General Linear Model: Examples

Biostatistics 653

Applied Statistics III: Longitudinal Data Analysis

Dental Data

- Changes in the distance (measured in mm) from the center of the pituitary gland to the pterygomaxillary fissure are important in orthodontic therapy

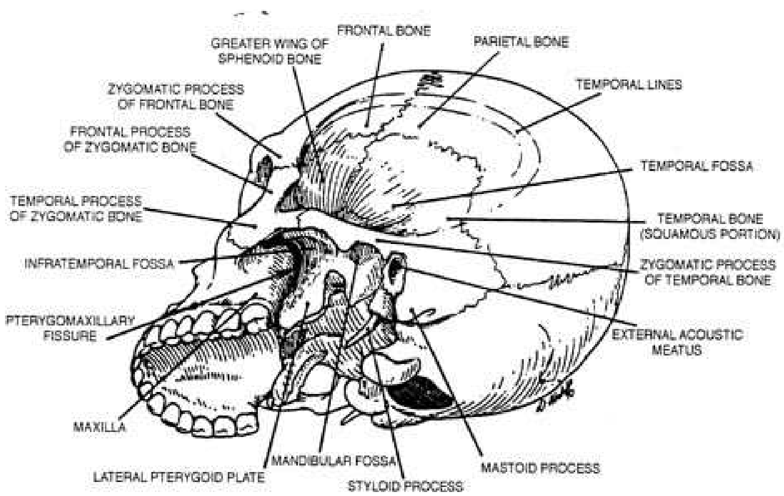
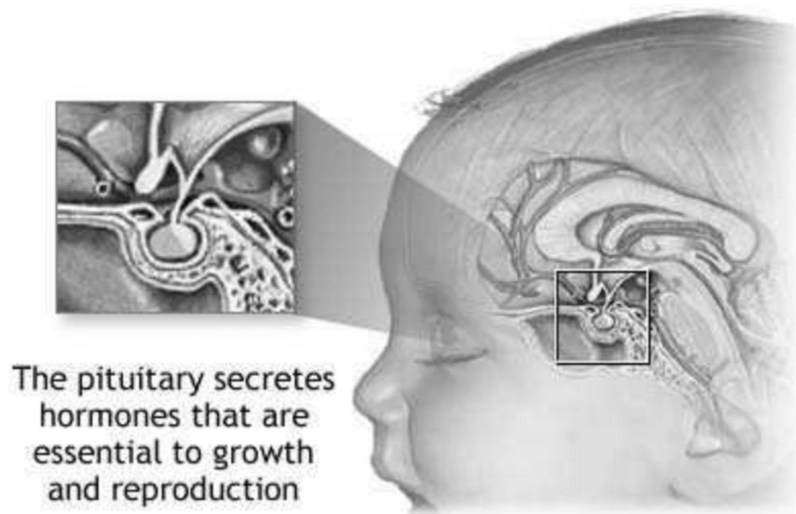
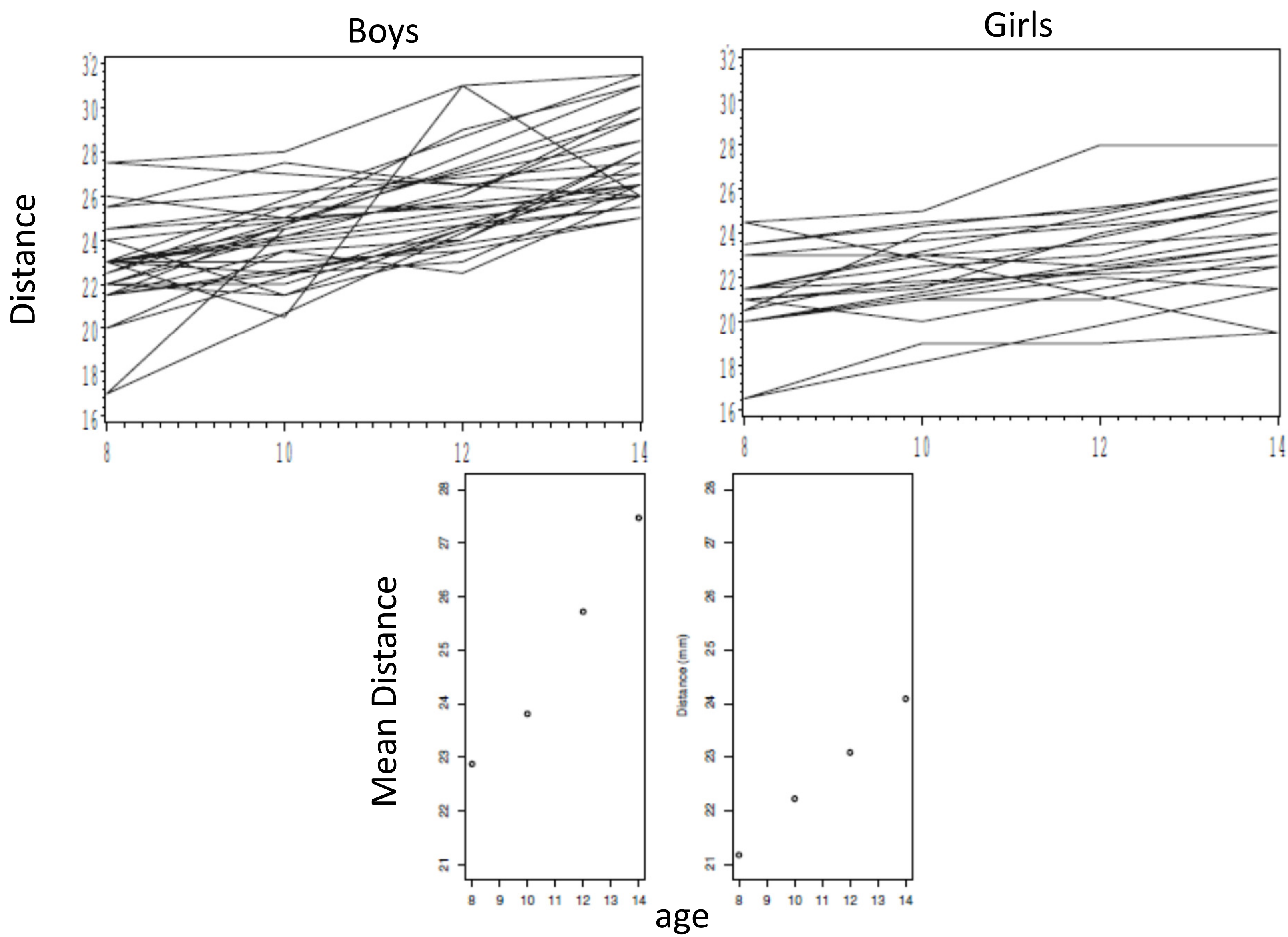


