Analysis of 2-level repeated measures data

Contents

This example illustrates how multilevel modeling may be used to recognize explicitly the hierarchical structure of repeated measurement data.

Five models will be fitted and discussed:

- A variance decomposition model
- Modeling linear growth
- Modeling non-linear growth
- Introducing a covariate when modeling non-linear growth
- A model with complex variation at level-1 of the hierarchy

1. Description of the data

The data set used contains repeated measurements on 82 striped mice and was obtained from the Department of Zoology at the University of Pretoria, South Africa (see du Toit, 1979). A number of male and female mice were released in an outdoor enclosure with nest boxes and sufficient food and water. They were allowed to multiply freely. Occurrence of birth was recorded daily and newborn mice were weighed weekly, from the end of the second week after birth until physical maturity was reached. The data set consists of the weights of 42 male and 40 female mice. For male mice, 9 repeated weight measurements are available and for the female mice 8 repeated measurements.

The first 11 observations from this data set, contained in **mouse.lsf**, and the variable names to be used are shown below.

The response variable WEIGHT contains the weight measurements (in grams) for all mice at the different times of measurement. The explanatory variables which may be used are the time points at which measurements were made (TIME), the squared values of these time points (TIMESQ), and the gender of the mice (GENDER). It is also assumed that the growth of the mice during this period can be adequately described with a parabolic function.

A hierarchical level-2 structure is incorporated where the individual mice are the level-2 units. Unique numbers identifying the mice are contained in the variable IDEN2, which will be used as the level-2 ID for the analysis. The variable IDEN1 identifies the occasions on which measurements for a particular mouse were made and will be used as the level-1 ID. From the description of the data set as given above, it follows that there are 82 level-2 units, with either 8 or 9 measurements nested within each level-2 unit.

The variable CONSTANT consists of a column of 1s and may be used as alternative to the intercept term automatically included in the model.

2. Variance decomposition

The simplest multilevel model is equivalent to a one-way ANOVA with random effects. Although this model is not interesting in itself, it is useful as a preliminary step in a multilevel analysis as it provides important information about the outcome variability at each of the levels of the hierarchy. It may also function as a baseline with which more sophisticated models may be compared.

Let the subscript i denote the i-th level-2 unit, in this case the i-th mouse. The subscript j refers to the j-th weight measurement for the i-th mouse. Using this notation, the one-way ANOVA model can be written as:

$$y_{ij} = \beta_0 + u_{ij} + e_{ij}$$

where u_{ij} denotes the random component on level 2 of the model. It is assumed that u_{ij} has an expected value of 0 and a variance of $\Phi_{(2)}$. The variance $\Phi_{(2)}$ may be interpreted as the "between-group" variability. Likewise, it is assumed that e_{ij} is $N(0, \Phi_{(1)})$ distributed. Thus $\Phi_{(1)}$ may be interpreted as the "within-group" variability.

This model is also known as a fully unconditional model (Bryk & Raudenbush, 1992), as no predictors are specified at either level of the hierarchy.

To specify the model, open the data file **mouse.lsf** and then select the **Multilevel** option from the main menu bar. On the **Title and Options** dialog box, enter a title for this analysis and leave all other options default. Click **Next** to move to the next dialog box.

Indicate that IDEN2 will serve as level-2 ID for this analysis on the **Identification Variables** dialog box. As there are no weights in these data, click **Next** twice to move on to the **Select Response and Fixed Variables** dialog box.

