Topic 19: Split-plot design

Montgomery: chapter -14

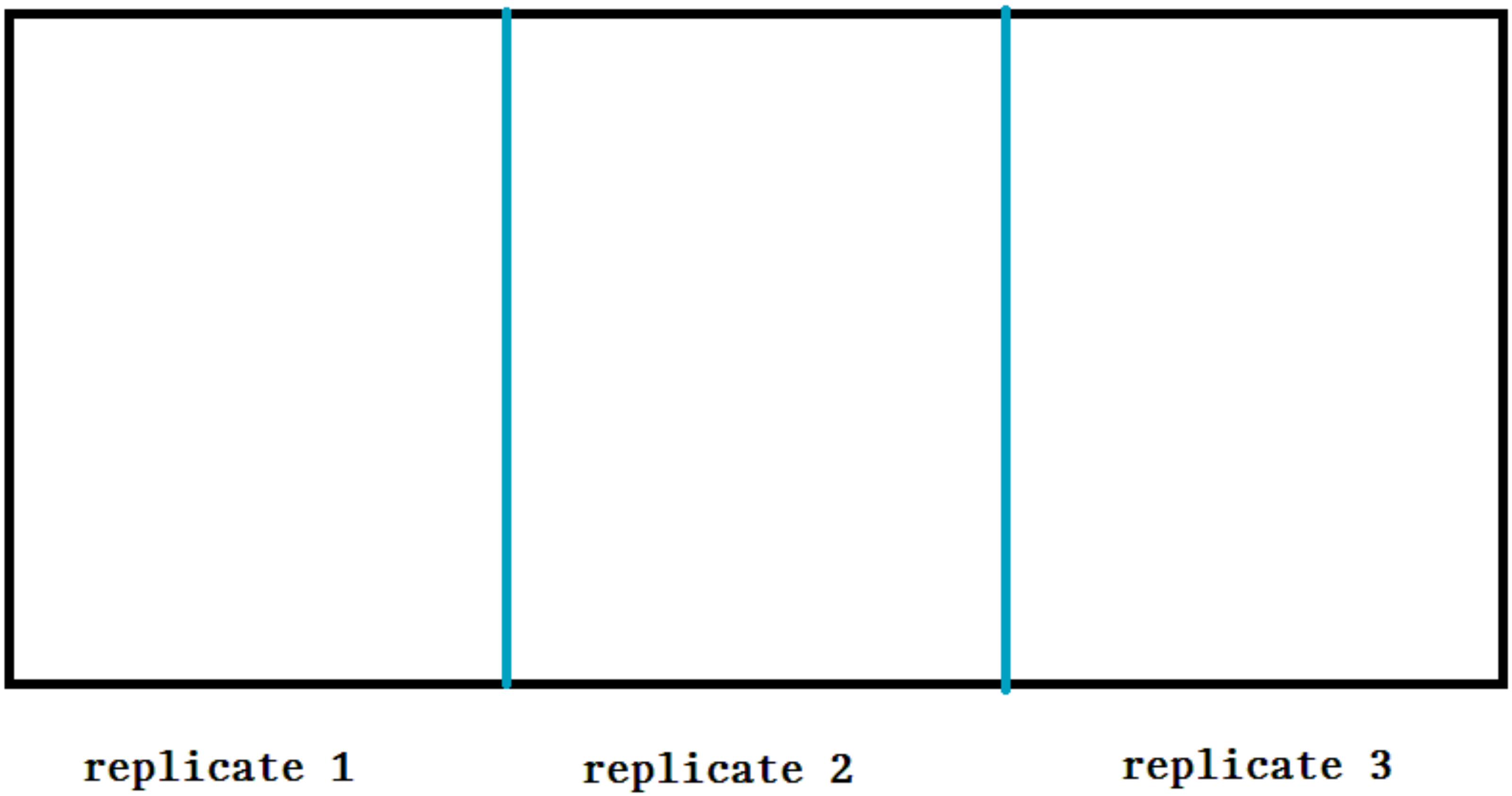
Prof. Lingling An
University of Arizona

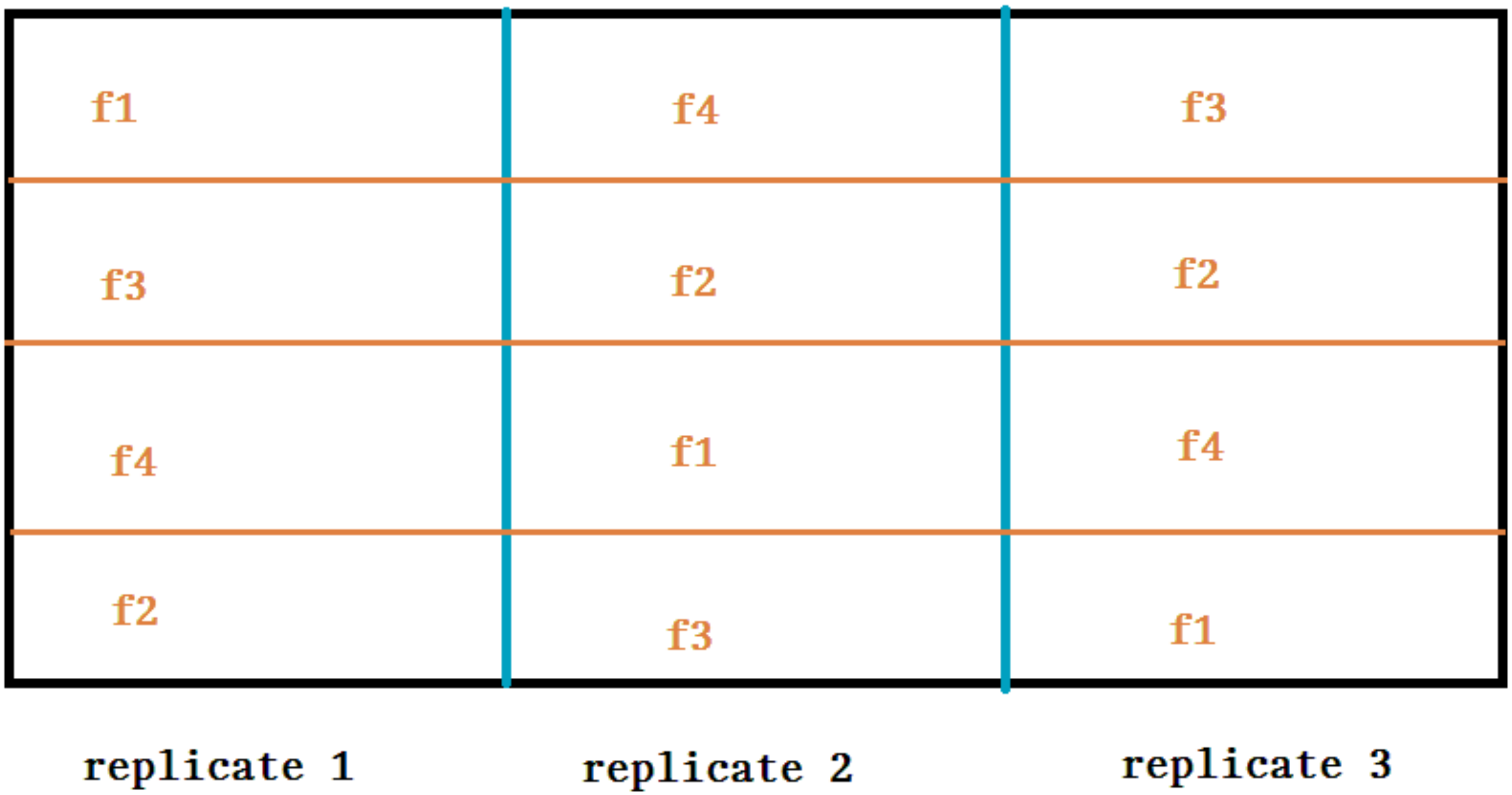