FIGURE 4



- Measure this distance at ages 8, 10, 12, and 14 in 27 children (16 boys and 11 girls

Dental Data



Dental Data

- For comparison purposes, we start by fitting an ordinary least squares regression to the data, obtaining intercept and slope estimates.

```
/* creating a unique subject identifier */
data proydataid; set proydata;
newid=_N_;
run;
/* convert data to format with one line per observation (multiple lines / subject) */

data proyuniv;
set proydataid;
time=8; dist=age8; output;
time=10; dist=age10; output;
time=12; dist=age12; output;
time=14; dist=age14; output;
drop age8 age10 age12 age14;
run;
```

```
/* FIRST, an OLS fit */
title 'OLS fit';
proc glm data=proyuniv;
class gender;
model dist=gender gender*time/noint solution;
run;
*************************************************************************
The GLM Procedure
Dependent Variable: dist
                      Sum of
Source      DF       Squares    Mean Square   F Value  P

Model        4   62715.99290    15678.99822   3078.04 <0.0001
Error      104     529.75710        5.09382

UncorrTot  108   63245.75000


R-Square     Coeff Var      Root MSE     dist Mean

0.422729      9.394892      2.256949      24.02315
```

Source	DF	Type I SS	Mean Square	F Value	P
GENDER	2	62468.52273	31234.26136	6131.80	<0.0001
time*GENDER	2	247.47017	123.73509	24.29	<0.0001

Source	DF	Type III SS	Mean Square	F Value	P
GENDER	2	1205.108665	602.554332	118.29	<0.0001
time*GENDER	2	247.470170	123.735085	24.29	<0.0001

Parameter		Estimate	Standard Error	t Value	Pr > |t|
GENDER	1	16.34062	1.416	11.54	<.0001
GENDER	2	17.37272	1.708	10.17	<.0001
time*GENDER	1	0.78437	0.126	6.22	<.0001
time*GENDER	2	0.47954	0.152	3.15	0.0021

Dental Data

- From this OLS t assuming independence of all measures, we obtain $\hat{\beta}_{0,B} = 16.34$, $\hat{\beta}_{1,B} = 0.78$, $\hat{\beta}_{0,G} = 17.37$, and $\hat{\beta}_{1,G} = 0.49$. The (incorrect) standard errors are given by 1.41, 0.13, 1.71, and 0.15, respectively.