For the fully unconditional model we select WEIGHT as response variable.
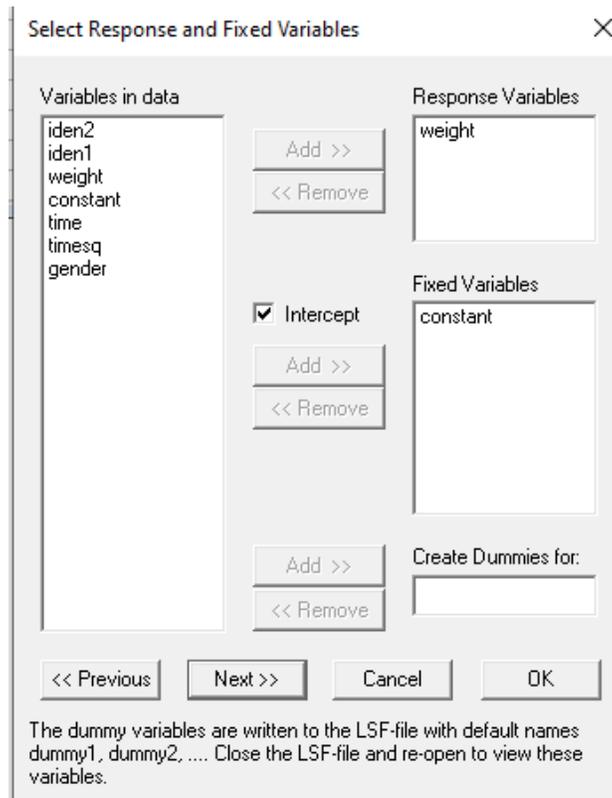
Note that the box for Intercept is checked on the dialog box shown below. If it is unclicked, the variable CONSTANT should be selected as random at both levels on the **Random Variables** dialog box.
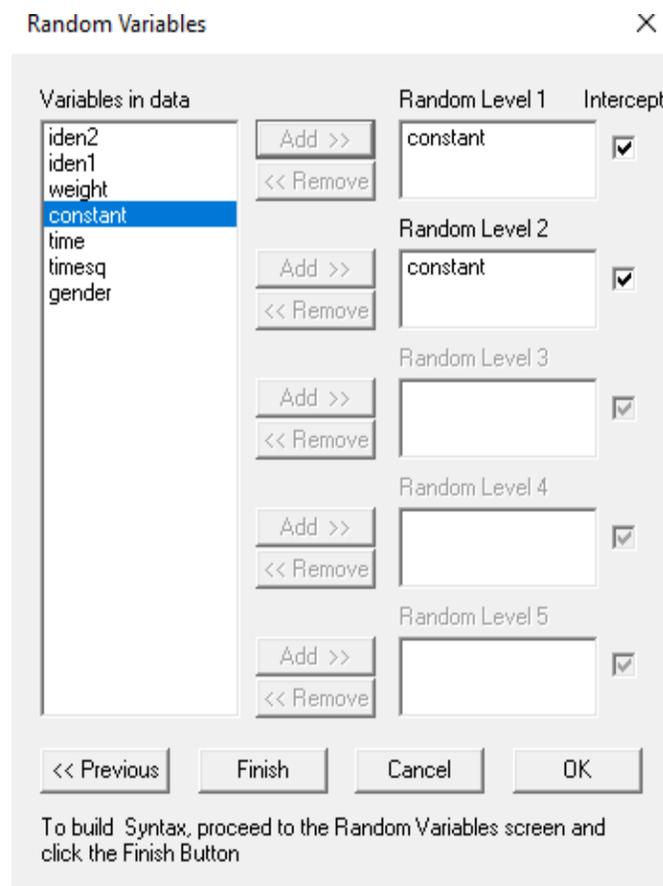
Click **Finish** to return to the main LISREL window where the completed syntax file **MOUSE.PRL** is now displayed.

```
L MOUSE.PRL                                                    [ — ][ □ ][ X ]

OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=STANDARD ;
TITLE=Mouse data: Variance decomposition;
SY='MOUSE.lsf';
ID2=iden2;
RESPONSE=weight;
FIXED=intcept;
RANDOM1=intcept;
RANDOM2=intcept;
```

With the exception of the use of the optional TITLE command, this input file is the most basic one which can be used for the analysis of a level-2 model. Note that in the OPTIONS command the default values of the options MAXITER, CONVERGE, and OUTPUT are used.

Run PRELIS with this input file. Convergence is achieved after 3 iterations, and the details of the last iteration, as given in the output file, are shown below.

```
NUMBER OF LEVEL 2 UNITS :        82
 NUMBER OF LEVEL 1 UNITS :       698


N2  :        1        2        3        4        5        6        7        8
N1  :        9        9        9        9        9        9        9        9


N2  :        9       10       11       12       13       14       15       16
N1  :        9        9        9        9        9        9        9        9


N2  :       17       18       19       20       21       22       23       24
N1  :        9        9        9        9        9        9        9        9


N2  :       25       26       27       28       29       30       31       32
N1  :        9        9        9        9        9        9        9        9


N2  :       33       34       35       36       37       38       39       40
N1  :        9        9        9        9        9        9        9        9


N2  :       41       42       43       44       45       46       47       48
N1  :        9        9        8        8        8        8        8        8


N2  :       49       50       51       52       53       54       55       56
N1  :        8        8        8        8        8        8        8        8


N2  :       57       58       59       60       61       62       63       64
N1  :        8        8        8        8        8        8        8        8


N2  :       65       66       67       68       69       70       71       72
N1  :        8        8        8        8        8        8        8        8


N2  :       73       74       75       76       77       78       79       80
N1  :        8        8        8        8        8        8        8        8


N2  :       81       82
N1  :        8        8
```

Mouse data: Variance decomposition

```
---------------------------------------
CONVERGENCE REACHED IN    3 ITERATIONS
---------------------------------------
```

Mouse data: Variance decomposition

ITERATION NUMBER   3

```
                    +----------------------+
                    | FIXED PART OF MODEL  |
                    +----------------------+
```

```
-------------------------------------------------------------------
  COEFFICIENTS            BETA-HAT      STD.ERR.     Z-VALUE      PR > |Z|
-------------------------------------------------------------------
  intcept                 28.63410      0.57021      50.21634     0.00000
```

```
                    +----------------------+
                    |   -2 LOG-LIKELIHOOD  |
                    +----------------------+
```

```
  DEVIANCE= -2*LOG(LIKELIHOOD) =    5425.49001592990
  NUMBER OF FREE PARAMETERS =              3
```

```
                    +----------------------+
                    | RANDOM PART OF MODEL |
                    +----------------------+
```

```
-------------------------------------------------------------------
  LEVEL 2                      TAU-HAT     STD.ERR.    Z-VALUE   PR > |Z|
-------------------------------------------------------------------
  intcept /intcept            11.32910     4.25185     2.66451   0.00771
```

```
-------------------------------------------------------------------
  LEVEL 1                      TAU-HAT     STD.ERR.    Z-VALUE   PR > |Z|
-------------------------------------------------------------------
  intcept /intcept           130.32083     7.42514    17.55130   0.00000
```

```
        LEVEL 2 COVARIANCE MATRIX

                intcept

  intcept       11.32910
```

```
       LEVEL 2 CORRELATION MATRIX

                intcept

  intcept       1.0000


          LEVEL 1 COVARIANCE MATRIX

                 intcept

  intcept       130.32083


          LEVEL 1 CORRELATION MATRIX

                intcept

  intcept       1.0000
```

