2 Multilevel Modeling

This chapter[1] presents and illustrates the newly added multilevel modeling features in LISREL 8.30.

It starts out with a brief and general introduction, followed by an explanation of the statistical theory behind multilevel analysis.

The remainder of the chapter is devoted to how to do actual multilevel analysis. First, through an overview of the syntax for creating an input file. Next, through a step-by-step discussion of several examples.

2.1 Introduction to Multilevel Modeling

The analysis of data with a hierarchical structure has been described in the literature under various names. It is known as hierarchical modeling, random coefficient modeling, latent curve modeling, growth curve modeling or multilevel modeling. The basic underlying structure of measurements nested within units at a higher level of the hierarchy is, however, common to all. In a growth model, with repeated measures, for example, the measurements or outcomes are nested within the experimental units (second level units) of the hierarchy. We can describe these outcomes as a sum of effects for the individual measurement and for the experimental unit for which the measurement was made. Regression coefficients may be present at some or all of the levels, and variance components at different levels of the hierarchy may also be obtained.

[1]Written by Matilda du Toit and Stephen du Toit, Scientific Software International, Chicago

Inference can be drawn from available data for such a model for the population means at any level. Hierarchical models are particularly useful in the modeling of data from complex surveys, as cluster or multi-stage sample designs are frequently used for populations with a hierarchical structure. Ignoring the hierarchical structure of data can have serious implications, as the use of alternatives such as aggregation and disaggregation of information to another level can induce high collinearity among predictors and large or biased standard errors for the estimates. Standard fixed parameter regression models do not allow for the exploration of variation between groups, which may be of interest in its own right. For a discussion of the effects of these alternatives, see Bryk & Raudenbush (1992), Longford (1987), and Rasbash (1993).

In contrast, multilevel or hierarchical modeling provides the opportunity to study variation at different levels of the hierarchy. Such a model can also include separate regression coefficients at different levels of the hierarchy that have no meaning without recognition of the hierarchical structure of the population. The dependence of repeated measurements belonging to one experimental unit in a typical growth curve analysis, for example, is taken into account with this approach. In addition, the data to be analyzed need not be balanced in nature. This has the advantage that estimates can also be units for which a very limited amount of information is available.

In the examples given here, we primarily focus on the application of this approach to growth curve models and repeated measurements data. To illustrate the wide applicability of this analysis tool, we also include examples of the analysis of data of an educational nature and data from the 1995 CPC survey.

2.2 Theoretical Background

In this section the concept of multilevel modeling is introduced. A fixed parameter linear regression model is considered first (p. 15), followed by a level-2 model (p. 17), and finally a general level-3 model (p. 19).

Next, a brief overview of estimation procedures that may be used for the analysis of unbalanced hierarchical data is given (p. 22). Then, statistical inference is discussed, starting on p. 27.

Survey data in the social sciences are usually of a categorical nature. In Section 2.2.6 the analysis of data with categorical response variables is considered.

2.2.1 A fixed parameter linear regression model

Consider the dental measurement data set first analyzed by Potthoff & Roy (1964). The data set contains the dental measurements of 11 girls and 16 boys at ages 8, 10, 12, and 14 years. Each measurement is the distance in millimeters between the center of the pituitary and the pterygomaxillary fissure.

Suppose we wish to investigate the relationship between the measurements $y_{ij}, j = 1,2,3,4$ for child i and the ages at which the measurement were taken. Denote these ages for individual i by x_{i1}, x_{i2}, x_{i3}, and x_{i4}.

Traditionally, a single linear equation may be estimated by pooling all 27 cases. The measurement may then be expressed as a linear function of the ages at which measurements are taken and could be written as

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + e_{ij} \ , \qquad j = 1,2,3,4. \tag{2.1}$$

Let

$$\mathbf{x}'_{ij} = [\ 1 \quad x_{ij}\ ]$$

and

$$\boldsymbol{\beta}' = [\ \beta_0 \quad \beta_1\ ]$$

From (2.1), the measurement for individual i can be rewritten as

$$y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{e}_{ij} \ , \qquad i = 1,2,\ldots,N\ ;\ j = 1,2,3,4\ .$$

where $N = 27$ denotes the total number of children for which measurements were available.