Example 1

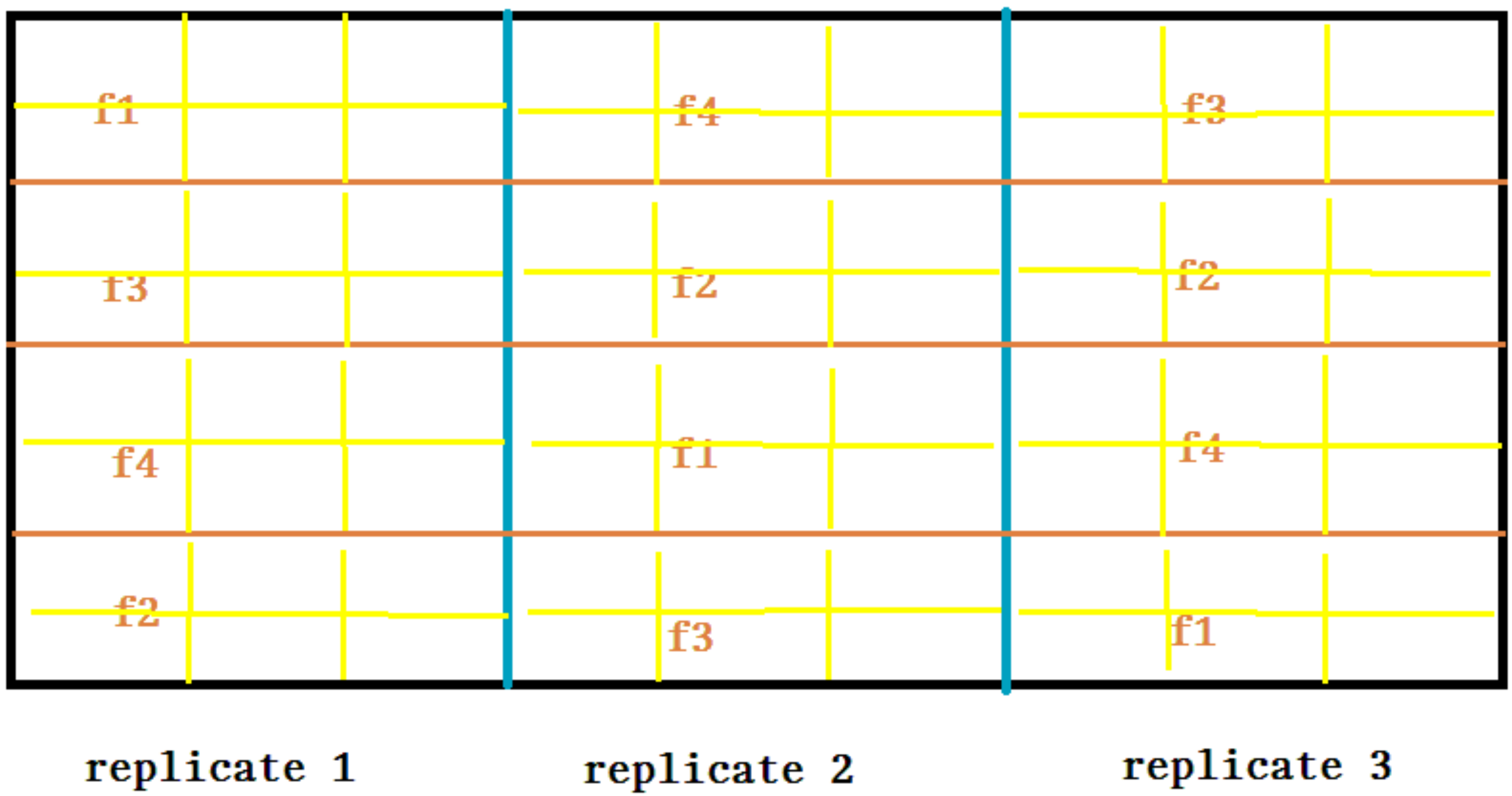
- Example 1: Study six corn varieties and four fertilizers and yield is the response. Three replicates are needed.

Method 1: completely randomized full factorial design, 24 level combinations of variety and fertilizer are applied to 24*3=72 pieces of land (each to three).

