Topic 18: Nested design

Montgomery: chapter -14

Prof. Lingling An
University of Arizona

Nested and Split-Plot Designs

- Text reference, Chapter 14
- These are **multifactor** experiments that have some important industrial/ agricultural applications
- Nested and split-plot designs frequently involve one or more **random** factors, so the methodology of Chapter 13 (expected mean squares, variance components) is important
- There are **many** variations of these designs – we consider only some basic situations

Crossed vs Nested Factors

- Factors A (a levels)and B (b levels) are considered crossed if

  Every combinations of A and B (ab of them) occurs.

  An example:

	Factor A			
Factor B	1	2	3	4
1	xx	xx	xx	xx
2	xx	xx	xx	xx
3	xx	xx	xx	xx

A	1			2			3			4		
B	1	2	3	1	2	3	1	2	3	1	2	3
	x	x	x	x	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x	x	x	x	x

- Factor B is considered nested under A (a levels) if
  1. under each fixed level (i) of A, B has b_i levels.
  2. the levels of B under the same level of A are comparable.
  3. under a level of A, the levels of B can be arbitrarily numbered.

A	1			2			3			4		
B	1	2	3	4	5	6	7	8	9	10	11	12
	x	x	x	x	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x	x	x	x	x

Material Purity Experiment

Consider a company that buys raw material in batches from three different suppliers. The purity of this raw material varies considerably, which causes problems in manufacturing the finished product. We wish to determine if the variability in purity is attributable to difference between the suppliers. Four batches of raw material are selected at random from each supplier, three determinations of purity are made on each batch. The data, after coding by subtracting 93 are given below.

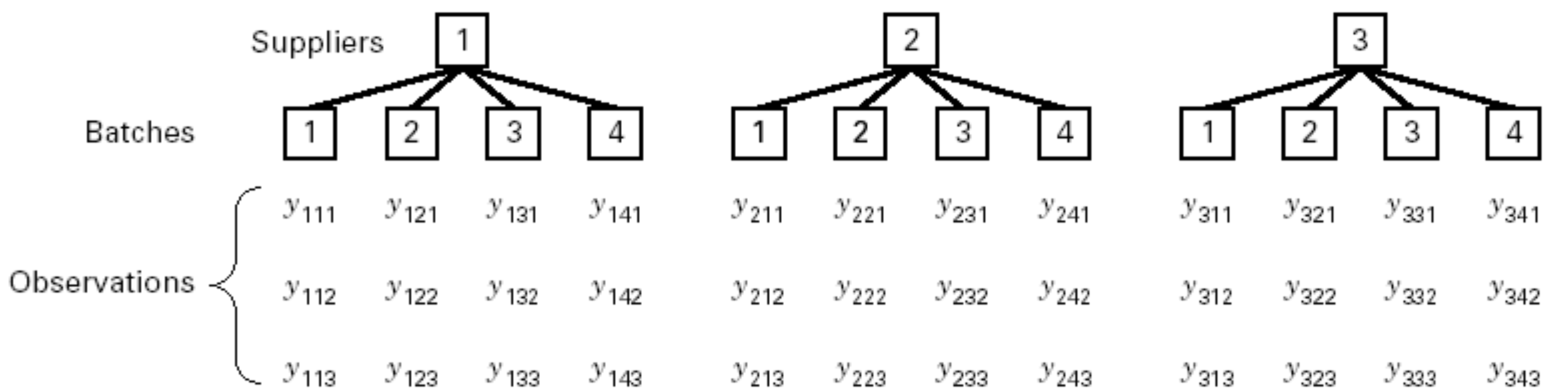


Figure 14-1 A two-stage nested design.

	Supplier 1				Supplier 2				Supplier 3			
Batches	1	2	3	4	1	2	3	4	1	2	3	4
	1	-2	-2	1	1	0	-1	0	2	-2	1	3
	-1	-3	0	4	-2	4	0	3	4	0	-1	2
	0	-4	1	0	-3	2	-2	2	0	2	2	1
$y_{ij.}$	0	-9	-1	5	-4	6	-3	5	6	0	2	6
$y_{i..}$	-5				4				14			

Other Examples for Nested Factors

1. Drug company interested in stability of product
   - Two manufacturing sites
   - Three batches from each site
   - Ten tablets from each batch
2. Stratified random sampling procedure
   - Randomly sample five states
   - Randomly select three counties
   - Randomly select two towns
   - Randomly select five households

