Statistical Analysis of Hierarchical Data

David Zucker

Hebrew University, Jerusalem, Israel

Unit 1

Linear Mixed Models

Examples of Hierarchical Data

Hierarchical data = subunits within units

- Students within schools
- Eyes within patients
- Measurements within subjects in repeated measures data

(longitudinal data)

Examples of Linear Mixed Models

1. One-way ANOVA With Random Effect

$Y_{ij} = \mu + a_i + \epsilon_{ij}$

$a_i \sim N(0, \sigma_a^2)$, $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$

2. A Fixed Group Effect and a Random Cluster Effect

$Y_{ij} = \mu + \beta X_i + a_i + \epsilon_{ij}$

$X_i = 1$ for treatment, $X_i = 0$ for control

3. More General Regression Effects

$Y_{ij} = \mu + \sum_{k=1}^{p} \beta_k X_{ijk} + a_i + \epsilon_{ij}$

4. More Than One Level

$Y_{ijk} = \mu + \sum_{l=1}^{p} \beta_j X_{ijkl} + a_i + b_{ij} + \epsilon_{ijk}$

Examples of Linear Mixed Models – Continued

5. Random Line With a Group Effect

$$Y_{ij} = \gamma_{1i} + \gamma_{2i} t_{ij} + \epsilon_{ij}$$

$$\gamma_{1i} = \beta_0^{(1)} + \beta_1^{(1)} X_i + b_{1i}$$

$$\gamma_{2i} = \beta_0^{(2)} + \beta_1^{(2)} X_i + b_{2i}$$

$$\mathbf{b} \sim N(\mathbf{0}, \mathbf{G})$$

$$Y_{ij} = [\beta_0^{(1)} + \beta_1^{(1)} X_i] + [\beta_0^{(2)} + \beta_1^{(2)} X_i] t_{ij} + [b_{1i} + b_{2i} t_{ij}] + \epsilon_{ij}$$

General Form of Linear Mixed Model

$$Y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{b}_i + \epsilon_{ij}$$

$$\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}^{(1)}))$$

$$\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i(\boldsymbol{\theta}^{(2)}))$$

in matrix form

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i$$

$$\mathbf{Y}_i \sim N(\mathbf{X}_i \boldsymbol{\beta}, \mathbf{V}_i(\boldsymbol{\theta}))$$

$$\mathbf{V}_i(\boldsymbol{\theta}) = \mathbf{Z}_i \mathbf{G}(\boldsymbol{\theta}^{(1)}) \mathbf{Z}_i^T + \mathbf{R}_i(\boldsymbol{\theta}^{(2)})$$

Random Line Example in General LMM Form

random line model

$$Y_{ij} = [\beta_0^{(1)} + \beta_1^{(1)} X_i] + [\beta_0^{(2)} + \beta_1^{(2)} X_i] t_{ij} + [b_{1i} + b_{2i} t_{ij}] + \epsilon_{ij}$$

representation in general LMM form

$$\begin{aligned}
Y_{ij} &= \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{b}_i + \epsilon_{ij} \\
\mathbf{x}_{ij}^T &= [1, X_i, t_{ij}, X_i t_{ij}] \\
\boldsymbol{\beta}^T &= [\beta_0^{(1)}, \beta_1^{(1)}, \beta_0^{(2)}, \beta_1^{(2)}] \\
\mathbf{z}_{ij}^T &= [1, t_{ij}] \\
\mathbf{b}_i &= [b_{1i}, b_{2i}]
\end{aligned}$$

Statistical Inference for Parameters

Estimation – Maximum Likelihood

Large Sample Distribution

$$\boldsymbol{\phi} = (\boldsymbol{\beta}, \boldsymbol{\theta})$$

$$\hat{\boldsymbol{\phi}} - \boldsymbol{\phi} \mathrel{\dot{\sim}} N(\mathbf{0}, \boldsymbol{\Omega}(\hat{\boldsymbol{\phi}}))$$

$$\boldsymbol{\Omega}(\boldsymbol{\phi}) = \text{inverse Fisher information matrix}$$

Inference for Individual Fixed-Effect Coefficients

Regard

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\Omega_{jj}^{\beta\beta}}}$$

as being approximately normally distributed,

or approximately t-distributed with d d.f.

Choice of d: Various methods available

- `DDFM` option in `SAS PROC MIXED`

  `DDFM=SAT` uses Satterthwaite approximation

- the SPSS `MIXED` procedure also uses the Satterthwaite approximation

Inference for Individual Coefficients – Continued

Testing $H_0 : \beta_j = 0$: Reject if

$$\left| \frac{\hat{\beta}_j}{\sqrt{\Omega^{\beta\beta}_{jj}}} \right| \geq t_d(1 - \tfrac{\alpha}{2})$$

Confidence interval for β_j

$$\hat{\beta}_j \pm t_d(1 - \tfrac{\alpha}{2})\sqrt{\Omega^{\beta\beta}_{jj}}$$

Testing a Single Linear Combination of Fixed Effects

Look at $\psi = \mathbf{c}^T \boldsymbol{\beta}$

$\hat{\psi} = \mathbf{c}^T \hat{\boldsymbol{\beta}} \mathrel{\dot{\sim}} N(\psi, \mathbf{c}^T \boldsymbol{\Omega}^{(\beta\beta)} \mathbf{c})$

Test $H_0 : \psi = 0$ using t-test with d d.f.

Confidence interval for ψ

$$\hat{\psi} \pm t_d(1 - \tfrac{\alpha}{2}) \sqrt{\mathbf{c}^T \boldsymbol{\Omega}^{(\beta\beta)} \mathbf{c})}$$

Examples:

1. Expected response for a given set of covariate values

2. Treatment contrast

Testing Several Linear Combinations

Look at $\boldsymbol{\phi} = \mathbf{L}^T\boldsymbol{\beta}$, $\mathbf{L}^T = m \times p$ matrix

$$\hat{\boldsymbol{\phi}} = \mathbf{L}^T\hat{\boldsymbol{\beta}} \dot{\sim} N(\boldsymbol{\phi}, \mathbf{L}^T\boldsymbol{\Omega}^{(\beta\beta)}\mathbf{L})$$

test $H_0 : \boldsymbol{\psi} = 0$ using F-test with d.f. m and d

$$F = \hat{\boldsymbol{\phi}}^T(\mathbf{L}^T\boldsymbol{\Omega}^{(\beta\beta)}\mathbf{L})^{-1}\hat{\boldsymbol{\phi}}/m$$

Example:

Test for overall group effect in multiple group study

REML Estimation of Covariance Parameters

Idea: Adjust the estimator of $\boldsymbol{\theta}$ to take into account the estimation of $\boldsymbol{\beta}$

Similar to

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \text{ vs. } \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

The Method:

Let $\mathbf{U}$ be an $n \times (n-p)$ matrix of full rank such that $\mathbf{U}^T\mathbf{X} = \mathbf{0}$

Then base estimation of $\boldsymbol{\theta}$ on $\mathbf{T} = \mathbf{U}^T\mathbf{Y}$

We have $\mathbf{T} \sim N(\mathbf{0}, \mathbf{U}^T\mathbf{V}(\boldsymbol{\theta})\mathbf{U})$

Software

- SAS: `PROC MIXED`
- SPSS: `MIXED`
- R: several procedures
  - `lme` function it `nlme` package
  - `lmer` function in `lme4` package
  - `lmerTest` package for Satterthwaite degrees of freedom

Example

Three groups of rats given different growth treatments

and then followed over time

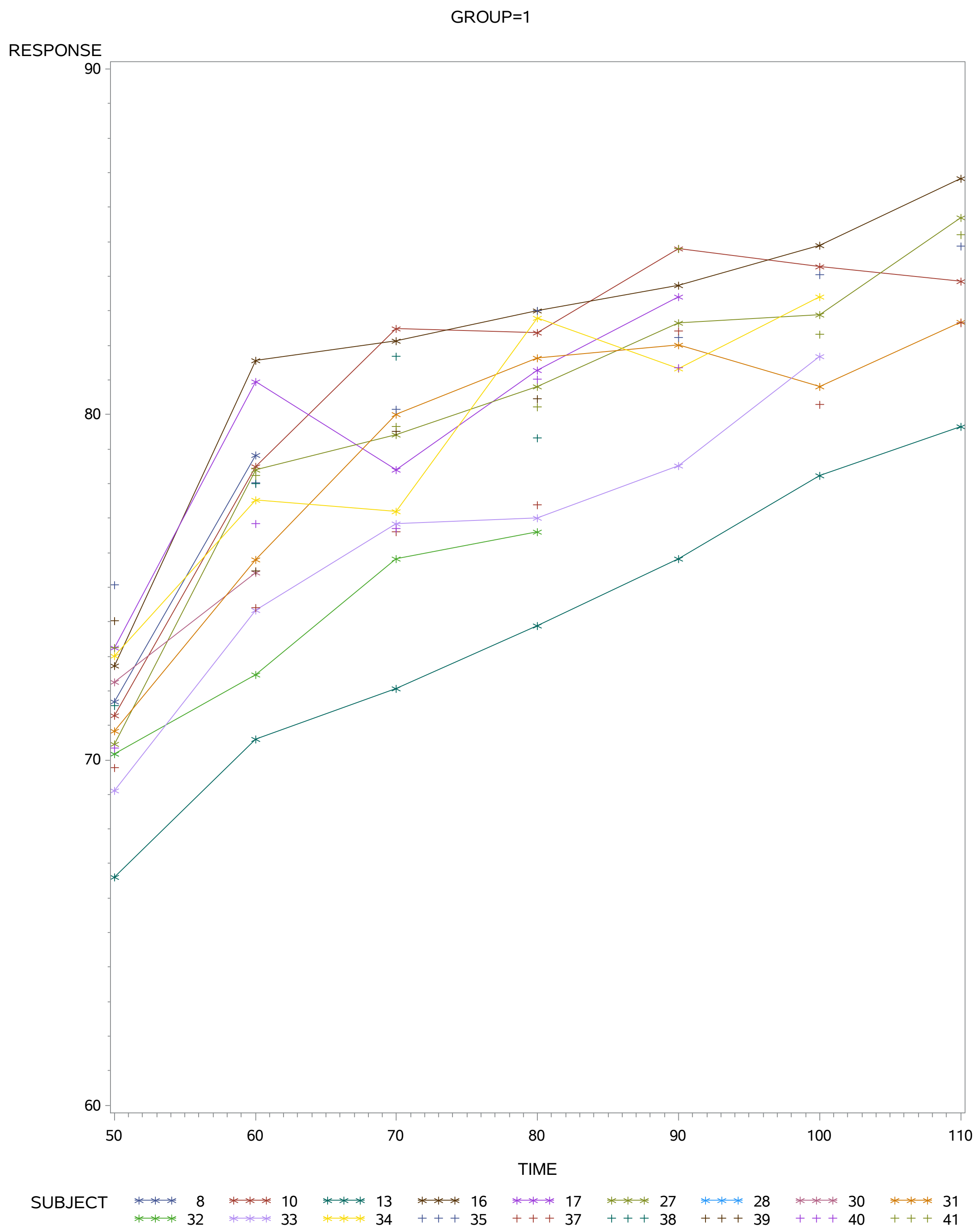
Analysis in `SAS PROC MIXED`

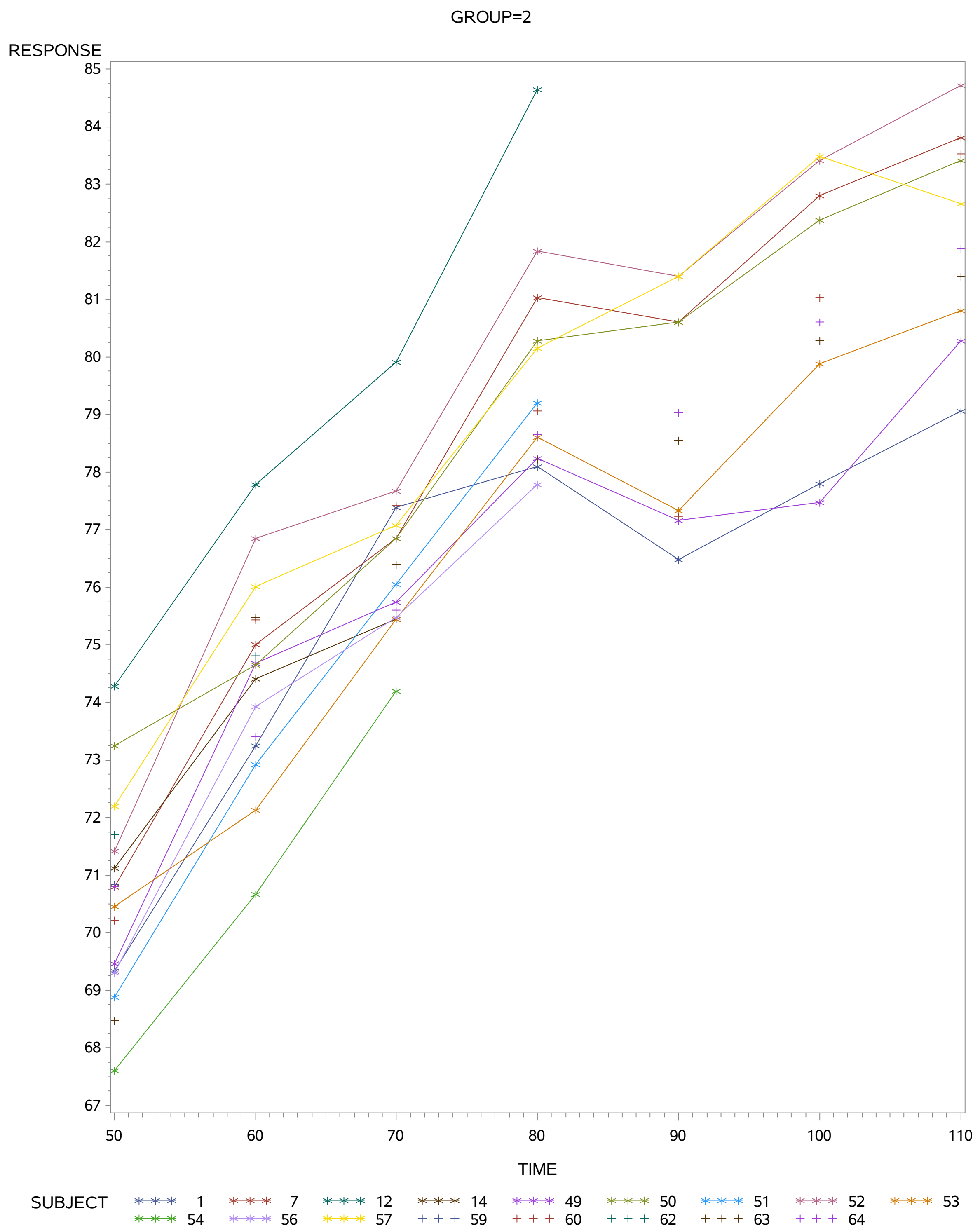
First model:

```
proc mixed method=reml ratio covtest;
class group;
model response = group time group*time / solution ddfm=sat;
random intercept time / subject=subject type=un;
run;
```

