Design of Engineering Experiments – Nested Designs

- Text reference, Chapter 14, Pg. 525
- These are **multifactor** experiments that have some important industrial 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

Two-Stage Nested Design

- Section 14-1 (pg. 525)
- In a nested design, the levels of one factor (*B*) is **similar** to but not **identical** to each other at different levels of another factor (*A*)
- Consider a company that purchases material from three suppliers
  - The material comes in batches
  - Is the purity of the material uniform?
- Experimental design
  - Select four batches at random from each supplier
  - Make three purity determinations from each batch

Two-Stage Nested Design

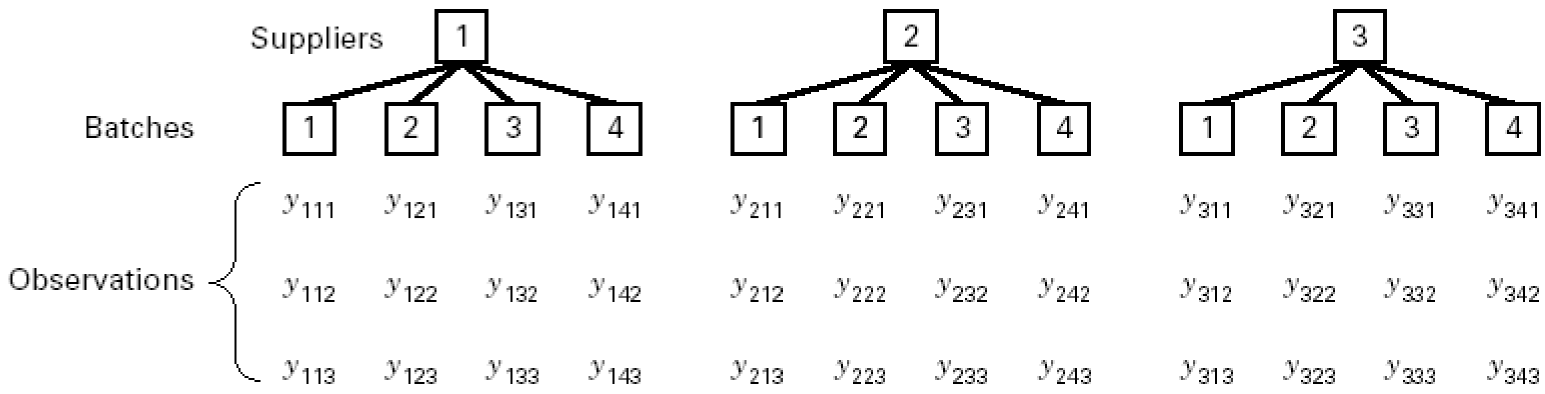


Figure 14-1 A two-stage nested design.

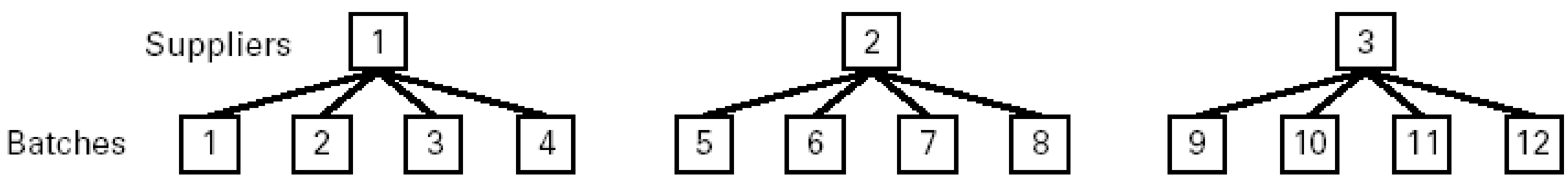


Figure 14-2 Alternate layout for the two-stage nested design.

