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
  - Topic 18: Mixed models
    A Two factor mixed model - analysis and SAS
    B Satterthwaite’s approx F-tests and CI

A. Two-Factor Random Model

\[ y_{ijk} = \mu + \tau_i + \beta_j + (\tau \beta)_{ij} + e_{ijk} \]  
\[ \tau_i \sim N(0, \sigma^2_{\tau}) \quad \beta_j \sim N(0, \sigma^2_{\beta}) \quad (\tau \beta)_{ij} \sim N(0, \sigma^2_{\tau \beta}) \]

- \( \text{Var}(y_{ijk}) = \sigma^2 + \sigma^2_{\tau} + \sigma^2_{\beta} + \sigma^2_{\tau \beta} \)
- Expected MS’s similar to one-factor random model
  \[ E(\text{MS}_E) = \sigma^2 \]
  \[ E(\text{MS}_A) = \sigma^2 + b\sigma^2_{\tau} + na\sigma^2_{\beta} \]
  \[ E(\text{MS}_B) = \sigma^2 + an\sigma^2_{\tau} + na\sigma^2_{\beta} \]
  \[ E(\text{MS}_{AB}) = \sigma^2 + na\sigma^2_{\tau \beta} \]

- EMS determine what MS to use in denominator
  \[ H_0 : \sigma^2_{\tau} = 0 \rightarrow \text{MS}_A/\text{MS}_{AB} \]
  \[ H_0 : \sigma^2_{\beta} = 0 \rightarrow \text{MS}_B/\text{MS}_{AB} \]
  \[ H_0 : \sigma^2_{\tau \beta} = 0 \rightarrow \text{MS}_{AB}/\text{MS}_E \]

- No hierarchical testing. Look at all tests

Estimating Variance Components

- Using ANOVA method
  \[ \hat{\sigma}^2_E = \text{MS}_E \]
  \[ \hat{\sigma}^2_{\tau} = (\text{MS}_A - \text{MS}_{AB})/bn \]
  \[ \hat{\sigma}^2_{\beta} = (\text{MS}_B - \text{MS}_{AB})/an \]
  \[ \hat{\sigma}^2_{\tau \beta} = (\text{MS}_{AB} - \text{MS}_E)/n \]

- 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
options nocenter ls=75;
data randr;
input part operator resp @@;
cards;
1 1 21 1 20 1 2 1 20 1 3 1 3 21
2 1 24 2 23 2 24 2 24 2 3 23 2 3 3 24
3 1 20 3 1 21 3 1 21 3 3 20 3 3 22
4 1 27 4 1 27 4 2 28 4 2 26 4 3 27 4 3 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;
run;
quit;

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

The Mixed Procedure

Type 1 Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>operator</td>
<td>2</td>
<td>2.616667</td>
<td>1.308333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>part</td>
<td>19</td>
<td>1185.425000</td>
<td>62.390789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>operator*part</td>
<td>38</td>
<td>27.050000</td>
<td>0.711842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>60</td>
<td>59.500000</td>
<td>0.991667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Type 1 Analysis of Variance (Error)

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>operator</td>
<td>2</td>
<td>2.616667</td>
<td>1.308333</td>
</tr>
<tr>
<td>part</td>
<td>19</td>
<td>1185.425000</td>
<td>62.390789</td>
</tr>
<tr>
<td>operator*part</td>
<td>38</td>
<td>27.050000</td>
<td>0.711842</td>
</tr>
<tr>
<td>Residual</td>
<td>60</td>
<td>59.500000</td>
<td>0.991667</td>
</tr>
</tbody>
</table>

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Source</th>
<th>Estimate</th>
<th>Error Value</th>
<th>Pr Z Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>operator</td>
<td>0.0149</td>
<td>0.0330</td>
<td>0.32</td>
<td>-0.0497</td>
<td>0.0795</td>
</tr>
<tr>
<td>part</td>
<td>10.2798</td>
<td>3.3738</td>
<td>3.05</td>
<td>3.6673</td>
<td>16.8924</td>
</tr>
<tr>
<td>operator*part</td>
<td>-0.1399</td>
<td>0.1219</td>
<td>-1.15</td>
<td>-0.2611</td>
<td>0.3789</td>
</tr>
<tr>
<td>Residual</td>
<td>0.9917</td>
<td>0.1811</td>
<td>5.48</td>
<td>0.0001</td>
<td>0.7143</td>
</tr>
</tbody>
</table>

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 |
operator*part | 38 | 27.050000 | 0.711842 | . | . |

Tests of Hypotheses for Random Model Analysis of Variance

Source | 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: MS(operator*part) | 60 | 59.500000 | 0.991667 | . | . |

The Mixed Procedure

Estimation Method REML

Iteration History

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Source</th>
<th>Estimate</th>
<th>Error Value</th>
<th>Pr Z Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>operator</td>
<td>0.0106</td>
<td>0.03286</td>
<td>0.32</td>
<td>0.3732</td>
<td>0.06</td>
</tr>
<tr>
<td>part</td>
<td>10.2513</td>
<td>3.3738</td>
<td>3.04</td>
<td>3.0012</td>
<td>0.05</td>
</tr>
<tr>
<td>operator*part</td>
<td>0.8832</td>
<td>0.1262</td>
<td>7.00</td>
<td>&lt;0.0001</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Convergence criteria met.

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Source</th>
<th>Cov Parm</th>
<th>Estimate</th>
<th>Error Value</th>
<th>Pr Z Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>operator</td>
<td>0.0149</td>
<td>0.0330</td>
<td>0.45</td>
<td>0.6510</td>
<td>0.05</td>
<td>0.0497</td>
</tr>
<tr>
<td>part</td>
<td>10.2798</td>
<td>3.3738</td>
<td>3.05</td>
<td>3.6673</td>
<td>0.05</td>
<td>16.8924</td>
</tr>
<tr>
<td>operator*part</td>
<td>-0.1399</td>
<td>0.1219</td>
<td>-1.15</td>
<td>-0.2611</td>
<td>0.3789</td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td>0.9917</td>
<td>0.1811</td>
<td>5.48</td>
<td>&lt;0.0001</td>
<td>0.05</td>
<td>1.4698</td>
</tr>
</tbody>
</table>

17-4

17-5

17-6

17-7
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
    \[ \hat{\sigma}^2 \pm 1.96(0.0330) = (-0.05, 0.08) \]
    \[ \hat{\sigma}^2 \pm 1.96(3.3738) = (3.67, 16.89) \]
- 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 \( \epsilon_{ijm} \), 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

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

\[ \tau_i \]
\[ \beta_j \]
\[ (\tau\beta)_{ij} \]
\[ \epsilon_{k(ij)} \]

2-Factor Random Model

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

\[ \tau_i \]
\[ \beta_j \]
\[ (\tau\beta)_{ij} \]
\[ \epsilon_{k(ij)} \]
2-Factor Mixed Model (A fixed)

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

\[ \tau_i \]
\[ \beta_j \]
\[ (\tau\beta)_{ij} \]
\[ \epsilon_{k(ij)} \]

3-Factor Mixed Model (A fixed)

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

\[ \tau_i \]
\[ \beta_j \]
\[ \delta_k \]
\[ (\tau\beta)_{ij} \]
\[ (\tau\delta)_{ik} \]
\[ (\beta\delta)_{jk} \]
\[ (\tau\beta\delta)_{ijk} \]
\[ \epsilon_{l(ijk)} \]

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