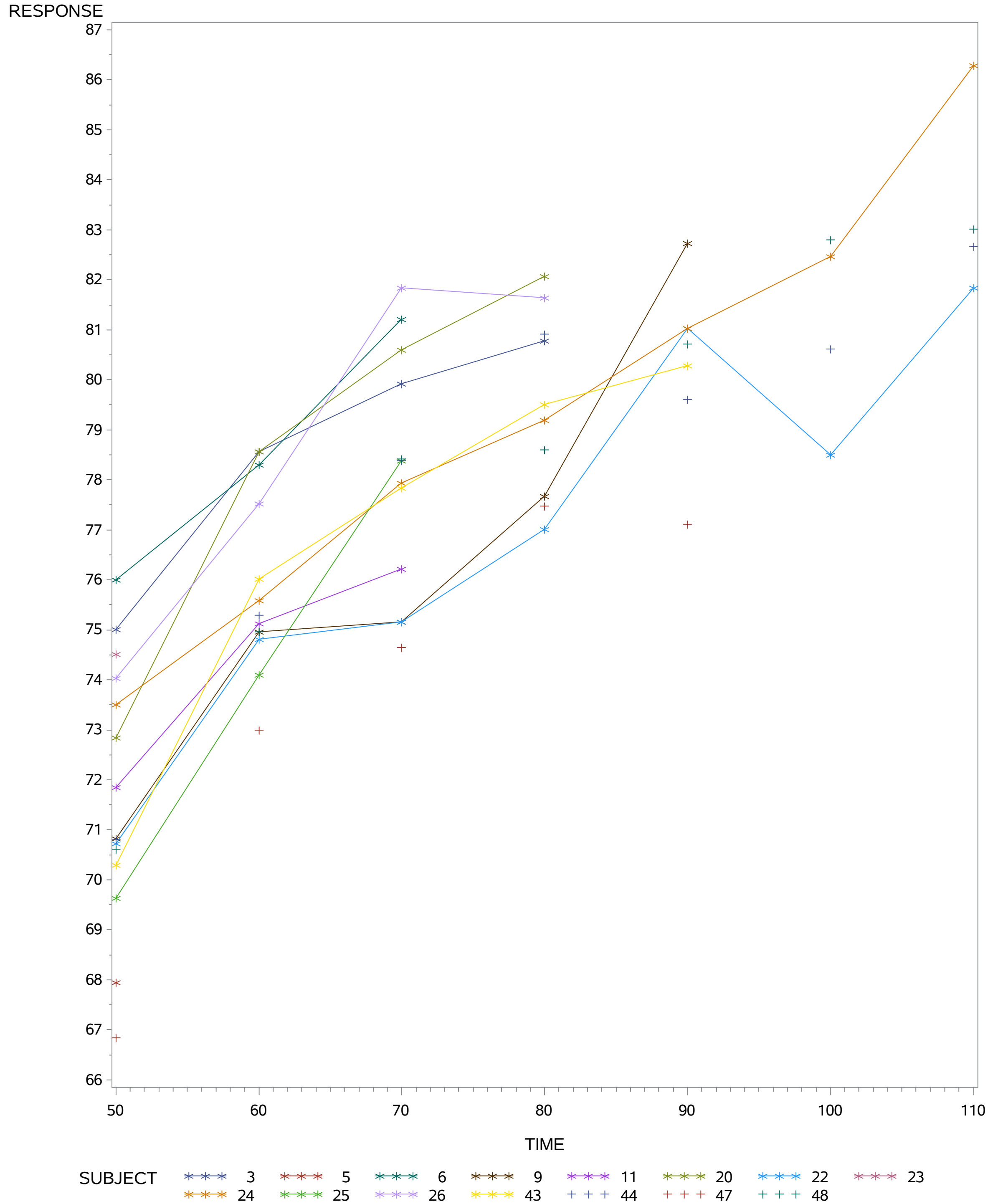
Second model:

```
proc mixed method=reml ratio covtest;
class group;
model response = group time group*time / solution ddfm=sat;
random intercept / subject=subject type=un;
run;
```


GROUP=1
RESPONSE
90
80
70
60
50
60
70
80
90
100
110
TIME
SUBJECT
8
10
13
16
17
27
28
30
31
32
33
34
35
37
38
39
40
41


GROUP=2
RESPONSE
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
50
60
70
80
90
100
110
TIME
SUBJECT
1
7
12
14
49
50
51
52
53
54
56
57
59
60
62
63
64

GROUP=3


RESPONSE
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
50
60
70
80
90
100
110
TIME
SUBJECT
3
5
6
9
11
20
22
23
24
25
26
43
44
47
48

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	RESPONSE
Covariance Structure	Unstructured
Subject Effect	SUBJECT
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
GROUP	3	1 2 3

Dimensions	
Covariance Parameters	4
Columns in X	8
Columns in Z per Subject	2
Subjects	50
Max Obs per Subject	7

Number of Observations	
Number of Observations Read	350
Number of Observations Used	252
Number of Observations Not Used	98

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1178.84457909	
1	4	1083.75040734	0.00936467
2	2	1081.86275760	0.00023405
3	1	1081.78756649	0.00000117
4	1	1081.78719582	0.00000000

Convergence criteria met.

The Mixed Procedure

Covariance Parameter Estimates						
Cov Parm	**Subject**	**Ratio**	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**
UN(1,1)	SUBJECT	0.8404	2.3413	2.3237	1.01	0.1568
UN(2,1)	SUBJECT	0.002706	0.007538	0.01623	0.46	0.6424
UN(2,2)	SUBJECT	0	0	.	.	.
Residual		1.0000	2.7860	0.2796	9.96	<.0001

Fit Statistics	
-2 Res Log Likelihood	1081.8
AIC (Smaller is Better)	1087.8
AICC (Smaller is Better)	1087.9
BIC (Smaller is Better)	1093.5

Null Model Likelihood Ratio Test		
DF	**Chi-Square**	**Pr > ChiSq**
2	97.06	<.0001

Solution for Fixed Effects						
Effect	**GROUP**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept		62.6627	0.9881	81	63.42	<.0001
GROUP	1	0.9350	1.2747	66.5	0.73	0.4658
GROUP	2	0.09478	1.2775	64.3	0.07	0.9411
GROUP	3	0	.	.	.	.
TIME		0.2052	0.01290	174	15.91	<.0001
TIME*GROUP	1	-0.00905	0.01610	179	-0.56	0.5749
TIME*GROUP	2	-0.01982	0.01605	177	-1.23	0.2187
TIME*GROUP	3	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
GROUP	2	57.6	0.37	0.6900
TIME	1	180	981.96	<.0001
TIME*GROUP	2	181	0.81	0.4473

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	RESPONSE
Covariance Structure	Unstructured
Subject Effect	SUBJECT
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
GROUP	3	1 2 3

Dimensions	
Covariance Parameters	2
Columns in X	8
Columns in Z per Subject	1
Subjects	50
Max Obs per Subject	7

Number of Observations	
Number of Observations Read	350
Number of Observations Used	252
Number of Observations Not Used	98

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1178.84457909	
1	2	1082.01242375	0.00000799
2	1	1082.00987356	0.00000000

Convergence criteria met.

The Mixed Procedure

Covariance Parameter Estimates						
Cov Parm	Subject	Ratio	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	SUBJECT	1.2194	3.3969	0.8530	3.98	<.0001
Residual		1.0000	2.7857	0.2795	9.97	<.0001

Fit Statistics	
-2 Res Log Likelihood	1082.0
AIC (Smaller is Better)	1086.0
AICC (Smaller is Better)	1086.1
BIC (Smaller is Better)	1089.8

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
1	96.83	<.0001

Solution for Fixed Effects						
Effect	GROUP	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		62.7378	1.0173	233	61.67	<.0001
GROUP	1	0.8562	1.3181	227	0.65	0.5166
GROUP	2	0.02159	1.3220	225	0.02	0.9870
GROUP	3	0	.	.	.	.
TIME		0.2039	0.01274	216	16.01	<.0001
TIME*GROUP	1	-0.00772	0.01592	214	-0.48	0.6284
TIME*GROUP	2	-0.01857	0.01587	215	-1.17	0.2433
TIME*GROUP	3	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
GROUP	2	221	0.32	0.7291
TIME	1	214	998.99	<.0001
TIME*GROUP	2	213	0.75	0.4759

Example – Analysis in SPSS

```
MIXED RESPONSE BY GROUP WITH TIME
   /FIXED=GROUP TIME GROUP*TIME | SSTYPE(3)
   /METHOD=REML
   /RANDOM=INTERCEPT | SUBJECT(SUBJECT) COVTYPE(VC)
   /PRINT=SOLUTION.
```

```
 DATASET ACTIVATE DataSet1.
MIXED RESPONSE BY GROUP WITH TIME
   /FIXED=GROUP TIME GROUP*TIME | SSTYPE(3)
   /METHOD=REML
   /RANDOM=INTERCEPT | SUBJECT(SUBJECT) COVTYPE(VC)
   /PRINT=SOLUTION.
```

Mixed Model Analysis

Notes

Output Created		10-FEB-2015 15:02:13
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	350
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the model.
Syntax		MIXED RESPONSE BY GROUP WITH TIME /FIXED=GROUP TIME GROUP*TIME | SSTYPE(3) /METHOD=REML /RANDOM=INTERCEPT | SUBJECT(SUBJECT) COVTYPE (VC) /PRINT=SOLUTION.
Resources	Processor Time	00:00:00.08
	Elapsed Time	00:00:00.23

[DataSet1]

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	GROUP	3		2	
	TIME	1		1	
	GROUP * TIME	3		2	
Random Effects	Intercept[b]	1	Variance Components	1	SUBJECT
Residual				1	
Total		9		8	

a. Dependent Variable: RESPONSE.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

Information Criteria[a]

-2 Restricted Log Likelihood	1082.010
Akaike's Information Criterion (AIC)	1086.010
Hurvich and Tsai's Criterion (AICC)	1086.059
Bozdogan's Criterion (CAIC)	1095.021
Schwarz's Bayesian Criterion (BIC)	1093.021

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: RESPONSE.

Fixed Effects

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	222.848	14592.710	.000
GROUP	2	221.235	.316	.729
TIME	1	213.775	999.033	.000
GROUP * TIME	2	213.326	.745	.476

a. Dependent Variable: RESPONSE.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% ... Lower Bound	95% ... Upper Bound
Intercept	62.737772	1.017263	233.116	61.673	.000	60.733568	64.741977
[GROUP=1]	.856228	1.318108	227.123	.650	.517	-1.741057	3.453512
[GROUP=2]	.021642	1.321985	225.125	.016	.987	-2.583405	2.626690
[GROUP=3]	0[b]	0	.	.	.	.	.
TIME	.203913	.012739	216.409	16.007	.000	.178805	.229022
[GROUP=1] * TIME	-.007717	.015921	214.444	-.485	.628	-.039099	.023664
[GROUP=2] * TIME	-.018571	.015871	214.813	-1.170	.243	-.049854	.012712
[GROUP=3] * TIME	0[b]	0	.	.	.	.	.

a. Dependent Variable: RESPONSE.

b. This parameter is set to zero because it is redundant.

Covariance Parameters

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error
Residual		2.785546	.279442
Intercept [subject = SUBJECT]	Variance	3.397755	.853334

a. Dependent Variable: RESPONSE.

Example – Analysis in R

```
library(lme4)
library(lmerTest)
grp1 = (GROUP==1)
grp2 = (GROUP==2)
as.numeric(grp1)
as.numeric(grp2)
grp1tim = grp1*TIME
grp2tim = grp2*TIME
reres1= lmer(RESPONSE ~ grp1 + grp2 + TIME + grp1tim +
  grp2tim + (1|SUBJECT), REML=TRUE)
summary(reres1, ddf="Satterthwaite")
```

```
Linear mixed model fit by REML t-tests use Satterthwaite approximations to degrees of freedom
[merModLmerTest]
Formula: RESPONSE ~ grp1 + grp2 + TIME + grp1tim + grp2tim + (1 | SUBJECT)

REML criterion at convergence: 1082

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.23579 -0.59830  0.03894  0.58385  2.25315

Random effects:
 Groups   Name        Variance Std.Dev.
 SUBJECT  (Intercept) 3.398    1.843
 Residual             2.786    1.669
Number of obs: 252, groups:  SUBJECT, 50

Fixed effects:
              Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  62.737772   1.017263 233.120000  61.673   <2e-16 ***
grp1TRUE      0.856228   1.318108 227.120000   0.650    0.517
grp2TRUE      0.021642   1.321985 225.130000   0.016    0.987
TIME          0.203913   0.012739 216.410000  16.007   <2e-16 ***
grp1tim      -0.007717   0.015921 214.440000  -0.485    0.628
grp2tim      -0.018571   0.015871 214.810000  -1.170    0.243
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Correlation of Fixed Effects:
         (Intr) g1TRUE g2TRUE TIME   grp1tm
grp1TRUE -0.772
grp2TRUE -0.769  0.594
TIME     -0.854  0.659  0.657
grp1tim   0.683 -0.842 -0.526 -0.800
grp2tim   0.685 -0.529 -0.838 -0.803  0.642
```

Example - Continued

Examining Contrasts

```
proc mixed;
class group;
model response = group time group*time / solution ddfm=sat;
random intercept / subject=subject;
estimate 'lincomb' group -0.5 -0.5 1 / cl alpha=0.05;
contrast 'lincomb1' group -0.5 -0.5 1;
contrast 'lincomb2' group -1 1 0, group -1 0 1;
run;
```

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
lincomb	-0.4389	1.1784	229	-0.37	0.7099

Alpha	Lower	Upper
0.05	-2.7608	1.8830

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
lincomb1	1	229	0.14	0.7099
lincomb2	2	222	0.32	0.7291

Contrasts in R – Direct Computation Using Formulas

```
reres1= lmer(RESPONSE ~ grp1 + grp2 + TIME + grp1tim + grp2tim +
  (1|SUBJECT), REML=TRUE)
beta = fixef(reres1)
omega = vcov(reres1)
l1 = c(0,-0.5,-0.5,0,0,0)
psi1 = t(l1) %*% beta
vpsi1 = t(l1) %*% omega %*% l1
sdpsi1 = sqrt(vpsi1)
cbind(psi1,sdpsi1)
l2t = rbind(c(0,1,0,0,0,0),c(0,0,1,0,0,0))
phi = l2t %*% beta
mdlmat = solve(l2t %*% omega %*% t(l2t))
f = t(phi) %*% mdlmat %*% phi / nrow(l2t)

Output:
> ans
       psi1   sdpsi1
1 -0.438935 1.178421
> f
1 x 1 Matrix of class "dgeMatrix"
          [,1]
[1,] 0.3163821
```

Inference for Covariance Parameters

Recall: $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \dot{\sim} N(\mathbf{0}, \boldsymbol{\Omega}^{\theta\theta})$

Wald Test of $H_0 : \theta_j = 0$: $Z_j^{Wald} = \hat{\theta}_j / \sqrt{\Omega_{jj}^{\theta\theta}}$

Wald CI: $\hat{\theta}_j \pm z_{1-\alpha/2}\sqrt{\Omega_{jj}^{\theta\theta}}$

Handling Positivity Constraints on Covariance Parameters

Satterthwaite approach: Use the approximation

$$\frac{\hat{\theta}_j}{\theta_j} \dot{\sim} \frac{\chi_\nu^2}{\nu}$$

with $\nu = 2(Z_j^{Wald})^2$ (matching distn of $\hat{\theta}_j$ on first two moments)

Transformation and delta method (Taylor approximation):

$$\log \hat{\theta}_j \dot{\sim} N(\log \theta_j, \Omega_{jj}^{\theta\theta}/\hat{\theta}_j^2)$$

Covariance Structures

Recall the model

$$\begin{aligned}\mathbf{Y}_i &= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i \\ \mathbf{b}_i &\sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}^{(1)})) \\ \boldsymbol{\epsilon}_i &\sim N(\mathbf{0}, \mathbf{R}_i(\boldsymbol{\theta}^{(2)}))\end{aligned}$$

Need to specify the forms of the matrices $\mathbf{G}(\boldsymbol{\theta}^{(1)})$ and $\mathbf{R}(\boldsymbol{\theta}^{(2)})$

RANDOM statement specifies structure of $\mathbf{G}(\boldsymbol{\theta}^{(1)})$

- `TYPE=VC` means the elements of $\mathbf{b}_i$ are independent
- `TYPE=UN` means that $\mathbf{G}(\boldsymbol{\theta}^{(1)})$ is unstructured (this is the usual choice)

REPEATED statement specifies structure of $\mathbf{R}(\boldsymbol{\theta}^{(2)})$

- Default structure of $\mathbf{R}(\boldsymbol{\theta}^{(2)})$ is $\sigma^2\mathbf{I}$
- Other structures are available for longitudinal or spatial data

Structures for $\mathbf{R}(\boldsymbol{\theta}^{(2)})$ for Longitudinal Data

Many structures offered in `PROC MIXED` are appropriate only for data sets where all individuals are measured at the same set of times (except for missing entries).