Two-Stage Nested Design Statistical Model and ANOVA

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

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

$$df : abn - 1 = a - 1 + a(b-1) + ab(n-1)$$

Because not every level of factor B appears with every level of factor A, there can be no interaction between A and B

Table 14-2 Analysis of Variance Table for the Two-Stage Nested Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
A	$bn \sum (\bar{y}_{i..} - \bar{y}_{...})^2$	$a - 1$	MS_A
B within A	$n \sum \sum (\bar{y}_{ij.} - \bar{y}_{i..})^2$	$a(b - 1)$	$MS_{B(A)}$
Error	$\sum \sum \sum (y_{ijk} - \bar{y}_{ij.})^2$	$ab(n - 1)$	MS_E
Total	$\sum \sum \sum (y_{ijk} - \bar{y}_{...})^2$	$abn - 1$	

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

Two-Stage Nested Design Example 14-1 (pg. 528)

Three suppliers, four batches (selected randomly) from each supplier, three samples of material taken (at random) from each batch

Experiment and data, Table 14-3

Data is coded

R balanced ANOVA will analyze nested designs

Mixed model, assume restricted form

R code and output

```
>aov1=aov(Purity~Supplier/Batches)
>(sum1=summary(aov1)[[1]])
```

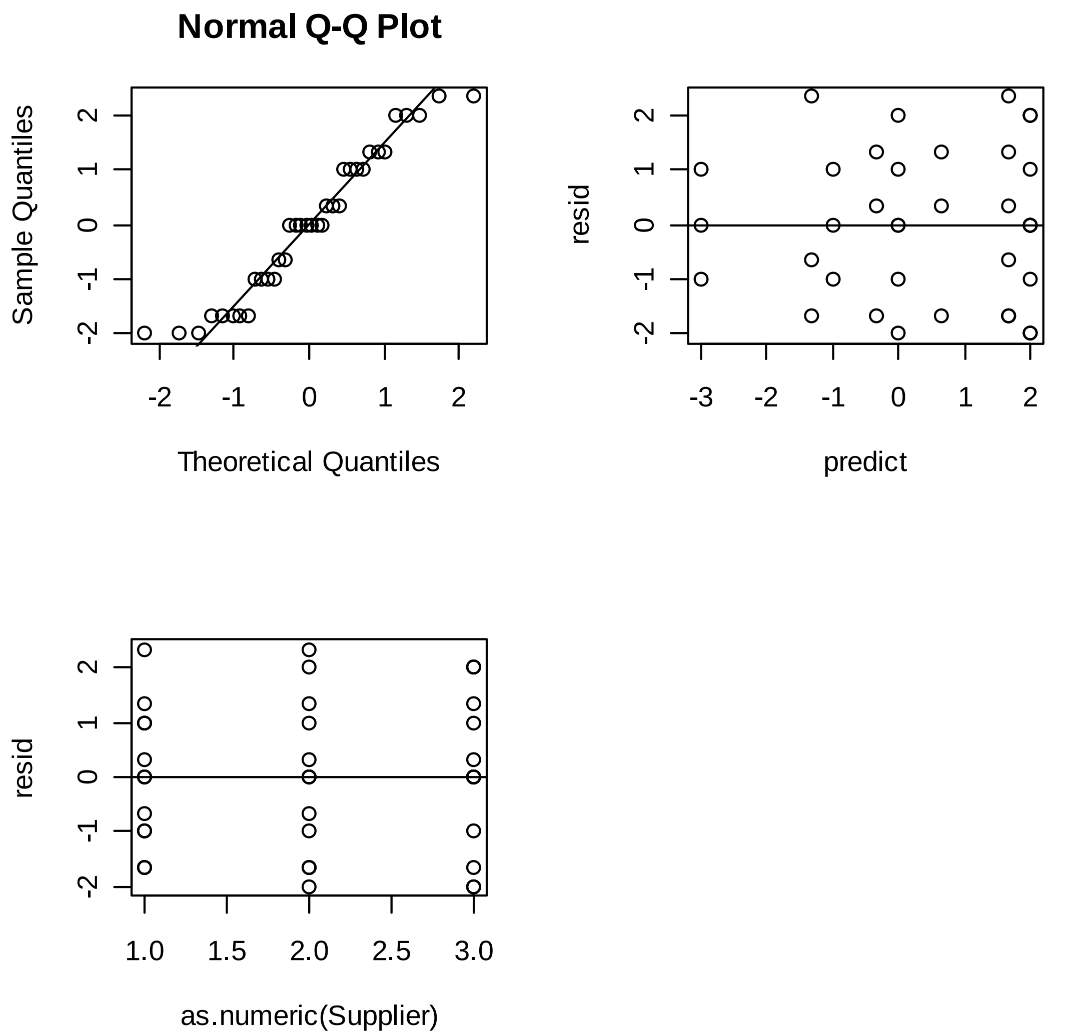
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Supplier	2	15.056	7.528	2.8526	0.07736 .
Supplier:Batches	9	69.917	7.769	2.9439	0.01667 *
Residuals	24	63.333	2.639		

