

Book: Contemporary Statistical Models  
 for the Plant and Soil Sciences.  
Oliver Schabenberger & Francis J. Pierce  
 Applications 479  
Taylor & Francis

experiment station research, the analysis of the interaction shows that for farms 3, 7, and 8 (and by implication farms in the target region that are alike) the new treatment holds promise.

Output 7.10.

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr >  t
Tx eff. on Farm 1	-1.9364	1.0117	17.6	-1.91	0.0720
Tx eff. on Farm 2	5.3249	1.0117	17.6	5.26	<.0001
Tx eff. on Farm 3	-4.1009	1.0117	17.6	-4.05	0.0008
Tx eff. on Farm 4	3.9200	1.0117	17.6	3.87	0.0011
Tx eff. on Farm 5	-0.4200	1.0117	17.6	-0.42	0.6831
Tx eff. on Farm 6	0.2409	1.0117	17.6	0.24	0.8145
Tx eff. on Farm 7	-7.7718	1.0117	17.6	-7.68	<.0001
Tx eff. Farm 8	-2.1535	1.0117	17.6	-2.13	0.0477

### 7.6.3 Nested Errors through Subsampling

Subsampling in designed experiments is the recording of multiple, independent observations on the experimental units. We term the experimental material on which subsamples are collected the observational units to distinguish them from the experimental units to which treatments are assigned. Subsamples are sometimes referred to as pseudo-replications, an unfortunate terminology, because they are not replications in the proper sense. The variation among subsamples from the same experimental unit expresses the homogeneity within the unit; we term this source of variation **observational error**. The proper error variation to compare treatment effects is variation among experimental units that received the same treatment, termed **experimental error variance**. In experimental designs with subsampling, care must be exercised to (i) not consider subsamples as replications of the treatments, (ii) separate experimental from observational error, and (iii) to perform tests of hypotheses properly. To demonstrate that subsamples do not substitute for treatment replication consider a  $3 \times 2$  factorial experiment with six experimental units. Assume that these units are pots containing four plants each. The data structure for such an experiment could be displayed as in Table 7.8.

Table 7.8. Generic data layout of  $3 \times 2$  factorial without replication and four subsamples per experimental unit

Plant	Factor Combination					
	$A_1B_1$	$A_1B_2$	$A_2B_1$	$A_2B_2$	$A_3B_1$	$A_3B_2$
1	3.5	5.0	5.5	7.0	5.5	6.0
2	4.0	5.5	6.0	9.0	4.5	6.5
3	3.0	4.0	5.0	8.0	8.5	9.5
4	4.5	3.5	5.0	6.5	7.0	7.0

It is tempting to analyze these data with a two-way factorial analysis of variance based on the linear model

$$Y_{ijk} = \mu_{ij} + e_{ijk}^* = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}^*, \quad [7.64]$$

assuming that  $e_{ijk}^*$  is the "experimental" error for the  $k^{\text{th}}$  experimental unit receiving level  $i$  of factor  $A$  and level  $j$  of factor  $B$ . As will be shown shortly,  $e_{ijk}^*$  in [7.64] is the observational error and the experimental error has been confounded with the treatment means. Since The SAS® System cannot know whether repeated values in the data set that share the same treatment assignment represent subsamples or replicates, an analysis of the data in Table 7.8 with model [7.64] will be *successful*. Using proc glm significant main effects of factors  $A$  and  $B$  ( $p = 0.0004$  and  $0.0144$ ) and a nonsignificant interaction are inferred (Output 7.11).

```

data noreps;
  input A B plant y;
  datalines;
1 1 1 3.5
1 1 2 4.0
1 1 3 3.0
1 1 4 4.5
1 2 1 5.0
1 2 2 5.5
1 2 3 4.0
... and so forth ...
;;
run;

proc glm data=noreps;
  class A B;
  model y = A B A*B;
run; quit;

```

#### Output 7.11.

#### The GLM Procedure

Class Level Information		
Class	Levels	Values
A	3	1 2 3
B	2	1 2
Number of observations		24

#### Dependent Variable: y

Source	DF	Sum of			F Value	Pr > F
		Squares	Mean Square			
Model	5	47.34375000	9.46875000		6.94	0.0009
Error	18	24.56250000	1.36458333			
Corrected Total	23	71.90625000				
R-Square	Coeff Var	Root MSE	y Mean			
0.658409	20.09727	1.168154	5.812500			
Source	DF	Type I SS	Mean Square	F Value	Pr > F	
A	2	34.56250000	17.28125000	12.66	0.0004	
B	1	10.01041667	10.01041667	7.34	0.0144	
A*B	2	2.77083333	1.38541667	1.02	0.3821	
Source	DF	Type III SS	Mean Square	F Value	Pr > F	
A	2	34.56250000	17.28125000	12.66	0.0004	
B	1	10.01041667	10.01041667	7.34	0.0144	
A*B	2	2.77083333	1.38541667	1.02	0.3821	