Using matrix notation, the set of regression equations given above may be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \ ,$$

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} , \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_N \end{bmatrix} , \quad \text{and} \quad \mathbf{e} = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_N \end{bmatrix} .$$

It is assumed that $e_{11}, e_{12}, \ldots$ are uncorrelated with mean zero and constant variance σ^2. Thus,

$$\mathsf{E}(\mathbf{e}) = \mathbf{0} \tag{2.2}$$

and

$$\mathsf{Cov}\,(\mathbf{e}, \mathbf{e}') = \sigma^2 \mathbf{I} \ . \tag{2.3}$$

Under the assumptions given by (2.2) and (2.3), the ordinary least squares estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is obtained as

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}' \ \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}$$

where

$$\mathsf{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$$

and

$$\mathsf{Cov}\,(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}') = \sigma^2 (\mathbf{X}' \ \mathbf{X})^{-1} \ .$$

2.2.2 A level-2 model

In this case the individual children are the level-2 units and the dental measurements at different ages the level-1 units. There are four level-1 units nested within each level-2 unit and there are 27 level-2 units.

The fact that the regression coefficients usually vary from one individual to another, may be accommodated by regarding the unknown regression parameters as random variables with mean $\boldsymbol{\beta}$ and covariance matrix $\boldsymbol{\Phi}$.

The model for the 27 level-2 units can then be defined as

$$\mathbf{y}_i = \mathbf{X}\mathbf{b}_i + \mathbf{e}_i \ , \qquad i = 1, 2, \ldots, 27 \tag{2.4}$$

where the 4×2 matrix $\mathbf{X}$ is given by

$$\mathbf{X} = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix}$$

with the first column denoting the intercept term and the second column giving the ages at which measurements were made.

It is assumed that $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{27}$ are a random sample from a multivariate normal distribution with

$$\mathsf{E}(\mathbf{b}_i) = \boldsymbol{\beta}$$

and

$$\mathsf{Cov}(\mathbf{b}_i, \mathbf{b}_i') = \boldsymbol{\Phi} \ .$$

The vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{27}$ are assumed to be independently and identically distributed as $N(\mathbf{0}, \sigma^2\mathbf{I})$ independent of $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{27}$. Under these assumptions it follows that

$$\mathsf{E}(\mathbf{y}_i) = \mathbf{X}\boldsymbol{\beta} \tag{2.5}$$

and

$$\mathsf{Cov}(\mathbf{y}_i, \mathbf{y}_i') = \mathbf{X}\boldsymbol{\Phi}\mathbf{X}' + \sigma^2\mathbf{I}\ . \tag{2.6}$$

If, however, the gender of an individual is to be taken into account, (2.4) can be rewritten as

$$\mathbf{y}_i = \mathbf{X}\mathbf{b}_i + \mathbf{e}_i\ , \tag{2.7}$$

where

$$\mathbf{b}_i = \begin{bmatrix} \beta_0 + u_{i1} \\ \beta_1 + u_{i2} \\ \beta_2 + 0 \end{bmatrix}$$

and where β_2 denotes the gender coefficient.

It is more convenient to write (2.7) as

$$\mathbf{y}_i = \mathbf{X}_{(f)}\boldsymbol{\beta} + \mathbf{X}_{(2)}\mathbf{u}_i + \mathbf{e}_i\ , \qquad i = 1, 2, \ldots, 27\ . \tag{2.8}$$

The matrix $\mathbf{X}_{(f)}$ is the design matrix for the fixed part of the model. If gender is coded '1' for boys and '-1' for girls, the matrix $\mathbf{X}_{(f)}$, in the case of a female, is given by

$$\mathbf{X}_{(f)} = \begin{bmatrix} 1 & 8 & -1 \\ 1 & 10 & -1 \\ 1 & 12 & -1 \\ 1 & 14 & -1 \end{bmatrix}\ .$$

The vector $\mathbf{b}_i$ defines a model with intercept and slope coefficients which are allowed to vary over the units. A fixed gender effect is also included. The vector $\boldsymbol{\beta}$ as given in (2.8) contains coefficients for the fixed part of the model, while the vector $\mathbf{u}_i$ contains those coefficients allowed to vary over level-2 units.