Method 2: Select three fields of large area. Each field is divided into four areas (four whole-plots), four fertilizers are randomly assigned to the four whole-plots. Each area is further divided into six subareas (sub-plots), and the six varieties are randomly planted in these sub-plots.


replicate 1
replicate 2
replicate 3


f1
f4
f3
f3
f2
f2
f4
f1
f4
f2
f3
f1
replicate 1
replicate 2
replicate 3


f1
f4
f3
f3
f2
f2
f4
f1
f4
f2
f3
f1
replicate 1
replicate 2
replicate 3

This leads to a split-plot design:

- whole-plot (treatment) factor: fertilizer
- sub-plot (treatment) factor: corn variety

Example 2

- Example 2: A paper manufacturer is investigating three different pulp preparation methods and four different cooking temperatures for the pulp and study their effect on the tensile strength of the paper. Three replicates are needed.
- Because the pilot plant is only capable of making 12 runs per day, so the experimenter decides to run one replicate on each of the three days and to consider the days as blocks.

- On any day, a batch of pulp is produced by one of the the three methods (a whole-plot). Then the batch is divided into four samples (four sub-plots), and each sample is cooked at one of the four temperatures. Then a second batch of pulp is made up using another of the three methods. This second batch is also divided into four samples that are tested at the four temperatures. The process is then repeated for the third method. The data is given below.

	Day 1			Day 2			Day 3		
Method	1	2	3	1	2	3	1	2	3
Temp									
200	30	34	29	28	31	31	31	35	32
225	35	41	26	32	36	30	27	40	34
250	27	38	33	40	42	32	41	39	39
275	36	42	36	41	40	40	40	44	45

- The split-plot is a multifactor experiment where it is not possible to completely randomize the order of the runs:
  - In replicate 1, select a pulp preparation method, prepare a batch
  - Divide the batch into four sections or samples, and assign one of the temperature levels to each
  - Repeat for each pulp preparation method
  - Conduct replicates 2 and 3 similarly

- Each replicate (sometimes called **blocks**) has been divided into three parts, called the **whole plots**
- Pulp preparation methods is the **whole plot treatment**
- Each whole plot has been divided into four **subplots** or **split-plots**
- Temperature is the **subplot treatment**
- Generally, the hard-to-change factor is assigned to the whole plots
- This design requires only 9 batches of pulp (assuming three replicates)

Split-Plot Structure

- factors are crossed (different than nested)
- randomization restriction (different than completely randomized)
- Information on factor effects from two levels (or strata).
- split-plot can be considered as two superimposed blocked designs:
  - A: whole-plot factor(a); B: sub-plot factor (b), r replicates
  - RCBD$_A$: number of trt: a, number of blk: r.
  - RCBD$_B$: number of trt: b, number of blk: ra.

  for whole-plots, subdivision to smaller sub-plots are ignored. For sub-plots, whole-plots considered blocks.
- More power for main subplot effect and interaction
- Should use design only for practical reasons
- Randomized factorial design more powerful if feasible

A typical Data Layout for split-plot design

	Block 1			Block 2			Block 3		
WP-Factor A	1	2	3	1	2	3	1	2	3
SP-Factor B									
1	y_{111}	y_{121}	.	.	.	.	.	.	y_{331}
2	y_{112}	y_{122}	.	.	.	.	.	.	y_{332}
3	y_{113}	y_{123}	.	.	.	.	.	.	y_{333}
4	y_{114}	y_{124}	.	.	.	.	.	.	y_{334}

In general:

y_{ijk} where i denotes Block i, j denotes the jth level of the whole-plot factor A, and k denotes the kth level of the sub-plot factor B.

Statistical Model I

$$y_{ijk} = \mu + r_i + \alpha_j + (r\alpha)_{ij} + \beta_k + (r\beta)_{ik} + (\alpha\beta)_{jk} + (r\alpha\beta)_{ijk} + \epsilon_{ijk}$$

$$i = 1, 2, \ldots, r, j = 1, 2, \ldots, a, k = 1, 2, \ldots, b$$

- r_i: block effects (random) $\sim N(0, \sigma_r^2)$
- α_j: whole-plot factor (A) main effects (fixed)
- $(r\alpha)_{ij}$: whole-plot error (random) $\sim$ normal with $\sigma^2_{r\alpha}$.