 # The F test for B is valid.

The F test for A

```
>(F.A<-sum1[1,3]/sum1[2,3])
>(p.A<-1-pf(F.A,sum1[1,1],sum1[2,1]))
>c(F.A, p.A)c
[1] 0.9690107 0.4157831
```

Diagnostic Checking



Variance Components

$$\hat{\sigma}^2 = MSE$$

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

```
##########variance components

> sigma2.E=sum1[3,3]

> # Estimate the variance of the nested factor B

> n<-3

> sigma2.B=(sum1[2,3]-sigma2.E)/n

> c(sigma2.B,sigma2.E)
[1] 1.709877 2.638889
```

Practical Interpretation – Example 14-1

- There is no difference in purity among suppliers, but significant difference in purity among batches (within suppliers)
- What are the practical implications of this conclusion?- request suppliers to reduce their batch to batch variation.
- Examine residual plots – pg. 532 – plot of residuals versus supplier is very important (why?)- check the assumption that the batch to batch variation are constant across all the suppliers.
- The variance component σ_β^2 is greater than zero as suggested by ANOVA

Variations of the Nested Design

- Staggered nested designs (Pg. 533)
  - Prevents too many degrees of freedom from building up at lower levels
  - See the supplemental text material for an example
- Several levels of nesting (pg. 534)
  - The alloy formulation example
  - This experiment has three stages of nesting
- Experiments with both nested and “crossed” or factorial factors (pg. 536)

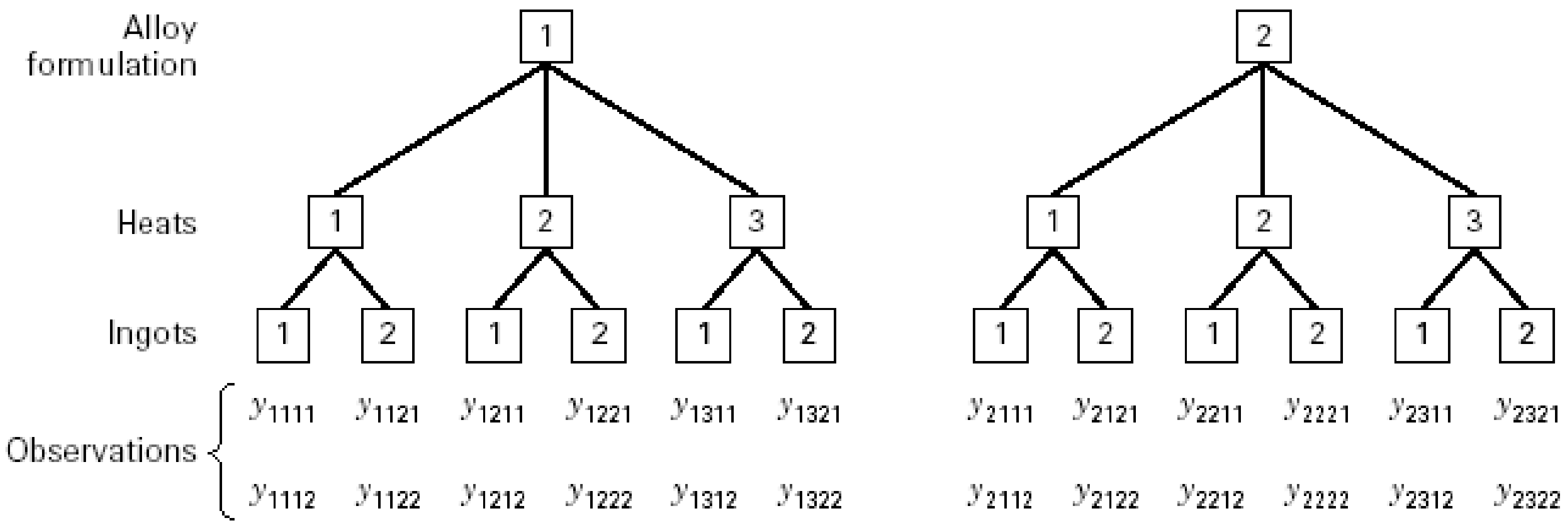


Figure 14-5 A three-stage nested design.