Statistical Model

- Two factor nested model

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

- Bracket notation represents nested factor
- Cannot include interaction
- Factors may be random or fixed
- Can use EMS algorithm to derive tests

Sum of Squares Decomposition

$$y_{ijk} = \bar{y}_{...} + (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{ij.} - \bar{y}_{i..}) + (y_{ijk} - \bar{y}_{ij.}).$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2 = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 + n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..})^2$$

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$$

$$\mathrm{SS_T} = \mathrm{SS_A} + \mathrm{SS_{B(A)}} + \mathrm{SS_E}$$

Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	SS_{A}	$a-1$	MS_{A}	
B(A)	$\text{SS}_{\text{B(A)}}$	$a(b-1)$	$\text{MS}_{\text{B(A)}}$	
Error	SS_{E}	$ab(n-1)$	MS_{E}	
Total	SS_{T}	$abn-1$		

$$\text{SS}_{\text{T}} = \sum\sum\sum y_{ijk}^2 - y_{...}^2/abn$$

$$\text{SS}_{\text{A}} = \frac{1}{bn}\sum y_{i..}^2 - y_{...}^2/abn$$

$$\text{SS}_{\text{B(A)}} = \frac{1}{n}\sum\sum y_{ij.}^2 - \frac{1}{bn}\sum y_{i..}^2$$

$$\text{SS}_{\text{E}} = \sum\sum\sum y_{ijk}^2 - \frac{1}{n}\sum\sum y_{ij.}^2$$

- Use EMS to define proper tests

Two-Factor Nested Model with Fixed Effects

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)}$$

where (1) $\sum_{i=1}^{a} \tau_i = 0$, (2) $\sum_{j=1}^{b} \beta_{j(i)} = 0$ for each i.

	F	F	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + \frac{bn\Sigma\tau_i^2}{a-1}$
$\beta_{j(i)}$	1	0	n	$\sigma^2 + \frac{n\Sigma\Sigma\beta_{j(i)}^2}{a(b-1)}$
$\epsilon_{k(ij)}$	1	1	1	σ^2

- Estimates: $\hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$; $\hat{\beta}_{j(i)} = \bar{y}_{ij.} - \bar{y}_{i..}$.
- Tests: $\mathrm{MS_A}/\mathrm{MS_E}$ for $\tau_i = 0$; $\mathrm{MS_{B(A)}}/\mathrm{MS_E}$ for $\beta_{j(i)} = 0$.

Two-Factor Nested Model with Random Effects

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)}$$

where $\tau_i \sim N(0, \sigma_\tau^2)$ and $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$.

	R	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	1	b	n	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\tau^2$
$\beta_{j(i)}$	1	1	n	$\sigma^2 + n\sigma_\beta^2$
$\epsilon_{k(ij)}$	1	1	1	σ^2

- Estimates: $\hat{\sigma}_\tau^2 = (\mathrm{MS_A} - \mathrm{MS_{B(A)}})/nb$; $\hat{\sigma}_\beta^2 = (\mathrm{MS_{B(A)}} - \mathrm{MS_E})/n$.
- tests: $\mathrm{MS_A}/\mathrm{MS_{B(A)}}$ for $\sigma_\tau^2 = 0$; $\mathrm{MS_{B(A)}}/\mathrm{MS_E}$ for $\sigma_\beta^2 = 0$.

Two-Factor Nested Model with Mixed Effects

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)}$$

where $\sum_{i=1}^{a} \tau_i = 0$, and $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$.

	F	R	R	
	a	b	n	
term	i	j	k	EMS
τ_i	0	b	n	$\sigma^2 + n\sigma_\beta^2 + \frac{bn\Sigma\tau_i^2}{a-1}$
$\beta_{j(i)}$	1	0	n	$\sigma^2 + n\sigma_\beta^2$
$\epsilon_{k(ij)}$	1	1	1	σ^2

- Estimates: $\hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$; $\hat{\sigma}_\beta^2 = (\mathrm{MS_{B(A)}} - \mathrm{MS_E})/n$.
- Tests: $\mathrm{MS_A}/\mathrm{MS_{B(A)}}$ for $\tau_i = 0$; $\mathrm{MS_{B(A)}}/\mathrm{MS_E}$ for $\sigma_\beta^2 = 0$.