Notice that the  $F$  statistics on which the  $p$ -values are based are obtained by dividing the main effects or interaction mean square by the mean square error of 1.3645. This mean square error is based on 18 degrees of freedom,  $4 - 1 = 3$  degrees of freedom for the subsamples in each of 6 experimental units. This analysis is clearly wrong, since the experimental error in this design has  $t(r - 1)$  degrees of freedom where  $t$  denotes the number of treatments and  $r$

the number of replications for each treatment. Since each of the  $t = 6$  treatments was assigned to only one pot, we have  $r = 1$  and  $t(r - 1) = 0$ . What SAS® terms the error source in this model is the observational error and  $\hat{\sigma}^2 = 1.36458$  is an estimate of the observational error variance. The correct model for the subsampling design contains separate random terms for experimental and observational error. In the two-factor design we obtain

$$\begin{aligned} Y_{ijkl} &= \mu_{ij} + e_{ijk} + d_{ijkl} \\ \text{Var}[e_{ijk}] &= \sigma_e^2, \text{Var}[d_{ijkl}] = \sigma_d^2, \end{aligned} \quad [7.65]$$

where  $k = 1, \dots, r$  indexes the replications,  $e_{ijk}$  is the experimental error as defined above, and  $d_{ijkl}$  is the observational (subsampling) error for subsample  $l = 1, \dots, n$  on replicate  $k$ .  $\sigma_e^2$  and  $\sigma_d^2$  are the experimental and observational error variances, respectively. If  $k = 1$ , as for the data in Table 7.8, the model becomes

$$Y_{ijl} = \mu_{ij} + e_{ij} + d_{ijl} = \mu_{ij}^* + d_{ijl}. \quad [7.66]$$

and the experimental error is now confounded with the treatments. This is model [7.64] where  $e_{ijk}^*$  is replaced with  $d_{ijl}$  and  $\mu_{ij}$  is replaced with  $\mu_{ij}^*$ . Because  $\mu_{ij}$  and  $e_{ij}$  in [7.66] have the same subscript the two sources of variability are confounded. The only random variation that can be estimated is the variance of  $d_{ijl}$ , the observational error. Finally, the observational error mean square is not the correct denominator for  $F$ -tests (Table 7.9). The statistic  $MS(\text{Treatment})/MS(\text{Obs. Error})$  thus is not a test statistic for the absence of treatment effects ( $f(\mu_{ij}^2) = 0$ ) but for the simultaneous absence of treatment effects and the experimental error, a nonsensical proposition.

Table 7.9. Expected mean squares in subsampling design without treatment replications (model [7.66])

Source of Variation	DF	$E[MS]$
Treatments + Experimental Error	$t - 1$	$\sigma_d^2 + n\sigma_e^2 + f(\mu_{ij}^2)$
Observational Error	$t(r - 1)$	$\sigma_d^2$

Table 7.10. Expected mean squares in completely randomized design with subsampling ( $t$  denotes number of treatments,  $r$  number of replicates, and  $n$  number of subsamples)

Source of Variation	DF	$E[MS]$
Treatments ( $Tx$ )	$t - 1$	$\sigma_d^2 + n\sigma_e^2 + f(\tau_i^2)$
Experimental Error ( $EE$ )	$t(r - 1)$	$\sigma_d^2 + n\sigma_e^2$
Observational Error ( $OE$ )	$tr(n - 1)$	$\sigma_d^2$

In subsampling designs with treatment replication, experimental and observational error variances are estimable and not confounded with effects. Table 7.10 displays the expected mean squares for  $t$  treatments in a completely randomized design with  $r$  replications per treatment and  $n$  subsamples per experimental unit for the linear model

$$\begin{aligned} Y_{ijk} &= \mu + \tau_i + e_{ij} + d_{ijk} \\ i &= 1, \dots, t; j = 1, \dots, r; k = 1, \dots, n. \end{aligned}$$

Notice that experimental error degrees of freedom are not affected by the number of subsamples, and that  $F_{obs} = MS(Tx)/MS(EE)$  is the test statistic for testing treatment effects.

The data in Table 7.11, taken from Steel, Torrie, and Dickey (1997, p. 159), represent a  $3 \times 2$  factorial treatment structure arranged in completely randomized design with  $r = 3$  replicates and  $n = 4$  subsamples per experimental unit. From a large group of plants four were randomly assigned to each of 18 pots. Six treatments were then randomly assigned to the pots such that each treatment was replicated three times. The treatments consisted of all possible combinations of three hours of daylight (8, 12, 16 hrs) and two levels of night temperatures (low, high). The outcome of interest was the stem growth of mint plants grown in nutrient solution under the assigned conditions. The experimental units are the pots since treatments were assigned to those. Stem growth was measured for each plant in a pot, hence there are four subsamples per experimental unit.

