.2	
3	9	Block 1	150	8	25	1	
14	10	Block 1	150	5	30	1.2	
6	11	Block 1	150	5	30	1	
4	12	Block 1	150	8	25	1	
5	13	Block 1	100	5	30	1	
15	14	Block 1	100	8	30	1.2	
11	15	Block 1	100	8	25	1.2	
8	16	Block 1	100	8	30	1	

Typically, most inadvertent split-plotting is not taken into account in the analysis. That is, the experimenter analyzes the data as if the experiment had been conducted in random order. Therefore, it is logical to ask about the impact of ignoring the inadvertent split-plotting. While this question has not been studied in detail, generally inadvertently running a split-plot and not properly accounting for it in the analysis probably does not have major impact **so long as the whole plot factor effects are large**. These factor effect estimates will probably have larger variances than the factor effects in the subplots, so part of the risk is that small differences in the whole-plot factors may not be detected. Obviously, the more systematic fashion in which the whole-plot factor temperature was varied in Table 2 also exposes the experimenter to confounding of temperature with some nuisance variable that is also changing with time. The most extreme case of this would occur if the first eight runs in the experiment were made with temperature at the low level (say), followed by the last eight runs with temperature at the high level.