The matrix $\mathbf{X}_{(2)}$ is the random parameter matrix on level 2 of the model and is given by

$$\mathbf{X}_{(2)} = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix} .$$

Let

$$\mathsf{E}(\mathbf{u}_i) = \mathbf{0}$$

and

$$\mathsf{Cov}\,(\mathbf{u}_i, \mathbf{u}_i') = \mathbf{\Phi}_{(2)}$$

while $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{27}$ are assumed to be identically and independently distributed as $N(\mathbf{0}, \sigma^2\mathbf{I})$. Then (*cf.* (2.5) and (2.6)),

$$\mathsf{E}(\mathbf{y}_i) = \mathbf{X}_{(f)}\boldsymbol{\beta} \tag{2.9}$$

and

$$\mathsf{Cov}\,(\mathbf{y}_i, \mathbf{y}_i') = \mathbf{X}_{(2)}\mathbf{\Phi}_{(2)}\mathbf{X}_{(2)}' + \sigma^2\mathbf{I} \,. \tag{2.10}$$

2.2.3 A general level-3 model

Consider the situation where a response variable y may deoend on a set of p predictors $x_1, x_2, \ldots, x_p$. The general level-3 model is defined as

$$\mathbf{y}_{ijk} = \mathbf{x}_{(f)ijk}'\boldsymbol{\beta} + \mathbf{x}_{(3)ijk}'\mathbf{v}_i + \mathbf{x}_{(2)ijk}'\mathbf{u}_i + \mathbf{x}_{(1)ijk}'\mathbf{e}_{ijk} \ , \tag{2.11}$$

where

$i = 1, 2, \ldots, N$ denotes level-3 units (*e.g.*, educational departments),

$j = 1, 2, \ldots, n_i$	denotes level-2 units (*e.g.*, schools), and
$k = 1, 2, \ldots, n_{ij}$	denotes level-1 units (*e.g.*, pupils).
$\mathbf{x}'_{(f)ijk} : 1 \times s$	is a typical row of the design matrix of the fixed part of the model, the elements being a subset of the p predictors.
$\mathbf{x}'_{(3)ijk} : 1 \times q$	is a typical row of the design matrix for the random part at level 3, the elements being a subset of the p predictors.
$\mathbf{x}'_{(2)ijk} : 1 \times m$	is a typical row of the design matrix for the random part at level 2, the elements being a subset of the p predictors.
$\mathbf{x}'_{(1)ijk} : 1 \times r$	is a typical row of the design matrix for the random part at level 1, the elements being a subset of the p predictors.
$\boldsymbol{\beta} : s \times 1$	is a vector of fixed, but unknown parameters to be estimated.

It is assumed that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are independently and identically distributed with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(3)}$. It is further assumed that $\mathbf{u}_{i1}, \mathbf{u}_{i2}, \ldots, \mathbf{u}_{in_i}$ are *i.i.d.* with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(2)}$, while $\mathbf{e}_{ij1}, \mathbf{e}_{ij2}, \ldots, \mathbf{e}_{ijn_{ij}}$ are *i.i.d.* with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(1)}$. Finally, it is assumed that $\mathbf{v}_i$, $\mathbf{u}_{ij}$, and $\mathbf{e}_{ijk}$ are independent.

Let

$$\mathbf{y}_i = \begin{bmatrix} \mathbf{y}_{i1} \\ \mathbf{y}_{i2} \\ \vdots \\ \mathbf{y}_{in_i} \end{bmatrix} \tag{2.12}$$

where $\mathbf{y}_{ij}$ denotes the $n_{ij} \times 1$ vector of responses for the i-th level-3 unit and the j-th level-2 unit. Note that $\mathbf{y}_{ij}$ can be expressed as