- Now, we consider fitting a model that accounts for the repeated measures. Because we are unsure about the proper covariance structure, and it might be the case that covariances are different by gender, we start by fitting two separate, purely exploratory, models: one model using the unstructured covariance among boys, and one using the unstructured covariance among girls.

Dental Data

- The purpose of fitting these separate models is just to investigate the covariance structure by gender. (We note that this is a purely exploratory analysis used to help us choose the covariance structure. In the coming weeks, we will discuss more formal methods for selecting the covariance structure.) In this case, we dene a new Zi* that is a 4 by 2 matrix with a column of 1's and t_i as the other column and a corresponding $\gamma_G = (\gamma_{0G}, \gamma_{1G})^T$ and $\gamma_B = (\gamma_{0B}, \gamma_{1B})^T$.

Dental Data

- For girls, we fit the model

$$Y_i = Z_i^* \gamma_G + \epsilon_i, \epsilon \sim N(0, \Sigma_G)$$

- For boys, we fit the model

$$Y_i = Z_i^* \gamma_B + \epsilon_i, \epsilon \sim N(0, \Sigma_B)$$

- where Σ_G, Σ_B are unstructured. The SAS code follows in the next page. Note we use ML estimation and the model-based (default) variance estimator.

```
title 'Unstructured gender-specific covariance';
proc mixed data=proyuniv method=ml;  by gender;
class newid;
model dist=time/solution;
repeated/type=un subject=newid r rcorr;
run;
*************************************************************
GENDER=1
```

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Unstructured
Subject Effect	newid
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	16	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Dimensions

Covariance Parameters	10
Columns in X	2
Columns in Z	0
Subjects	16
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	64
Number of Observations Used	64
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	287.18814467	
1	2	264.37833982	0.00000565
2	1	264.37792193	0.00000000

Convergence criteria met.

GENDER=1

The Mixed Procedure

Estimated R Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	5.7813	2.0152	3.3585	1.4987
2	2.0152	4.4035	2.0982	2.6472
3	3.3585	2.0982	6.6064	3.0421
4	1.4987	2.6472	3.0421	4.0783

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.3994	0.5434	0.3086
2	0.3994	1.0000	0.3890	0.6247
3	0.5434	0.3890	1.0000	0.5861
4	0.3086	0.6247	0.5861	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	newid	5.7813
UN(2,1)	newid	2.0152
UN(2,2)	newid	4.4035
UN(3,1)	newid	3.3585
UN(3,2)	newid	2.0982
UN(3,3)	newid	6.6064
UN(4,1)	newid	1.4987
UN(4,2)	newid	2.6472
UN(4,3)	newid	3.0421
UN(4,4)	newid	4.0783

Fit Statistics

-2 Log Likelihood	264.4
AIC (smaller is better)	288.4
AICC (smaller is better)	294.5
BIC (smaller is better)	297.6

GENDER=1

The Mixed Procedure

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
9	22.81	0.0066

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	15.8282	1.1179	15	14.16	<.0001
time	0.8340	0.09274	15	8.99	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
time	1	15	80.86	<.0001

GENDER=2

The Mixed Procedure

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Unstructured
Subject Effect	newid
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	11	17 18 19 20 21 22 23 24 25 26 27

Dimensions

Covariance Parameters	10
Columns in X	2
Columns in Z	0
Subjects	11
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	44
Number of Observations Used	44
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	190.75564656	
1	2	130.64154698	0.00000000

Estimated R Matrix for newid 17

Row	Col1	Col2	Col3	Col4
1	4.1129	3.0512	3.9496	3.9689
2	3.0512	3.2894	3.6632	3.7080
3	3.9496	3.6632	5.0966	4.9788
4	3.9689	3.7080	4.9788	5.4076