**Table 7.11.** One-week stem growth of mint plants data from Steel et al. (1997, p. 159)

Plant	Low Night Temperature								
	8 hrs			12 hrs			16 hrs		
	Pot 1	Pot 2	Pot 3	Pot 1	Pot 2	Pot 3	Pot 1	Pot 2	Pot 3
1	3.5	2.5	3.0	5.0	3.5	4.5	5.0	5.5	5.5
2	4.0	4.5	3.0	5.5	3.5	4.0	4.5	6.0	4.5
3	3.0	5.5	2.5	4.0	3.0	4.0	5.0	5.0	6.5
4	4.5	5.0	3.0	3.5	4.0	5.0	4.5	5.0	5.5

Plant	High Night Temperature								
	8 hrs			12 hrs			16 hrs		
	Pot 1	Pot 2	Pot 3	Pot 1	Pot 2	Pot 3	Pot 1	Pot 2	Pot 3
1	8.5	6.5	7.0	6.0	6.0	6.5	7.0	6.0	11.0
2	6.0	7.0	7.0	5.5	8.5	6.5	9.0	7.0	7.0
3	9.0	8.0	7.0	3.5	4.5	8.5	8.5	7.0	9.0
4	8.5	6.5	7.0	7.0	7.5	7.5	8.5	7.0	8.0

Copyright © 1997 by The McGraw-Hill Companies, Inc. Reproduced from Table 7.8 of Steel, R.G.D., Torrie, J.H., and Dickey, D.A. (1997), *Principles and Procedures of Statistics. A Biometrical Approach*, McGraw-Hill, New York, with permission.

The linear model for these data is

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} + d_{ijkl} \quad [7.67]$$

$$i = 1, \dots, a = 3; j = 1, \dots, b = 2; k = 1, \dots, r = 3; l = 1, \dots, n = 4$$

$$\text{Var}[e_{ijk}] = \sigma_e^2, \text{Var}[d_{ijkl}] = \sigma_d^2,$$

where the  $e_{ijk}$  and  $d_{ijkl}$  are zero-mean uncorrelated random variables. The analysis of variance is shown in Table 7.12.

Table 7.12. Analysis of variance of mint plants data

Source	df
Hours	$a - 1 = 2$
Temperature	$b - 1 = 1$
Hours $\times$ Temperature	$(a - 1)(b - 1) = 2$
Experimental Error	$ab(r - 1) = 6(3 - 1) = 12$
Observational Error	$abr(n - 1) = 18(4 - 1) = 54$
Total	$abr n - 1 = 71$

The analysis of variance can be obtained with proc glm of The SAS® System (Output 7.12):

```
proc glm data=mintstems;
  class hour night pot;
  model growth = hour night hour*night pot(hour*night);
run; quit;
```

The sequential (Type I) and partial (Type III) sums of squares are identical because the design is orthogonal. Notice that the source denoted Error is again the observational error as can be seen from the associated degrees of freedom and the experimental error is modeled as pot(hour\*night). The *F* statistics calculated by proc glm are obtained by dividing the mean square of a source of variability by the mean square for the Error source; hence they use the observational error mean square as a denominator and are incorrect. The two error mean square estimates in Output 7.12 are

$$\hat{\sigma}_d^2 = 0.9340$$

$$\hat{\sigma}_d^2 + n\hat{\sigma}_e^2 = 2.1527,$$

hence dividing by the observational mean square error is detrimental in two ways. The *F* statistic is inflated and the *p*-value is calculated from a distribution with incorrect (too many) degrees of freedom.

The correct tests can be obtained in two ways with proc glm. One can add a random statement indicating which terms of the model statement are random variables and The SAS® System will construct the appropriate test statistics based on the formulas of expected mean squares. Alternatively one can use the test statement if the correct error term is known. The two methods lead to the following procedure calls (output not shown).

```
proc glm data=mintstems;
  class hour night pot;
  model growth = hour night hour*night pot(hour*night);
  random pot(hour*night) / test;
run; quit;

proc glm data=mintstems;
  class hour night pot;
  model growth = hour night hour*night pot(hour*night);
  test h=hour night hour*night e=pot(hour*night);
run; quit;
```

**Output 7.12.**

The GLM Procedure					
Class Level Information					
Class	Levels	Values			
hour	3	8 12 16			
night	2	Hig Low			
pot	3	1 2 3			
Number of observations 72					
Dependent Variable: growth					
Sum of					
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	17	205.4756944	12.0868056	12.94	<.0001
Error	54	50.4375000	0.9340278		
Corrected Total	71	255.9131944			
R-Square	Coeff Var	Root MSE	growth Mean		
0.802912	16.70696	0.966451	5.784722		
Source	DF	Type I SS	Mean Square	F Value	Pr > F
hour	2	22.2986111	11.1493056	11.94	<.0001
night	1	151.6701389	151.6701389	162.38	<.0001
hour*night	2	5.6736111	2.8368056	3.04	0.0562
pot (hour*night)	12	25.8333333	2.1527778	2.30	0.0186
Source	DF	Type III SS	Mean Square	F Value	Pr > F
hour	2	22.2986111	11.1493056	11.94	<.0001
night	1	151.6701389	151.6701389	162.38	<.0001
hour*night	2	5.6736111	2.8368056	3.04	0.0562
pot (hour*night)	12	25.8333333	2.1527778	2.30	0.0186

A more elegant approach is to use proc mixed which is specifically designed for mixed models. The statements to analyze the two-way factorial with subsampling are

```
proc mixed data=mintstems;
  class hour night pot;
  model growth = hour night hour*night;
  random pot(hour*night);
run; quit;
```

or

```
proc mixed data=mintstems;
  class hour night pot;
  model growth = hour night hour*night;
  random intercept / subject=pot(hour*night);
run; quit;
```

The two versions of proc mixed code differ only in the form of the random statement and yield identical results. The second form explicitly defines the experimental units pot(hour\*night) as clusters and the columns of the  $Z_i$  matrix as having an intercept only. The residual maximum likelihood estimates of the variance components  $\sigma_e^2$  and  $\sigma_d^2$  are  $\hat{\sigma}_e^2 = 0.3047$  and  $\hat{\sigma}_d^2 = 0.9340$ , respectively (Output 7.13).

