Example - 2-Factor Nested Model (Mixed)

τ_i $\beta_{j(i)}$ $\epsilon_{k(ij)}$

Example

A company is interested in testing the uniformity of their film-coated pain tablets. A random sample of three batches were collected from each of their two blending sites. Five tablets were assayed from each batch.

Site	1			2		
Batch	1	2	3	4	5	6
	5.03	4.64	5.10	5.05	5.46	4.90
	5.10	4.73	5.15	4.96	5.15	4.95
	5.25	4.82	5.20	5.12	5.18	4.86
	4.98	4.95	5.08	5.12	5.18	4.86
	5.05	5.06	5.14	5.05	5.11	5.07

- What are the factors?
- Are any nested?
- Which are random and which are fixed?

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Site	.01825	1	.01825	
Batch(Site)	.45401	4	.11350	
Error	.29020	24	.01209	
Total	.76246	29		

$y_{11.} = 25.41$ $\quad y_{12.} = 24.20$ $\quad y_{13.} = 25.67$

$y_{21.} = 25.30$ $\quad y_{22.} = 26.08$ $\quad y_{23.} = 24.64$

$$\sum\sum\sum y_{ijk}^2 = 763.8188$$

$\mathrm{SS_T} = 763.8188 - 151.3^2/30 = .76247$

$\mathrm{SS_A} = (75.28^2 + 76.02^2)/15 - 151.3^2/30 = .01825$

$\mathrm{SS_{B(A)}} = (25.41^2 + 24.20^2 + \ldots + 24.64^2)/5 - 763.07459 = .45401$

$\mathrm{SS_E} = 763.8188 - 763.5286 = .2902$

Results

- Site: $F = .01825/.1135 = .1608$. There is not enough evidence to suggest that the two coating sites are different.
- Batch: $F = .1135/.0121 = 9.39$. Compare to $F_{4,24}$. There is significant batch-to-batch variability.

$$\hat{\sigma}^2 = .0121 \qquad \hat{\sigma}^2_\beta = \frac{.1135 - .0121}{5} = .0203$$

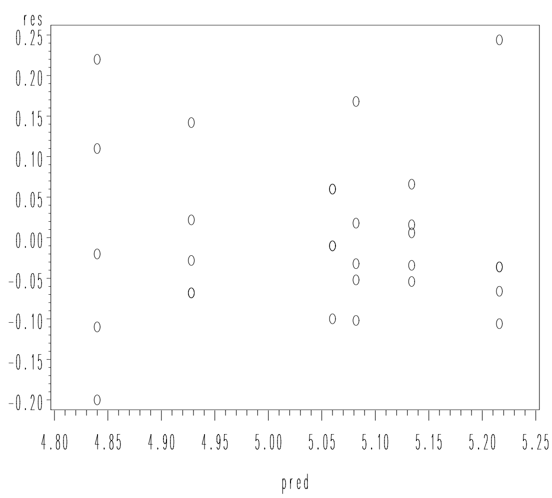
- Batch variability is $.0203/(.0203 + .0121) = 62.7\%$ of the total variability. It appears that efforts should be made to eliminate the batch-to-batch variability. Investigate what goes into coating a batch and see where the variability could be.

SAS Program

```
options nocenter ls=75;

data new;
infile "coating.dat";
input site batch tablet resp;

proc glm;
class site batch;
model resp=site batch(site);
random batch;
test h=site e=batch(site);
output out=new1 p=pred r=res;

symbol1 v=circle;
proc gplot;
plot res*pred;

proc glm data=new;
class site batch;
model resp=site batch site*batch;
random batch site*batch;
test h=site e=batch*site;
run;
```



```
Dependent Variable: RESP
                                    Sum of          Mean
Source                   DF        Squares        Square  F Value    Pr > F
Model                     5      0.4722667     0.0944533     7.81    0.0002
Error                    24      0.2902000     0.0120917
Corrected Total          29      0.7624667

Source                   DF      Type I SS   Mean Square  F Value    Pr > F
SITE                      1      0.0182533     0.0182533     1.51    0.2311
BATCH(SITE)               4      0.4540133     0.1135033     9.39    0.0001


Tests of Hypotheses using the Type III MS for
BATCH(SITE) as an error term

Source                   DF    Type III SS   Mean Square  F Value    Pr > F
SITE                      1      0.0182533     0.0182533     0.16    0.7089

--------------------------------------------------------------------------

Dependent Variable: RESP
                                   Sum of          Mean
Source                   DF        Squares        Square  F Value    Pr > F
Model                     5      0.4722667     0.0944533     7.81    0.0002
Error                    24      0.2902000     0.0120917
Corrected Total          29      0.7624667

Source                   DF      Type I SS   Mean Square  F Value    Pr > F
SITE                      1      0.0182533     0.0182533     1.51    0.2311
BATCH                     2      0.0115267     0.0057633     0.48    0.6266
SITE*BATCH                2      0.4424867     0.2212433    18.30    0.0001
```

Nested Model as Factorial

- Suppose we treat design as two factor factorial
- Naively interpret SAS results
  - Significant batch*site variability
  - No longer significant batch-to-batch variability
- What does interaction mean?
- We're assuming batch 1 effect similar across sites
- Can't separate interaction from main effect
- Notice $\mathrm{SS}_{AB} + \mathrm{SS}_B = \mathrm{SS}_{B(A)}$

  $\mathrm{df}_{AB} + \mathrm{df}_B = \mathrm{df}_{B(A)}$
- Could use factorial results and properly analyze model