Estimated R Correlation Matrix for newid 17

Row	Col1	Col2	Col3	Col4
1	1.0000	0.8295	0.8627	0.8416
2	0.8295	1.0000	0.8946	0.8792
3	0.8627	0.8946	1.0000	0.9484
4	0.8416	0.8792	0.9484	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	newid	4.1129
UN(2,1)	newid	3.0512
UN(2,2)	newid	3.2894
UN(3,1)	newid	3.9496
UN(3,2)	newid	3.6632
UN(3,3)	newid	5.0966
UN(4,1)	newid	3.9689
UN(4,2)	newid	3.7080
UN(4,3)	newid	4.9788
UN(4,4)	newid	5.4076

Fit Statistics

-2 Log Likelihood	130.6
AIC (smaller is better)	154.6
AICC (smaller is better)	164.7
BIC (smaller is better)	159.4

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
9	60.11	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	17.4220	0.6930	10	25.14	<.0001
time	0.4823	0.06144	10	7.85	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
time	1	10	61.62	<.0001

Dental Data

- What do you think about an appropriate covariance structure based on these two separate models?

- Now, we will try fitting a single model with separate covariances for each gender, assuming the compound symmetry structure for each. We will use a coding scheme that allows us to estimate group-specific slopes and intercepts directly.

- We fit the model using maximum likelihood. We use the options *r* and *rcorr* to print the covariance and correlation matrices, and we also obtain estimated mean distances for girls and boys at age 11.

Dental Data

- Thus, we fit the model
$$Y_i = X_i\beta + \epsilon_i, \epsilon_i \sim N(0, \Sigma_i)$$
- In this case, for girls:
$$\Sigma_i = \sigma_G^2 \begin{pmatrix} 1 & \rho_G & \rho_G & \rho_G \\ \rho_G & 1 & \rho_G & \rho_G \\ \rho_G & \rho_G & 1 & \rho_G \\ \rho_G & \rho_G & \rho_G & 1 \end{pmatrix}$$

for boys:
$$\Sigma_i = \sigma_B^2 \begin{pmatrix} 1 & \rho_B & \rho_B & \rho_B \\ \rho_B & 1 & \rho_B & \rho_B \\ \rho_B & \rho_B & 1 & \rho_B \\ \rho_B & \rho_B & \rho_B & 1 \end{pmatrix}$$

- SAS code for this model follows, using ML estimation and the model-based standard errors

```
title 'CS with Gender-Specific Covariance';
proc mixed data=proyuniv method=ml;
class newid gender;
model dist=gender gender*time/noint solution;
repeated/type=cs subject=newid r rcorr group=gender;
estimate 'girl at 11' gender 0 1 gender*time 0 11;
estimate 'boy at 11' gender 1 0 gender*time 11 0;
run;
```

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Compound Symmetry
Subject Effect	newid
Group Effect	GENDER
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
GENDER	2	1 2

Dimensions

Covariance Parameters	4
Columns in X	4
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	478.24175986	
1	1	408.81297228	0.00000000

Convergence criteria met.

Estimated R Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	5.2041	2.4463	2.4463	2.4463
2	2.4463	5.2041	2.4463	2.4463
3	2.4463	2.4463	5.2041	2.4463
4	2.4463	2.4463	2.4463	5.2041

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.4701	0.4701	0.4701
2	0.4701	1.0000	0.4701	0.4701
3	0.4701	0.4701	1.0000	0.4701
4	0.4701	0.4701	0.4701	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	newid	GENDER 1	2.7577
CS	newid	GENDER 1	2.4463
Variance	newid	GENDER 2	0.5900
CS	newid	GENDER 2	3.8804

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

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	69.43	<.0001

Solution for Fixed Effects

Effect	GENDER	Estimate	Standard Error	DF	t Value	Pr > |t|
GENDER	1	16.3406	1.1130	25	14.68	<.0001
GENDER	2	17.3727	0.8311	25	20.90	<.0001
time*GENDER	1	0.7844	0.09283	79	8.45	<.0001
time*GENDER	2	0.4795	0.05179	79	9.26	<.0001

Covariance Matrix for Fixed Effects

Row	Effect	GENDER	Col1	Col2	Col3	Col4
1	GENDER	1	1.2388		-0.09480	
2	GENDER	2		0.6907		-0.02950
3	time*GENDER	1	-0.09480		0.008618	
4	time*GENDER	2		-0.02950		0.002682

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GENDER	2	25	326.26	<.0001
time*GENDER	2	79	78.57	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
girl at 11	22.6477	0.6051	79	37.43	<.0001
boy at 11	24.9688	0.4427	79	56.40	<.0001

Dental Data

- Note that in this model, we obtain $\hat{\beta}_{0,B} = 16.34$, $\hat{\beta}_{1,B} = 0.78$, $\hat{\beta}_{0,G} = 17.37$, and $\hat{\beta}_{1,G} = 0.48$. The standard errors are given by 1.11, 0.09, 0.83, and 0.05, respectively.