**Output 7.13.****The Mixed Procedure****Model Information**

Data Set	WORK.MINTSTEMS
Dependent Variable	growth
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
hour	3	8 12 16
night	2	Hig Low
pot	3	1 2 3

Dimensions	
Covariance Parameters	2
Columns in X	12
Columns in Z	18
Subjects	1
Max Obs Per Subject	72
Observations Used	72
Observations Not Used	0
Total Observations	72

**Covariance Parameter Estimates**

Cov Parm	Subject	Estimate
Intercept	pot(hour*night)	0.3047
Residual		0.9340

**Fit Statistics**

Res Log Likelihood	-103.9
Akaike's Information Criterion	-105.9
Schwarz's Bayesian Criterion	-106.8
-2 Res Log Likelihood	207.7

**Type 3 Tests of Fixed Effects**

Effect	Num DF	Den DF	F Value	Pr > F
hour	2	12	5.18	0.0239
night	1	12	70.45	<.0001
hour*night	2	12	1.32	0.3038

The latter estimate is labeled as Residual in the Covariance Parameter Estimates table. Since the data are completely balanced these estimates are identical to the method of moment estimator implied by the analysis of variance. From  $\hat{\sigma}_d^2 + n\hat{\sigma}_e^2 = 2.1527$  and  $\hat{\sigma}_d^2 = 0.934$  one obtains the moment estimator of the experimental error variance as

$$\hat{\sigma}_e^2 = (2.1527 - 0.9340)/4 = 0.3047.$$

Results of the main effects and interaction tests are shown in the Type 3 Tests of Fixed Effects table. The F statistics are identical to those obtained in proc glm if one uses the correct mean square error term there. For example from Output 7.12 one obtains

$$F_{obs} = \frac{MS(Hour)}{MS(EE)} = \frac{11.1493}{2.1527} = 5.18$$

$$F_{obs} = \frac{MS(Night)}{MS(EE)} = \frac{151.670}{2.1527} = 70.45$$

$$F_{obs} = \frac{MS(Hour \times Night)}{MS(EE)} = \frac{2.8368}{2.1527} = 1.32.$$

These are the  $F$  statistics shown in Output 7.13. Also notice that the denominator degrees of freedom are set to the correct degrees of freedom associated with the experimental error (Table 7.12).

The marginal correlation structure in the subsampling model [7.67] is compound symmetric, observational errors are nested within experimental errors. Adding the `v=list` option to the `random` statement of `proc mixed` requests a printout of the (estimated) marginal variance-covariance matrices of the clusters (subjects) in `list`. For example,

```
random intercept / subject=pot(hour*night) v=1;
```

requests a printout of the variance-covariance matrix for the first cluster (Output 7.14). It is easy to verify that this matrix is of the form

$$\hat{\sigma}_e^2 \mathbf{J}_4 + \hat{\sigma}_d^2 \mathbf{I}_4.$$

#### Output 7.14.

Estimated V Matrix for pot(hour\*night) 1 8 High

Row	Col1	Col2	Col3	Col4
1	1.2387	0.3047	0.3047	0.3047
2	0.3047	1.2387	0.3047	0.3047
3	0.3047	0.3047	1.2387	0.3047
4	0.3047	0.3047	0.3047	1.2387

The same analysis can thus be obtained by modeling the marginal variance-covariance matrix  $\text{Var}[\mathbf{Y}_i]$  directly as a compound symmetric matrix. Replacing the `random` statement with a `repeated` statement and choosing the appropriate covariance structure (`type=cs`), the statements

```
proc mixed data=mintstems noitprint;
  class hour night pot;
  model growth = hour night hour*night;
  repeated / sub=pot(hour*night) type=cs r=1;
run; quit;
```

lead to the same results as in Output 7.13 and Output 7.14, only the covariance parameter `Intercept` in Output 7.13 has been renamed to `cs` (Output 7.15).

