# Topic 27 - Mixed Effects Models

STAT 525 - Fall 2013



# Outline

- Two-way mixed effects
- Three-way models
- Computing Expected Mean Squares





# Two-Factor Mixed Effects Model

- One factor random and one factor fixed
- Assume A fixed and B random
- Parameter assumptions/restrictions are now:

  $\sum \alpha_i = 0$ and $\beta \sim \mathrm{N}(0, \sigma^2_\beta)$

  $(\alpha\beta)_{ij} \sim \mathrm{N}(0, \sigma^2_{\alpha\beta})$
- Known as the **unrestricted** mixed model
- SAS uses this model in its procedures





# Unrestricted Mixed Model

- Same partition of total sum of squares as two-way random
- Different assumptions/restrictions alter EMS

  E(MSE)$=\sigma^2$

  E(MSA) $= \sigma^2 + bn\sum \alpha_i^2/(a-1) + n\sigma^2_{\alpha\beta}$

  E(MSB) $= \sigma^2 + an\sigma^2_\beta + n\sigma^2_{\alpha\beta}$

  E(MSAB) $= \sigma^2 + n\sigma^2_{\alpha\beta}$

# Hypothesis Tests

- Tests require different MS in denom

  $H_0 : \alpha_1 = \alpha_2 = \ldots = 0 \rightarrow$ MSA/MSAB

  $H_0 : \sigma^2_\beta = 0 \rightarrow$ MSB/MSAB

  $H_0 : \sigma^2_{\alpha\beta} = 0 \rightarrow$ MSAB/MSE

- Variance Estimates (Using ANOVA method)

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

  $\hat{\sigma}^2_\beta = (\text{MSB} - \text{MSAB})/\text{an}$

  $\hat{\sigma}^2_{\alpha\beta} = (\text{MSAB} - \text{MSE})/\text{n}$





# Multiple Comparisons

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

$$\begin{aligned}
\overline{Y}_{i..} &= \mu + \alpha_i + \overline{\beta}_{.} + \overline{(\alpha\beta)}_{i.} + \overline{\varepsilon}_{i..} \\
\text{Var}(\overline{Y}_{i..}) &= \sigma^2_\beta/b + \sigma^2_{\alpha\beta}/b + \sigma^2/bn \\
\overline{Y}_{i..} - \overline{Y}_{i'..} &= \alpha_i - \alpha_{i'} + \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{i'.} + \overline{\varepsilon}_{i..} - \overline{\varepsilon}_{i'..} \\
\text{Var}(\overline{Y}_{i..} - \overline{Y}_{i'..}) &= 2\sigma^2_{\alpha\beta}/b + 2\sigma^2/bn \\
&= 2(n\sigma^2_{\alpha\beta} + \sigma^2)/bn
\end{aligned}$$

- For pairwise comparison, use 2MSAB/bn
- For $\overline{Y}_{i..}$, use approximate method





# Other Two-Way Mixed Model

- Consider interaction a hybrid of random and fixed effect
- Parameter assumptions/restrictions are now:
  1. $\sum \alpha_i = 0$ and $\beta \sim \text{N}(0, \sigma^2_\beta)$
  2. $(\alpha\beta)_{ij} \sim \text{N}(0, (\text{a}-1)\sigma^2_{\alpha\beta}/\text{a})$
  3. $\sum (\alpha\beta)_{ij} = 0$ for $\beta$ level $j$
- Known as **restricted** mixed effects model





# Restricted Mixed Model

- The $(a-1)/a$ is used to simplify the EMS

  E(MSE)=$\sigma^2$

  E(MSA) $= \sigma^2 + bn\sum \alpha_i^2/(a-1) + n\sigma^2_{\alpha\beta}$

  E(MSB) $= \sigma^2 + an\sigma^2_\beta$

  E(MSAB) $= \sigma^2 + n\sigma^2_{\alpha\beta}$

- Because of (3), not all $(\alpha\beta)_{ij}$ indep

$$\text{Cov}((\alpha\beta)_{\text{ij}}, (\alpha\beta)_{\text{i'j}}) = -\frac{1}{\text{a}}\sigma^2_{\alpha\beta}$$