- β_k: sub-plot factor (B) main effects (fixed)
- $(r\beta)_{ik}$: block-B interaction (random) $\sim$ normal with $\sigma^2_{r\beta}$.
- $(\alpha\beta)_{jk}$ Interaction between A and B (fixed)
- $(r\alpha\beta)_{ijk}$: sub-plot error (random) $\sim$ normal with $\sigma^2_{r\alpha\beta}$
- ϵ_{ijk}: random error $\sim N(0, \sigma^2)$

Sum of Squares

- $\mathrm{SS_r} = ab\sum_i(\bar{y}_{i..} - \bar{y}_{...})^2$, df=r-1.
- $\mathrm{SS_A} = rb\sum_j(\bar{y}_{.j.} - \bar{y}_{...})^2$, df=a-1.
- $\mathrm{SS_{rA}} = b\sum_{i,j}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$, df=(r-1)(a-1)
- $\mathrm{SS_B} = ar\sum_k(\bar{y}_{..k} - \bar{y}_{...})^2$, df=(b-1)
- $\mathrm{SS_{rB}} = a\sum_{i,k}(\bar{y}_{i.k} - \bar{y}_{i..} - \bar{y}_{..k} + \bar{y}_{...})^2$ df=(r-1)(b-1)
- $\mathrm{SS_{AB}} = r\sum_{j,k}(\bar{y}_{.jk} - \bar{y}_{.j.} - \bar{y}_{..k} + \bar{y}_{...})^2$ df=(a-1)(b-1)
- $\mathrm{SS_{rAB}} = \sum_{i,j,k}(y_{ijk} - \bar{y}_{ij.} - \bar{y}_{i.k} - \bar{y}_{.jk} + \bar{y}_{i..} + \bar{y}_{.j.} + \bar{y}_{..k} - \bar{y}_{...})^2$, df=(r-1)(a-1)(b-1).
- $\mathrm{SS_E} =$?

Expected mean squares (restricted)

		r	a	b	1	
		R	F	F	R	
	term	i	j	k	h	$E(MS)$
	r_i	1	a	b	1	$\sigma^2 + ab\sigma_r^2$
whole plot	α_j	r	0	b	1	$\sigma^2 + b\sigma_{r\alpha}^2 + \frac{rb\Sigma\alpha_j^2}{a-1}$
	$(r\alpha)_{ij}$	1	0	b	1	$\sigma^2 + b\sigma_{r\alpha}^2$
	β_k	r	a	0	1	$\sigma^2 + a\sigma_{r\beta}^2 + \frac{ra\Sigma\beta_k^2}{b-1}$
	$(r\beta)_{ik}$	1	a	0	1	$\sigma^2 + a\sigma_{r\beta}^2$
subplot	$(\alpha\beta)_{jk}$	r	0	0	1	$\sigma^2 + \sigma_{r\alpha\beta}^2 + \frac{r\Sigma\Sigma(\alpha\beta)_{jk}^2}{(a-1)(b-1)}$
	$(r\alpha\beta)_{ijk}$	1	0	0	1	$\sigma^2 + \sigma_{r\alpha\beta}^2$
	ϵ_{ijk}	1	1	1	1	σ^2 (not estimable)

There are two error structures; the whole-plot error and the subplot error

Estimates and tests of fixed effects

- $\hat{\alpha}_j = \bar{y}_{.j.} - \bar{y}_{...}$ for $j = 1, 2, \ldots, a$
- $\hat{\beta}_k = \bar{y}_{..k} - \bar{y}_{...}$ for $k = 1, 2, \ldots, b$
- $(\hat{\alpha\beta})_{jk} = \bar{y}_{.jk} - \bar{y}_{.j.} - \bar{y}_{..k} + \bar{y}_{...}$.
- Test $\alpha_j = 0$, $F_0 = \mathrm{MS_A}/\mathrm{MS_{rA}}$
- Test $\beta_k = 0$, $F_0 = \mathrm{MS_B}/\mathrm{MS_{rB}}$
- Test $(\alpha\beta)_{jk} = 0$, $F_0 = \mathrm{MS_{AB}}/\mathrm{MS_{rAB}}$.

