Topic 28 - Mixed Model Example

-Random Regression-

STAT 525 - Fall 2013

Mixed versus Fixed Effects Model

- Similar to GLM but data are permitted
  - To be correlated
  - To have nonconstant variance
- Need to model covariance arises when
  - Experimental units are grouped or clustered either by design or grouped in terms of spatial orientation
  - Repeat measurements on the same experimental unit

General Mixed Effect Linear Model

- Can be expressed as

$$Y = X\beta + Z\delta + \varepsilon$$

  X and Z based on trt/design structure

  β is a vector of fixed-effect parameters

  δ is a vector of random-effect parameters

  ε is the error vector

- δ and ε assumed MVN and uncorrelated
  - means 0
  - covariance matrices G and R

DF Adjustments

- Various studies have looked at behavior of mixed model procedures on relatively "small" experiments and found two key issues
  1. Degrees of freedom adjustment
  2. Standard error bias
- Giesbrecht and Burns (1985) generalized Satterthwaite's degrees of freedom approximation for all linear models
- Both KR and SATTERTH adjust df in same manner

Standard Error Bias

- Harville (1984) showed that when estimated variance/covariance parameters are used the SEs are biased downward
  - This does not occur in balanced designs without covariance
  - Only slightly occurs in unbalanced designs without covariance
  - The more complex the design with covariance...the more severe the bias
- Kenward and Roger (1997) derived bias correction term

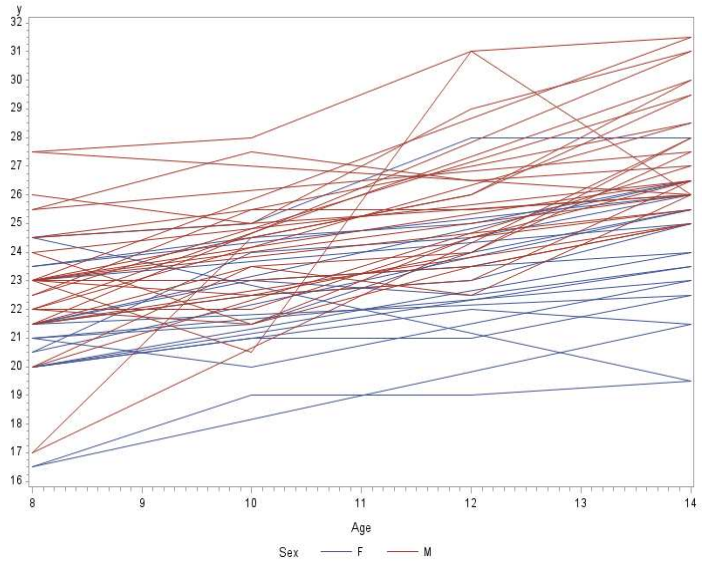
Example - Growth Curve

(Potthoff and Roy, 1964)

- Dental growth measurements of distance (mm) from pituitary gland to the pteryomaxillary fissure
- Measurements taken on 11 girls and 16 boys

```
data growth;
input Person Sex $ y1 y2 y3 y4;
y=y1; Age=8; Age1=-3; output; y=y2; Age=10;Age1=-1; output;
y=y3; Age=12;Age1=1; output;  y=y4; Age=14;Age1=3; output;
drop y1-y4;
datalines;
 1 F 21.0 20.0 21.5 23.0
 2 F 21.0 21.5 24.0 25.5
 3 F 20.5 24.0 24.5 26.0
 :
25 M 22.5 25.5 25.5 26.0
26 M 23.0 24.5 26.0 30.0
27 M 22.0 21.5 23.5 25.0
;
```


Profiles Plot



Linear Mixed Models

- General positive linear trend over age
- Males' distances typically larger than females' distances
- Will consider various linear models that describe distance as a function of age and sex
  - Two-factor ANOVA with factors AGE and SEX
  - Sex-specific regression lines