In the first part of the abbreviated output file shown here, the data summary for the hierarchical structure is given. The first 42 level-2 units are the male mice and the last 40 the female mice.

From the random part of the output it can be seen that the variation over measurements (level 1) is large and overwhelms the variation between the mice (level 2). The so-called intraclass correlation can be calculated as:

$$\hat{\rho} = \frac{\hat{\Phi}_{(2)}}{\hat{\Phi}_{(2)} + \hat{\Phi}_{(1)}} = \frac{11.32910}{11.32910 + 130.32083} = 0.0799$$

indicating that about 8 percent of the variance in weight measurements is between mice. The value of $-2\ln L$ (likelihood function) at convergence is 5425.4900.

3. Modeling linear growth

The variance decomposition model may now be extended by including the variable TIME as a fixed effect in the model. The model thus becomes

$$y_{ij} = \beta_0 + \beta_1 TIME_{ij} + u_{ij} + e_{ij}$$

with the variable TIME used as predictor of the response measurements.

The only change to the input file (**mouse2.prl**) is in the FIXED command, which now includes the variable TIME.

```
L  MOUSE2.prl                                    ─  ▢  ✕

OPTIONS Maxiter=30 Conv=0.0001;
TITLE=Mouse data: linear growth model one;
SY=MOUSE.LSF;
ID2=iden2;
RESPONSE=weight;
FIXED=intcept time;
RANDOM1=intcept;
RANDOM2=intcept;
```

Details of the last iteration for this model are:

```
ITERATION NUMBER      3

                   +-----------------------+
                   | FIXED PART OF MODEL   |
                   +-----------------------+


-----------------------------------------------------------------------
  COEFFICIENTS           BETA-HAT       STD.ERR.      Z-VALUE      PR > |Z|
-----------------------------------------------------------------------
  intcept                 9.09586        0.60387      15.06258      0.00000
  time                    4.09218        0.06258      65.39108      0.00000


                   +-----------------------+
                   |   -2 LOG-LIKELIHOOD   |
                   +-----------------------+

  DEVIANCE= -2*LOG(LIKELIHOOD) =    4137.57876020825
  NUMBER OF FREE PARAMETERS =              4


                   +-----------------------+
                   | RANDOM PART OF MODEL  |
                   +-----------------------+


--------------------------------------------------------------------
  LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE    PR > |Z|
--------------------------------------------------------------------
  intcept /intcept              20.69397      3.53655      5.85146    0.00000


--------------------------------------------------------------------
  LEVEL 1                        TAU-HAT      STD.ERR.     Z-VALUE    PR > |Z|
--------------------------------------------------------------------
  intcept /intcept              16.46288      0.93806     17.54996    0.00000
```

```
        LEVEL 2 COVARIANCE MATRIX

                intcept

 intcept        20.69397



        LEVEL 2 CORRELATION MATRIX

                intcept

 intcept         1.0000



        LEVEL 1 COVARIANCE MATRIX

                intcept

 intcept        16.46288



        LEVEL 1 CORRELATION MATRIX

                intcept

 intcept         1.0000
```

Both the fixed effects are highly significant, indicating significant variation in the intercepts and effect of time of measurement on the response variable over the different mice. An expected increase of 4.0922 grams in weight is expected for each increase of a week in TIME.

The log-likelihood value for this model is 4137.5788, compared to the value of 5424.4900 for the fully unconditional model. This reduction in the log-likelihood value indicates considerable variation between mice in their linear growth rates and also that the model fitted here explains more of the variation in the data than the previous one.

The random variation on level 2 of the model is higher and that on level-1 lower than in the fully unconditional model. The amount of variation in weight between mice is now calculated as

$$\hat{\rho} = \frac{\hat{\Phi}_{(2)}}{\hat{\Phi}_{(2)} + \hat{\Phi}_{(1)}} = \frac{20.69379}{20.69379 + 16.46288} = 0.5569,$$

that is 5 percent.

From the reduction in the level-1 variance component it can e seen that the variable TIME accounts for a considerable part of the variance previously noted on this level.

