Random and Mixed Factorial Designs I

Design of Experiments - Montgomery
Chapter 12

Factorial Experiments with Random Effects

- Topics 14-16 have focused on fixed effects
  - Always use MSE in denominator of F-test
  - Use MSE in linear combinations and CIs
- Not always true when random factors present
  - May use interaction MS or combination of MS's
- Will now use EMS as guide for tests

- Presentation of material
  - Topic 17: How and why use EMS
    - A Two factor random model - analysis and SAS
    - B How to calculate EMS
  - Topic18 : Mixed models
    - A Two factor mixed model - analysis and SAS
    - B Satterthwaite's approx F-tests and CI

A. Two-Factor Random Model

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

$$\tau_i \sim \mathrm{N}(0, \sigma^2_\tau) \quad \beta_j \sim \mathrm{N}(0, \sigma^2_\beta) \quad (\tau\beta)_{ij} \sim \mathrm{N}(0, \sigma^2_{\tau\beta})$$

- $\mathrm{Var}(y_{ijk}) = \sigma^2 + \sigma^2_\tau + \sigma^2_\beta + \sigma^2_{\tau\beta}$
- Expected MS's similar to one-factor random model

  $\mathrm{E(MS_E)} = \sigma^2$

  $\mathrm{E(MS_A)} = \sigma^2 + bn\sigma^2_\tau + n\sigma^2_{\tau\beta}$

  $\mathrm{E(MS_B)} = \sigma^2 + an\sigma^2_\beta + n\sigma^2_{\tau\beta}$

  $\mathrm{E(MS_{AB})} = \sigma^2 + n\sigma^2_{\tau\beta}$

- EMS determine what MS to use in denominator

  $H_0 : \sigma^2_\tau = 0 \rightarrow \mathrm{MS_A/MS_{AB}}$

  $H_0 : \sigma^2_\beta = 0 \rightarrow \mathrm{MS_B/MS_{AB}}$

  $H_0 : \sigma^2_{\tau\beta} = 0 \rightarrow \mathrm{MS_{AB}/MS_E}$

- No hierarchical testing. Look at all tests

Estimating Variance Components

- Using ANOVA method

  $\hat{\sigma}^2 = \mathrm{MS_E}$

  $\hat{\sigma}^2_\tau = (\mathrm{MS_A} - \mathrm{MS_{AB}})/bn$

  $\hat{\sigma}^2_\beta = (\mathrm{MS_B} - \mathrm{MS_{AB}})/an$

  $\hat{\sigma}^2_{\tau\beta} = (\mathrm{MS_{AB}} - \mathrm{MS_E})/n$

- Sometimes results in negative estimates
- Proc Varcomp and Proc Mixed compute estimates
- Can use different estimation procedures

  ANOVA method - Method = type1

  RMLE method - Method = reml (default)

- Proc Mixed

  Variance component estimates

  Hypothesis tests and confidence intervals

Gauge Capability Example in Text 12-2

```
options nocenter ls=75;

data randr;
 input part operator resp @@;
 cards;
1 1 21 1 1 20 1 2 20 1 2 20 1 3 19 1 3 21
2 1 24 2 1 23 2 2 24 2 2 24 2 3 23 2 3 24
3 1 20 3 1 21 3 2 19 3 2 21 3 3 20 3 3 22
4 1 27 4 1 27 4 2 28 4 2 26 4 3 27 4 3 28
.
.
;

proc glm;
 model resp=operator|part;
 random operator part operator*part / test;
 test H=operator E=operator*part;
 test H=part E=operator*part;


proc mixed cl maxiter=20 covtest method=type1;
 class operator part;
 model resp = ;
 random operator part operator*part;


proc mixed cl maxiter=20 covtest;
 class operator part;
 model resp = ;
 random operator part operator*part;
run;
quit;
```


Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	59	1215.091667	20.594774	20.77	<.0001
Error	60	59.500000	0.991667		
Corrected Total	119	1274.591667			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.32	0.2750
part	19	1185.425000	62.390789	62.92	<.0001
operator*part	38	27.050000	0.711842	0.72	0.8614

Source	Type III Expected Mean Square
operator	Var(Error) + 2 Var(operator*part) + 40 Var(operator)
part	Var(Error) + 2 Var(operator*part) + 6 Var(part)
operator*part	Var(Error) + 2 Var(operator*part)

Tests of Hypotheses Using the Type III MS for operator*part as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.84	0.1730
part	19	1185.425000	62.390789	87.65	<.0001

Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: resp

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	2	2.616667	1.308333	1.84	0.1730
part	19	1185.425000	62.390789	87.65	<.0001
Error	38	27.050000	0.711842		

Error: MS(operator*part)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator*part	38	27.050000	0.711842	0.72	0.8614
Error: MS(Error)	60	59.500000	0.991667		

The Mixed Procedure

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square
operator	2	2.616667	1.308333
part	19	1185.425000	62.390789
operator*part	38	27.050000	0.711842
Residual	60	59.500000	0.991667

Type 1 Analysis of Variance

Source	Expected Mean Square	Error Term	Error DF
operator	Var(Residual) + 2 Var(operator*part) + 40 Var(operator)	MS(operator*part)	38
part	Var(Residual) + 2 Var(operator*part) + 6 Var(part)	MS(operator*part)	38
operator*part	Var(Residual) + 2 Var(operator*part)	MS(Residual)	60
Residual	Var(Residual)	.	.