Mixed Models Considered

```
/* linear regression w/ unstructured covariance matrix */
proc mixed data=growth;
  class person sex;
  model y = sex age(sex) / noint s;
  repeated / type=un sub=person r=1;
  estimate 'diff in ints' sex 1 -1;
  estimate 'diff in slopes' age(sex) 1 -1;
run;

/* linear regression w/ compound symmetry covariance matrix */
proc mixed data=growth;
   class person sex;
   model y = sex age(sex) / noint s;
   repeated / type=cs sub=person r=1;
run;

 /* random regression coefficients model */
proc mixed data=growth;
   class person sex;
   model y = sex age(sex) / noint s;
   random intercept age / type=un sub=person g v=1;
run;
```


Mixed Models Considered

```
 /* linear regression w/ compound symmetry - heterogeneity across sex */
proc mixed data=growth;
   class person sex;
   model y = sex age(sex) / noint s;
   repeated / type=cs sub=person group=sex r=1,12;
run;

 /* linear regression w/ toeplitz*/
proc mixed data=growth;
   class person sex;
   model y = sex age(sex) / noint s;
   repeated / type=toep sub=person r=1;
run;

 /* two-way ANOVA w/ CS - heterogeneity across sex*/
proc mixed data=growth;
   class person sex age;
   model y = sex|age / s;
   repeated / type=cs group=sex sub=person r=1;
run;
```


Unstructered

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	5.4252	2.7092	3.8411	2.7151
2	2.7092	4.1906	2.9745	3.3137
3	3.8411	2.9745	6.2632	4.1332
4	2.7151	3.3137	4.1332	4.9862

Fit Statistics

-2 Res Log Likelihood	424.5
AIC (smaller is better)	444.5
AICC (smaller is better)	446.9
BIC (smaller is better)	457.5

Solution for Fixed Effects

Effect	Sex	Estimate	Standard Error	DF	t Value	Pr > |t|
Sex	F	17.4254	1.2612	25	13.82	<.0001
Sex	M	15.8423	1.0457	25	15.15	<.0001
Age(Sex)	F	0.4764	0.1066	25	4.47	0.0001
Age(Sex)	M	0.8268	0.08843	25	9.35	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
diff in ints	1.5831	1.6384	25	0.97	0.3432
diff in slopes	-0.3504	0.1385	25	-2.53	0.0181

Compound Symmetry

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	5.2207	3.2986	3.2986	3.2986
2	3.2986	5.2207	3.2986	3.2986
3	3.2986	3.2986	5.2207	3.2986
4	3.2986	3.2986	3.2986	5.2207

Fit Statistics

-2 Res Log Likelihood	433.8
AIC (smaller is better)	437.8
AICC (smaller is better)	437.9
BIC (smaller is better)	440.3

Solution for Fixed Effects

Effect	Sex	Estimate	Standard Error	DF	t Value	Pr > |t|
Sex	F	17.3727	1.1835	104	14.68	<.0001
Sex	M	16.3406	0.9813	104	16.65	<.0001
Age(Sex)	F	0.4795	0.09347	79	5.13	<.0001
Age(Sex)	M	0.7844	0.07750	79	10.12	<.0001

Random Regression Coefficients

Estimated G Matrix

Row	Effect	Person	Col1	Col2
1	Intercept	1	5.7864	-0.2896
2	Age	1	-0.2896	0.03252

Estimated V Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	4.9502	3.1751	3.1162	3.0574
2	3.1751	4.9625	3.3176	3.3888
3	3.1162	3.3176	5.2351	3.7202
4	3.0574	3.3888	3.7202	5.7679

Fit Statistics

-2 Res Log Likelihood	432.6
AIC (smaller is better)	440.6
AICC (smaller is better)	441.0
BIC (smaller is better)	445.8

Solution for Fixed Effects

Effect	Sex	Estimate	Standard Error	DF	t Value	Pr > |t|
Sex	F	17.3727	1.2284	25	14.14	<.0001
Sex	M	16.3406	1.0185	25	16.04	<.0001
Age(Sex)	F	0.4795	0.1037	25	4.62	<.0001
Age(Sex)	M	0.7844	0.08600	25	9.12	<.0001