- We also fit the model in an alternative parameterization in order to get tests and other quantities of interest more easily.

```
title 'CS by Gender Covariance: Tests';
proc mixed data=proyuniv method=ml;
class newid gender;
model dist=gender time gender*time/solution chisq covb;
repeated/type=cs subject=newid r rcorr group=gender;
run;
```

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Compound Symmetry
Subject Effect	newid
Group Effect	GENDER
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
GENDER	2	1 2

Dimensions

Covariance Parameters	4
Columns in X	6
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	478.24175986	
1	1	408.81297228	0.00000000

Convergence criteria met.

Estimated R Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	5.2041	2.4463	2.4463	2.4463
2	2.4463	5.2041	2.4463	2.4463
3	2.4463	2.4463	5.2041	2.4463
4	2.4463	2.4463	2.4463	5.2041

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.4701	0.4701	0.4701
2	0.4701	1.0000	0.4701	0.4701
3	0.4701	0.4701	1.0000	0.4701
4	0.4701	0.4701	0.4701	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	newid	GENDER 1	2.7577
CS	newid	GENDER 1	2.4463
Variance	newid	GENDER 2	0.5900
CS	newid	GENDER 2	3.8804

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

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	69.43	<.0001

Solution for Fixed Effects

Effect	GENDER	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		17.3727	0.8311	25	20.90	<.0001
GENDER	1	-1.0321	1.3890	25	-0.74	0.4644
GENDER	2	0	.	.	.	.
time		0.4795	0.05179	79	9.26	<.0001
time*GENDER	1	0.3048	0.1063	79	2.87	0.0053
time*GENDER	2	0	.	.	.	.

Covariance Matrix for Fixed Effects

Row	Effect	GENDER	Col1	Col2	Col3	Col4
1	Intercept		0.6907	-0.6907		-0.02950
2	GENDER	1	-0.6907	1.9294		0.02950
3	GENDER	2				
4	time		-0.02950	0.02950		0.002682
5	time*GENDER	1	0.02950	-0.1243		-0.00268
6	time*GENDER	2				

Covariance Matrix for Fixed Effects

Row	Col5	Col6
1	0.02950	
2	-0.1243	
3		
4	-0.00268	
5	0.01130	
6		

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
GENDER	1	25	0.55	0.55	0.4575	0.4644
time	1	79	141.37	141.37	<.0001	<.0001
time*GENDER	1	79	8.22	8.22	0.0041	0.0053

Dental Data

- In this framework, we can more easily conduct the test of parallelism and other tests of interest. In addition, we can obtain estimates and standard errors of average distances at age 11 by hand (or by programming).

- We can also conduct a LRT of the test of parallelism. First, we can get the likelihood of the reduced model. In this case, we fit the model using reference cell coding, using

$$Y_i = Z_i'\alpha' + \epsilon_i, \epsilon_i \sim N(0, \Sigma_i)$$

Dental Data

- Here, Z_i' contains the first three columns of Zi (eliminating the interaction term's column) so that

for girls: $\mathrm{Z}_i{}' = \sigma_G^2 \begin{pmatrix} 1 & 0 & 8 \\ 1 & 0 & 10 \\ 1 & 0 & 12 \\ 1 & 0 & 14 \end{pmatrix}$; for boys: $\mathrm{Z}_i{}' = \sigma_G^2 \begin{pmatrix} 1 & 1 & 8 \\ 1 & 1 & 10 \\ 1 & 1 & 12 \\ 1 & 1 & 14 \end{pmatrix}$

and $\alpha' = (\alpha_0', \alpha_1', \alpha_2')^T$.

```
/* fitting reduced model to illustrate lrt */
proc mixed data=proyuniv method=ml;
class newid gender;
model dist=gender time/solution;
repeated/type=cs subject=newid r rcorr group=gender;
run;
```

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Compound Symmetry
Subject Effect	newid
Group Effect	GENDER
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
GENDER	2	1 2

Dimensions

Covariance Parameters	4
Columns in X	4
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	480.68362161	
1	4	416.64891361	0.00045640
2	1	416.59716984	0.00000276
3	1	416.59686755	0.00000000

Convergence criteria met.