It is expected that the linear growth rate may vary from mouse to mouse around its mean value, rather than be fixed. The random component on level-2 of the hierarchy is thus extended to include TIME as well (**mouse3.prl**).

```
L MOUSE3.prl                                    [ — ][ ▢ ][ ✕ ]
│OPTIONS ;
│TITLE=Mouse data: linear growth model two;
│SY=MOUSE.LSF;
│ID2=iden2;
│RESPONSE=weight;
│FIXED=intcept time;
│RANDOM1=intcept;
│RANDOM2=intcept time;
```

After running PRELIS on this input file to fit the model to the data, we find the output for this model in the **mouse3.out** file:

```
----------------------------------------
 CONVERGENCE REACHED IN    4 ITERATIONS
----------------------------------------


 Mouse data: linear growth model two

 ITERATION NUMBER    4


                    +-----------------------+
                    | FIXED PART OF MODEL   |
                    +-----------------------+


---------------------------------------------------------------------------
  COEFFICIENTS          BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
---------------------------------------------------------------------------
  intcept               9.20384       0.46702      19.70746       0.00000
  time                  4.05978       0.12358      32.85144       0.00000



                    +-----------------------+
                    |   -2 LOG-LIKELIHOOD   |
                    +-----------------------+

  DEVIANCE= -2*LOG(LIKELIHOOD) =    3873.66054782870
  NUMBER OF FREE PARAMETERS =               6
```

```
                    +----------------------+
                    | RANDOM PART OF MODEL  |
                    +----------------------+


--------------------------------------------------------------------
  LEVEL 2                      TAU-HAT      STD.ERR.    Z-VALUE   PR > |Z|
--------------------------------------------------------------------
  intcept /intcept             12.85320     2.80949     4.57492   0.00000
  time    /intcept             -1.64430     0.59236    -2.77584   0.00551
  time    /time                 1.07389     0.19582     5.48415   0.00000


--------------------------------------------------------------------
  LEVEL 1                      TAU-HAT      STD.ERR.    Z-VALUE   PR > |Z|
--------------------------------------------------------------------
  intcept /intcept              8.90088     0.54469    16.34126   0.00000


          LEVEL 2 COVARIANCE MATRIX

                   intcept          time

  intcept        12.85320
  time           -1.64430       1.07389



          LEVEL 2 CORRELATION MATRIX

                   intcept          time

  intcept          1.0000
  time            -0.4426       1.0000


          LEVEL 1 COVARIANCE MATRIX

                   intcept

  intcept         8.90088


          LEVEL 1 CORRELATION MATRIX

                   intcept

  intcept         1.0000
```

Once again a considerable reduction in the value of the function $-2\ln L$ is noted. The estimates for the fixed effects in the model stayed fairly constant. On level 2 of the model we see that all three elements of the covariance matrix of random parameters $\Phi_{(2)}$ are significant at a 5% level. The correlation between the intercept term and the TIME term is given as -0.4426. The level-1 or error variance is further reduced to 8.90088. It can thus be concluded that the inclusion of the variable TIME significantly reduced the variation between measurements, i.e., the level- units. The total variation on a particular level of the hierarchy may also be calculated. In this case, the total variance on level 2 is the variance of the sum of the two random coefficients associated with the intercept and TIME and may be written as:

$$Var\left(Intercept_{ij} + TIME_{ij}\Phi_{(2)TIME,TIME}\right) = \Phi_{(2)INTERCEPT,INTERCEPT} + 2\Phi_{(2)TIME,INTERCEPT}\left(TIME_{ij}\right) + \Phi_{(2)TIME,TIME}\left(TIME_{ij}\right)^2$$

$$= 12.85320 + 2(-1.64430)\left(TIME_{ij}\right) + 1.07389\left(TIME_{ij}\right)^2$$

We can thus write the total variation at level 2 as a quadratic function of the variable TIME. A graph of this total variance against the nine time points is given in the figure below. The increase in variance over time is to be expected with data of this nature. It could also be an indication that the assumption that a parabolic function can adequately describe this phase in the development of the mice may not be valid and that other functions should be considered.
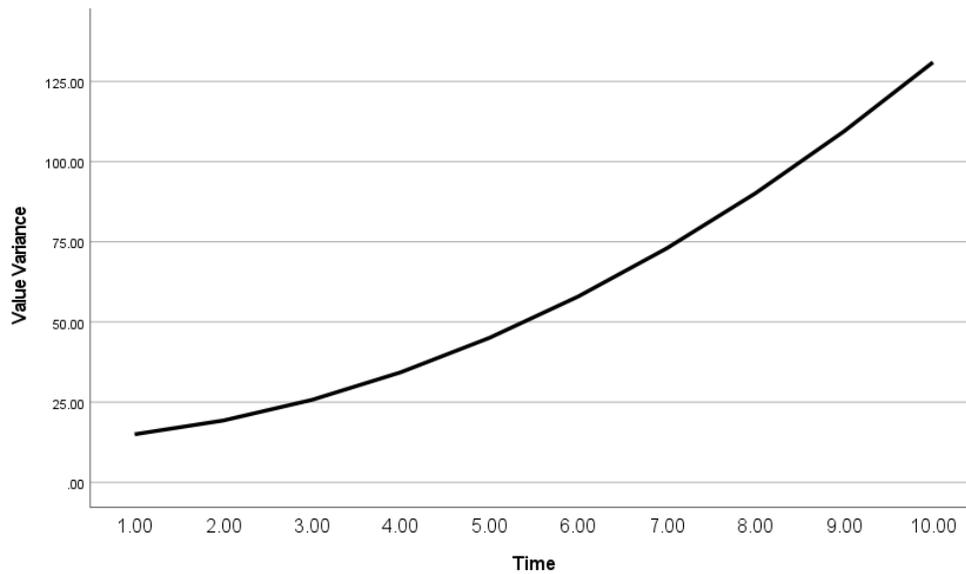
4. Modeling non-linear growth