E.g. `VC` structure: $\epsilon_{ij} \sim N(0, \sigma_j^2)$ independent across j

Some available structures (referred to as "spatial") are appropriate for more general measurement schedules

E.g., `SP(EXP)` structure: $\text{Cov}(\epsilon_{ij}, \epsilon_{ik}) = \sigma^2 \exp(-d_{ijk}/\theta)$
In time series terms, this is an AR(1) structure

Example of Analysis with `SP(EXP)` Structure

```
proc mixed covtest cl;
class group;
model response = group time group*time / solution ddfm=sat;
random intercept / subject=subject;
repeated / subject=subject type=sp(exp)(time);
run;
```

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	RESPONSE
Covariance Structures	Variance Components, Spatial Exponential
Subject Effects	SUBJECT, SUBJECT
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
GROUP	3	1 2 3

Dimensions	
Covariance Parameters	3
Columns in X	8
Columns in Z per Subject	1
Subjects	50
Max Obs per Subject	7

Number of Observations	
Number of Observations Read	350
Number of Observations Used	252
Number of Observations Not Used	98

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1178.84457909	
1	2	1082.01282746	0.00036877
2	1	1082.00373938	0.00000623
3	1	1081.99149279	0.00002403
4	1	1080.94833145	0.00488863
5	1	1078.20005697	0.01134414
6	2	1074.91220775	0.25432100
7	2	1072.08307174	0.05731147
8	2	1068.63552431	0.00035089

The Mixed Procedure

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
9	1	1068.52277297	0.00000160
10	1	1068.52227929	0.00000000

Convergence criteria met.

Covariance Parameter Estimates								
Cov Parm	**Subject**	**Estimate**	**Standard Error**	**Z Value**	**Pr > Z**	**Alpha**	**Lower**	**Upper**
Intercept	SUBJECT	2.5356	0.9674	2.62	0.0044	0.05	1.3524	6.3735
SP(EXP)	SUBJECT	11.5649	3.7322	3.10	0.0010	0.05	6.7053	24.5531
Residual		3.6941	0.7055	5.24	<.0001	0.05	2.6244	5.5858

Fit Statistics	
-2 Res Log Likelihood	1068.5
AIC (Smaller is Better)	1074.5
AICC (Smaller is Better)	1074.6
BIC (Smaller is Better)	1080.3

Solution for Fixed Effects						
Effect	**GROUP**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept		61.9196	1.2379	74	50.02	<.0001
GROUP	1	0.5195	1.6106	70.7	0.32	0.7480
GROUP	2	0.01386	1.6168	70	0.01	0.9932
GROUP	3	0	.	.	.	.
TIME		0.2133	0.01634	56.9	13.05	<.0001
TIME*GROUP	1	-0.00456	0.02064	52.7	-0.22	0.8261
TIME*GROUP	2	-0.01861	0.02062	52	-0.90	0.3709
TIME*GROUP	3	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
GROUP	2	68.4	0.08	0.9258
TIME	1	50.5	650.86	<.0001
TIME*GROUP	2	49.6	0.51	0.6062

Heterogeneous Variances

- Allow variance of random effect to depend on group
- Allow variance of error term to depend on group

```
proc mixed cl covtest;
class group;
model response = group time group*time / solution ddfm=sat;
random intercept / subject=subject group=group;
repeated / subject=subject group=group;
run;
```

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	RESPONSE
Covariance Structure	Variance Components
Subject Effects	SUBJECT, SUBJECT
Group Effects	GROUP, GROUP
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
GROUP	3	1 2 3

Dimensions	
Covariance Parameters	6
Columns in X	8
Columns in Z per Subject	3
Subjects	50
Max Obs per Subject	7

Number of Observations	
Number of Observations Read	350
Number of Observations Used	252
Number of Observations Not Used	98

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1178.84457909	
1	2	1080.35972524	0.00936978
2	1	1076.81077043	0.00171905
3	1	1076.20164930	0.00010287
4	1	1076.16814216	0.00000050
5	1	1076.16798582	0.00000000

Convergence criteria met.

The Mixed Procedure

Covariance Parameter Estimates									
Cov Parm	Subject	Group	Estimate	Standard Error	Z Value	Pr > Z	Alpha	Lower	Upper
Intercept	SUBJECT	GROUP 1	3.9316	1.5947	2.47	0.0068	0.05	2.0290	10.6267
Intercept	SUBJECT	GROUP 2	2.3423	1.0213	2.29	0.0109	0.05	1.1605	6.9602
Intercept	SUBJECT	GROUP 3	4.2192	1.9509	2.16	0.0153	0.05	2.0195	13.6506
Residual	SUBJECT	GROUP 1	3.5616	0.5742	6.20	<.0001	0.05	2.6582	5.0213
Residual	SUBJECT	GROUP 2	2.2920	0.3793	6.04	<.0001	0.05	1.6985	3.2635
Residual	SUBJECT	GROUP 3	2.2963	0.4647	4.94	<.0001	0.05	1.6014	3.5688

Fit Statistics	
-2 Res Log Likelihood	1076.2
AIC (Smaller is Better)	1088.2
AICC (Smaller is Better)	1088.5
BIC (Smaller is Better)	1099.6

Solution for Fixed Effects						
Effect	GROUP	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		62.6838	0.9758	56.9	64.24	<.0001
GROUP	1	0.9096	1.3516	133	0.67	0.5021
GROUP	2	0.07338	1.2293	115	0.06	0.9525
GROUP	3	0	.	.	.	.
TIME		0.2046	0.01165	52.5	17.57	<.0001
TIME*GROUP	1	-0.00835	0.01587	123	-0.53	0.5996
TIME*GROUP	2	-0.01918	0.01446	104	-1.33	0.1875
TIME*GROUP	3	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
GROUP	2	129	0.31	0.7372
TIME	1	181	1056.64	<.0001
TIME*GROUP	2	119	0.93	0.3969

Comparing Models

Can compare models using likelihood ratio test:

1. Run larger model, extract $-2\log L$
2. Run smaller model, extract $-2\log L$
3. Compute G^2 = difference in the value of $-2\log L$
4. Test significance based on the chi-square distn with r d.f.,

r = the number of added parameters in the larger model

Important note:

- If we are changing fixed effects, LR test is valid **only** with ML
- If we are changing **only** random effects, can use ML or REML

Example:

Model with homogenous variance structure: $-2\log L = 1082.0$

Model with heterogeneous variance structure: $-2\log L = 1076.2$

$G^2 = 5.8$, $r = 4$, p-value $= 0.21$

Nested Random Effects – Example

2 groups (fixed effect)

4 schools per group (1st level random effect)

2 classes per school (2nd level random effect)

5 students per class

```
proc mixed ratio covtest;
class group school class;
model score = group / solution ddfm=sat;
random intercept / subject=school v;
random intercept / subject=class(school);
run;
```

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	score
Covariance Structure	Variance Components
Subject Effects	school, class(school)
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	Levels	Values
group	2	1 2
school	8	1 2 3 4 5 6 7 8
class	2	1 2

Dimensions	
Covariance Parameters	3
Columns in X	3
Columns in Z per Subject	3
Subjects	8
Max Obs per Subject	10

Number of Observations	
Number of Observations Read	80
Number of Observations Used	80
Number of Observations Not Used	0

Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	386.89928016	
1	1	365.14616302	0.00000000

Convergence criteria met.

The Mixed Procedure

Estimated V Matrix for school 1										
Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col10
1	8.3841	3.7580	3.7580	3.7580	3.7580	3.1312	3.1312	3.1312	3.1312	3.1312
2	3.7580	8.3841	3.7580	3.7580	3.7580	3.1312	3.1312	3.1312	3.1312	3.1312
3	3.7580	3.7580	8.3841	3.7580	3.7580	3.1312	3.1312	3.1312	3.1312	3.1312
4	3.7580	3.7580	3.7580	8.3841	3.7580	3.1312	3.1312	3.1312	3.1312	3.1312
5	3.7580	3.7580	3.7580	3.7580	8.3841	3.1312	3.1312	3.1312	3.1312	3.1312
6	3.1312	3.1312	3.1312	3.1312	3.1312	8.3841	3.7580	3.7580	3.7580	3.7580
7	3.1312	3.1312	3.1312	3.1312	3.1312	3.7580	8.3841	3.7580	3.7580	3.7580
8	3.1312	3.1312	3.1312	3.1312	3.1312	3.7580	3.7580	8.3841	3.7580	3.7580
9	3.1312	3.1312	3.1312	3.1312	3.1312	3.7580	3.7580	3.7580	8.3841	3.7580
10	3.1312	3.1312	3.1312	3.1312	3.1312	3.7580	3.7580	3.7580	3.7580	8.3841

Covariance Parameter Estimates						
Cov Parm	Subject	Ratio	Estimate	Standard Error	Z Value	Pr > Z
Intercept	school	0.6769	3.1312	2.2890	1.37	0.0857
Intercept	class(school)	0.1355	0.6268	0.7930	0.79	0.2147
Residual		1.0000	4.6261	0.8178	5.66	<.0001

Fit Statistics	
-2 Res Log Likelihood	365.1
AIC (Smaller is Better)	371.1
AICC (Smaller is Better)	371.5
BIC (Smaller is Better)	371.4

Solution for Fixed Effects						
Effect	group	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		2.3298	0.9883	6	2.36	0.0565
group	1	0.3980	1.3977	6	0.28	0.7854
group	2	0	.	.	.	.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
group	1	6	0.08	0.7854

"Estimation" of Random Effects

$$\begin{bmatrix} \mathbf{Y}_i \\ \mathbf{b}_i \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{X}_i\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{V}_i & \mathbf{Z}_i\mathbf{G} \\ \mathbf{G}\mathbf{Z}_i^T & \mathbf{G} \end{bmatrix} \right)$$

$$\mathbf{b}_i|\mathbf{Y}_i \sim N(\mathbf{D}_i(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta}), \mathbf{Q}_i)$$

$$\hat{\mathbf{b}}_i = \mathbf{D}_i(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})$$

Without the normality assumption, the above predictor is Best Linear Unbiased Predictor (BLUP), where best means minimum mean square error

With estimates substituted for unknown parameters (as above), the term Empirical BLUP (EBLUP) is used

Estimation of Random Effects – Continued

Ignoring the estimation error in estimating the random effects parameters $\boldsymbol{\theta}$, we can write

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\mathbf{b}} - \mathbf{b} \end{bmatrix} \sim N(\mathbf{0}, \mathbf{C})$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}$$

where the $\mathbf{C}_{rs}$ are matrices that can be written down

Prediction Intervals

$$\mathbf{L}^T \begin{bmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\mathbf{b}} - \mathbf{b} \end{bmatrix} \sim N(\mathbf{0}, \mathbf{L}^T \mathbf{C} \mathbf{L})$$

Prediction interval for

$$\mathbf{L}^T \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{b} \end{bmatrix}$$

is given by

$$\mathbf{L}^T \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} \pm z_{1-\alpha/2} [\mathbf{L}^T \mathbf{C} \mathbf{L}]^{1/2}$$

Example – Gaucher's Disease Study

Study of growth of children with Gaucher's disease

Measurements are standardized height values

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_3 X_i + \beta_4 X_i t_{ij} + b_{i1} + b_{i2} t_{ij} + \epsilon_j$$

X_i = group indicator (0/1)

```
proc mixed method=reml ratio covtest;
model y = sex t sex*t / solution covb ddfm=satterth outpred=op;
random int t / subject=id type=un;
estimate 'pred1'
    intercept 1
    t 1.5
    sex 1
    sex*t 1.5
     |
    intercept 1
     t 1.5
    / e
      sub 1 0 0 0 0 0 0
          0 0 0 0 0 0 0
          0 0 0 0 0 0 0
          0 0 0 0 0 0 0;
```

```
estimate 'pred2'
    intercept 1
    t 3.5
    sex 1
    sex*t 3.5
     |
    intercept 1
     t 3.5
    / e
      sub 1 0 0 0 0 0 0
          0 0 0 0 0 0 0
          0 0 0 0 0 0 0
          0 0 0 0 0 0 0;
run;

proc print data=op;
run;
```

Obs	id	t	sex	y
1	1	0.50	1	-0.34000
2	1	1.00	1	0.00000
3	1	1.50	1	0.50000
4	1	2.06	1	0.25000
5	2	0.30	1	0.31500
6	2	0.58	1	0.04500
7	2	1.50	1	0.26500
8	2	2.06	1	0.29500
9	2	2.58	1	0.16500
10	3	0.80	1	0.07500
11	3	1.20	1	0.21500
12	3	1.85	1	-0.02500
13	3	2.20	1	0.69500
14	3	2.50	1	0.92500
15	4	0.60	1	-0.57140
16	4	1.60	1	-0.22140
17	4	2.50	1	0.09857
18	4	3.10	1	0.53860
19	5	0.35	0	0.07500
20	5	0.55	0	0.34500
21	5	0.83	0	0.58500
22	5	1.00	0	0.53500
23	5	1.50	0	0.63500
24	5	2.00	0	0.52500
25	6	1.10	1	-0.28750
26	6	2.13	1	0.68250
27	6	2.50	1	1.03250
28	6	3.00	1	1.31250
29	7	0.60	1	-1.53500
30	7	1.00	1	-1.28500
31	7	1.40	1	-0.88500
32	7	2.00	1	-0.51500
33	7	2.50	1	-0.13500
34	8	0.60	0	-0.31670
35	8	2.40	0	0.00333
36	8	2.60	0	0.23330
37	8	3.40	0	0.14330
38	8	3.60	0	0.40330