Estimated R Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	5.4838	2.3530	2.3530	2.3530
2	2.3530	5.4838	2.3530	2.3530
3	2.3530	2.3530	5.4838	2.3530
4	2.3530	2.3530	2.3530	5.4838

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.4291	0.4291	0.4291
2	0.4291	1.0000	0.4291	0.4291
3	0.4291	0.4291	1.0000	0.4291
4	0.4291	0.4291	0.4291	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	newid	GENDER 1	3.1308
CS	newid	GENDER 1	2.3530
Variance	newid	GENDER 2	0.6211
CS	newid	GENDER 2	3.8726

Fit Statistics

-2 Log Likelihood	416.6
AIC (smaller is better)	430.6
AICC (smaller is better)	431.7
BIC (smaller is better)	439.7

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	64.09	<.0001

Solution for Fixed Effects

Effect	GENDER	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		16.6218	0.7945	25	20.92	<.0001
GENDER	1	2.3210	0.7498	25	3.10	0.0048
GENDER	2	0	.	.	.	.
time		0.5478	0.04681	80	11.70	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GENDER	1	25	9.58	0.0048
time	1	80	136.97	<.0001

Dental Data

- Based on the output, we conduct the likelihood ratio test for the hypothesis of parallelism as $T_{LRT} = 2\log(\hat{L}_{FULL} - \hat{L}_{RED}) = 416.6 - 408.8 = 7.8$, which we compare to a chi-square critical value on 1 degree of freedom ($1.96^2 = 3.84$), leading us to reject the hypothesis of parallelism. We note that this test can be done only if the models were estimated using ML. While the REML covariance estimator typically has better performance than the ML estimator, it does not make sense to do mean model selection based on log-likelihoods from which we have "eliminated" the contribution of the mean, because this "elimination" depends on the mean parameterization of the model at hand. We will discuss selection of the mean model for longitudinal data in more detail soon.

Dental Data

- Now, suppose that we don't believe our covariance structure and would like to use the sandwich variance estimator.

- Note that now it is no longer necessary for our assumed structure of Σ_i to be correct in order to get proper standard errors for our estimates of β.

```
/* now, suppose we don't believe the covariance structure;*/
/* get sandwich est */
title 'CS by Gender Covariance: Sandwich Variance Est';
proc mixed data=proyuniv empirical method=ml;
class newid gender;
model dist=gender gender*time/noint solution covb;
repeated/type=cs subject=newid r rcorr group=gender;

estimate 'girl at 11' gender 0 1 gender*time 0 11;
estimate 'boy at 11' gender 1 0 gender*time 11 0;
run;
```

The Mixed Procedure

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Compound Symmetry
Subject Effect	newid
Group Effect	GENDER
Estimation Method	ML
Residual Variance Method	None
Fixed Effects SE Method	Empirical
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
GENDER	2	1 2

Dimensions

Covariance Parameters	4
Columns in X	4
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	478.24175986	
1	1	408.81297228	0.00000000

Convergence criteria met.

Estimated R Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	5.2041	2.4463	2.4463	2.4463
2	2.4463	5.2041	2.4463	2.4463
3	2.4463	2.4463	5.2041	2.4463
4	2.4463	2.4463	2.4463	5.2041

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.4701	0.4701	0.4701
2	0.4701	1.0000	0.4701	0.4701
3	0.4701	0.4701	1.0000	0.4701
4	0.4701	0.4701	0.4701	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	newid	GENDER 1	2.7577
CS	newid	GENDER 1	2.4463
Variance	newid	GENDER 2	0.5900
CS	newid	GENDER 2	3.8804

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

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	69.43	<.0001

Solution for Fixed Effects

Effect	GENDER	Estimate	Standard Error	DF	t Value	Pr > |t|
GENDER	1	16.3406	1.1715	25	13.95	<.0001
GENDER	2	17.3727	0.7252	25	23.96	<.0001
time*GENDER	1	0.7844	0.09835	79	7.98	<.0001
time*GENDER	2	0.4795	0.06313	79	7.60	<.0001