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(ij)} + \varepsilon_{(ijk)l} \begin{cases} i = 1,2,\ldots,a \\ j = 1,2,\ldots b \\ k = 1,2,\ldots,c \\ l = 1,2,\ldots,n \end{cases}$$

Table 14-7 Analysis of Variance for the Three-Stage Nested Design

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

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

```
>aov1=aov(Hardness~Alloy/Heat/Ingot)
> sum1=summary(aov1)[[1]]
> sum1
                    Df   Sum Sq  Mean Sq  F value   Pr(>F)
Alloy                1    315.4   315.4    1.7672   0.208447
Alloy:Heat           4   6453.8  1613.5    9.0411   0.001317 **
Alloy:Heat:Ingot     6   2226.3   371.0    2.0792   0.132167
Residuals           12   2141.5   178.5

# F-test for C is valide

# F tests for A and B
> F.A.B=sum1[1:2,3]/sum1[3,3]
> p.A.B=1-pf(F.A.B,sum1[1:2,1],sum1[3,1])
> c(F.A.B,p.A.B)
[1] 0.84997193 4.34845592 0.39212351 0.05450405
```

	F value	p-value
Alloy	0.84997193	0.39212351
Alloy:Heat	4.34845592	0.05450405

 # No effect is significant. Maybe eliminate C %in% B first

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	$\sigma^2 + 6\sigma^2_{\gamma} + 24\sum \beta_j^2$
$\gamma_{k(j)}$	$\sigma^2 + 6\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	$\sigma^2 + 2\sigma^2_{\tau\gamma} + 4\sum\sum (\tau\beta)_{ij}^2$
$(\tau\gamma)_{ik(j)}$	$\sigma^2 + 2\sigma^2_{\tau\gamma}$
$\epsilon_{(ijk)l}$	σ^2

Fixture and Layout are crossed, Operator nested within Layout

> aov1=aov(time~Fixture*(Layout/Operator))

> sum1=summary(aov1)[[1]]

> sum1

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Fixture	2	82.792	41.396	17.7411	1.862e-05 ***
Layout	1	4.083	4.083	1.7500	0.198344
Layout:Operator	6	71.917	11.986	**5.1369**	**0.001606 ****
Fixture:Layout	2	19.042	9.521	4.0804	0.029828 *
Fixture:Layout:Operator	12	65.833	5.486	**2.3512**	**0.036043** *
Residuals	24	56.000	2.333		

F tests for Operator%in%Layout, Fixture:(Operator%in%Layout) are valid

```
#F test for Fixture and Fixture:Layout
> F.F.FL=sum1[c(1,4),3]/sum1[5,3]
> p.F.FL=1-pf(F.F.FL,sum1[c(1,4),1],sum1[5,1])
> c(F.F.FL,p.F.FL)
[1] 7.545569620 1.735443038 0.007553076 0.217769146
```

	F value	p-vaue
Fixture	7.545569620	0.007553076
Fixture: Layout	1.735443038	0.217769146

```
# F test for Layout
> F.L=sum1[2,3]/sum1[3,3]
> p.L=1-pf(F.L,sum1[2,1],sum1[3,1])
> c(F.L,p.L)
[1] 0.3406721 0.5807041
```

 # It seems that Layout has no effect

```
# Reanalyze data. Since O%in%L, we create a new factor called LO

LO=factor(paste(unclass(Layout),unclass(Operator),sep=""))

aov2=aov(time~Fixture*LO)

sum2=summary(aov2)[[1]]

sum2

# We reach a two-way crosted
# factorial mixed model with F fixed and LO random

# The F test for Fixture needs calculation

F.F=sum2[1,3]/sum2[3,3]

p.F=1-pf(F.F,sum2[1,1],sum2[3,1])

(F.F,p.F)
```

```
> # Reanalyze data. Since O%in%L, we create a new factor called LO
>
> LO=factor(paste(unclass(Layout),unclass(Operator),sep=""))
> aov2=aov(time~Fixture*LO)
➢sum2=summary(aov2)[[1]]

> sum2
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Fixture	2	82.792	41.396	17.7411	1.862e-05	***
LO	7	76.000	10.857	**4.6531**	**0.002074**	**
Fixture:LO	14	84.875	6.062	**2.5982**	**0.019222**	*
Residuals	24	56.000	2.333			

```
> # We reach a two-way crosted
> # factorial mixed model with F fixed and LO random
> # The F test for Fixture needs calculation
>
> F.F=sum2[1,3]/sum2[3,3]
>
> p.F=1-pf(F.F,sum2[1,1],sum2[3,1])
>
> c(F.F,p.F)
[1] 6.828178694 0.008517878
```