Obs	id	t	sex	y
39	9	0.40	0	-0.00400
40	9	0.95	0	-0.09400
41	9	1.40	0	0.59600
42	9	1.95	0	0.69600
43	9	2.10	0	0.91600
44	9	2.60	0	1.20600
45	9	3.20	0	1.43600
46	10	0.14	0	-0.17000
47	10	0.24	0	-0.04000
48	10	0.34	0	0.26000
49	10	0.64	0	0.00000
50	10	1.09	0	0.18000
51	10	1.34	0	0.22000
52	10	1.59	0	0.48000
53	10	1.84	0	0.39000
54	10	2.09	0	0.54000
55	10	2.34	0	0.69000
56	11	0.30	1	0.37000
57	11	0.45	1	0.34000
58	11	0.75	1	0.30000
59	11	1.05	1	0.54000
60	11	1.30	1	0.46000
61	11	1.60	1	0.46000
62	11	1.75	1	0.64000
63	11	2.05	1	0.61000
64	11	2.45	1	0.67000
65	11	2.73	1	0.67000
66	11	3.05	1	0.66000
67	11	3.35	1	0.63000
68	12	0.25	1	0.12000
69	12	0.51	1	0.16000
70	12	0.61	1	0.24000
71	12	0.85	1	0.32000
72	12	1.15	1	0.67000
73	12	1.25	1	0.74000
74	12	1.35	1	0.99000
75	12	1.65	1	0.91000
76	12	1.85	1	1.04000

Obs	id	t	sex	y
77	12	2.15	1	0.96000
78	12	2.55	1	1.00000
79	12	3.10	1	0.94000
80	13	0.35	0	-0.29330
81	13	0.60	0	-0.56330
82	13	0.85	0	-0.88330
83	13	1.10	0	-0.86330
84	13	1.40	0	-0.63330
85	13	1.90	0	-0.79330
86	13	2.20	0	-0.45330
87	13	2.50	0	-0.09333
88	13	3.10	0	-0.02333
89	13	3.35	0	-0.12330
90	13	3.70	0	-0.16330
91	14	0.25	0	0.27220
92	14	0.50	0	0.33220
93	14	0.93	0	0.29220
94	15	0.65	1	-0.17250
95	15	0.85	1	-0.31250
96	15	1.55	1	0.76750
97	16	1.08	1	-0.00500
98	16	1.70	1	0.13500
99	16	2.08	1	0.30500
100	16	2.50	1	0.73500
101	16	2.66	1	0.84500
102	16	2.75	1	1.06500
103	16	2.90	1	1.05500
104	16	3.15	1	1.24500
105	16	3.60	1	1.62500
106	16	4.00	1	1.63500
107	16	4.45	1	1.89500
108	17	1.00	0	-0.31000
109	17	1.50	0	-0.50000
110	17	2.00	0	-0.60000
111	18	0.60	0	0.57000
112	18	1.60	0	0.79000
113	19	0.50	0	0.04000
114	19	1.00	0	-0.16000

Obs	id	t	sex	y
115	19	1.50	0	-0.21000
116	20	0.49	0	-0.41000
117	20	1.49	0	-0.06000
118	21	0.50	0	0.07000
119	21	1.00	0	0.18000
120	22	0.50	0	0.58000
121	22	1.00	0	0.81000
122	22	1.50	0	1.00000
123	22	2.00	0	1.15000
124	22	2.50	0	1.11000
125	23	0.50	0	0.26000
126	23	1.50	0	0.47000
127	24	0.50	0	0.26000
128	24	1.50	0	0.42000
129	25	0.50	0	0.23000
130	25	1.00	0	0.50000
131	26	0.50	0	-0.11000
132	26	1.00	0	0.58000
133	27	0.40	1	0.17000
134	27	0.85	1	0.36000
135	28	1.50	1	0.10000
136	28	2.00	1	0.10000

The Mixed Procedure

Model Information	
Data Set	WORK.INDAT
Dependent Variable	y
Covariance Structure	Unstructured
Subject Effect	id
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Dimensions	
Covariance Parameters	4
Columns in X	4
Columns in Z per Subject	2
Subjects	28
Max Obs per Subject	12

Number of Observations	
Number of Observations Read	136
Number of Observations Used	136
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	200.01624430	
1	2	46.17001886	0.03532186
2	1	41.64421400	0.01747033
3	1	39.40644048	0.00645686
4	1	38.60542669	0.00130806
5	1	38.45426928	0.00007377
6	1	38.44644479	0.00000028
7	1	38.44641604	0.00000000

Convergence criteria met.

The Mixed Procedure

Covariance Parameter Estimates						
Cov Parm	Subject	Ratio	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	9.0101	0.2551	0.07993	3.19	0.0007
UN(2,1)	id	-2.2526	-0.06379	0.03255	-1.96	0.0501
UN(2,2)	id	1.7965	0.05087	0.02189	2.32	0.0101
Residual		1.0000	0.02832	0.004402	6.43	<.0001

Fit Statistics	
-2 Res Log Likelihood	38.4
AIC (Smaller is Better)	46.4
AICC (Smaller is Better)	46.8
BIC (Smaller is Better)	51.8

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
3	161.57	<.0001

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-0.06080	0.1377	28.2	-0.44	0.6621
sex	-0.4348	0.2104	27.6	-2.07	0.0482
t	0.2298	0.07465	22.8	3.08	0.0054
sex*t	0.2228	0.1073	19.5	2.08	0.0513

Covariance Matrix for Fixed Effects					
Row	Effect	Col1	Col2	Col3	Col4
1	Intercept	0.01895	-0.01895	-0.00632	0.006317
2	sex	-0.01895	0.04426	0.006317	-0.01395
3	t	-0.00632	0.006317	0.005573	-0.00557
4	sex*t	0.006317	-0.01395	-0.00557	0.01152

The Mixed Procedure

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
sex	1	27.6	4.27	0.0482
t	1	22.8	9.48	0.0054
sex*t	1	19.5	4.31	0.0513

Coefficients for pred1

Effect	Subject	Row1
Intercept		1
sex		1
t		1.5
sex*t		1.5
Intercept	1	1
t	1	1.5
Intercept	2	
t	2	
Intercept	3	
t	3	
Intercept	4	
t	4	
Intercept	5	
t	5	
Intercept	6	
t	6	
Intercept	7	
t	7	
Intercept	8	
t	8	
Intercept	9	
t	9	
Intercept	10	
t	10	
Intercept	11	
t	11	
Intercept	12	
t	12	
Intercept	13	

The Mixed Procedure

Coefficients for pred1		
Effect	**Subject**	**Row1**
t	13	
Intercept	14	
t	14	
Intercept	15	
t	15	
Intercept	16	
t	16	
Intercept	17	
t	17	
Intercept	18	
t	18	
Intercept	19	
t	19	
Intercept	20	
t	20	
Intercept	21	
t	21	
Intercept	22	
t	22	
Intercept	23	
t	23	
Intercept	24	
t	24	
Intercept	25	
t	25	
Intercept	26	
t	26	
Intercept	27	
t	27	
Intercept	28	
t	28	

The Mixed Procedure

Coefficients for pred2		
Effect	**Subject**	**Row1**
Intercept		1
sex		1
t		3.5
sex*t		3.5
Intercept	1	1
t	1	3.5
Intercept	2	
t	2	
Intercept	3	
t	3	
Intercept	4	
t	4	
Intercept	5	
t	5	
Intercept	6	
t	6	
Intercept	7	
t	7	
Intercept	8	
t	8	
Intercept	9	
t	9	
Intercept	10	
t	10	
Intercept	11	
t	11	
Intercept	12	
t	12	
Intercept	13	
t	13	
Intercept	14	
t	14	
Intercept	15	
t	15	
Intercept	16	

The Mixed Procedure

Coefficients for pred2		
Effect	**Subject**	**Row1**
t	16	
Intercept	17	
t	17	
Intercept	18	
t	18	
Intercept	19	
t	19	
Intercept	20	
t	20	
Intercept	21	
t	21	
Intercept	22	
t	22	
Intercept	23	
t	23	
Intercept	24	
t	24	
Intercept	25	
t	25	
Intercept	26	
t	26	
Intercept	27	
t	27	
Intercept	28	
t	28	

Estimates					
Label	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
pred1	0.2040	0.08764	95.9	2.33	0.0220
pred2	1.0763	0.2896	96	3.72	0.0003

Obs	id	t	sex	y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
1	1	0.50	1	-0.34000	-0.23214	0.12588	107.791	0.05	-0.48165	0.01738	-0.10786
2	1	1.00	1	0.00000	-0.01407	0.08893	95.098	0.05	-0.19062	0.16247	0.01407
3	1	1.50	1	0.50000	0.20399	0.08764	95.892	0.05	0.03002	0.37796	0.29601
4	1	2.06	1	0.25000	0.44823	0.12876	106.230	0.05	0.19295	0.70350	-0.19823
5	2	0.30	1	0.31500	0.13964	0.11673	99.522	0.05	-0.09196	0.37124	0.17536
6	2	0.58	1	0.04500	0.15819	0.10008	96.860	0.05	-0.04044	0.35682	-0.11319
7	2	1.50	1	0.26500	0.21913	0.07462	88.034	0.05	0.07083	0.36743	0.04587
8	2	2.06	1	0.29500	0.25622	0.09181	95.239	0.05	0.07397	0.43848	0.03878
9	2	2.58	1	0.16500	0.29067	0.12189	100.186	0.05	0.04885	0.53248	-0.12567
10	3	0.80	1	0.07500	-0.03951	0.12097	107.423	0.05	-0.27931	0.20030	0.11451
11	3	1.20	1	0.21500	0.14223	0.09113	101.513	0.05	-0.03855	0.32300	0.07277
12	3	1.85	1	-0.02500	0.43754	0.07594	88.895	0.05	0.28665	0.58844	-0.46254
13	3	2.20	1	0.69500	0.59656	0.09150	98.988	0.05	0.41500	0.77812	0.09844
14	3	2.50	1	0.92500	0.73286	0.11323	104.702	0.05	0.50835	0.95738	0.19214
15	4	0.60	1	-0.57140	-0.58950	0.13736	103.416	0.05	-0.86191	-0.31709	0.01810
16	4	1.60	1	-0.22140	-0.17006	0.08691	92.551	0.05	-0.34266	0.00253	-0.05134
17	4	2.50	1	0.09857	0.20743	0.09535	91.952	0.05	0.01806	0.39680	-0.10886
18	4	3.10	1	0.53860	0.45910	0.12762	98.255	0.05	0.20585	0.71234	0.07950
19	5	0.35	0	0.07500	0.28726	0.10139	104.883	0.05	0.08623	0.48830	-0.21226
20	5	0.55	0	0.34500	0.33248	0.08647	100.369	0.05	0.16093	0.50402	0.01252
21	5	0.83	0	0.58500	0.39578	0.07170	90.115	0.05	0.25334	0.53822	0.18922
22	5	1.00	0	0.53500	0.43421	0.06801	86.835	0.05	0.29903	0.56940	0.10079
23	5	1.50	0	0.63500	0.54725	0.08406	102.405	0.05	0.38052	0.71399	0.08775
24	5	2.00	0	0.52500	0.66029	0.12414	106.702	0.05	0.41419	0.90640	-0.13529
25	6	1.10	1	-0.28750	-0.13987	0.13790	107.857	0.05	-0.41321	0.13348	-0.14763
26	6	2.13	1	0.68250	0.64861	0.08265	89.146	0.05	0.48438	0.81284	0.03389
27	6	2.50	1	1.03250	0.93185	0.09048	91.916	0.05	0.75215	1.11155	0.10065
28	6	3.00	1	1.31250	1.31461	0.12178	103.511	0.05	1.07310	1.55612	-0.00211
29	7	0.60	1	-1.53500	-1.43989	0.11621	105.026	0.05	-1.67032	-1.20947	-0.09511
30	7	1.00	1	-1.28500	-1.17163	0.08911	98.096	0.05	-1.34847	-0.99479	-0.11337
31	7	1.40	1	-0.88500	-0.90336	0.07475	88.283	0.05	-1.05191	-0.75481	0.01836
32	7	2.00	1	-0.51500	-0.50096	0.08989	97.859	0.05	-0.67935	-0.32257	-0.01404
33	7	2.50	1	-0.13500	-0.16563	0.12521	105.005	0.05	-0.41390	0.08264	0.03063
34	8	0.60	0	-0.31670	-0.27804	0.14566	102.382	0.05	-0.56694	0.01086	-0.03866
35	8	2.40	0	0.00333	0.08146	0.07435	88.568	0.05	-0.06628	0.22920	-0.07813
36	8	2.60	0	0.23330	0.12140	0.07456	86.988	0.05	-0.02679	0.26960	0.11190
37	8	3.40	0	0.14330	0.28118	0.09608	90.081	0.05	0.09031	0.47206	-0.13788
38	8	3.60	0	0.40330	0.32113	0.10505	91.635	0.05	0.11247	0.52979	0.08217

Obs	id	t	sex	y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
39	9	0.40	0	-0.00400	-0.09029	0.11362	97.852	0.05	-0.31577	0.13519	0.08629
40	9	0.95	0	-0.09400	0.21052	0.08499	94.610	0.05	0.04179	0.37925	-0.30452
41	9	1.40	0	0.59600	0.45664	0.06826	89.728	0.05	0.32102	0.59226	0.13936
42	9	1.95	0	0.69600	0.75745	0.06389	86.371	0.05	0.63045	0.88444	-0.06145
43	9	2.10	0	0.91600	0.83949	0.06642	87.121	0.05	0.70747	0.97150	0.07651
44	9	2.60	0	1.20600	1.11295	0.08379	91.637	0.05	0.94653	1.27937	0.09305
45	9	3.20	0	1.43600	1.44111	0.11496	95.513	0.05	1.21289	1.66932	-0.00511
46	10	0.14	0	-0.17000	-0.05791	0.08687	94.295	0.05	-0.23039	0.11456	-0.11209
47	10	0.24	0	-0.04000	-0.02747	0.08163	93.572	0.05	-0.18957	0.13463	-0.01253
48	10	0.34	0	0.26000	0.00297	0.07663	92.733	0.05	-0.14921	0.15516	0.25703
49	10	0.64	0	0.00000	0.09430	0.06356	89.498	0.05	-0.03198	0.22059	-0.09430
50	10	1.09	0	0.18000	0.23130	0.05306	85.309	0.05	0.12580	0.33680	-0.05130
51	10	1.34	0	0.22000	0.30741	0.05409	86.418	0.05	0.19988	0.41494	-0.08741
52	10	1.59	0	0.48000	0.38351	0.06001	89.562	0.05	0.26428	0.50274	0.09649
53	10	1.84	0	0.39000	0.45962	0.06958	92.665	0.05	0.32145	0.59780	-0.06962
54	10	2.09	0	0.54000	0.53573	0.08152	94.896	0.05	0.37389	0.69757	0.00427
55	10	2.34	0	0.69000	0.61184	0.09495	96.353	0.05	0.42338	0.80030	0.07816
56	11	0.30	1	0.37000	0.32842	0.08503	90.623	0.05	0.15951	0.49733	0.04158
57	11	0.45	1	0.34000	0.34885	0.07910	90.263	0.05	0.19171	0.50598	-0.00885
58	11	0.75	1	0.30000	0.38970	0.06806	89.288	0.05	0.25448	0.52493	-0.08970
59	11	1.05	1	0.54000	0.43056	0.05864	87.906	0.05	0.31402	0.54710	0.10944
60	11	1.30	1	0.46000	0.46461	0.05267	86.527	0.05	0.35991	0.56930	-0.00461
61	11	1.60	1	0.46000	0.50546	0.04870	85.138	0.05	0.40864	0.60229	-0.04546
62	11	1.75	1	0.64000	0.52589	0.04829	84.834	0.05	0.42987	0.62191	0.11411
63	11	2.05	1	0.61000	0.56675	0.05076	85.256	0.05	0.46583	0.66766	0.04325
64	11	2.45	1	0.67000	0.62122	0.05977	86.997	0.05	0.50242	0.74003	0.04878
65	11	2.73	1	0.67000	0.65936	0.06875	88.189	0.05	0.52273	0.79598	0.01064
66	11	3.05	1	0.66000	0.70294	0.08065	89.237	0.05	0.54269	0.86319	-0.04294
67	11	3.35	1	0.63000	0.74379	0.09283	89.941	0.05	0.55936	0.92822	-0.11379
68	12	0.25	1	0.12000	0.23814	0.08334	92.123	0.05	0.07263	0.40365	-0.11814
69	12	0.51	1	0.16000	0.33192	0.07173	90.872	0.05	0.18943	0.47440	-0.17192
70	12	0.61	1	0.24000	0.36798	0.06761	90.257	0.05	0.23368	0.50229	-0.12798
71	12	0.85	1	0.32000	0.45455	0.05882	88.464	0.05	0.33767	0.57142	-0.13455
72	12	1.15	1	0.67000	0.56275	0.05100	85.994	0.05	0.46137	0.66413	0.10725
73	12	1.25	1	0.74000	0.59882	0.04945	85.362	0.05	0.50050	0.69714	0.14118
74	12	1.35	1	0.99000	0.63489	0.04854	84.963	0.05	0.53838	0.73140	0.35511
75	12	1.65	1	0.91000	0.74309	0.04980	85.551	0.05	0.64409	0.84209	0.16691
76	12	1.85	1	1.04000	0.81523	0.05376	87.049	0.05	0.70838	0.92208	0.22477