■ TABLE 14.1

Expected Mean Squares in the Two-Stage Nested Design

$E(MS)$	A Fixed B Fixed	A Fixed B Random	A Random B Random
$E(MS_A)$	$\sigma^2 + \dfrac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + \dfrac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\tau^2$
$E(MS_{B(A)})$	$\sigma^2 + \dfrac{n\sum\sum \beta_{j(i)}^2}{a(b-1)}$	$\sigma^2 + n\sigma_\beta^2$	$\sigma^2 + n\sigma_\beta^2$
$E(MS_E)$	σ^2	σ^2	σ^2

SAS Code for Purity Experiment

```
option nocenter ps=40 ls=72;
data purity;
input supp batch resp@@;
datalines;
1 1 1 1 1 -1 1 1 0
1 2 -2 1 2 -3 1 2 -4
1 3 -2 1 3 0 1 3 1
1 4 1 1 4 4 1 4 0
2 1 1 2 1 -2 2 1 -3
2 2 0 2 2 4 2 2 2
2 3 -1 2 3 0 2 3 -2
2 4 0 2 4 3 2 4 2
3 1 2 3 1 4 3 1 0
3 2 -2 3 2 0 3 2 2
3 3 1 3 3 -1 3 3 2
3 4 3 3 4 2 3 4 1
;
```

```
/* if both supp and batch are random*/
proc mixed method=type1;
class supp batch;
model resp=;
random supp batch(supp);
run;
```

Type 1 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
supp	2	15.055556	7.527778	Var(Residual) + 3 Var(batch(supp)) + 12 Var(supp)	MS(batch(supp))	9	0.97	0.4158
batch(supp)	9	69.916667	7.768519	Var(Residual) + 3 Var(batch(supp))	MS(Residual)	24	2.94	0.0167
Residual	24	63.333333	2.638889	Var(Residual)	.	.	.	.

Covariance Parameter Estimates	
Cov Parm	Estimate
supp	-0.02006
batch(supp)	1.7099
Residual	2.6389

```
/* if only batch is random*/
proc mixed method=type1;
class supp batch;
model resp=supp;
random batch(supp);
run;
```

Type 1 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
supp	2	15.055556	7.527778	Var(Residual) + 3 Var(batch(supp)) + Q(supp)	MS(batch(supp))	9	0.97	0.4158
batch(supp)	9	69.916667	7.768519	Var(Residual) + 3 Var(batch(supp))	MS(Residual)	24	2.94	0.0167
Residual	24	63.333333	2.638889	Var(Residual)	.	.	.	.

Covariance Parameter Estimates	
Cov Parm	Estimate
batch(supp)	1.7099
Residual	2.6389

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
supp	2	9	0.97	0.4158

or

- Estimates:

$$\hat{\tau}_1 = \bar{y}_{1..} - \bar{y}_{...} = -28/36$$

$$\hat{\tau}_2 = \bar{y}_{2..} - \bar{y}_{...} = -1/36$$

$$\hat{\tau}_3 = \bar{y}_{3..} - \bar{y}_{...} = -29/36$$

$$\hat{\sigma}_\tau^2 = MS_E = 2.64$$

$$\hat{\sigma}_\beta^2 = \frac{MS_{B(A)} - MS_E}{n} = \frac{7.77 - 2.64}{3} = 1.71$$

- Hypothesis test

$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$:

$$F_0 = .97, \text{ P-value} = 0.4158, \text{ Accept } H_0$$

$H_0 : \sigma_\beta^2 = 0$:

$$F_0 = 2.94, \text{ P-value} = 0.0167, \text{ Reject } H_0$$

- Suppliers are not different, variability due to batches.

```
/* if both supp and batch are fixed*/
proc mixed method=type1;
class supp batch;
model resp=supp batch(supp);
run;
```

Type 1 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
supp	2	15.055556	7.527778	Var(Residual) + Q(supp,batch(supp))	MS(Residual)	24	2.85	0.0774
batch(supp)	9	69.916667	7.768519	Var(Residual) + Q(batch(supp))	MS(Residual)	24	2.94	0.0167
Residual	24	63.333333	2.638889	Var(Residual)	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
supp	2	24	2.85	0.0774
batch(supp)	9	24	2.94	0.0167