SAS Code – proc GLM

```
data paper;
input block method temp resp@@;
datalines;
1 1 1 30 1 1 2 35 1 1 3 37 1 1 4 36
1 2 1 34 1 2 2 41 1 2 3 38 1 2 4 42
1 3 1 29 1 3 2 26 1 3 3 33 1 3 4 36
2 1 1 28 2 1 2 32 2 1 3 40 2 1 4 41
2 2 1 31 2 2 2 36 2 2 3 42 2 2 4 40
2 3 1 31 2 3 2 30 2 3 3 32 2 3 4 40
3 1 1 31 3 1 2 37 3 1 3 41 3 1 4 40
3 2 1 35 3 2 2 40 3 2 3 39 3 2 4 44
3 3 1 32 3 3 2 34 3 3 3 39 3 3 4 45
;
proc glm data=paper;
class block method temp;
```

```
model resp=block method block*method temp block*temp
 method*temp block*method*temp;
random block block*method block*temp block*method*temp;
test h=method e=block*method;
test h=temp e=block*temp;
test h=method*temp e=block*method*temp;
run;
```

Output

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	35	822.9722222	23.5134921	.	.
Error	0	0.0000000	.		
CoTotal	35	822.9722222			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
block	2	77.5555556	38.7777778	.	.
method	2	128.3888889	64.1944444	.	.
block*method	4	36.2777778	9.0694444	.	.
temp	3	434.0833333	144.6944444	.	.
block*temp	6	20.6666667	3.4444444	.	.
method*temp	6	75.1666667	12.5277778	.	.
blo*meth*tmp	12	50.8333333	4.2361111	.	.

Tests Using the Type III MS for block*method as Error Ter

Source	DF	Type III SS	Mean Square	F Value	Pr >
method	2	128.3888889	64.1944444	7.08	0.0485

Tests Using the Type III MS for block*temp as Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr >
temp	3	434.0833333	144.6944444	42.01	0.0002

Tests Using the Type III MS for block*method*temp as E.Te

Source	DF	Type III SS	Mean Square	F Value	Pr
method*temp	6	75.1666667	12.52777778	2.96	0.05

SAS code – proc mixed

```
proc mixed data=paper method=type1;

class block method temp;

model resp=method temp method*temp;

random block block*method block*temp
block*method*temp;

run;
```

default is "REML"

output

The Mixed Procedure

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	Expected Mean Square
method	2	128.388889	64.194444	Var(Residual) + Var(block*method*temp) + 4 Var(block*method) + Q(method,method*temp)
temp	3	434.083333	144.694444	Var(Residual) + Var(block*method*temp) + 3 Var(block*temp) + Q(temp,method*temp)
method*temp	6	75.166667	12.527778	Var(Residual) + Var(block*method*temp) + Q(method*temp)
block	2	77.555556	38.777778	Var(Residual) + Var(block*method*temp) + 3 Var(block*temp) + 4Var(block*method) + 12 Var(block)

Source	DF	Sum of Squares	Mean Square	Expected Mean Square
block*method	4	36.277778	9.069444	Var(Residual) + Var(block*method*temp) + 4 Var(block*method)
block*temp	6	20.666667	3.444444	Var(Residual) + Var(block*method*temp) + 3 Var(block*temp)
block*method*temp	12	50.833333	4.236111	Var(Residual) + Var(block*method*temp)
Residual	0	0	1.110223E-12	0

The Mixed Procedure

Type 1 Analysis of Variance

Source	Error Term	Error DF	F Value	Pr > F
method	MS(block*method)	4	7.08	0.0485
temp	MS(block*temp)	6	42.01	0.0002
method*temp	MS(block*method*temp)	12	2.96	0.0520
block	MS(block*method) + MS(block*temp) - MS(block*method*temp)	2.8507	4.68	0.1256
block*method	MS(block*method*temp)	12	2.14	0.1382
block*temp	MS(block*method*temp)	12	0.81	0.5797
block*method*temp	MS(Residual)	0	3.82E12	.
Residual	.	.	.	.

Covariance Parameter Estimates

Cov Parm	Estimate
block	2.5417
block*method	1.2083
block*temp	-0.2639
block*method*temp	4.2361
Residual	1.11E-12

■ **TABLE 14.16**

Analysis of Variance for the Split-Plot Design Using the Tensile Strength Data from Table 14.14

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Replicates (or blocks)	77.55	2	38.78		
Preparation method (*A*)	128.39	2	64.20	7.08	0.05
Whole plot error (replicates (or blocks) × *A*)	36.28	4	9.07		
Temperature (*B*)	434.08	3	144.69	41.94	<0.01
Replicates (or blocks) × *B*	20.67	6	3.45		
AB	75.17	6	12.53	2.96	0.05
Subplot error (replicates (or blocks) × *AB*)	50.83	12	4.24		
Total	822.97	35			