Obs	id	t	sex	y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
77	12	2.15	1	0.96000	0.92343	0.06321	89.496	0.05	0.79784	1.04903	0.03657
78	12	2.55	1	1.00000	1.06771	0.07992	91.814	0.05	0.90897	1.22645	-0.06771
79	12	3.10	1	0.94000	1.26608	0.10671	93.461	0.05	1.05418	1.47798	-0.32608
80	13	0.35	0	-0.29330	-0.71907	0.08643	90.210	0.05	-0.89077	-0.54737	0.42577
81	13	0.60	0	-0.56330	-0.67329	0.07753	89.736	0.05	-0.82733	-0.51925	0.10999
82	13	0.85	0	-0.88330	-0.62751	0.06934	89.084	0.05	-0.76529	-0.48973	-0.25579
83	13	1.10	0	-0.86330	-0.58173	0.06213	88.218	0.05	-0.70519	-0.45826	-0.28157
84	13	1.40	0	-0.63330	-0.52679	0.05531	86.927	0.05	-0.63672	-0.41686	-0.10651
85	13	1.90	0	-0.79330	-0.43523	0.05039	85.097	0.05	-0.53543	-0.33504	-0.35807
86	13	2.20	0	-0.45330	-0.38030	0.05213	85.006	0.05	-0.48395	-0.27665	-0.07300
87	13	2.50	0	-0.09333	-0.32536	0.05712	85.659	0.05	-0.43892	-0.21180	0.23203
88	13	3.10	0	-0.02333	-0.21549	0.07386	87.468	0.05	-0.36229	-0.06870	0.19216
89	13	3.35	0	-0.12330	-0.16971	0.08248	88.078	0.05	-0.33362	-0.00581	0.04641
90	13	3.70	0	-0.16330	-0.10562	0.09548	88.743	0.05	-0.29535	0.08411	-0.05768
91	14	0.25	0	0.27220	0.24575	0.11457	105.447	0.05	0.01859	0.47292	0.02645
92	14	0.50	0	0.33220	0.28154	0.09639	91.230	0.05	0.09009	0.47300	0.05066
93	14	0.93	0	0.29220	0.34310	0.11400	89.757	0.05	0.11662	0.56958	-0.05090
94	15	0.65	1	-0.17250	-0.19738	0.11513	107.626	0.05	-0.42559	0.03083	0.02488
95	15	0.85	1	-0.31250	-0.04405	0.09970	97.637	0.05	-0.24192	0.15382	-0.26845
96	15	1.55	1	0.76750	0.49262	0.13083	86.073	0.05	0.23255	0.75269	0.27488
97	16	1.08	1	-0.00500	-0.10584	0.10165	96.687	0.05	-0.30760	0.09591	0.10084
98	16	1.70	1	0.13500	0.27650	0.07548	94.798	0.05	0.12664	0.42635	-0.14150
99	16	2.08	1	0.30500	0.51083	0.06224	92.213	0.05	0.38721	0.63445	-0.20583
100	16	2.50	1	0.73500	0.76984	0.05253	87.723	0.05	0.66545	0.87423	-0.03484
101	16	2.66	1	0.84500	0.86851	0.05082	86.044	0.05	0.76748	0.96953	-0.02351
102	16	2.75	1	1.06500	0.92401	0.05043	85.290	0.05	0.82373	1.02428	0.14099
103	16	2.90	1	1.05500	1.01651	0.05075	84.496	0.05	0.91560	1.11741	0.03849
104	16	3.15	1	1.24500	1.17068	0.05381	84.574	0.05	1.06369	1.27767	0.07432
105	16	3.60	1	1.62500	1.44818	0.06569	87.266	0.05	1.31762	1.57874	0.17682
106	16	4.00	1	1.63500	1.69485	0.08055	89.789	0.05	1.53483	1.85487	-0.05985
107	16	4.45	1	1.89500	1.97236	0.09978	91.791	0.05	1.77418	2.17053	-0.07736
108	17	1.00	0	-0.31000	-0.41450	0.12514	102.431	0.05	-0.66271	-0.16629	0.10450
109	17	1.50	0	-0.50000	-0.43485	0.09482	90.924	0.05	-0.62321	-0.24649	-0.06515
110	17	2.00	0	-0.60000	-0.45521	0.12791	99.266	0.05	-0.70899	-0.20142	-0.14479
111	18	0.60	0	0.57000	0.53638	0.14323	107.122	0.05	0.25244	0.82031	0.03362
112	18	1.60	0	0.79000	0.75098	0.14126	102.046	0.05	0.47079	1.03118	0.03902
113	19	0.50	0	0.04000	-0.10048	0.12784	102.945	0.05	-0.35403	0.15306	0.14048
114	19	1.00	0	-0.16000	-0.09294	0.09482	90.697	0.05	-0.28130	0.09542	-0.06706

Obs	id	t	sex	y	Pred	StdErrPred	DF	Alpha	Lower	Upper	Resid
115	19	1.50	0	-0.21000	-0.08541	0.12527	95.023	0.05	-0.33411	0.16329	-0.12459
116	20	0.49	0	-0.41000	-0.35686	0.14385	107.180	0.05	-0.64203	-0.07169	-0.05314
117	20	1.49	0	-0.06000	-0.05586	0.14037	101.725	0.05	-0.33429	0.22257	-0.00414
118	21	0.50	0	0.07000	0.06743	0.12901	107.876	0.05	-0.18828	0.32315	0.00257
119	21	1.00	0	0.18000	0.18080	0.12311	108.264	0.05	-0.06323	0.42482	-0.00080
120	22	0.50	0	0.58000	0.63128	0.12154	104.999	0.05	0.39029	0.87228	-0.05128
121	22	1.00	0	0.81000	0.77071	0.08821	97.276	0.05	0.59565	0.94578	0.03929
122	22	1.50	0	1.00000	0.91014	0.07415	87.779	0.05	0.76278	1.05751	0.08986
123	22	2.00	0	1.15000	1.04957	0.08902	96.723	0.05	0.87289	1.22626	0.10043
124	22	2.50	0	1.11000	1.18900	0.12272	103.969	0.05	0.94565	1.43236	-0.07900
125	23	0.50	0	0.26000	0.24388	0.14380	107.177	0.05	-0.04119	0.52894	0.01612
126	23	1.50	0	0.47000	0.45743	0.14045	101.750	0.05	0.17884	0.73603	0.01257
127	24	0.50	0	0.26000	0.23245	0.14380	107.177	0.05	-0.05261	0.51752	0.02755
128	24	1.50	0	0.42000	0.42345	0.14045	101.750	0.05	0.14486	0.70205	-0.00345
129	25	0.50	0	0.23000	0.28072	0.12901	107.876	0.05	0.02501	0.53644	-0.05072
130	25	1.00	0	0.50000	0.40931	0.12311	108.264	0.05	0.16529	0.65334	0.09069
131	26	0.50	0	-0.11000	0.11455	0.12901	107.876	0.05	-0.14116	0.37027	-0.22455
132	26	1.00	0	0.58000	0.32290	0.12311	108.264	0.05	0.07888	0.56693	0.25710
133	27	0.40	1	0.17000	0.14564	0.12773	107.503	0.05	-0.10754	0.39883	0.02436
134	27	0.85	1	0.36000	0.31920	0.12087	109.496	0.05	0.07966	0.55874	0.04080
135	28	1.50	1	0.10000	0.02262	0.12328	107.464	0.05	-0.22176	0.26701	0.07738
136	28	2.00	1	0.10000	0.19942	0.12911	106.014	0.05	-0.05656	0.45540	-0.09942

Example – Fay-Herriot Model for Sample Surveys

$$Y_i = \mu + b_i + \epsilon_i$$

$$b_i \sim N(0, \tau^2),$$

$$\epsilon_i \sim N(0, \sigma_i^2), \quad \sigma_i^2 \text{ known}$$

$$\hat{\mu} = \sum_{i=1}^{n} c_i Y_i / \sum_{i=1}^{n} c_i, \quad c_i = \tau^2/(\tau^2 + \sigma_i^2)$$

$$\hat{b}_i = c_i(Y_i - \hat{\mu})$$

$$\hat{Y}_i = \hat{\mu} + \hat{b}_i = (1 - c_i)\hat{\mu} + c_i Y_i$$

Fay-Herriot Model – Continued

```
data fh_example;
input area y sig2;
sig = sqrt(sig2);
cards;
[SNIP]

proc mixed;
model y = / solution ddfm=bw;
random intercept sig  / subject=area type=vc solution;
parms (1) (1) (1e-6) / hold=2,3; run;

data newdat; set fh_example;
muhat = -1.4257;
tau2hat = 0.02865;
wt = tau2hat/(tau2hat+sig2);
bhat = wt*(y-muhat);
ypred = (1-wt)*muhat + wt*y;
proc print; run;
```

The Mixed Procedure

Model Information	
Data Set	WORK.FH_EXAMPLE
Dependent Variable	y
Covariance Structure	Variance Components
Subject Effect	area
Estimation Method	REML
Residual Variance Method	Parameter
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Dimensions	
Covariance Parameters	3
Columns in X	1
Columns in Z per Subject	2
Subjects	22
Max Obs per Subject	1

Number of Observations	
Number of Observations Read	22
Number of Observations Used	22
Number of Observations Not Used	0

Parameter Search

CovP1	CovP2	CovP3	Res Log Like	-2 Res Log Like
1.0000	1.0000	1E-6	-22.5604	45.1209

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
1	4	15.68964193	.
2	1	14.23081332	0.02773574
3	1	13.83348999	0.00304341
4	1	13.79294980	0.00004770
5	1	13.79235186	0.00000001
6	1	13.79235169	0.00000000

Convergence criteria met.

The Mixed Procedure

Covariance Parameter Estimates		
Cov Parm	**Subject**	**Estimate**
Intercept	area	0.02865
sig	area	1.0000
Residual		1E-6

Fit Statistics	
-2 Res Log Likelihood	13.8
AIC (Smaller is Better)	15.8
AICC (Smaller is Better)	16.0
BIC (Smaller is Better)	16.9

PARMS Model Likelihood Ratio Test		
DF	**Chi-Square**	**Pr > ChiSq**
1	31.33	<.0001

Solution for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-1.4257	0.06238	21	-22.86	<.0001

Solution for Random Effects						
Effect	**Subject**	**Estimate**	**Std Err Pred**	**DF**	**t Value**	**Pr > |t|**
Intercept	1	0.1148	0.1523	21	0.75	0.4593
sig	1	1.3768	0.4655	21	2.96	0.0075
Intercept	2	-0.08043	0.1513	21	-0.53	0.6006
sig	2	-0.9311	0.4784	21	-1.95	0.0651
Intercept	3	0.01229	0.1508	21	0.08	0.9358
sig	3	0.1396	0.4853	21	0.29	0.7764
Intercept	4	0.1555	0.1508	21	1.03	0.3142
sig	4	1.7669	0.4853	21	3.64	0.0015
Intercept	5	0.03509	0.1468	21	0.24	0.8134
sig	5	0.3529	0.5315	21	0.66	0.5140
Intercept	6	-0.06729	0.1468	21	-0.46	0.6514
sig	6	-0.6767	0.5315	21	-1.27	0.2169
Intercept	7	-0.09372	0.1468	21	-0.64	0.5301

The Mixed Procedure

Solution for Random Effects						
Effect	Subject	Estimate	Std Err Pred	DF	t Value	Pr > |t|
sig	7	-0.9425	0.5315	21	-1.77	0.0907
Intercept	8	0.04373	0.1452	21	0.30	0.7663
sig	8	0.4208	0.5484	21	0.77	0.4514
Intercept	9	0.01937	0.1452	21	0.13	0.8952
sig	9	0.1864	0.5484	21	0.34	0.7373
Intercept	10	0.02780	0.1435	21	0.19	0.8482
sig	10	0.2549	0.5670	21	0.45	0.6577
Intercept	11	0.1243	0.1426	21	0.87	0.3931
sig	11	1.1149	0.5756	21	1.94	0.0663
Intercept	12	0.08255	0.1426	21	0.58	0.5688
sig	12	0.7403	0.5756	21	1.29	0.2124
Intercept	13	0.1402	0.1373	21	1.02	0.3188
sig	13	1.1051	0.6253	21	1.77	0.0917
Intercept	14	-0.1281	0.1373	21	-0.93	0.3612
sig	14	-1.0101	0.6253	21	-1.62	0.1211
Intercept	15	-0.2262	0.1360	21	-1.66	0.1110
sig	15	-1.7300	0.6368	21	-2.72	0.0129
Intercept	16	0.01829	0.1335	21	0.14	0.8923
sig	16	0.1324	0.6576	21	0.20	0.8424
Intercept	17	0.04705	0.1318	21	0.36	0.7247
sig	17	0.3285	0.6711	21	0.49	0.6296
Intercept	18	-0.01343	0.1306	21	-0.10	0.9190
sig	18	-0.09140	0.6805	21	-0.13	0.8944
Intercept	19	0.05239	0.1272	21	0.41	0.6845
sig	19	0.3322	0.7060	21	0.47	0.6428
Intercept	20	-0.08748	0.1272	21	-0.69	0.4991
sig	20	-0.5547	0.7060	21	-0.79	0.4408
Intercept	21	-0.07042	0.1231	21	-0.57	0.5734
sig	21	-0.4113	0.7346	21	-0.56	0.5815
Intercept	22	-0.1062	0.1192	21	-0.89	0.3829
sig	22	-0.5745	0.7602	21	-0.76	0.4582