Different CS for each sex

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	4.8870	4.2786	4.2786	4.2786
2	4.2786	4.8870	4.2786	4.2786
3	4.2786	4.2786	4.8870	4.2786
4	4.2786	4.2786	4.2786	4.8870

Estimated R Matrix for Person 12

Row	Col1	Col2	Col3	Col4
1	5.4571	2.6407	2.6407	2.6407
2	2.6407	5.4571	2.6407	2.6407
3	2.6407	2.6407	5.4571	2.6407
4	2.6407	2.6407	2.6407	5.4571

Fit Statistics

-2 Res Log Likelihood	414.7
AIC (smaller is better)	422.7
AICC (smaller is better)	423.1
BIC (smaller is better)	427.8

Solution for Fixed Effects

Effect	Sex	Estimate	Standard Error	DF	t Value	Pr > |t|
Sex	F	17.3727	0.8587	27.6	20.23	<.0001
Sex	M	16.3406	1.1287	60	14.48	<.0001
Age(Sex)	F	0.4795	0.05259	32	9.12	<.0001
Age(Sex)	M	0.7844	0.09382	47	8.36	<.0001

Toeplitz

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	5.2826	3.3659	3.6804	2.5285
2	3.3659	5.2826	3.3659	3.6804
3	3.6804	3.3659	5.2826	3.3659
4	2.5285	3.6804	3.3659	5.2826

Fit Statistics

-2 Res Log Likelihood	429.4
AIC (smaller is better)	437.4
AICC (smaller is better)	437.8
BIC (smaller is better)	442.6

Solution for Fixed Effects

Effect	Sex	Estimate	Standard Error	DF	t Value	Pr > |t|
Sex	F	17.4089	1.2932	41.2	13.46	<.0001
Sex	M	16.2704	1.0732	41.2	15.17	<.0001
Age(Sex)	F	0.4759	0.1046	28.1	4.55	<.0001
Age(Sex)	M	0.7973	0.08674	28.1	9.19	<.0001

ANOVA w/ Heterogenity CS

Estimated R Matrix for Person 1

Row	Col1	Col2	Col3	Col4
1	4.9159	4.2689	4.2689	4.2689
2	4.2689	4.9159	4.2689	4.2689
3	4.2689	4.2689	4.9159	4.2689
4	4.2689	4.2689	4.2689	4.9159

Estimated R Matrix for Person 12

1	5.4901	2.6297	2.6297	2.6297
2	2.6297	5.4901	2.6297	2.6297
3	2.6297	2.6297	5.4901	2.6297
4	2.6297	2.6297	2.6297	5.4901

Fit Statistics

-2 Res Log Likelihood	406.4
AIC (smaller is better)	414.4
AICC (smaller is better)	414.8
BIC (smaller is better)	419.5

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Sex	1	19.6	8.80	0.0077
Age	3	68.4	45.08	<.0001
Sex*Age	3	68.4	3.01	0.0360

Warning!!

- Appears ANOVA model is best fitting but the fixed effects structure has changed.
- This is not accounted for when using REML estimation.
- Need to compare this model to others using ML estimates

```
/* Two-way ANOVA w/ CS by gender*/
proc mixed data=growth method=ml;
   class person sex age;
   model y = sex|age / noint s ddfm=kr;
   repeated / type=cs sub=person group=sex r=1;
run;

 /* linear regression w/ compound symmetry - heterogeneity across sex */
proc mixed data=growth method=ml;
   class person sex;
   model y = sex age(sex) / noint s ddfm=kr;
   repeated / type=cs sub=person group=sex r=1,12;
run;
```


Comparison

```
***Linear Regression

                Fit Statistics
      -2 Log Likelihood               408.8
      AIC (smaller is better)         424.8
      AICC (smaller is better)        426.3
      BIC (smaller is better)         435.2

***ANOVA

                Fit Statistics
      -2 Log Likelihood               407.4
      AIC (smaller is better)         429.4
      AICC (smaller is better)        432.1
      BIC (smaller is better)         443.6
```

Linear Regression Model best fitting model *