Statistical Model II

$$y_{ijk} = \mu + r_i + \alpha_j + (r\alpha)_{ij} + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

- r_i: block effects (random) $\sim N(0, \sigma_r^2)$
- α_j: whole-plot factor (A) main effects (fixed)
- $(r\alpha)_{ij}$: whole plot error $\sim$ normal with $\sigma_{r\alpha}^2$
- β_k: sub-plot factor (B) main effects (fixed)
- $(\alpha\beta)_{jk}$: A and B interaction (fixed)
- ϵ_{ijk}: sub-plot error $N(0, \sigma_\epsilon^2)$.

- Expected mean square

Term	$E(MS)$
r_i	$\sigma_\epsilon^2 + ab\sigma_r^2$
α_j(A)	$\sigma_\epsilon^2 + b\sigma_{r\alpha}^2 + \frac{rb\Sigma\alpha_j^2}{a-1}$
$(r\alpha)_{ij}$	$\sigma_\epsilon^2 + b\sigma_{r\alpha}^2$ (whole plot error)
β_k (B)	$\sigma_\epsilon^2 + \frac{ra\Sigma\alpha_j^2}{b-1}$
$(\alpha\beta)_{jk}$ (AB)	$\sigma_\epsilon^2 + \frac{r\Sigma\Sigma(\alpha\beta)_{jk}^2}{(a-1)(b-1)}$
ϵ_{ijk}	σ_ϵ^2 (subplot error)

SAS code

```
proc glm data=paper;
class block method temp;
model resp=block method block*method temp method*temp;
random block block*method;
test h=method e=block*method;
run;

proc mixed data=paper method=type1;
class block method temp;
model resp=method temp method*temp;
random block block*method;
run;
```

Output - glm

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	17	751.4722222	44.2042484	11.13	<.0001
Error	18	71.5000000	3.9722222		
Corrected Total	35	822.9722222			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
block	2	77.5555556	38.7777778	9.76	0.0013
~~method~~	~~2~~	~~128.3888889~~	~~64.1944444~~	~~16.16~~	~~<.0001~~
block*method	4	36.2777778	9.0694444	2.28	0.1003
temp	3	434.0833333	144.6944444	36.43	<.0001
method*temp	6	75.1666667	12.5277778	3.15	0.0271

Output - glm

Dependent Variable: resp

Tests of Hypotheses Using the Type III MS for block*method as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
method	2	128.3888889	64.1944444	7.08	0.0485

Output - mixed

Type 1 Analysis of Variance

Source	Error Term	Error DF	F Value	Pr > F
method	MS(block*method)	4	7.08	0.0485
temp	MS(Residual)	18	36.43	<.0001
method*temp	MS(Residual)	18	3.15	0.0271
~~block~~	~~MS(block*method)~~	~~4~~	~~4.28~~	~~0.1016~~
block*method	MS(Residual)	18	2.28	0.1003
Residual	.	.	.	.

General Split-Plot Designs

- Can have $>$ one whole-plot factor and $>$ one subplot factor with various blocking schemes.
- split-plot design consists of two superimposed blocked design