Empirical Covariance Matrix for Fixed Effects

Row	Effect	GENDER	Col1	Col2	Col3	Col4
1	GENDER	1	1.3724		-0.1067	
2	GENDER	2		0.5259		-0.02918
3	time*GENDER	1	-0.1067		0.009672	
4	time*GENDER	2		-0.02918		0.003986

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GENDER	2	25	384.22	<.0001
time*GENDER	2	79	60.65	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
girl at 11	22.6477	0.6051	79	37.43	<.0001
boy at 11	24.9688	0.4427	79	56.40	<.0001

Dental Data

- Note that in this model, we obtain $\hat{\beta}_{0,B} = 16.34$, $\hat{\beta}_{1,B} = 0.78$, $\hat{\beta}_{0,G} = 17.37$, and $\hat{\beta}_{1,G} = 0.48$. The standard errors are given by 1.17, 0.10, 0.73, and 0.06, respectively.

- We may also wish to compare the results to those we would have obtained using REML estimation. The REML estimators should have better properties than ML estimators in this small data set.

```
title 'CS by Gender Covariance: REML';
proc mixed data=proyuniv;
class newid gender;
model dist=gender gender*time/noint solution covb;
repeated/type=cs subject=newid r rcorr group=gender;
estimate 'girl at 11' gender 0 1 gender*time 0 11;
estimate 'boy at 11' gender 1 0 gender*time 11 0;
run;
```

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Compound Symmetry
Subject Effect	newid
Group Effect	GENDER
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
GENDER	2	1 2

Dimensions

Covariance Parameters	4
Columns in X	4
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	483.55911746	
1	1	414.66636550	0.00000000

Estimated R Matrix for newid 1

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

Estimated R Correlation Matrix for newid 1

Row	Col1	Col2	Col3	Col4
1	1.0000	0.4839	0.4839	0.4839
2	0.4839	1.0000	0.4839	0.4839
3	0.4839	0.4839	1.0000	0.4839
4	0.4839	0.4839	0.4839	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	newid	GENDER 1	2.8164
CS	newid	GENDER 1	2.6407
Variance	newid	GENDER 2	0.6085
CS	newid	GENDER 2	4.2786

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

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	68.89	<.0001

Solution for Fixed Effects

Effect	GENDER	Estimate	Standard Error	DF	t Value	Pr > |t|
GENDER	1	16.3406	1.1287	25	14.48	<.0001
GENDER	2	17.3727	0.8587	25	20.23	<.0001
time*GENDER	1	0.7844	0.09382	79	8.36	<.0001
time*GENDER	2	0.4795	0.05259	79	9.12	<.0001

Covariance Matrix for Fixed Effects

Row	Effect	GENDER	Col1	Col2	Col3	Col4
1	GENDER	1	1.2740		-0.09681	
2	GENDER	2		0.7374		-0.03042
3	time*GENDER	1	-0.09681		0.008801	
4	time*GENDER	2		-0.03042		0.002766

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GENDER	2	25	309.43	<.0001
time*GENDER	2	79	76.53	<.0001

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
girl at 11	22.6477	0.6347	79	35.68	<.0001
boy at 11	24.9688	0.4572	79	54.61	<.0001

Dental Data

- Note that in this model, we obtain $\hat{\beta}_{0,B} = 16.34$, $\hat{\beta}_{1,B} = 0.78$, $\hat{\beta}_{0,G} = 17.37$, and $\hat{\beta}_{1,G} = 0.48$. The standard errors are given by 1.13, 0.09, 0.86, and 0.05, respectively.

Summary of Model Fitting Results for Dental Data

Model	$\widehat{\beta}_{0G}$	$\widehat{\beta}_{0B}$	$\widehat{\beta}_{1G}$	$\widehat{\beta}_{1B}$
OLS	17.37	16.34	0.47	0.78
	(1.71)	(1.42)	(0.15)	(0.13)
Gender-Specific CS				
(a) ML, model-based SE	17.37	16.34	0.48	0.78
	(0.83)	(1.11)	(0.05)	(0.09)
(b) ML, sandwich SE	17.37	16.34	0.48	0.78
	(0.73)	(1.18)	(0.60)	(0.10)
(c) REML, model-based SE	17.37	16.34	0.48	0.78
	(0.86)	(1.13)	(0.05)	(0.09)