NOTE: If $X_i \sim N(0, \sigma^2)$ then $\begin{cases} X_i - \overline{X} \sim N(0, \frac{n-1}{n}\sigma^2) \\ \text{Cov}(\text{X}_\text{i} - \overline{\text{X}}, \text{X}_\text{j} - \overline{\text{X}}) = -\frac{1}{\text{n}}\sigma^2 \end{cases}$

# Hypothesis Tests

- Tests require different MS in denom

  $H_0 : \alpha_1 = \alpha_2 = \ldots = 0 \rightarrow$ MSA/MSAB

  $H_0 : \sigma^2_\beta = 0 \rightarrow$ MSB/MSE

  $H_0 : \sigma^2_{\alpha\beta} = 0 \rightarrow$ MSAB/MSE

- Variance Estimates (Using ANOVA method)

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

  $\hat{\sigma}^2_\beta =$ (MSB − MSE)/an

  $\hat{\sigma}^2_{\alpha\beta} =$ (MSAB − MSE)/n





# Multiple Comparisons

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

$$\begin{aligned}
\overline{Y}_{i..} &= \mu + \alpha_i + \overline{\beta}_{.} + \overline{(\alpha\beta)}_{i.} + \overline{\varepsilon}_{i..} \\
\mathrm{Var}(\overline{Y}_{i..}) &= \sigma^2_\beta/b + (a-1)\sigma^2_{\alpha\beta}/ab + \sigma^2/bn \\
\overline{Y}_{i..} - \overline{Y}_{i'..} &= \alpha_i - \alpha_{i'} + \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{i'.} + \overline{\varepsilon}_{i..} - \overline{\varepsilon}_{i'..} \\
\mathrm{Var}(\overline{Y}_{i..} - \overline{Y}_{i'..}) &= 2\sigma^2_{\alpha\beta}/b + 2\sigma^2/bn \\
&= 2(n\sigma^2_{\alpha\beta} + \sigma^2)/bn
\end{aligned}$$

- For pairwise comparison, use 2MSAB/bn
- For $\overline{Y}_{i..}$, use approximate method





# Unrestricted versus Restricted Models

- Test of $H_0 : \sigma^2_\beta = 0$
  - Test over MSAB or MSE
  - Unrestricted considered more conservative test because of DF
- Standard error of Factor A treatment means
  - Use different standard error (slides 6 and 10)
- To decide which model is appropriate, suppose you ran experiment again and sampled some of the same levels of the random effect. Does this mean that the interaction effects for these levels are the same as before? Yes: Restricted No: Unrestricted





# Example from Montgomery

- Want to assess variability in a measurement system
- Twenty parts selected from production process
- Gauge used by 3 operators to measure parts
- Each part measured twice by each operator
- Will consider operators fixed
- Will investigate both restricted and unrestricted results

Gauge Capability Example in Text 12-3

```
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;
 class operator part;
 model resp=operator|part;
 random part operator*part / test;
 means operator / tukey lines E=operator*part;
 lsmeans operator / adjust=tukey E=operator*part tdiff stderr;

proc mixed alpha=.05 cl covtest;
 class operator part;
 model resp=operator / ddfm=kr;
 random part operator*part;
 lsmeans operator / alpha=.05 cl diff adjust=tukey;
run;
quit;
```





```
                                      Sum of
Source                    DF         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) + Q(operator)
part                    Var(Error) + 2 Var(operator*part) + 6 Var(part)
operator*part           Var(Error) + 2 Var(operator*part)

Tests of Hypotheses for Mixed Model Analysis of Variance

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
```