Source	F Value	Pr > F
operator	1.84	0.1730
part	87.65	<.0001
operator*part	0.72	0.8614

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
operator	0.0149	0.0330	0.45	0.6510	0.05	-0.0497	0.0795
part	10.2798	3.3738	3.05	0.0023	0.05	3.6673	16.8924
operator*part	-0.1399	0.1219	-1.15	0.2511	0.05	-0.3789	0.0990
Residual	0.9917	0.1811	5.48	<.0001	0.05	0.7143	1.4698

The Mixed Procedure

Estimation Method REML

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	624.67452320	
1	3	409.39453674	0.00003340
2	1	409.39128078	0.00000004
3	1	409.39127700	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
operator	0.0106	0.03286	0.32	0.3732	0.05	0.001103	3.737E12
part	10.2513	3.3738	3.04	0.0012	0.05	5.8888	22.1549
operator*part	0	.	.	.	.	.	.
Residual	0.8832	0.1262	7.00	<.0001	0.05	0.6800	1.1938

Confidence Intervals for Variance Components

- Can use asymptotic variance estimates to form CI
- Known as Wald's approximate CI
- Mixed: option CL=WALD or METHOD=TYPE1

  Use standard normal → 95% CI uses 1.96

  $\hat{\sigma}^2_\beta \pm 1.96(.0330) = (-0.05, 0.08)$

  $\hat{\sigma}^2_\tau \pm 1.96(3.3738) = (3.67, 16.89)$

- In general Proc Mixed uses Satterthwaite CI

  Default method - REML

  Versions $<$ 6.12 computed Wald CI

  Current uses Satterthwaite's Approximation

  Will discuss this CI construction in next topic

B. Rules For Expected Mean Squares

- In models so far, EMS fairly straightforward
- Could show EMS using brute force expectation method
- For mixed models, good to have formal procedure
- Montgomery describes procedure for **restricted** model

  0 Write the error term in the model as $\epsilon_{(ij..)m}$, where m represents the replication subscript

  1 Write each variable term in model as a row heading in a two-way table

  2 Write the subscripts in the model as column headings. Over each subscript write F if factor fixed and R if random. Over this, write down the levels of each subscript

  3 For each row, copy the number of observations under each subscript, providing the subscript does not appear in the row variable term

  4 For any bracketed subscripts in the model, place a 1 under those subscripts that are inside the brackets

  5 Fill in remaining cells with a 0 (if subscript represents a fixed factor) or a 1 (if random factor).

  6 To find the expected mean square of any term (row), cover the entries in the columns that contain non-bracketed subscript letters in this term in the model. For those rows with at least the same subscripts, multiply the remaining numbers to get coefficient for corresponding term in the model.

2-Factor Fixed Model

$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$

τ_i
β_j
$(\tau\beta)_{ij}$
$\epsilon_{k(ij)}$

2-Factor Random Model

$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$

τ_i
β_j
$(\tau\beta)_{ij}$
$\epsilon_{k(ij)}$

2-Factor Mixed Model (A fixed)

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

τ_i
β_j
$(\tau\beta)_{ij}$
$\epsilon_{k(ij)}$

3-Factor Mixed Model (A fixed)

$$y_{ijk} = \mu + \tau_i + \beta_j + \delta_k + (\tau\beta)_{ij} + (\tau\delta)_{ik} + (\beta\delta)_{jk} + \epsilon_{ijkl}$$

τ_i
β_j
δ_k
$(\tau\beta)_{ij}$
$(\tau\delta)_{ik}$
$(\beta\delta)_{jk}$
$(\tau\beta\delta)_{ijk}$
$\epsilon_{l(ijk)}$

Construction of Hasse Diagram

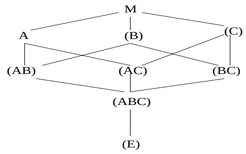
- Described in Oehlert (2000)
- Used for both **restricted** and **unrestricted** models
- Provide graphical view of design
- Shows nested/crossed and random/fixed structure
- Every term in model is a node
- Terms/nodes placed in layered structure

  Term U is above term V if all terms in U are in V

- Join nodes based on nested/crossed structure
- Brackets placed around random terms

3-Factor Mixed Model

- Denominator for U is leading eligible random term(s)
- Leading: Closest connected random term below U
- Eligible:
  - Unrestricted : Any random term possible
  - Restricted : Any without fixed factor not in U



Restricted Model:
A: Leading random terms are AB and AC → approximate test
B: Leading random term is BC because AB has fixed factor A
BC: Leading term is E because ABC has fixed factor A

Unrestricted Model:
A: Leading random terms are AB and AC → approximate test
B: Leading random term is AB and BC → approximate test
BC: Leading term is ABC