Other Scenarios for Nested Factors

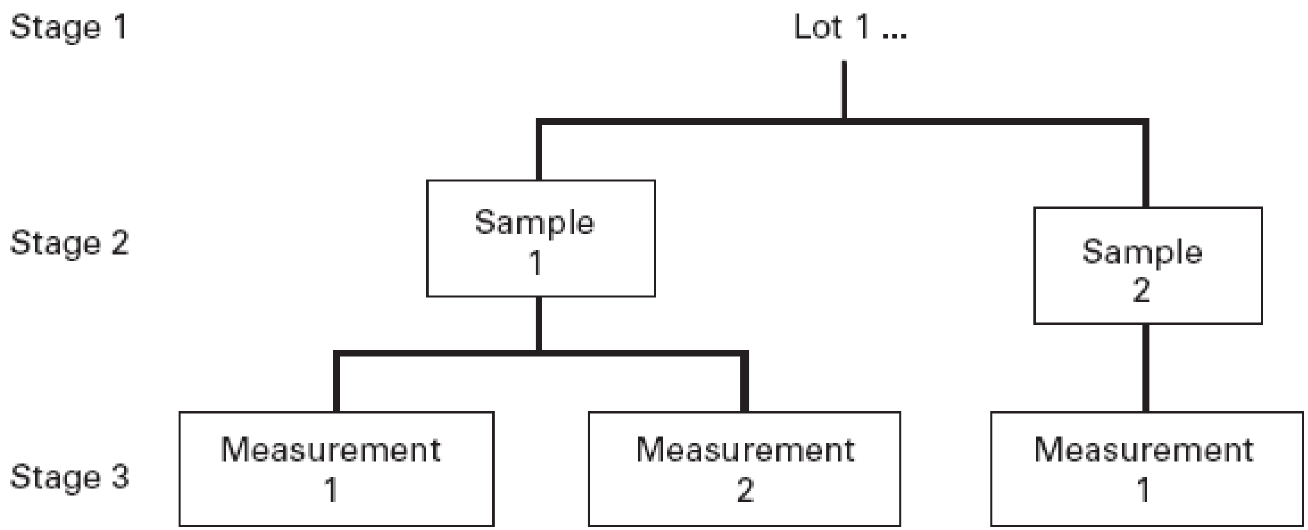
- Staggered Nested Designs
- General m-Stage Nested Designs

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(ij)} + \epsilon_{l(ijk)}$$

- Designs with Both Nested and Factorial Factors

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(j)} + (\tau\beta)_{ij} + (\tau\gamma)_{ik(j)} + \epsilon_{l(ijk)}$$

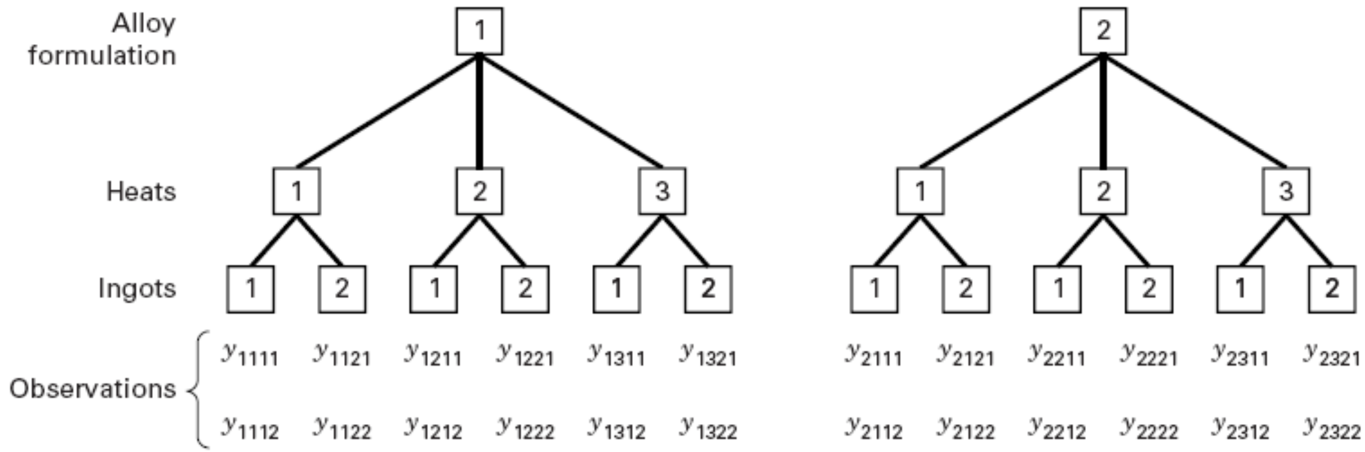
- Sections 14.2, 14.3 in Montgomery.