Obs	area	y	x	sig2	sig	muhat	tau2hat	wt	bhat	ypred
1	1	-0.838	0.112	0.118	0.34351	-1.4257	0.02865	0.19536	0.11481	-1.31089
2	2	-1.815	0.206	0.110	0.33166	-1.4257	0.02865	0.20664	-0.08044	-1.50614
3	3	-1.368	0.104	0.106	0.32558	-1.4257	0.02865	0.21277	0.01228	-1.41342
4	4	-0.695	0.168	0.106	0.32558	-1.4257	0.02865	0.21277	0.15547	-1.27023
5	6	-1.289	0.169	0.083	0.28810	-1.4257	0.02865	0.25661	0.03508	-1.39062
6	7	-1.688	0.211	0.083	0.28810	-1.4257	0.02865	0.25661	-0.06731	-1.49301
7	8	-1.791	0.195	0.083	0.28810	-1.4257	0.02865	0.25661	-0.09374	-1.51944
8	9	-1.266	0.221	0.076	0.27568	-1.4257	0.02865	0.27377	0.04372	-1.38198
9	10	-1.355	0.077	0.076	0.27568	-1.4257	0.02865	0.27377	0.01936	-1.40634
10	11	-1.331	0.195	0.069	0.26268	-1.4257	0.02865	0.29339	0.02778	-1.39792
11	12	-1.015	0.185	0.066	0.25690	-1.4257	0.02865	0.30269	0.12432	-1.30138
12	13	-1.153	0.202	0.066	0.25690	-1.4257	0.02865	0.30269	0.08254	-1.34316
13	14	-1.036	0.108	0.051	0.22583	-1.4257	0.02865	0.35970	0.14017	-1.28553
14	15	-1.782	0.204	0.051	0.22583	-1.4257	0.02865	0.35970	-0.12816	-1.55386
15	16	-2.031	0.072	0.048	0.21909	-1.4257	0.02865	0.37378	-0.22625	-1.65195
16	17	-1.380	0.142	0.043	0.20736	-1.4257	0.02865	0.39986	0.01827	-1.40743
17	18	-1.313	0.136	0.040	0.20000	-1.4257	0.02865	0.41733	0.04703	-1.37867
18	19	-1.457	0.172	0.038	0.19494	-1.4257	0.02865	0.42986	-0.01345	-1.43915
19	20	-1.313	0.202	0.033	0.18166	-1.4257	0.02865	0.46472	0.05237	-1.37333
20	21	-1.614	0.087	0.033	0.18166	-1.4257	0.02865	0.46472	-0.08751	-1.51321
21	22	-1.565	0.177	0.028	0.16733	-1.4257	0.02865	0.50574	-0.07045	-1.49615
22	23	-1.621	0.072	0.024	0.15492	-1.4257	0.02865	0.54416	-0.10627	-1.53197

Unit 2
Nonlinear Mixed Models

Example 1 – Hospital Comparisons

Want to compare hospitals with respect to surgical mortality rates
Response variable: alive or dead
Have data on a number of hospitals
Want to analyze variation across hospitals
Want to account for patient-level factors ("case mix")

Example 2 – Toxicology Study

female mice assigned to different doses of a chemical
fetuses examined for malformations

Example 3 – Booster Seat Study (HU Public Health School)

11 neighborhoods, average of 50 children sampled per neighborhood

Response variable: child uses booster seat (yes/no)

interested in
- effect of neighborhood-level variables
- effect of child-level variables
- variation across neighborhoods

Logistic Regression With Random Effect

The Setup

i = cluster ($i = 1, \ldots, n$)
j = individual within cluster ($j = 1, \ldots, J_i$)
Y_{ij} = 0–1 binary response
X_{ijk} = value of explanatory variable k for individual ij ($k = 1, \ldots, p$)
We regard the X_{ijk}'s as fixed quantities.
Clusters are assumed independent.

The Concept

- Random effect b_i used to express within-cluster dependence
- Responses within each cluster assumed conditionally independent given the b_i's

Software

- SAS:
  - `PROC NLMIXED`
  - `PROC GLMMIX`
  - `NLMIXED` is more flexible
- SPSS: `GENLINMIXED`
  - Estimation is not by maximum likelihood, but instead by "penalized quasi-likelihood", which is an inferior method
- R: `glmer` function in `lme4` package

The SAS procedures are superior to the R functions in terms of the options offered.

Model Specification

$$p(\mathbf{x}_{ij}, \boldsymbol{\beta}, b_i) = \Pr(Y_{ij} = 1 | b_i) = \frac{\exp(\boldsymbol{\beta}^T \mathbf{x}_{ij} + b_i)}{1 + \exp(\boldsymbol{\beta}^T \mathbf{x}_{ij} + b_i)}$$

$$\begin{aligned} f_{\mathbf{Y}_i|b_i}(\mathbf{y}_i|b_i) &= \Pr(Y_{ij} = y_{ij}, j = 1, \ldots, J_i|b_i) \\ &= \prod_{j=1}^{J_i} p(\mathbf{x}_{ij}, \boldsymbol{\beta}, b_i)^{y_{ij}} (1 - p(\mathbf{x}_{ij}, \boldsymbol{\beta}, b_i))^{1-y_{ij}} \end{aligned}$$

$$f_{b_i}(b_i) = \sigma_b^{-1} \phi(b_i/\sigma_b) \quad (\text{i.e. } b_i \sim N(0, \sigma_b^2))$$

Leads to

$$\begin{aligned} f_{\mathbf{Y}_i}(\mathbf{y}_i) &= \Pr(Y_{ij} = y_{ij}, j = 1, \ldots, J_i) \\ &= \int_{-\infty}^{\infty} \left[\prod_{j=1}^{J_i} p(\mathbf{x}_{ij}, \boldsymbol{\beta}, b_i)^{y_{ij}} (1 - p(\mathbf{x}_{ij}, \boldsymbol{\beta}, b_i))^{1-y_{ij}} \right] \sigma_b^{-1} \phi(b_i/\sigma_b) db_i \\ &= \int_{-\infty}^{\infty} \left[\prod_{j=1}^{J_i} p(\mathbf{x}_{ij}, \boldsymbol{\beta}, \sigma_b \xi_i)^{y_{ij}} (1 - p(\mathbf{x}_{ij}, \boldsymbol{\beta}, \sigma_b \xi_i))^{1-y_{ij}} \right] \phi(\xi_i) d\xi_i \end{aligned}$$

Computation of Integrals

- Laplace Approximation: valid if J_i is large for all i
- Adaptive Gaussian Quadrature (AGQ)
- Simulation (offered only by SAS and not by R)

Statistical Inference for Parameters

Estimation of $\boldsymbol{\phi} = (\boldsymbol{\beta}, \sigma_b^2)$: Maximum Likelihood (or REML)

Large Sample Distribution

$$\hat{\boldsymbol{\phi}} - \boldsymbol{\phi} \mathrel{\dot{\sim}} N(\mathbf{0}, \boldsymbol{\Omega}(\hat{\boldsymbol{\phi}}))$$

$$\boldsymbol{\Omega}(\boldsymbol{\phi}) = \text{inverse Fisher information matrix}$$

Types of Inference

- Inference for Individual Fixed-Effect Coefficients
- Testing a Single Linear Combination of Fixed Effects
- Testing Several Linear Combinations

Similar to LMM case, with similar syntax in `SAS PROC NLMIXED`

Also:

- Wald test for significance of σ_b^2
- Confidence interval for σ_b^2

"Estimation" of Random Effects

By Bayes' theorem, the conditional density of b_i given $\mathbf{Y}_i$ is given by

$$f_{b_i|\mathbf{Y}_i}(b_i|\mathbf{Y}_i) = \frac{f_{\mathbf{Y}_i|b_i}(\mathbf{Y}_i|b_i)f_{b_i}(b_i)}{f_{\mathbf{Y}_i}(\mathbf{Y}_i)}$$

Expressions for the quantities on the right side have been given previously.

In principle, we could "estimate" b_i using

$$\hat{b}_i = E[b_i|\mathbf{Y}_i] = \frac{\int b_i f_{\mathbf{Y}_i|b_i}(\mathbf{Y}_i|b_i)f_{b_i}(b_i)db_i}{f_{\mathbf{Y}_i}(\mathbf{Y}_i)}$$

and we could use $f_{b_i|\mathbf{Y}_i}(b_i|\mathbf{Y}_i)$ to compute an empirical Bayes 95% credible interval for b_i.

`SAS PROC NLMIXED` and the R package `lme4` use an approximate method, valid for large J_i.

Estimation of Random Effects – Continued

Let b_i^* be the value of b that maximizes $f_{b_i|\mathbf{Y}_i}(b|\mathbf{Y}_i)$ (posterior mode).

Define $\bar{g}_i(b) = J_i^{-1} \log(f_{b_i}(b) f_{\mathbf{Y}_i|b_i}(\mathbf{Y}_i|b))$

Finally, define $v_i = -\bar{g}_i''(b_i^*)^{-1}$

The approximation is then

$$b_i|\mathbf{Y}_i \dot{\sim} N(b_i^*, J_i^{-1} v_i)$$

Estimation of Random Effects – Continued

Now, $b_i^* = b_i^*(\boldsymbol{\psi})$. In practice, we have to substitute $\hat{\boldsymbol{\psi}}$ for $\boldsymbol{\psi}$.

`SAS PROC MIXED` offers a correction for this.

The R function `lmer4` does not.

Defining $\mathbf{c} = \partial b_i^* / \partial \boldsymbol{\psi}$, the correction is as follows:

$$\text{Cov}\left(\begin{bmatrix} \hat{\boldsymbol{\psi}} \\ b_i \end{bmatrix} \middle| \mathbf{Y}_i \right) = \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{\Omega}\mathbf{c} \\ \mathbf{c}^T\boldsymbol{\Omega} & J_i^{-1} v_i + \mathbf{c}^T \Omega \mathbf{c} \end{bmatrix}$$

Using this result and the delta method, we can compute credible intervals for functions of $\boldsymbol{\psi}$ and b_i, such as $p(\mathbf{x}, \boldsymbol{\beta}, b_i)$ for a given $\mathbf{x}$.

Example 1

Mortality rates in 12 hospitals performing cardiac surgery in babies

```
proc nlmixed;
parms beta=-2.5 sigb2=1;
bounds sigb2 >=0;
odds = exp(beta+bi);
p = odds/(1+odds);
model d ~ binomial(n,p);
random bi ~ normal(0,sigb2) subject=hospital out=odat1;
predict p out=odat2;
run;

proc print data=odat1;
proc print data=odat2;
run;
```

The NLMIXED Procedure

Specifications	
Data Set	WORK.INDAT3
Dependent Variable	d
Distribution for Dependent Variable	Binomial
Random Effects	bi
Distribution for Random Effects	Normal
Subject Variable	hospital
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	12
Observations Not Used	0
Total Observations	12
Subjects	12
Max Obs per Subject	1
Parameters	2
Quadrature Points	1

Parameters

beta	sigb2	NegLogLike
-2.5	1	43.1068224

Iteration History

Iter		Calls	NegLogLike	Diff	MaxGrad	Slope
1		11	40.1506943	2.956128	6.823605	-22.055
2		16	39.964399	0.186295	9.1092	-81.3191
3		23	39.086557	0.877842	7.066766	-3.73841
4		26	38.6612433	0.425314	0.240901	-1.92463
5		28	38.6608621	0.000381	0.170556	-0.00113
6		31	38.6607362	0.000126	0.014656	-0.0003
7		34	38.6607343	1.88E-6	0.001063	-4.05E-6
8		37	38.6607343	9.998E-9	3.532E-6	-2E-8

NOTE: GCONV convergence criterion satisfied.

The NLMIXED Procedure

Fit Statistics	
-2 Log Likelihood	77.3
AIC (smaller is better)	81.3
AICC (smaller is better)	82.7
BIC (smaller is better)	82.3

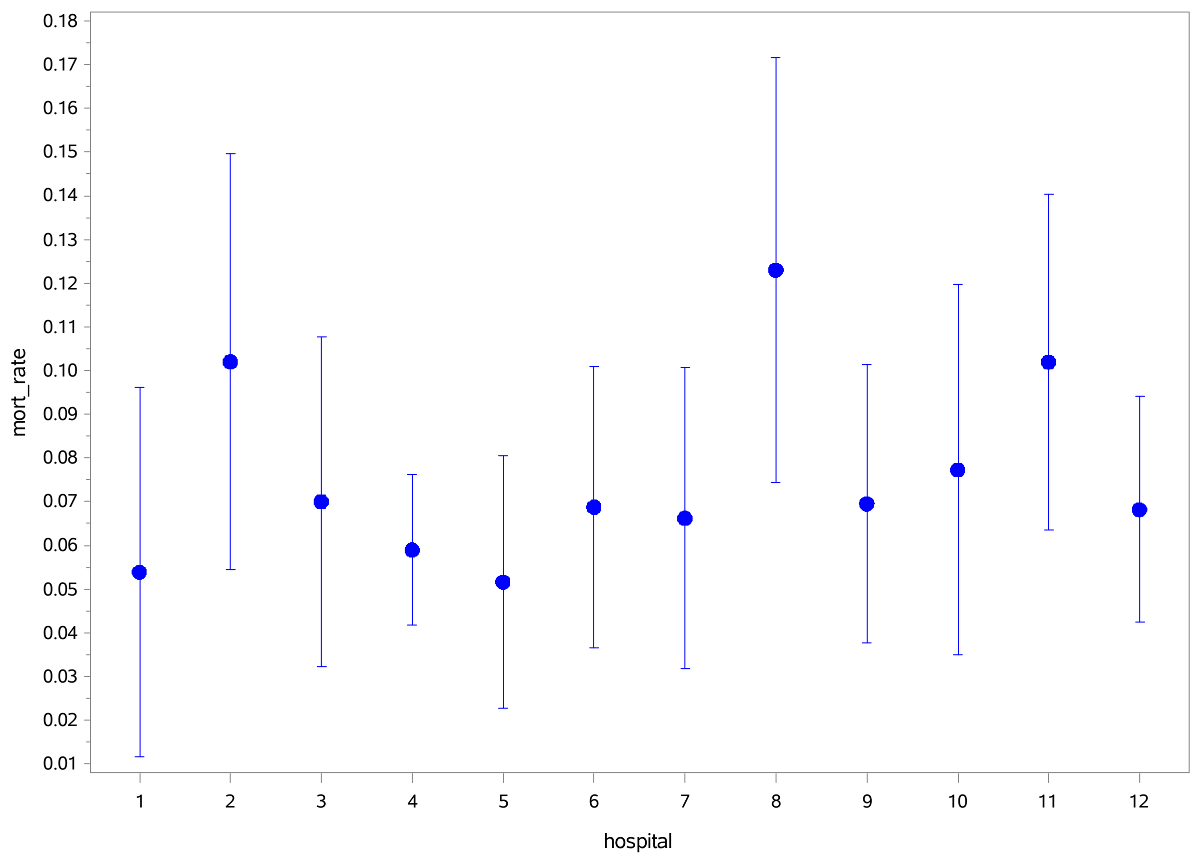
Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
beta	-2.5438	0.1347	11	-18.88	<.0001	0.05	-2.8404	-2.2473	-1.45E-7
sigb2	0.1272	0.08585	11	1.48	0.1664	0.05	-0.06172	0.3162	-3.53E-6