**Output 7.15.****The Mixed Procedure****Model Information**

Data Set	WORK.MINTSTEMS
Dependent Variable	growth
Covariance Structure	Compound Symmetry
Subject Effect	pot(hour*night)
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

**Class Level Information**

Class	Levels	Values
hour	3	8 12 16
night	2	Hig Low
pot	3	1 2 3

**Dimensions**

Covariance Parameters	2
Columns in X	12
Columns in Z	0
Subjects	18
Max Obs Per Subject	4
Observations Used	72
Observations Not Used	0
Total Observations	72

**Estimated R Matrix for pot(hour\*night) 1 8 Hig**

Row	Col1	Col2	Col3	Col4
1	1.2387	0.3047	0.3047	0.3047
2	0.3047	1.2387	0.3047	0.3047
3	0.3047	0.3047	1.2387	0.3047
4	0.3047	0.3047	0.3047	1.2387

**Covariance Parameter Estimates**

Cov Parm	Subject	Estimate
CS	pot(hour*night)	0.3047
Residual		0.9340

**Fit Statistics**

Res Log Likelihood	-103.9
Akaike's Information Criterion	-105.9
Schwarz's Bayesian Criterion	-106.8
-2 Res Log Likelihood	207.7

**Type 3 Tests of Fixed Effects**

Effect	Num DF	Den DF	F Value	Pr > F
hour	2	12	5.18	0.0239
night	1	12	70.45	<.0001
hour*night	2	12	1.32	0.3038

### 7.6.4 Recovery of Inter-Block Information in Incomplete Block Designs

Incompleteness of block designs can have many causes. Some are by design, others by *accident*. Among the accidental causes are destruction or loss of experimental units and discarding of erroneous measurements. Frequently incompleteness is a design feature if the size of the blocks is such that not all treatments can be accommodated. Since calculations for incomplete designs are considerably more involved than for completely balanced designs, experimental plans were developed that ensure some sort of balance in the treatment allocation to either reduce computational burden and/or to ensure a certain precision in treatment comparisons. Incomplete block designs in this category are known as balanced incomplete block designs (BIBs), partially balanced incomplete block designs (PBIBs) and various special cases thereof, such as the lattice designs (see, e.g. Yates 1936, 1940; Bose and Nair 1939; Cochran and Cox 1957 as some of the historically significant references on the subject). We will not discuss the various forms of incomplete block designs here in detail, but the basic issues that come to bear when not all treatments are allocated in every block. Hoshmand (1994, Ch. 4.3) discusses various forms of agronomic lattice designs, which are special cases of BIBs or PBIBs.

To illustrate the problem that arises in incomplete block designs consider the following treatment layout in a BIB with  $t = 5$  treatments in  $b = 10$  blocks of size  $k = 3$ .

**Table 7.13.** A balanced incomplete block design (BIB) (Treatments that appear in a particular block are marked as  $x$ )

Block	Treatment				
	A	B	C	D	E
1		$x$	$x$		$x$
2	$x$			$x$	$x$
3		$x$		$x$	$x$
4			$x$	$x$	$x$
5	$x$	$x$		$x$	
6	$x$	$x$			$x$
7	$x$	$x$	$x$		
8		$x$	$x$	$x$	
9	$x$		$x$		$x$
10	$x$		$x$	$x$	

This design is balanced in two ways. Each treatment is replicated the same number of times (6) throughout the experiment and each pair of treatments appears the same number of times (3) within a block. For example, treatments  $B$  and  $C$  appear in block 1, 7, 8. Treatments  $A$  and  $B$  appear in blocks 5, 6, and 7. As a result, all treatment comparisons will be made with the same precision in the experiment. However, because of the incompleteness, block and treatment effects are not orthogonal. Whether block effects are removed or not prior to assessing treatment effects is critical. To see this consider a comparison of treatments  $A$  and  $B$ . The naïve approach is to base this comparison on the two arithmetic averages  $\bar{y}_A$  and  $\bar{y}_B$ . Their difference is not an estimate of the treatment effect; however, since these are

averages calculated over different blocks.  $\bar{y}_A$  is calculated from information in blocks 2, 5, 6, 7, 9, 10 and  $\bar{y}_B$  from information in blocks 1, 3, 4, 5, 6, 7, 8. The difference  $\bar{y}_A - \bar{y}_B$  carries not only information about differences between the treatments but also about block effects. To obtain a fair comparison of the treatments unaffected by the block effects, the treatment sum of squares must be adjusted for the block effects and treatment means are not estimated as arithmetic averages. A statistical model for the design in Table 7.13 is  $Y_{ij} = \mu + \rho_i + \tau_j + e_{ij}$  where the  $\rho_i$  are block effects ( $i = 1, \dots, 10$ ),  $\tau_j$  are the treatment effects ( $j = 1, \dots, 5$ ) and  $e_{ij}$  are independent experimental errors with mean 0 and variance  $\sigma^2$ . The only difference between this linear model and one for a randomized *complete* block design is that not all combinations  $ij$  are possible. The appropriate estimate of the mean of the  $j^{\text{th}}$  treatment in the incomplete design is  $\hat{\mu} + \hat{\tau}_j$  where carets denote the least squares estimate.  $\hat{\mu} + \hat{\tau}_j$  is also known as the least squares mean for treatment  $j$ . In fact, these estimates are always appropriate. In a balanced design it turns out that the least squares estimates are identical to the arithmetic averages. The question thus should not be when one should use the least squares means for treatment comparisons, but when one can rely on arithmetic means.