■ **FIGURE 14.4** A three-stage staggered nested design

General m-Stage Nested Designs

■ **FIGURE 14.5** A three-stage nested design



$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(ij)} + \epsilon_{l(ijk)}$$

■ TABLE 14.7

Analysis of Variance for the Three-Stage Nested Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
A	$bcn \sum_i (\bar{y}_{i...} - \bar{y}_{....})^2$	$a - 1$	MS_A
B (within A)	$cn \sum_i \sum_j (\bar{y}_{ij..} - \bar{y}_{....})^2$	$a(b - 1)$	$MS_{B(A)}$
C (within B)	$n \sum_i \sum_j \sum_k (\bar{y}_{ijk.} - \bar{y}_{....})^2$	$ab(c - 1)$	$MS_{C(B)}$
Error	$\sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{ijk.})^2$	$abc(n - 1)$	MS_E
Total	$\sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{....})^2$	$abcn - 1$	

■ TABLE 14.8

Expected Mean Squares for a Three-Stage Nested Design with *A* and *B* Fixed and *C* Random

Model Term	Expected Mean Square
τ_i	$\sigma^2 + n\sigma_\gamma^2 + \frac{bcn \sum \tau_i^2}{a-1}$
$\beta_{j(i)}$	$\sigma^2 + n\sigma_\gamma^2 + \frac{cn \sum \sum \beta_{j(i)}^2}{a(b-1)}$
$\gamma_{k(ij)}$	$\sigma^2 + n\sigma_\gamma^2$
$\epsilon_{l(ijk)}$	σ^2

Example 14.2 Nested and Factorial Factors

■ **TABLE 14.9**
Assembly Time Data for Example 14.2

	Layout 1				Layout 2				
Operator	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	$y_{i...}$
Fixture 1	22	23	28	25	26	27	28	24	404
	24	24	29	23	28	25	25	23	
Fixture 2	30	29	30	27	29	30	24	28	447
	27	28	32	25	28	27	23	30	
Fixture 3	25	24	27	26	27	26	24	28	401
	21	22	25	23	25	24	27	27	
Operator totals, $y_{.jk.}$	149	150	171	149	163	159	151	160	
Layout totals, $y_{.j..}$	619				633				$1252 = y_{....}$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(j)} + (\tau\beta)_{ij} + (\tau\gamma)_{ik(j)} + \varepsilon_{(ijk)l} \quad \begin{cases} i = 1,2,3 \\ j = 1,2 \\ j = 1,2,3,4 \\ l = 1,2 \end{cases}$$

Example 14.2 – Expected Mean Squares

Assume that fixtures and layouts are fixed, operators are random – gives a **mixed** model (use restricted form)

TABLE 14.10

Expected Mean Squares for Example 14.2

Model Term	Expected Mean Square
τ_i	$\sigma^2 + 2\sigma^2_{\tau\gamma} + 8\sum \tau_i^2$
β_j	$\sigma^2 + 6\sigma^2_{\gamma} + 24\sum \beta_j^2$
$\gamma_{k(j)}$	$\sigma^2 + 6\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	$\sigma^2 + 2\sigma^2_{\tau\gamma} + 4\sum\sum (\tau\beta)^2_{ij}$
$(\tau\gamma)_{ik(j)}$	$\sigma^2 + 2\sigma^2_{\tau\gamma}$
$\epsilon_{(ijk)l}$	σ^2

■ TABLE 14.11

Analysis of Variance for Example 14.2

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Fixtures (F)	82.80	2	41.40	7.54	0.01
Layouts (L)	4.08	1	4.09	0.34	0.58
Operators (within layouts), $O(L)$	71.91	6	11.99	5.15	<0.01
FL	19.04	2	9.52	1.73	0.22
$FO(L)$	65.84	12	5.49	2.36	0.04
Error	56.00	24	2.33		
Total	299.67	47			

Last slide

- Read Sections: 14.1-14.3