Obs	hospital	Effect	Estimate	StdErrPred	DF	tValue	Probt	Alpha	Lower	Upper
1	1	bi	-0.32212	0.35256	11	-0.91364	0.38049	0.05	-1.09810	0.45387
2	2	bi	0.36896	0.25538	11	1.44475	0.17641	0.05	-0.19313	0.93105
3	3	bi	-0.04242	0.26151	11	-0.16219	0.87409	0.05	-0.61800	0.53317
4	4	bi	-0.22589	0.17785	11	-1.27011	0.23025	0.05	-0.61734	0.16556
5	5	bi	-0.36742	0.26318	11	-1.39609	0.19022	0.05	-0.94667	0.21183
6	6	bi	-0.06153	0.23582	11	-0.26091	0.79898	0.05	-0.58057	0.45751
7	7	bi	-0.10216	0.25335	11	-0.40324	0.69450	0.05	-0.65979	0.45546
8	8	bi	0.57953	0.23632	11	2.45230	0.03211	0.05	0.05939	1.09968
9	9	bi	-0.05005	0.23243	11	-0.21531	0.83347	0.05	-0.56163	0.46154
10	10	bi	0.06403	0.26967	11	0.23745	0.81667	0.05	-0.52951	0.65758
11	11	bi	0.36835	0.21995	11	1.67472	0.12215	0.05	-0.11575	0.85245
12	12	bi	-0.07068	0.20563	11	-0.34373	0.73753	0.05	-0.52326	0.38190

Obs	hospital	n	d	Pred	StdErrPred	DF	tValue	Probt	Alpha	Lower	Upper
1	1	47	0	0.05386	0.019228	11	2.80127	0.017236	0.05	0.011542	0.09618
2	2	148	18	0.10203	0.021640	11	4.71477	0.000635	0.05	0.054399	0.14966
3	3	119	8	0.07003	0.017137	11	4.08645	0.001800	0.05	0.032311	0.10775
4	4	810	46	0.05898	0.007853	11	7.51058	0.000012	0.05	0.041697	0.07627
5	5	211	8	0.05160	0.013149	11	3.92433	0.002375	0.05	0.022660	0.08054
6	6	196	13	0.06879	0.014642	11	4.69846	0.000652	0.05	0.036567	0.10102
7	7	148	9	0.06624	0.015696	11	4.21979	0.001437	0.05	0.031688	0.10078
8	8	215	31	0.12300	0.022081	11	5.57040	0.000168	0.05	0.074401	0.17160
9	9	207	14	0.06953	0.014456	11	4.80986	0.000545	0.05	0.037715	0.10135
10	10	97	8	0.07729	0.019231	11	4.01877	0.002020	0.05	0.034958	0.11961
11	11	256	29	0.10197	0.017422	11	5.85299	0.000110	0.05	0.063627	0.14032
12	12	360	24	0.06821	0.011723	11	5.81824	0.000116	0.05	0.042407	0.09401

Estimated Hospital-Specific Mortality Rate with 95% Bayesian Credible Interval



Surgery Example in R

```
library(lme4)
indat = read.csv("surg.csv",header=T,sep=",")
result = glmer(cbind(D,N-D) ~ (1|HOSPITAL), family=binomial,
  nAGQ=20, data=indat)
summary(result)
re = ranef(result,condVar=T)
b = re$HOSPITAL
p = predict(result,type="response")
re.var = as.vector(attr(b,"postVar"))
re.sd = sqrt(re.var)
cbind(b,re.sd,p)
```

```
Generalized linear mixed model fit by maximum likelihood (Adaptive
  Gauss-Hermite Quadrature, nAGQ = 20) [glmerMod]
 Family: binomial  ( logit )
Formula: cbind(D, N - D) ~ (1 | HOSPITAL)
   Data: indat

     AIC      BIC   logLik deviance df.resid
    31.5     32.5    -13.8     27.5       10

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.6345 -0.2637 -0.1174  0.2921  0.9417

Random effects:
 Groups   Name        Variance Std.Dev.
 HOSPITAL (Intercept) 0.1279   0.3577
Number of obs: 12, groups:  HOSPITAL, 12

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.544      0.135  -18.84   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

	(Intercept)	re.sd	p
1	-0.32345970	0.3129944	0.05378772
2	0.36978892	0.2162551	0.10209384
3	-0.04247421	0.2534751	0.07001647
4	-0.22597469	0.1376661	0.05897052
5	-0.36822979	0.2348369	0.05155415
6	-0.06158514	0.2215679	0.06878226
7	-0.10235849	0.2427728	0.06621615
8	0.58043864	0.1795419	0.12308623
9	-0.05007175	0.2171526	0.06952338
10	0.06428085	0.2605212	0.07729493
11	0.36895330	0.1788346	0.10201726
12	-0.07069031	0.1804992	0.06820135

Example 2 – Toxicology Study

probability of defect as a function of dose
doses are 0, 0.025, 0.05, 0.1, 0.15

```
proc nlmixed;
parms bet0=-1.38 bet1=9.10 tau2=1;
bounds tau2 >=0;
lodds = bet0 + bet1*dose + bi;
odds = exp(lodds);
p = odds/(1+odds);
model defect ~ binomial(1,p);
random bi ~ normal(0,tau2) subject=dam_id;
predict p out=odat2;
run;
```

The NLMIXED Procedure

Specifications	
Data Set	WORK.NPT2
Dependent Variable	y
Distribution for Dependent Variable	Binomial
Random Effects	bi
Distribution for Random Effects	Normal
Subject Variable	DAM_ID
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	1827
Observations Not Used	0
Total Observations	1827
Subjects	125
Max Obs per Subject	37
Parameters	3
Quadrature Points	5

Parameters

bet0	bet1	tau2	NegLogLike
-1.38	9.1	1	929.030878

Iteration History

Iter		Calls	NegLogLike	Diff	MaxGrad	Slope
1		4	911.588833	17.44204	15.29279	-302.197
2		7	908.137894	3.450939	17.95067	-27.4676
3		11	903.519457	4.618437	2.259132	-4.84905
4		15	902.98579	0.533667	1.018553	-0.37769
5		17	902.941843	0.043947	0.290373	-0.06491
6		20	902.928328	0.013515	0.299979	-0.01852
7		24	902.8218	0.106527	1.484275	-0.00686
8		26	902.747875	0.073925	0.478848	-0.09353
9		29	902.729429	0.018446	0.034158	-0.03429
10		32	902.7292	0.000229	0.016071	-0.00042
11		35	902.729194	5.356E-6	0.000127	-9.31E-6
12		38	902.729194	7.419E-9	1.628E-6	-1.48E-8

The NLMIXED Procedure

NOTE: GCONV convergence criterion satisfied.

Fit Statistics	
-2 Log Likelihood	1805.5
AIC (smaller is better)	1811.5
AICC (smaller is better)	1811.5
BIC (smaller is better)	1819.9

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
bet0	-2.0077	0.2820	124	-7.12	<.0001	0.05	-2.5659	-1.4495	1.628E-6
bet1	6.5952	3.4556	124	1.91	0.0586	0.05	-0.2443	13.4347	-2.58E-8
tau2	3.0885	0.6145	124	5.03	<.0001	0.05	1.8722	4.3048	-3.54E-7

Toxicology Example – Analysis in R

```
library(lme4)
indat = read.csv("ntp.csv",header=T,sep=",")
result = glmer(defect ~ DOSE + (1|DAM_ID),
  family=binomial, nAGQ=20, data=indat)
summary(result)
```

Output

```
Generalized linear mixed model fit by maximum likelihood
(Adaptive Gauss-Hermite Quadrature, nAGQ = 20) ['glmerMod']
 Family: binomial  ( logit )
Formula: defect ~ DOSE + (1 | DAM_ID)
   Data: indat

     AIC      BIC   logLik deviance df.resid
  1811.1   1827.6   -902.5   1805.1     1824

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5519 -0.5103 -0.2249  0.5872  3.5621

Random effects:
 Groups Name        Variance Std.Dev.
 DAM_ID (Intercept) 3.132    1.77
Number of obs: 1827, groups:  DAM_ID, 125

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.0118     0.2839  -7.087 1.37e-12 ***
DOSE          6.5795     3.4777   1.892   0.0585 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Correlation of Fixed Effects:
     (Intr)
DOSE -0.734
```

Toxicology Example Extended

Now let the random effect variance depend on dose.

```
proc nlmixed;
parms bet0=-2 bet1=6.6 a1=1.13 a2=0;
lodds = bet0 + bet1*dose + bi;
odds = exp(lodds);
p = odds/(1+odds);
tau2 = exp(a1 + a2*dose);
model defect ~ binomial(1,p);
random bi ~ normal(0,tau2) subject=dam_id;
predict p out=odat2;
run;
```

In `lme4`, you can run this model only with integrals computed using Laplace approximation.

The NLMIXED Procedure

Specifications	
Data Set	WORK.ODAT2
Dependent Variable	y
Distribution for Dependent Variable	Binomial
Random Effects	bi
Distribution for Random Effects	Normal
Subject Variable	DAM_ID
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	1827
Observations Not Used	0
Total Observations	1827
Subjects	125
Max Obs per Subject	37
Parameters	4
Quadrature Points	7

Parameters

bet0	bet1	a1	a2	NegLogLike
-2	6.6	1.13	0	902.526618

Iteration History

Iter		Calls	NegLogLike	Diff	MaxGrad	Slope
1		4	902.4271	0.099518	2.815935	-5.30947
2		8	899.132534	3.294566	8.32031	-26.9486
3		11	896.302489	2.830045	1.523043	-575.851
4		13	896.250694	0.051795	0.345652	-0.10623
5		17	896.230605	0.020089	0.297446	-0.02289
6		20	896.225542	0.005062	0.020699	-0.00754
7		23	896.22539	0.000153	0.003255	-0.00028
8		26	896.225389	1.591E-7	0.000013	-3.2E-7

NOTE: GCONV convergence criterion satisfied.

The NLMIXED Procedure

Fit Statistics	
-2 Log Likelihood	1792.5
AIC (smaller is better)	1800.5
AICC (smaller is better)	1800.5
BIC (smaller is better)	1811.8

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
bet0	-1.9933	0.2211	124	-9.01	<.0001	0.05	-2.4310	-1.5557	-8.12E-7
bet1	7.6877	4.1166	124	1.87	0.0642	0.05	-0.4602	15.8356	1.522E-6
a1	0.1289	0.3513	124	0.37	0.7142	0.05	-0.5664	0.8243	0.000013
a2	14.3928	4.2770	124	3.37	0.0010	0.05	5.9275	22.8581	2.701E-6

Poisson Regression with Random Effect

$$
\begin{aligned}
Y_{ij}|b_i &\sim \text{Poi}(\lambda_{ij}) \\
\lambda_{ij} &= \exp(\mathbf{x}_{ij}\boldsymbol{\beta} + b_i)
\end{aligned}
$$

Theory similar to logistic regression case

Example – Epilepsy Study

- drug vs. placebo in epilepsy patients
- Y_{ij} = number of seizures patient i suffered in week j

Epilepsy Example Continued

```
data test;
set test;
time=studyweek/10;
proc nlmixed;
parms beta0=0 beta1=0 beta2=0 beta3=0 sig2=1;
eta = beta0 + beta1*trt + beta2*time + beta3*trt*time + b;
lambda = exp(eta);
model nseizw ~ poisson(lambda);
random b ~ normal(0,sig2) subject = id;
run;
```

The NLMIXED Procedure

Specifications	
Data Set	WORK.TEST
Dependent Variable	nseizw
Distribution for Dependent Variable	Poisson
Random Effects	b
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	1419
Observations Not Used	0
Total Observations	1419
Subjects	89
Max Obs per Subject	27
Parameters	5
Quadrature Points	1

Parameters

beta0	beta1	beta2	beta3	sig2	NegLogLike
0	0	0	0	1	3162.88423

Iteration History

Iter		Calls	NegLogLike	Diff	MaxGrad	Slope
1		5	3155.48892	7.395312	54.56929	-790.214
2		7	3138.58751	16.90141	12.69424	-309.143
3		10	3137.9469	0.640608	11.21897	-6.4826
4		14	3136.65714	1.289756	4.370378	-5.93578
5		16	3136.35491	0.302235	1.982679	-1.40835
6		19	3136.22996	0.124948	0.314055	-0.21557
7		22	3136.22814	0.001821	0.093257	-0.00365
8		25	3136.22804	0.000099	0.006952	-0.00012
9		28	3136.22804	2.255E-7	0.000061	-4.61E-7

NOTE: GCONV convergence criterion satisfied.

The NLMIXED Procedure

Fit Statistics	
-2 Log Likelihood	6272.5
AIC (smaller is better)	6282.5
AICC (smaller is better)	6282.5
BIC (smaller is better)	6294.9

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
beta0	0.8179	0.1675	88	4.88	<.0001	0.05	0.4850	1.1507	-0.00006
beta1	-0.1704	0.2384	88	-0.71	0.4768	0.05	-0.6442	0.3035	0.000027
beta2	-0.1429	0.04404	88	-3.24	0.0017	0.05	-0.2304	-0.05536	0.000022
beta3	0.02289	0.06167	88	0.37	0.7114	0.05	-0.09966	0.1455	-3.26E-6
sig2	1.1537	0.1838	88	6.28	<.0001	0.05	0.7885	1.5190	-0.00003

Poisson Regression with Multiple Random Effects

$$\begin{aligned}
&Y_{ij}|\mathbf{b}_i \sim \text{Poi}(\lambda_{ij}) \\
&\lambda_{ij} = \exp(\mathbf{x}_{ij}\boldsymbol{\beta} + \mathbf{z}_{ij}\mathbf{b}_i) \\
&\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}))
\end{aligned}$$

A similar extension can be carried out with logistic regression

Note: For models with multiple random effects, `lmer4` offers only Laplace approximation for evaluating the integrals, while `NLMIXED` offers AGQ and Simulation.

Epilepsy Example Revisited

```
proc nlmixed
parms beta0=0.82 beta1=-0.17 beta2=-0.14 beta3=0.023
  sig2b1=1.15 sig2b2=1 rho=0;
eta = beta0 + beta1*trt + beta2*time + beta3*trt*time
  + b0 + b1*time;
lambda = exp(eta);
model nseizw ~ poisson(lambda);
cov = sqrt(sig2b1*sig2b2)*rho;
```