We illustrate the effect of nonorthogonality with data from a balanced incomplete block design reported by Cochran and Cox (1957, p. 448). Thirteen hybrids of corn were arranged in a field experiment in blocks of size  $k = 4$  such that each pair of treatments appeared once in a block throughout the experiment and each treatment is replicated four times. This arrangement requires  $b = 13$  blocks.

Table 7.14. Experimental layout of BIB in Cochran and Cox (1957, p. 448)<sup>†</sup>  
(showing yield of corn in pounds per plot)

Block	Hybrid												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1			25.3			19.9			29.0		24.6		
2		23.0	19.8				33.3					22.7	
3													16.2
4	27.3				27.0			35.6		19.3	31.7	26.6	
5							23.4	30.5	30.8	32.4			17.4
6			30.6	32.4	27.2								
7	34.7					31.1			25.7				32.8
8			34.4				33.3				30.5		
9	38.2	32.9	37.3									36.9	
10		28.7		30.7						31.3			
11	36.6				31.1			31.1		26.9		35.3	
12	31.8						33.7		27.8		28.4		
13		30.3						31.5	39.3			41.1	
											26.7		

<sup>†</sup>Cochran, W.G. and Cox, G.M. (1957), *Experimental Design*, 2<sup>nd</sup> Edition. Copyright © 1957 by John Wiley and Sons, Inc. This material is used by permission of John Wiley and Sons, Inc.

We obtain the analysis of variance for these data with proc glm (Output 7.16).

```
proc glm data=cornnyld;
class block hybrid;
model yield = block hybrid;
lsmeans hybrid / stderr;
means hybrid;
run; quit;
```

**Output 7.16.**

The GLM Procedure

## Class Level Information

Class	Levels	Values
block	13	1 2 3 4 5 6 7 8 9 10 11 12 13
hybrid	13	1 2 3 4 5 6 7 8 9 10 11 12 13

Number of observations 52

Dependent Variable: yield corn yield in pounds per plot

Source	DF	Sum of Squares		Mean Square	F Value	Pr > F
		Model	Error			
	24	1017.929231	538.217500	42.413718	2.13	0.0298
Corrected Total	51	1556.146731		19.933981		

R-Square	Coeff Var	Root MSE	yield Mean
0.654134	14.99302	4.464749	29.77885

Source	DF	Type I SS		Mean Square	F Value	Pr > F
		block	hybrid			
	12	689.3842308	328.5450000	57.4486859	2.88	0.0109
	12			27.3787500	1.37	0.2378

Source	DF	Type III SS		Mean Square	F Value	Pr > F
		block	hybrid			
	12	475.2650000	328.5450000	39.6054167	1.99	0.0677
	12			27.3787500	1.37	0.2378

## Least Squares Means

hybrid	yield LSMEAN	Standard Error	Pr >  t
1	33.0019231	2.4586721	<.0001
2	28.2711538	2.4586721	<.0001
3	30.2173077	2.4586721	<.0001
4	28.1019231	2.4586721	<.0001
5	29.9557692	2.4586721	<.0001
6	27.1019231	2.4586721	<.0001
7	29.7250000	2.4586721	<.0001
8	33.7173077	2.4586721	<.0001
9	29.0173077	2.4586721	<.0001
10	28.0250000	2.4586721	<.0001
11	24.5250000	2.4586721	<.0001
12	30.0865385	2.4586721	<.0001
13	35.3788462	2.4586721	<.0001

Level of hybrid	N	-----yield-----	
		Mean	Std Dev
1	4	35.3250000	2.75121185
2	4	29.8000000	2.40277617
3	4	30.0000000	6.92194578
4	4	28.0500000	5.50424079
5	4	30.7250000	2.55783111
6	4	28.0750000	6.08187197
7	4	31.7750000	6.57133929
8	4	31.8000000	3.38526218
9	4	28.1000000	2.25831796
10	4	28.1750000	8.00848508
11	4	22.4250000	5.01489448
12	4	27.9000000	4.06939799
13	4	34.9750000	6.09555849

What SAS® terms Type I SS and Type III SS are sequential and partial sums of squares, respectively. Sequential sums of squares are the sum of squares contributions of sources given that variability of the previously listed sources has been accounted for. The sequential block sum of squares of 689.38 is the sum of squares among block averages and the sequential hybrid sum of squares of 328.54 is the contribution of the treatment variability after adjusting for block effects. Inferences about treatment effects are to be based on the partial sums of squares. The nonorthogonality of this design is evidenced by the fact that the Type I SS and the Type III SS differ. In an orthogonal design, the two sets of sums of squares would be identical. Whenever the design is nonorthogonal great care must be exercised to estimate treatment means properly. The list of least squares means shows the estimates  $\hat{\mu} + \hat{\tau}_j$  that are adjusted for the block effects. Notice that all least squares means have the same standard error, since every treatment is replicated the same number of times. The final part of the output shows the result of the MEANS statement. These are the arithmetic sample averages of the observations for a particular treatment which do not estimate treatment means unbiasedly unless every treatment appears in every block. One must not base treatment comparisons on these quantities in a nonorthogonal design. The column STD DEV is the standard deviation of the four observations for each treatment. It is not the standard deviation of a treatment based on the analysis of variance.