In data of this nature, it is unlikely that the increase in weight measurement will be linear for all mice over the time period concerned. A nonlinear component may be introduced in the model discussed in the previous section by adding a quadratic term (the variable TIMESQ) to the model. The model previously given is thus extended to:

$$y_{ij} = \beta_0 + \beta_1 TIME_{ij} + \beta_2 TIMESQ_{ij} + u_{0j} + u_{1j}TIME_{ij} + u_{2j}TIMESQ_{ij} + e_{ij}$$

The addition of the variable TIMESQ in this case leads to the following changes in the FIXED and RANDOM2 commands contained in the input file:

```
FIXED=intcept time timesq;
RANDOM2=intcept time timesq;
```

In order to obtain the empirical Bayes residuals for the level-2 models and the fitted values for each observation, the option OUTPUT = ALL is added to the OPTIONS command. The complete input file **mouse4.prl** is now:

```
L MOUSE4.prl                                        ▯  ▫  ✕

OPTIONS OUTPUT=ALL ;
TITLE=Mouse data: non-linear growth model;
SY=MOUSE.LSF;
ID2=iden2;
RESPONSE=weight;
FIXED=intcept time timesq;
RANDOM1=intcept;
RANDOM2=intcept time timesq;
|
```

Convergence of the iterative procedure was reached in five iterations, producing the following output.

```
ITERATION NUMBER      5


                    +------------------------+
                    | FIXED PART OF MODEL    |
                    +------------------------+


------------------------------------------------------------------------
  COEFFICIENTS            BETA-HAT      STD.ERR.      Z-VALUE      PR > |Z|
------------------------------------------------------------------------
  intcept                  4.16213       0.45748       9.09788      0.00000
  time                     6.90560       0.30980      22.29056      0.00000
  timesq                  -0.29629       0.02960     -10.01009      0.00000


                    +------------------------+
                    |   -2 LOG-LIKELIHOOD    |
                    +------------------------+


  DEVIANCE= -2*LOG(LIKELIHOOD) =    3400.93248789783
  NUMBER OF FREE PARAMETERS =            10


                    +------------------------+
                    | RANDOM PART OF MODEL   |
                    +------------------------+


------------------------------------------------------------------------
  LEVEL 2                          TAU-HAT      STD.ERR.      Z-VALUE    PR > |Z|
------------------------------------------------------------------------
  intcept /intcept               11.72906       2.70296       4.33933    0.00001
  time    /intcept               -5.86552       1.58228      -3.70699    0.00021
  time    /time                   6.59372       1.23130       5.35509    0.00000
  timesq  /intcept                0.38927       0.14067       2.76733    0.00565
  timesq  /time                  -0.55909       0.11275      -4.95880    0.00000
  timesq  /timesq                 0.05814       0.01124       5.17424    0.00000
```

```
--------------------------------------------------------------------------------
  LEVEL 1                           TAU-HAT      STD.ERR.    Z-VALUE   PR > |Z|
--------------------------------------------------------------------------------
  intcept /intcept                  3.07863      0.20465    15.04328   0.00000


          LEVEL 2 COVARIANCE MATRIX

                  intcept           time        timesq

  intcept         11.72906
  time            -5.86552       6.59372
  timesq           0.38927      -0.55909       0.05814



          LEVEL 2 CORRELATION MATRIX

                  intcept        time      timesq

  intcept          1.0000
  time            -0.6670      1.0000
  timesq           0.4714     -0.9030      1.0000


          LEVEL 1 COVARIANCE MATRIX

                  intcept

  intcept         3.07863


          LEVEL 1 CORRELATION MATRIX

                  intcept

  intcept         1.0000
```

The fixed effects are all highly significant. There is an expected decrease of 0.2963 grams for each unit increase in the squared value of the time points. On the other hand, there is an estimated increase of 6.9056 grams with every increase of one week in time.

The expected value of the weights of the mice at time point number 2 may thus be calculated as

$$\text{Expected WEIGHT}_{i2} = 4.1621 + 2.00(6.9056) + 4.00(0.2963)$$
$$= 16.7881 \; grams.$$

From the random part of the model it can be seen that all the estimates of the random coefficients at level 2 of the model are significant. This also holds for all the interaction terms at this level of the model. The correlation between TIME and TIMESQ is rather high, at -0.9030.

Variation over measurements, on level 1 of the model, has been drastically reduced through the inclusion of the variable TIMESQ in the analysis. When comparing $-2\ln L$ for this model with that obtained for the linear growth model, a reduction of 473.7280 is noted, indicating that the inclusion of the variable TIMESQ significantly improved the fit of the model.