$$\mathbf{y}_{ij} = \mathbf{X}_{(f)ij}\boldsymbol{\beta} + \mathbf{X}_{(3)ij}\mathbf{v}_i + \mathbf{X}_{(2)ij}\mathbf{u}_{ij} + \begin{bmatrix} \mathbf{x}'_{(1)ij1}\mathbf{e}_{ij1} \\ \mathbf{x}'_{(1)ij2}\mathbf{e}_{ij2} \\ \vdots \\ \mathbf{x}'_{(1)ijn_{ij}}\mathbf{e}_{ijn_{ij}} \end{bmatrix} , \qquad (2.13)$$

where

$$\mathbf{X}_{(f)ij} = \begin{bmatrix} \mathbf{x}'_{(f)ij1} \\ \mathbf{x}'_{(f)ij2} \\ \vdots \\ \mathbf{x}'_{(f)ijn_{ij}} \end{bmatrix} , \qquad \mathbf{X}_{(3)ij} = \begin{bmatrix} \mathbf{x}'_{(3)ij1} \\ \mathbf{x}'_{(3)ij2} \\ \vdots \\ \mathbf{x}'_{(3)ijn_{ij}} \end{bmatrix} , \quad \text{and}$$

$$\mathbf{X}_{(2)ij} = \begin{bmatrix} \mathbf{x}'_{(2)ij1} \\ \mathbf{x}'_{(2)ij2} \\ \vdots \\ \mathbf{x}'_{(2)ijn_{ij}} \end{bmatrix} .$$

Under the distributional assumptions given above, it follows that

$$\mathsf{E}(\mathbf{y}_i) = \mathbf{X}_{(f)i}\boldsymbol{\beta} , \qquad (2.14)$$

where

$$\mathbf{X}_{(f)i} = \begin{bmatrix} \mathbf{X}_{(f)i1} \\ \mathbf{X}_{(f)i2} \\ \vdots \\ \mathbf{X}_{(f)in_i} \end{bmatrix} .$$

Also

$$\mathsf{Cov}\,(\mathbf{y}_i, \mathbf{y}'_i) = \mathbf{X}_{(3)i}\boldsymbol{\Phi}_{(3)}\mathbf{X}'_{(3)i} + \boldsymbol{\Lambda}_i , \qquad (2.15)$$

where

$$\mathbf{\Lambda}_i = \begin{bmatrix} \mathbf{\Lambda}_{i1} + \mathbf{D}_{i1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{i2} + \mathbf{D}_{i2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{\Lambda}_{in_i} + \mathbf{D}_{in_i} \end{bmatrix} \tag{2.16}$$

with

$$\mathbf{\Lambda}_{ij} = \mathbf{X}_{(2)ij}\mathbf{\Phi}_{(3)}\mathbf{X}'_{(2)ij}$$

and

$$\mathbf{D}_{ij} = \begin{bmatrix} \lambda_{ij1} & 0 & \ldots & 0 \\ 0 & \lambda_{ij2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{ijn_{ij}} \end{bmatrix} \tag{2.17}$$

with

$$\boldsymbol{\lambda}_{ijk} = \mathbf{x}'_{(1)ijk}\mathbf{\Phi}_{(1)}\mathbf{x}_{(1)ijk} \ .$$

2.2.4 Parameter estimation

The model given by (2.13) introduced in the Section *A general level-3 model* (page 19) may be written as

$$\mathbf{y}_i = \mathbf{X}_{(f)i}\boldsymbol{\beta} + \mathbf{X}_{(3)i}\mathbf{v}_i + \sum_{j=1}^{n_i}\mathbf{Z}_{(2)ij}\mathbf{u}_{ij} + \sum_{j=1}^{n_i}\sum_{k=1}^{n_{ij}}\mathbf{U}_{(1)ijk}\mathbf{e}_{ijk} \tag{2.18}$$

where

$$\mathbf{X}_{(3)i} = \begin{bmatrix} \mathbf{X}_{(3)i1} \\ \mathbf{X}_{(3)i2} \\ \vdots \\ \mathbf{X}_{(3)in_i} \end{bmatrix} , \tag{2.19}$$

$$\mathbf{Z}_{(2)ij} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{X}_{(2)ij} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} , \tag{2.20}$$

$$\mathbf{U}_{(1)ijk} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{X}_{(1)ijk} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} , \tag{2.21}$$