NOTE: If restricted model, test of part: F: 62.92, Pr > F: < .0001





```
Tukey's Studentized Range (HSD) Test for resp
Error Degrees of Freedom                 38
Error Mean Square                  0.711842
Critical Value of Studentized Range  3.44902
Minimum Significant Difference        0.4601

           Mean      N    operator
A       22.6000     40    3
A       22.3000     40    1
A       22.2750     40    2
                              Standard                  LSMEAN
operator     resp LSMEAN         Error    Pr > |t|      Number
1             22.3000000     0.1334018      <.0001           1
2             22.2750000     0.1334018      <.0001           2
3             22.6000000     0.1334018      <.0001           3

   Least Squares Means for Effect operator
   t for H0: LSMean(i)=LSMean(j) / Pr > |t|
           Dependent Variable: resp
i/j             1             2             3
   1                   0.132514      -1.59017
                         0.9904        0.2622
   2     -0.13251                    -1.72269
           0.9904                      0.2100
   3     1.590173      1.722688
           0.2622        0.2100
```





# Standard Error Issue

- The SE($\overline{Y}_{i.}$) on the previous page is incorrect. For the unrestricted model, it should be $\sqrt{(\sigma^2 + n\sigma^2_{\tau\beta} + n\sigma^2_{\beta})/bn}$, which can be estimated by $\sqrt{((a-1)MSAB + MSB)/(abn)} = .7292$.
- This is due to GLM being a fixed effects procedure
- We can calculate the correct error from the output but we need to know it is wrong. we'd also need to approximate the degrees of freedom.
- For restricted mixed model, we HAVE TO compute the SE by hand.

The Mixed Procedure

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
|---|---|---|---|
| 0 | 1 | 622.27805725 | |
| 1 | 2 | 409.45998838 | 0.00002843 |
| 2 | 1 | 409.45716449 | 0.00000003 |
| 3 | 1 | 409.45716136 | 0.00000000 |

Convergence criteria met.

Covariance Parameter Estimates

| Cov Parm | Estimate | Standard Error | Z Value | Pr Z | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|
| 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 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| operator | 2 | 38 | 1.48 | 0.2401 |
| operator | 2 | 98 | 1.48 | 0.2324 *** KR adjustment |

Unrestricted Model

$\mathrm{Var}(\overline{y}_{1..}) = (\sigma^2 + n\sigma^2_{\tau\beta} + n\sigma^2_{\beta})/bn = (0.8832 + 2(0) + 2(10.2513))/40 = 0.7312.$

$\mathrm{Var}(\overline{y}_{1..} - \overline{y}_{2..}) = 2(\sigma^2 + n\sigma^2_{\tau\beta})/bn = .8832/20 = 0.2101$

These estimates are slightly different than GLM because of the zero variance estimate.





Least Squares Means

| Effect | operator | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| operator | 1 | 22.3000 | 0.7312 | 20.1 | 30.50 | <.0001 |
| operator | 2 | 22.2750 | 0.7312 | 20.1 | 30.46 | <.0001 |
| operator | 3 | 22.6000 | 0.7312 | 20.1 | 30.91 | <.0001 |

Differences of Least Squares Means

| Effect | operator | _operator | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|---|
| operator | 1 | 2 | 0.02500 | 0.2101 | 38 | 0.12 | 0.9059 |
| operator | 1 | 3 | -0.3000 | 0.2101 | 38 | -1.43 | 0.1616 |
| operator | 2 | 3 | -0.3250 | 0.2101 | 38 | -1.55 | 0.1302 |
| operator | 1 | 2 | 0.02500 | 0.2101 | 98 | 0.12 | 0.9055** |
| operator | 1 | 3 | -0.3000 | 0.2101 | 98 | -1.43 | 0.1566** |
| operator | 2 | 3 | -0.3250 | 0.2101 | 98 | -1.55 | 0.1252** |

Differences of Least Squares Means

| Effect | operator | _operator | Lower | Upper | Adj Lower | Adj Upper |
|---|---|---|---|---|---|---|
| operator | 1 | 2 | -0.4004 | 0.4504 | -0.4875 | 0.5375 |
| operator | 1 | 3 | -0.7254 | 0.1254 | -0.8125 | 0.2125 |
| operator | 2 | 3 | -0.7504 | 0.1004 | -0.8375 | 0.1875 |
| operator | 1 | 2 | -0.3920 | 0.4420 | -0.4751 | 0.5251** |
| operator | 1 | 3 | -0.7170 | 0.1170 | -0.8001 | 0.2001** |
| operator | 2 | 3 | -0.7420 | 0.09201 | -0.8251 | 0.1751** |