The empirical Bayes residuals and their variances for the first five mice, as given in the file **mouse4.ba2**, is given in the table below.

```
MOUSE4.BA2 - Notepad                                    —    □    ✕

File  Edit  Format  View  Help
        1        1        5.9047        3.3066        intcept
        1        2       -3.6104        0.73621       time
        1        3        0.36730       0.73356E-02   timesq
        2        1        0.16302       3.3066        intcept
        2        2       -1.1345        0.73621       time
        2        3        0.13707       0.73356E-02   timesq
        3        1       -3.1320        3.3066        intcept
        3        2        2.3674        0.73621       time
        3        3       -0.13993       0.73356E-02   timesq
        4        1        0.16759       3.3066        intcept
        4        2       -0.18088       0.73621       time
        4        3        0.65171E-01   0.73356E-02   timesq
        5        1       -5.5456        3.3066        intcept
        5        2        0.96351       0.73621       time
        5        3        0.85922E-01   0.73356E-02   timesq
```

From this information the empirical Bayes estimates for any of the level-2 units may be computed. For the first five mice, these estimates are given in the following table.

Mouse no.	Fixed effect	Empirical Bayes estimates
1	Intercept	$4.1621 + 5.9047 = 10.0668$
	TIME	$6.9056 - 3.6104 = 3.2952$
	TIMESQ	$-0.2963 + 0.3673 = 0.0710$
2	Intercept	$4.1621 + 0.1630 = 4.3251$
	TIME	$6.9056 - 1.1345 = 5.7711$
	TIMESQ	$-0.2963 + 0.1371 = 0.1592$
3	Intercept	$4.1621 - 3.1320 = 1.0301$
	TIME	$6.9056 + 2.3674 = 9.2730$
	TIMESQ	$-0.2963 - 0.1399 = -0.4362$
4	Intercept	$4.1621 + 0.1676 = 4.3297$
	TIME	$6.9056 - 0.1809 = 6.7247$
	TIMESQ	$-0.2963 + 0.0652 = 0.2311$
5	Intercept	$4.1621 - 5.5456 = 01.3835$
	TIME	$6.9056 + 0.9635 = 7.8691$
	TIMESQ	$-0.2963 + 0.0859 = 0.2104$

The expected value of the weight of the first five mice at time point number 2 using the empirical Bayes estimates may thus be calculated as:

$$\text{WEIGHT}_{12} = 10.0668 + 2.00(3.2952) + 4.00(0.0710) = 16.9412 \text{ grams}$$
$$\text{WEIGHT}_{22} = 4.3251 + 2.00(5.7711) + 4.00(0.1592) = 15.2305 \text{ grams}$$
$$\text{WEIGHT}_{32} = 1.0301 + 2.00(9.2730) + 4.00(-0.4362) = 17.8313 \text{ grams}$$
$$\text{WEIGHT}_{42} = 4.3297 + 2.00(6.7247) + 4.00(0.2311) = 16.8547 \text{ grams}$$
$$\text{WEIGHT}_{52} = 9.7077 + 2.00(7.8961) + 4.00(-1.3835) = 19.9659 \text{ grams}$$

In the case of mice numbers 1, 3, 4, and 5, the estimated weights thus obtained are higher than previously calculated, while mouse number 2 is below the previously calculated value of 16.7881 units.

Finally, the residuals for the first 45 observations, that is the first five male mice, are considered. The following is an extract from the output file **mouse4.res**:

File	Edit	Format	View	Help				
	1	1	1	15.000	13.433	1.5670	
	2	1	2	17.000	16.941	0.58731E-01	
	3	1	3	23.000	20.592	2.4085	
	4	1	4	24.000	24.384	-0.38379	
	5	1	5	26.000	28.318	-2.3181	
	6	1	6	31.000	32.394	-1.3944	
	7	1	7	37.000	36.613	0.38726	
	8	1	8	42.000	40.973	1.0269	
	9	1	9	46.000	45.475	0.52453	
	10	2	1	11.000	9.9370	1.0630	
	11	2	2	14.000	15.230	-1.2305	
	12	2	3	20.000	20.205	-0.20547	
	13	2	4	24.000	24.862	-0.86203	
	14	2	5	29.000	29.200	-0.20016	
	15	2	6	35.000	33.220	1.7802	
	16	2	7	36.000	36.921	-0.92109	
	17	2	8	41.000	40.304	0.69611	
	18	2	9	43.000	43.368	-0.36826	
	19	3	1	11.000	9.8669	1.1331	
	20	3	2	16.000	17.831	-1.8312	
	21	3	3	24.000	24.923	-0.92315	
	22	3	4	30.000	31.143	-1.1426	
	23	3	5	39.000	36.490	2.5104	
	24	3	6	39.000	40.964	-1.9642	
	25	3	7	48.000	44.566	3.4336	
	26	3	8	47.000	47.296	-0.29607	
	27	3	9	48.000	49.153	-1.1533	
	28	4	1	13.000	10.823	2.1767	
	29	4	2	16.000	16.855	-0.85468	
	30	4	3	21.000	22.424	-1.4238	

The largest residual for the male mice is for observation number 25 which represents the fifth measurement for mouse 4, where a residual of 3.4336 is encountered.