```
random b0 b1 ~ normal([0, 0],[sig2b1, cov, sig2b2]) subject = id;
run;
```

The NLMIXED Procedure

Specifications	
Data Set	WORK.TEST
Dependent Variable	nseizw
Distribution for Dependent Variable	Poisson
Random Effects	b0 b1
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	1419
Observations Not Used	0
Total Observations	1419
Subjects	89
Max Obs per Subject	27
Parameters	7
Quadrature Points	1

Parameters

beta0	beta1	beta2	beta3	sig2b1	sig2b2	rho	NegLogLike
0.82	-0.17	-0.14	0.023	1.15	1	0	3056.54883

Iteration History

Iter		Calls	NegLogLike	Diff	MaxGrad	Slope
1		7	3040.33604	16.21279	24.98738	-186.863
2		10	3037.22671	3.109324	32.99192	-471.213
3		13	3036.20612	1.02059	19.26209	-20.9337
4		15	3034.37597	1.830152	13.37774	-31.2469
5		18	3034.02949	0.346478	7.818866	-5.92656
6		20	3033.77157	0.257928	5.664302	-2.88764
7		23	3033.64438	0.127183	1.402573	-2.24292
8		26	3033.60435	0.040031	1.853876	-0.27971
9		29	3033.58653	0.017818	0.272274	-0.04496
10		33	3033.58575	0.000788	0.089876	-0.00306
11		36	3033.58563	0.000118	0.0199	-0.0003
12		39	3033.58563	1.095E-6	0.001163	-2.08E-6

NOTE: GCONV convergence criterion satisfied.

The NLMIXED Procedure

Fit Statistics	
-2 Log Likelihood	6067.2
AIC (smaller is better)	6081.2
AICC (smaller is better)	6081.3
BIC (smaller is better)	6098.6

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
beta0	0.8945	0.1786	87	5.01	<.0001	0.05	0.5395	1.2494	-0.00048
beta1	-0.2443	0.2544	87	-0.96	0.3395	0.05	-0.7499	0.2612	-0.00044
beta2	-0.2715	0.09921	87	-2.74	0.0075	0.05	-0.4686	-0.07426	-0.0004
beta3	0.1067	0.1393	87	0.77	0.4458	0.05	-0.1702	0.3836	-0.00064
sig2b1	1.2707	0.2198	87	5.78	<.0001	0.05	0.8339	1.7076	0.000169
sig2b2	0.2373	0.05556	87	4.27	<.0001	0.05	0.1269	0.3477	-0.00116
rho	-0.3339	0.1312	87	-2.55	0.0127	0.05	-0.5946	-0.07317	0.000101

Unit 3
Generalized Estimating Equation Modelling

Random Effects Vs. GEE Modeling

Clustered data Y_{ij} with i=cluster and j=observation within cluster.

Random Effects Approach

$$\xi_{ij} = \mathbf{x}_{ij}^T\boldsymbol{\beta} + b_i$$
$$E[Y_{ij}|b_i] = g(\xi_{ij}) \text{ and } \mathrm{Var}(Y_{ij}|b_i) = \phi v(\xi_{ij})$$
$$Y_{ij}\text{'s conditionally independent given } b_i$$
$$b_i = \text{cluster-specific random effect} \sim N(0, \sigma_b^2)$$

GEE (Marginal Modeling) Approach

$$\eta_{ij} = \mathbf{x}_{ij}^T\boldsymbol{\beta}$$
$$E[Y_{ij}] = g(\eta_{ij}) \text{ and } \mathrm{Var}(Y_{ij}) = \phi v(\eta_{ij})$$

Analysis is done in a way that takes into account within-cluster dependence.

For logistic regression we have

- $g(\xi) = e^{\xi}/(1 + e^{\xi})$
- $v(\xi) = e^{\xi}/(1 + e^{\xi})^2$

Simple Example

Model: $E[Y_{ij}] = \beta X_{ij}$ and $\text{Var}(Y_{ij}) = \phi X_{ij}$

Estimation Criterion: $S = \sum_{i=1}^{n} \sum_{j=1}^{J_i} X_{ij}^{-1}(Y_{ij} - \beta X_{ij})^2$

Estimating Equation: $\sum_{i=1}^{n} \sum_{j=1}^{J_i}(Y_{ij} - \beta X_{ij}) = 0$

Estimator: $\hat{\beta} = \sum_{i=1}^{n} \sum_{j=1}^{J_i} Y_{ij} \,/\, \sum_{i=1}^{n} \sum_{j=1}^{J_i} X_{ij}$

This is the weighted least squares approach we'd use if the Y_{ij}'s were independent across i and j

Simple Example – Continued

Variance Estimation:

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^{n} \sum_{j=1}^{J_i} (Y_{ij} - \beta X_{ij})}{\sum_{i=1}^{n} \sum_{j=1}^{J_i} X_{ij}} \doteq \frac{n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{J_i} (Y_{ij} - \beta X_{ij})}{E\left[\sum_{j=1}^{J_i} X_{ij}\right]}$$

$$\text{Var}(\hat{\beta}) \doteq \frac{1}{n} \left(E\left[\sum_{j=1}^{J_i} X_{ij} \right] \right)^{-2} \text{Var} \left(\sum_{j=1}^{J_i} (Y_{ij} - \beta X_{ij}) \right)$$

$$= \frac{1}{n} \left(E\left[\sum_{j=1}^{J_i} X_{ij} \right] \right)^{-2} E\left[\left(\sum_{j=1}^{J_i} (Y_{ij} - \beta X_{ij}) \right)^2 \right]$$

$$\widehat{\text{Var}}(\hat{\beta}) = \frac{1}{n} \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{J_i} X_{ij} \right)^{-2} \left[\frac{1}{n} \sum_{i=1}^{n} \left(\sum_{j=1}^{J_i} (Y_{ij} - \hat{\beta} X_{ij}) \right)^2 \right]$$

Resembles the theory of stratified ratio estimators in Cochran's book.

General Theory of Independence Estimating Equations (IEE)

Estimation Criterion: $S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \sum_{j=1}^{J_i} v(\mathbf{x}_{ij}^T \boldsymbol{\beta})^{-1} (Y_{ij} - g(\mathbf{x}_{ij}^T \boldsymbol{\beta}))^2$

Momentarily ignoring the dependence of the weights $v(\mathbf{x}_{ij}^T \boldsymbol{\beta})^{-1}$ on $\boldsymbol{\beta}$ and differentiating, we get the estimating function

$$\mathbf{U}(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{U}_i(\boldsymbol{\beta})$$

where

$$\begin{aligned}
\mathbf{U}_i(\boldsymbol{\beta}) &= \mathbf{X}_i^T \mathbf{H}_i \mathbf{V}_i^{-1} (\mathbf{Y}_i - \mathbf{m}_i(\boldsymbol{\beta})) \\
\mathbf{m}_i(\boldsymbol{\beta}) &= [g(\mathbf{x}_{i1}^T \boldsymbol{\beta}), \ldots, g(\mathbf{x}_{iJ_i}^T \boldsymbol{\beta})]^T \\
\mathbf{V}_i &= \mathrm{diag}(v(\mathbf{x}_{i1}^T \boldsymbol{\beta}), \ldots, v(\mathbf{x}_{iJ_i}^T \boldsymbol{\beta})) \\
\mathbf{H}_i &= \mathrm{diag}(g'(\mathbf{x}_{i1}^T \boldsymbol{\beta}), \ldots, g'(\mathbf{x}_{iJ_i}^T \boldsymbol{\beta}))
\end{aligned}$$

The estimator $\hat{\boldsymbol{\beta}}$ is defined to be the solution to $\mathbf{U}(\boldsymbol{\beta}) = \mathbf{0}$.

For logistic and Poisson regression, the above is equivalent to maximum likelihood assuming all the observations are independent.

General IEE Theory – Continued

Denote by $\boldsymbol{\beta}^\circ$ the true value of $\boldsymbol{\beta}$

We have

$$\sqrt{n}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\circ) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{Q}}^{-1}\boldsymbol{\Omega}\bar{\mathbf{Q}}^{-1})$$
$$\bar{\mathbf{Q}} = E[\mathbf{X}_i^T\mathbf{H}_i\mathbf{V}_i^{-1}\mathbf{H}_i\mathbf{X}_i], \quad \boldsymbol{\Omega} = \mathrm{Cov}(\mathbf{U}_i(\boldsymbol{\beta}^\circ))$$

We can estimate $\bar{\mathbf{Q}}$ by $\tilde{\mathbf{Q}} = n^{-1}\sum_i \mathbf{X}_i^T\mathbf{H}_i\mathbf{V}_i^{-1}\mathbf{H}_i\mathbf{X}_i$ and $\boldsymbol{\Omega}$ by

$$\hat{\boldsymbol{\Omega}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{U}_i(\hat{\boldsymbol{\beta}})\mathbf{U}_i(\hat{\boldsymbol{\beta}})^T$$
$$= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i^T\mathbf{H}_i\mathbf{V}_i^{-1}(\mathbf{Y}_i - \mathbf{m}_i(\hat{\boldsymbol{\beta}}))(\mathbf{Y}_i - \mathbf{m}_i(\hat{\boldsymbol{\beta}}))^T\mathbf{V}_i^{-1}\mathbf{H}_i\mathbf{X}_i$$

The expression $\tilde{\mathbf{Q}}\hat{\boldsymbol{\Omega}}\tilde{\mathbf{Q}}$ is called the **robust sandwich variance estimator**.

• in the estimation scheme based on $S(\boldsymbol{\beta})$ and $\mathbf{U}(\boldsymbol{\beta})$, we act as if the Y_{ij}'s are independent
• the estimate of the covariance matrix of the estimator accounts for the within-cluster dependence

Generalized Estimating Equations (GEE)

Idea: try to improve the efficiency of the estimation by incorporating a guess of the correlation structure of $\mathbf{Y}_i$

Notation:
$\mathbf{V}_i$ = the "working" covariance matrix of $\mathbf{Y}_i$
$\mathbf{W}_i = \text{diag}(v(\mathbf{x}_{i1}^T\boldsymbol{\beta}), \ldots, v(\mathbf{x}_{iJ_i}^T\boldsymbol{\beta}))$ (the previous $\mathbf{V}_i$)
$\mathbf{R}_i(\boldsymbol{\theta})$ = the "working" correlation matrix of $\mathbf{Y}_i$
$\mathbf{V}_i(\boldsymbol{\theta}) = \mathbf{W}_i^{1/2}\mathbf{R}_i(\boldsymbol{\theta})\mathbf{W}_i^{1/2}$

Generalized Estimating Equations – Continued

The Estimation Scheme:

1. Start with an initial estimate of $\boldsymbol{\theta}$
2. Compute $\mathbf{V}_i(\boldsymbol{\theta})$
3. Compute $\hat{\boldsymbol{\beta}}$ as in IEE, but with the new $\mathbf{V}_i(\boldsymbol{\theta})$
4. Compute the residuals $e_{ij} = (Y_{ij} - g(\mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}))/v(\mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}})^{1/2}$
5. Compute an updated estimate of $\boldsymbol{\theta}$ based on the e_{ij}'s
6. Iterate

An alternate scheme is also possible, with $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ estimated simultaneously.

Remarks:

- it is NOT assumed that $\mathbf{R}_i(\boldsymbol{\theta})$ is the correct correlation matrix
- nonetheless, the resulting estimator of $\boldsymbol{\beta}$ is consistent
- moreover, the estimator of $\mathrm{Cov}(\sqrt{n}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\circ))$ is consistent

Correlation Structures for Binary Data

In the case of binary data, we have (writing $\mu_{ij} = g(\eta_{ij})$)

$$\text{Corr}(Y_{ij}, Y_{il}) = \frac{\Pr(Y_{ij} = 1, Y_{il} = 1) - \mu_{ij}\mu_{il}}{[\mu_{ij}(1 - \mu_{ij})\mu_{il}(1 - \mu_{il})]^{1/2}}$$

Now,

$$Pr(Y_{ij} = 1, Y_{il} = 1) \leq \min(\mu_{ij}, \mu_{il})$$
$$Pr(Y_{ij} = 1, Y_{il} = 1) = 1 - \Pr(\{Y_{ij} = 0\} \cup \{Y_{il} = 0\}) \geq \mu_{ij} + \mu_{il} - 1$$

These inequalities put awkward constraints on the possible values of $\text{Corr}(Y_{ij}, Y_{il})$.

`SAS PROC GENMOD` allows the working within-cluster dependence to be expressed instead in terms of the odds ratio

$$\text{OR}_{ijl} = \frac{\Pr(Y_{ij} = 1 | Y_{il} = 1) / \Pr(Y_{ij} = 0 | Y_{il} = 1)}{\Pr(Y_{ij} = 1 | Y_{il} = 0) / \Pr(Y_{ij} = 0 | Y_{il} = 0)}$$

Example

Patients in each of two centers are randomly assigned to groups receiving active treatment or a placebo.

During treatment, respiratory status (coded here as 0=poor, 1=good) is determined for each of four visits.

The variables center, treatment, sex, and baseline (baseline respiratory status) are classification variables with two levels.

The variable age (age at time of entry into the study) is a continuous variable.

Example – SAS Code

```
proc logistic descending;
class id_no;
model outcome=center2 active l_age baseline;
run;

proc genmod descending;
class id_no;
model outcome=center2 active l_age baseline / dist=bin;
repeated subject=id_no / type=ind;
run;

proc genmod descending;
class id_no;
model outcome=center2 active l_age baseline / dist=bin;
repeated subject=id_no / logor=exch;
```

LOGISTIC ANALYSIS

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.2101	0.9271	1.7036	0.1918
center2	1	0.6985	0.2345	8.8745	0.0029
active	1	1.2438	0.2296	29.3464	<.0001
l_age	1	-0.7789	0.2719	8.2024	0.0042
baseline	1	1.8175	0.2384	58.1302	<.0001

FIRST GEE ANALYSIS

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	1.2101	1.2048	-1.1514	3.5715	1.00	0.3152
center2	0.6985	0.3406	0.0309	1.3661	2.05	0.0403
active	1.2438	0.3255	0.6057	1.8818	3.82	0.0001
l_age	-0.7789	0.3634	-1.4911	-0.0666	-2.14	0.0321
baseline	1.8175	0.3389	1.1533	2.4817	5.36	<.0001

SECOND GEE ANALYSIS

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	1.1011	1.1836	-1.2188	3.4210	0.93	0.3522
center2	0.6935	0.3379	0.0312	1.3558	2.05	0.0401
active	1.2526	0.3217	0.6221	1.8831	3.89	<.0001
l_age	-0.7479	0.3558	-1.4452	-0.0506	-2.10	0.0355
baseline	1.8128	0.3359	1.1544	2.4712	5.40	<.0001
Alpha1	1.7060	0.2736	1.1698	2.2422	6.24	<.0001

Random Effects Vs. GEE Modeling Revisited

Random Effects Approach

$$E[Y_{ij}|b_i] = g(\mathbf{x}_{ij}^T\boldsymbol{\beta} + b_i), \quad \text{Var}(Y_{ij}|b_i) = \phi v(\mathbf{x}_{ij}^T\boldsymbol{\beta} + b_i)$$

Y_{ij}'s conditionally independent given b_i

b_i = cluster-specific random effect $\sim N(0, \sigma_b^2)$

$$E[Y_{ij}] = \int g(\mathbf{x}_{ij}^T\boldsymbol{\beta} + b_i)\sigma_b^{-1}\phi(b_i/\sigma_b)db_i$$

GEE (Marginal Modeling) Approach

$$E[Y_{ij}] = g(\mathbf{x}_{ij}^T\boldsymbol{\beta}), \quad \text{Var}(Y_{ij}) = \phi v(\mathbf{x}_{ij}^T\boldsymbol{\beta})$$

In general, the $\boldsymbol{\beta}$ being estimated in the GEE approach is not the same as the $\boldsymbol{\beta}$ being estimated in the random effects approach.

Exceptions:

- linear model case: $g(u) = u$
- multiplicative model case: $g(u) = e^u$

Random Effects Vs. GEE Modeling Revisited – Cont'd

Example:

- Binary Y
- Single Binary Covariate X (exposed/unexposed)
- $g(u) = e^u/(1 + e^u)$

Random Effects Model: β represents the odds ratio between a randomly selected unexposed individual and a randomly selected exposed individual *from the same cluster as the unexposed individual*

Marginal Model: β represents the odds ratio between a randomly selected unexposed individual and a randomly selected exposed individual *from the general population*