# Nobound Option

```
proc mixed alpha=.05 cl nobound;
 class operator part;
 model resp=operator / ddfm=kr;
 random part operator*part;
 lsmeans operator / alpha=.05 cl diff adjust=tukey;
```

Covariance Parameter Estimates

| Cov Parm | Estimate | Alpha | Lower | Upper |
|---|---|---|---|---|
| part | 10.2798 | 0.05 | 3.6673 | 16.8924 |
| operator*part | -0.1399 | 0.05 | -0.3789 | 0.09903 |
| Residual | 0.9917 | 0.05 | 0.7143 | 1.4698 |

Type 3 Tests of Fixed Effects

| Effect | NumDF | DenDF | F Value | Pr > F |
|---|---|---|---|---|
| operator | 2 | 38 | 1.84 | 0.1730 |

Least Squares Means

| Effect | operator | Estimate | StdError | DF | t Value | Pr > \|t\| | Alpha |
|---|---|---|---|---|---|---|---|
| operator | 1 | 22.3000 | 0.7292 | 19.9 | 30.58 | <.0001 | 0.05 |
| operator | 2 | 22.2750 | 0.7292 | 19.9 | 30.55 | <.0001 | 0.05 |
| operator | 3 | 22.6000 | 0.7292 | 19.9 | 30.99 | <.0001 | 0.05 |

Differences of Least Squares Means

| Effect | operator | _operator | Estimate | StdError | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|---|
| operator | 1 | 2 | 0.02500 | 0.1887 | 38 | 0.13 | 0.8953 |
| operator | 1 | 3 | -0.3000 | 0.1887 | 38 | -1.59 | 0.1201 |
| operator | 2 | 3 | -0.3250 | 0.1887 | 38 | -1.72 | 0.0931 |





# Multifactor Models

- 3-Factor: Can have 0,1,2, or 3 random effects
- Use EMS to determine tests
- In some cases, no straightforward test exists. In other words, there is no single MS for the denominator/numerator
- Must perform approximate $F$ test
- SAS Mixed and random statement use Satterthwaite or Kenward-Roger approximation

# Satterthwaite's Approximate F-test

- Involves a linear combination of mean squares
  - To test certain factor, choose numerator and denominator such that the difference in MS is a multiple of the effect of interest
  - Ratio approximately F where

$$\mathrm{F}_{\mathrm{p,q}} = \frac{\mathrm{MS_r} \pm \ldots \pm \mathrm{MS_s}}{\mathrm{MS_u} \pm \ldots \pm \mathrm{MS_v}}$$

$$p = \frac{(\mathrm{MS_r} \pm \ldots \pm \mathrm{MS_s})^2}{\mathrm{MS_r^2/f_r} + \ldots + \mathrm{MS_s^2/f_s}}$$

$$q = \frac{(\mathrm{MS_u} \pm \ldots \pm \mathrm{MS_v})^2}{\mathrm{MS_u^2/f_u} + \ldots + \mathrm{MS_v^2/f_v}}$$

  - $f_i$ is the degrees of freedom associated with $\mathrm{MS}_i$
  - No MS in both num and denom (indep)
  - Caution when subtraction is used





# Construction of Hasse Diagram

- Described in Oehlert (2000)
- Used for determining tests
- Every term in model is a node
- Terms/nodes placed in layered structure

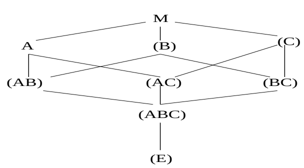
  U is above V if all terms in U are in V
- Join nodes based on structure
- Brackets placed around random terms





# 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





# Background Reading

- KNNL Section 25.2-25.6
- KNNL Chapter 22