A plot of the residuals against the observation numbers is shown below. For a more detailed discussion of the analysis of residuals in a multilevel context, the user is referred to Goldstein (1987, pp. 21-26).
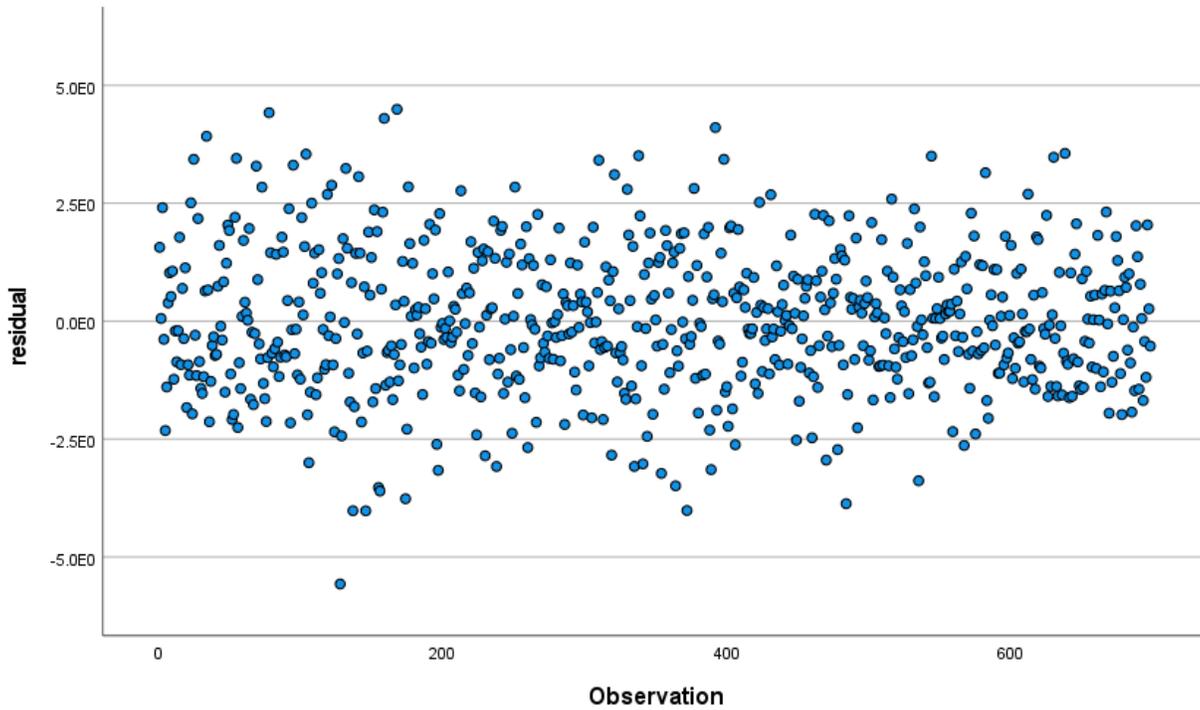
5. Introducing a covariate while modeling non-linear growth

In this example we want to determine whether there is a significant difference between the growth pattern of the male and female mice, as modeled in the non-linear growth model discussed previously. This can be determined by adding the gender of the mice as covariate to the model fitted in the previous section.

The variable GENDER is introduced as covariate by modifying the FIXED command to the following:

```
FIXED=intcept time timesq gender gender*time gender*timesq;
```

Edit the previous file accordingly and save it under a different name (**mouse5.prl**). Run PRELIS with this input file to fit the model to the data. Convergence is reached after 4 iterations and the following output is obtained.

```
ITERATION NUMBER     4


                    +-----------------------+
                    | FIXED PART OF MODEL   |
                    +-----------------------+


-------------------------------------------------------------------------
  COEFFICIENTS          BETA-HAT      STD.ERR.      Z-VALUE      PR > |Z|
-------------------------------------------------------------------------
  intcept                4.19133       0.44972       9.31996      0.00000
  time                   6.87771       0.29532      23.28885      0.00000
  timesq                -0.29399       0.02900     -10.13902      0.00000
  gender                -0.81492       0.44972      -1.81207      0.06998
  gender  *time          0.87235       0.29532       2.95389      0.00314
  gender  *timesq       -0.05947       0.02900      -2.05087      0.04028
```

```
                    +----------------------+
                    |   -2 LOG-LIKELIHOOD   |
                    +----------------------+


  DEVIANCE= -2*LOG(LIKELIHOOD) =     3389.53831478832
  NUMBER OF FREE PARAMETERS =                13



                    +----------------------+
                    | RANDOM PART OF MODEL  |
                    +----------------------+



--------------------------------------------------------------------------
  LEVEL 2                       TAU-HAT      STD.ERR.    Z-VALUE   PR > |Z|
--------------------------------------------------------------------------
  intcept /intcept              11.07339     2.60146     4.25660   0.00002
  time    /intcept              -5.16047     1.46703    -3.51763   0.00044
  time    /time                  5.83687     1.11334     5.24265   0.00000
  timesq  /intcept               0.34138     0.13349     2.55735   0.01055
  timesq  /time                 -0.50761     0.10447    -4.85897   0.00000
  timesq  /timesq                0.05464     0.01069     5.11098   0.00000


--------------------------------------------------------------------------
  LEVEL 1                       TAU-HAT      STD.ERR.    Z-VALUE   PR > |Z|
--------------------------------------------------------------------------
  intcept /intcept               3.07844     0.20464    15.04338   0.00000


        LEVEL 2 COVARIANCE MATRIX

                 intcept        time        timesq

  intcept        11.07339
  time           -5.16047     5.83687
  timesq          0.34138    -0.50761      0.05464



        LEVEL 2 CORRELATION MATRIX

                 intcept      time      timesq

  intcept         1.0000
  time           -0.6419    1.0000
  timesq          0.4389   -0.8989    1.0000



        LEVEL 1 COVARIANCE MATRIX

                 intcept

  intcept         3.07844
```

```
      LEVEL 1 CORRELATION MATRIX

            intcept

 intcept      1.0000
```