  Whole Plot

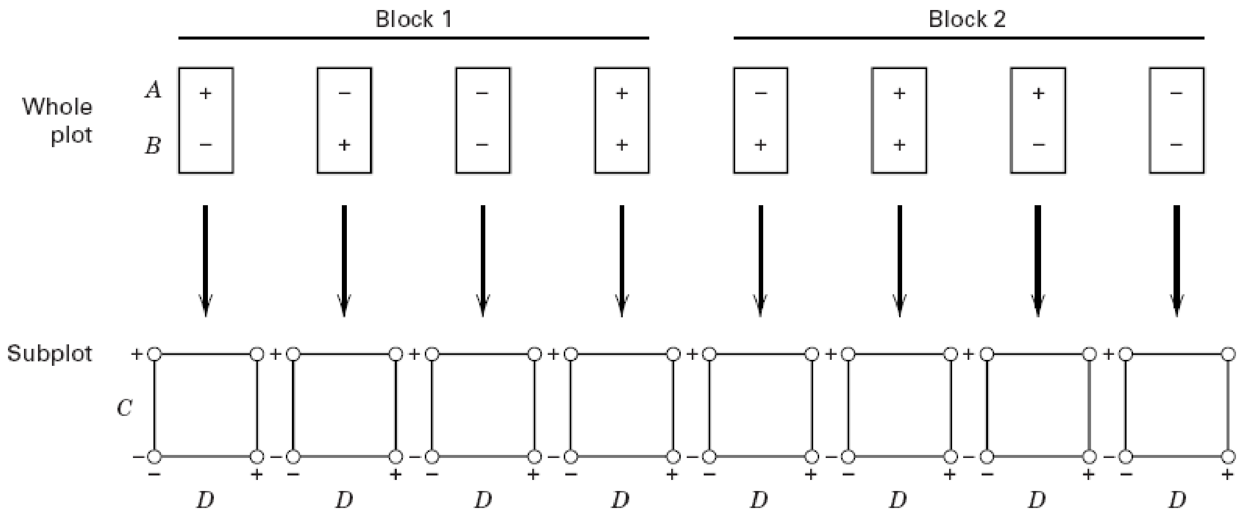
  – CRD, RCBD, Factorial D, BIBD, etc.

  Subplot

  – RCBD, BIBD, Factorial Design, etc.
- Analysis of Covariance

More than two factors – see page 627

A & *B* (gas flow & temperature) are hard to change; *C* & *D* (time and wafer position) are easy to change.



■ **FIGURE 14.7** A split-plot design with four design factors, two in the whole plot and two in the subplot

Other Variations

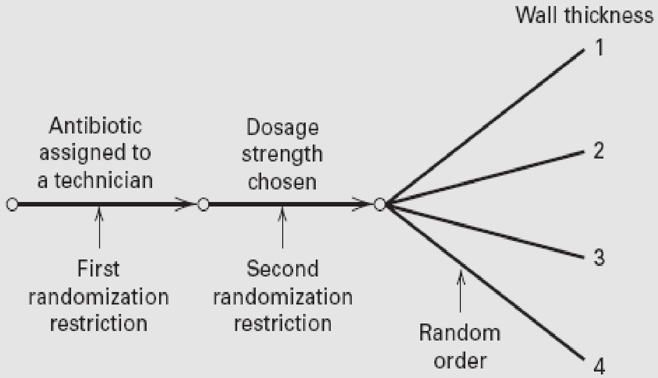
- Split-split-plot design

  1. randomization restriction can occur at any number of levels within the experiment

  2. two-level: split-split-plot design

- Strip-split-plot design (or Criss cross design, or Split-block design)

probably will have to id this vs random factor
the replication has become a factor here
(it's not inside the factors)



		Technician								
		1			2			3		
Blocks	Dosage strength	1	2	3	1	2	3	1	2	3
1	Wall thicknesses	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
2	Wall thicknesses	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
3	Wall thicknesses	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
4	Wall thicknesses	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4

FIGURE 14.10 A split-split-plot design

Absorption times of antibiotic capsule

- 3 technicians
- 3 dosage strengths
- 4 capsule wall thickness
- 4 replicates/ days
- A split-split-plot design
- Two randomization restrictions present within each replicate

- Strip-split-plot design (or Criss cross design, or Split-block design)

  Example: we want to compare the yield of a certain crop under different systems of soil preparation (A : a_1, a_2, a_3, a_4) and different density of seeding (B: b_1, b_2, b_3, b_4, b_5). Both operations (tilling and seeding) are done mechanically and it is impossible to perform both on small pieces of land. The arrangement shown below (strip-split-plot design) is then replicated r times, each time using different randomizations for A and B.

Strip plots

	b1	*b4*	*b2*	*b3*	*b5*
a4	a_4b_1	a_4b_4	a_4b_2	a_4b_3	a_4b_5
a1	a_1b_1	a_1b_4	a_1b_2	a_1b_3	a_1b_5
a2	a_2b_1	a_2b_4	a_2b_2	a_2b_3	a_2b_5
a3	a_3b_1	a_3b_4	a_3b_2	a_3b_3	a_3b_5

Whole plots

Last slide

- Read Sections: 14.4-14.5