An analysis of an incomplete block design such as the PROC GLM analysis above is termed an intra-block analysis that obtains treatment information by comparing block-adjusted least squares estimates. Yates (1936, 1940) coined the term along with the term inter-block analysis that also recovers treatment information contained in the block totals (averages). In incomplete block designs contrasts of block averages also contain contrasts among the treatments. To see this consider blocks 1 and 3 in Table 7.13. The first block contains treatments B, C, and E, the third block contains treatments B, D, and E. If  $\bar{Y}_{1.}$  denotes the average in block 1 and  $\bar{Y}_{3.}$  the average in block 3, then we have

$$\begin{aligned} E[\bar{Y}_{1.}] &= \mu + \rho_1 + \frac{1}{3}(\tau_B + \tau_C + \tau_E) \\ E[\bar{Y}_{3.}] &= \mu + \rho_3 + \frac{1}{3}(\tau_B + \tau_D + \tau_E). \end{aligned}$$

The difference of the block averages contains information about the treatments, namely,  $E[\bar{Y}_{1.} - \bar{Y}_{3.}] = \rho_1 - \rho_3 + \frac{1}{3}(\tau_C - \tau_D)$ . Unfortunately, this is not just a contrast among treatments, but involves the effects of the two blocks. The solution to uncovering the inter-block information is to let the block effects be random (with mean 0) since then  $E[\bar{Y}_{1.} - \bar{Y}_{3.}] = 0 - 0 + \frac{1}{3}(\tau_C - \tau_D) = \frac{1}{3}(\tau_C - \tau_D)$ , a contrast between treatment effects. The linear mixed model for the incomplete block design now becomes

$$Y_{ij} = \mu + \rho_i + \tau_j + e_{ij}, \quad e_{ij} \sim (0, \sigma^2), \quad \rho_j \sim (0, \sigma_\rho^2),$$

where the  $e_{ij}$  and  $\rho_j$  are independent. The term  $\rho_1 - \rho_3 + \frac{1}{3}(\tau_C - \tau_D)$  now represents the conditional (narrow inference space) comparison of the two block means and the unconditional (broad inference space) comparison is  $E[\bar{Y}_{1.} - \bar{Y}_{3.}] = E[E[\bar{Y}_{1.} - \bar{Y}_{3.} | \rho]] = \frac{1}{3}(\tau_C - \tau_D)$ .

For the corn hybrid experiment of Cochran and Cox (1957, p. 448) the inter-block analysis is carried out with the following PROC MIXED statements.

```

proc mixed data=cornylid;
  class block hybrid;
  model yield = hybrid;
  random block;
  lsmeans hybrid ;
  estimate 'hybrid 1 broad' intercept 1 hybrid 1;
  estimate 'hybrid 1 narrow' intercept 13 hybrid 13 |
    block 1 1 1 1 1 1 1 1 1 1 / divisor=13;
  estimate 'hybrid 1 in block 1' intercept 1 hybrid 1 | block 1;
  estimate 'hybrid 1 in block 2' intercept 1 hybrid 1 | block 0 1;
  estimate 'hybrid 1 vs hybrid2' hybrid 1 -1;
run; quit;

```

The inter-block analysis is invoked by moving the `block` term from the `model` statement to the `random` statement. The `estimate` statements are not necessary unless one wants to estimate treatment means in the narrow or intermediate inference spaces (§7.3). The `lsmeans` statement requests block-adjusted estimates of the hybrid means in the broad inference space. On Output 7.17 we notice that the  $F$  statistic for hybrid differences in the mixed analysis ( $F_{obs} = 1.67$ ) has changed from the intra-block analysis in `proc glm` ( $F_{obs} = 1.37$ ). This reflects the additional treatment information recovered by the inter-block analysis. Furthermore, the estimates of the treatment means have changed as compared to the least squares means reported by `proc glm`. The additional information recovered from block averages surfaces here again. In the example of §7.3 it was noted that the estimates of factor means would be the same if all factors would have been considered fixed and only the standard errors would differ between the fixed effects and mixed effects analysis. This statement was correct there because the design was completely balanced and hence orthogonal. In a nonorthogonal incomplete block design both the estimates of the treatment means as well as their standard errors differ between the fixed effects and mixed effects analysis.

#### Output 7.17.

##### The Mixed Procedure

Model Information	
Data Set	WORK.CORNYLD
Dependent Variable	yield
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class	Level Information	
	Levels	Values
block	13	1 2 3 4 5 6 7 8 9 10 11 12 13
hybrid	13	1 2 3 4 5 6 7 8 9 10 11 12 13

Dimensions	
Covariance Parameters	2
Columns in X	14
Columns in Z	13
Subjects	1
Max Obs Per Subject	52
Observations Used	52
Observations Not Used	0
Total Observations	52

Covariance Parameter Estimates	
Cov Parm	Estimate
block	6.0527
Residual	19.9340

## Output 7.17 (continued).

Fit Statistics						
Res Log Likelihood				-126.8		
Akaike's Information Criterion				-128.8		
Schwarz's Bayesian Criterion				-129.4		
-2 Res Log Likelihood				253.6		