The coefficients of GENDER*TIME and GENDER*TIMESQ are significant at a 5 percent level, but the coefficient for GENDER*Intercept = GENDER is not. Only small changes are noticeable when the random part of the model fitted is compared to the corresponding section of the output obtained in the third example.

When comparing the values of $-2\ln L$ for the two models, a reduction of 11.39 is noted. It can thus be concluded that, although the gender of the mice has no significance on the intercept, there are significant differences between the growth patterns of male and female mice prior to physical maturity.

6. Complex variation at level 1 of the model

In the final example of a level 2 model, the linear growth model fitted previously is extended to include complex variation on both levels of the hierarchy. The term "complex variation" refers to the existence of two or more random variables at the same level of the hierarchy. We include the variable TIME in this model to illustrate such a model.

We modify the RANDOM1 command previously used to include TIME as shown in **mouse6.prl**:

```
OPTIONS MAXITER=25 ;
TITLE=Mouse data: complex variation ;
SY=MOUSE.LSF;
ID2=iden2;
RESPONSE=weight;
FIXED=intcept time;
RANDOM1=Intcept time;
RANDOM2=intcept time;
```

The output obtained for this model is given below. This model needed 18 iterations for convergence. The MAXITER option of the OPTIONS command was used to increase the number of iterations from the default of 10 to 25.

```
                   +----------------------+
                   | FIXED PART OF MODEL  |
                   +----------------------+


-----------------------------------------------------------------------------
  COEFFICIENTS           BETA-HAT      STD.ERR.      Z-VALUE      PR > |Z|
-----------------------------------------------------------------------------
  intcept                8.67635       0.46342      18.72259      0.00000
  time                   4.30965       0.13067      32.98166      0.00000


                   +----------------------+
                   |   -2 LOG-LIKELIHOOD  |
                   +----------------------+

 DEVIANCE= -2*LOG(LIKELIHOOD) =    3822.83149832815
 NUMBER OF FREE PARAMETERS =               8
```

```
                    +----------------------+
                    | RANDOM PART OF MODEL  |
                    +----------------------+


  ------------------------------------------------------------------
   LEVEL 2                        TAU-HAT     STD.ERR.    Z-VALUE   PR > |Z|
  ------------------------------------------------------------------
   intcept /intcept               13.35035    2.77510     4.81076   0.00000
   time    /intcept               -1.61832    0.61801    -2.61861   0.00883
   time    /time                   1.17651    0.21925     5.36601   0.00000


  ------------------------------------------------------------------
   LEVEL 1                        TAU-HAT     STD.ERR.    Z-VALUE   PR > |Z|
  ------------------------------------------------------------------
   intcept /intcept               13.73386    2.13539     6.43156   0.00000
   time    /intcept               -2.91233    0.54471    -5.34654   0.00000
   time    /time                   0.83668    0.12967     6.45252   0.00000


          LEVEL 2 COVARIANCE MATRIX

                  intcept         time

   intcept        13.35035
   time           -1.61832      1.17651



          LEVEL 2 CORRELATION MATRIX

                 intcept        time

   intcept         1.0000
   time           -0.4083      1.0000



          LEVEL 1 COVARIANCE MATRIX

                  intcept         time

   intcept        13.73386
   time           -2.91233      0.83668



          LEVEL 1 CORRELATION MATRIX

                 intcept        time

   intcept         1.0000
   time           -0.8591      1.0000
```

From the random part of the output it can be seen that there is an increase in the variance of the intercept term on level 1 of the model when the coefficient for the variable TIME is also allowed to vary randomly over level 1 of the model. The error 1 variance, however, is also a function of the covariance between the intercept and TIME. All coefficients are highly significant.

When the two values of $-2\ln L$ are compared for these models, a decrease of 50.43 is noted. The addition of the coefficient for the variable TIME on level 1 of the model thus seems to lead to an improved fit compared with the linear growth model. This implies that the level-1 error variances are heteroscedastic.

7. Conclusions

In the five examples discussed here, various models were considered for the analysis of repeated measurement data with a level-2 hierarchical structure. These models included a variance decomposition model, two linear growth models and a non-linear growth model. The inclusion of a covariate and the possibility of complex level-1 variation were also considered.

When the respective $-2\ln L$ values of these models are compared, the nonlinear model with a covariate included had the lowest value, namely, 3389.5383. It would thus appear that the growth of the 82 mice up to physical maturity can best be described by a parabola with the gender of the mice as a covariate. From the scatterplot shown previously, however, it seems as if other nonlinear functions for the modeling of the growth of the mice can also be considered.