Type 3 Tests of Fixed Effects						
Effect	Num DF	Den DF	F Value	Pr > F		
				1.67	0.1293	
hybrid	12	27				

Label	Estimates					
	Estimate	Error	DF	t Value	Pr >  t	
hybrid 1 broad	34.1712	2.4447	27	13.98	<.0001	
hybrid 1 narrow	34.1712	2.3475	27	14.56	<.0001	
hybrid 1 in block 1	32.6735	2.9651	27	11.02	<.0001	
hybrid 1 in block 2	31.2751	2.9651	27	10.55	<.0001	
hybrid 1 in block 7	34.1635	2.6960	27	12.67	<.0001	
hybrid 1 vs hybrid2	5.1305	3.3331	27	1.54	0.1354	

Least Squares Means						
Effect	hybrid	Standard				
		Estimate	Error	DF	t Value	Pr >  t
hybrid	1	34.1712	2.4447	27	13.98	<.0001
hybrid	2	29.0406	2.4447	27	11.88	<.0001
hybrid	3	30.1079	2.4447	27	12.32	<.0001
hybrid	4	28.0758	2.4447	27	11.48	<.0001
hybrid	5	30.3429	2.4447	27	12.41	<.0001
hybrid	6	27.5917	2.4447	27	11.29	<.0001
hybrid	7	30.7568	2.4447	27	12.58	<.0001
hybrid	8	32.7523	2.4447	27	13.40	<.0001
hybrid	9	28.5556	2.4447	27	11.68	<.0001
hybrid	10	28.1005	2.4447	27	11.49	<.0001
hybrid	11	23.4680	2.4447	27	9.60	<.0001
hybrid	12	28.9860	2.4447	27	11.86	<.0001
hybrid	13	35.1756	2.4447	27	14.39	<.0001

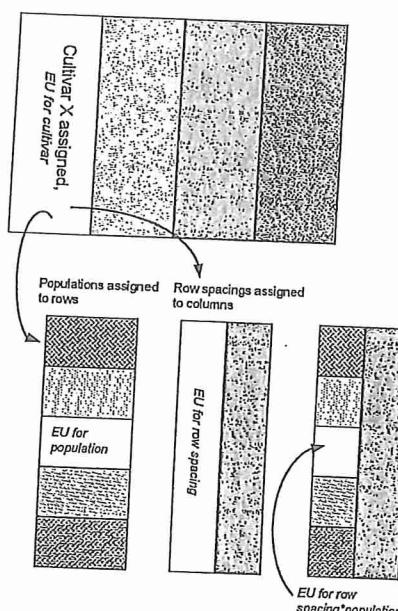
The first two estimate statements produce the hybrid 1 estimate in the broad and narrow inference space and show that the lsmeans statement operates in the broad inference space. The third through fifth estimate statements show how to estimate the hybrid mean in a particular plot. Notice that hybrid 1 did not appear in blocks 1 or 2 in the experiment but did in block 7. Nevertheless, we are able to predict how well the hybrid would have done in blocks 1 and 2, although this prediction is less precise than prediction of the hybrid's performance in blocks where the hybrid was observed. In an intra-block analysis where blocks are fixed, it is not possible to differentiate a hybrid's performance by block.

### 7.6.5 A Split-Strip-Plot Experiment for Soybean Yield

An experiment was conducted at the Tidewater Agricultural Research and Extension Center in Suffolk, Virginia to investigate how soybean yield response depended on soybean cultivar, row spacing, and plant population. The three factors and their levels considered in the experiment were

\* Cultivar (AG3601, AG3701, AG4601, AG4701)

- Plant Population (60, 120, 180, 240, 300 thousand per acre)
- Row spacing (9", 18").



**Figure 7.18.** Experimental layout for a single block in the soybean yield experiment. Cultivars (varieties) were assigned to the large experimental units (plots), row spacings and population densities to perpendicular strips within the plots. The experiment was brought to our attention and the data were made kindly available by Dr. David Holshouser, Tidewater Agricultural Research and Extension Center, Virginia Polytechnic Institute and State University. Used with permission.

Although the experiment was conducted in four site-years, we consider only a single site-year here. At each site four replications of the cultivars were arranged in a randomized block design. Because of technical limitations, it was decided to apply the row spacing and population densities in strips within a cultivar experimental unit (plot). It was determined at random which side (strip) of the plot received 9" spacing. Then the population densities were assigned randomly to five strips running perpendicular to the row spacing strips. Figure 7.18 displays a schematic layout of one of the four blocks in the experiment.

The factors Row Spacing and Population Density are a split of the experimental unit to which a cultivar is assigned, but are not arranged in a  $2 \times 5$  factorial structure. Considering the cultivar experimental units, Row Spacing and Population Density form a strip-plot (split-block) design with 16 blocks (replications). Each cultivar experimental unit serves as a replicate for the split-block design of the other two factors. We call this design a split-strip-plot design.

There are experimental units of four different sizes in this experiment, hence the linear model will contain four different experimental error sources of variability associated with the