and where $\mathbf{v}_i$, $\mathbf{u}_{ij}$, and $\mathbf{e}_{ijk}$ denote the random parameter vectors on level 3, level 2, and level 1 of the model. It will be convenient to replace the double subscript jk with the single subscript l where $l = 1, 2, \ldots, n_i^*$ with

$$n_i^* = \sum_{j=1}^{n_i} n_{ij} \ .$$

Thus, (2.18) can be rewritten as

$$\mathbf{y}_i = \mathbf{X}_{(3)i}\mathbf{v}_i + \sum_{j=1}^{n_i} \mathbf{Z}_{(2)ij}\mathbf{u}_{ij} + \sum_{l=1}^{n_i^*} \mathbf{U}_{(1)il}\mathbf{e}_{il} \ . \tag{2.22}$$

Under the distributional assumptions given in Section *A general level-3 model*, it follows that

$$\mathsf{E}(\mathbf{y}_i) = \mathbf{X}_{(f)i}\boldsymbol{\beta} \ ,$$

where

$$\mathbf{X}_{(f)i} = \begin{bmatrix} \mathbf{X}_{(f)i1} \\ \mathbf{X}_{(f)i2} \\ \vdots \\ \mathbf{X}_{(f)in_i} \end{bmatrix} . \tag{2.23}$$

It also follows that

$$\begin{aligned}\mathsf{Cov}\,(\mathbf{y}_i, \mathbf{y}_i') &= \boldsymbol{\Sigma}_i \\ &= \mathbf{X}_{(3)i}\boldsymbol{\Phi}_{(3)}\mathbf{X}_{(3)i}' + \sum_{j=1}^{n_i} \mathbf{Z}_{(2)ij}\boldsymbol{\Phi}_{(2)}\mathbf{Z}_{(2)ij}' + \\ &\quad + \sum_{l=1}^{n_i^*} \mathbf{U}_{(1)il}\boldsymbol{\Phi}_{(1)}\mathbf{U}_{(1)il}' \ .\end{aligned}$$

Suppose that $\hat{\boldsymbol{\Phi}}_{(3)}$, $\hat{\boldsymbol{\Phi}}_{(2)}$, and $\hat{\boldsymbol{\Phi}}_{(1)}$ are consistent estimators of $\boldsymbol{\Phi}_{(3)}$, $\boldsymbol{\Phi}_{(2)}$, and $\boldsymbol{\Phi}_{(1)}$, respectively, so that

$$\mathbf{V}_i = \mathbf{X}_{(3)i}\hat{\boldsymbol{\Phi}}_{(3)}\mathbf{X}_{(3)i}' + \sum_{j=1}^{n_i} \mathbf{Z}_{(2)ij}\hat{\boldsymbol{\Phi}}_{(2)}\mathbf{Z}_{(2)ij}' + \sum_{l=1}^{n_i^*} \mathbf{U}_{(1)il}\hat{\boldsymbol{\Phi}}_{(1)}\mathbf{U}_{(1)il}'$$

is a consistent estimator of $\boldsymbol{\Sigma}_i$.

The generalized least squares estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is obtained as the minimum of the quadratic function

$$\mathbf{Q}_f = \sum_{i=1}^{N} [\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta}]' \mathbf{V}_i^{-1} [\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta}]$$

with solution

$$\hat{\boldsymbol{\beta}} = \left[\sum_{i=1}^{N} \mathbf{X}'_{(f)i} \mathbf{V}_i^{-1} \mathbf{X}_{(f)i} \right]^{-1} \left[\sum_{i=1}^{N} \mathbf{X}'_{(f)i} \mathbf{V}_i^{-1} \mathbf{y}_i \right] . \tag{2.24}$$

In order to estimate $\boldsymbol{\Phi}_{(3)}$, $\boldsymbol{\Phi}_{(2)}$, and $\boldsymbol{\Phi}_{(1)}$ let

$$\mathbf{y}_i^* = \mathsf{vecs}\,(\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta})(\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta})' , \tag{2.25}$$

then

$$\mathsf{E}(\mathbf{y}_i^*) = \mathsf{vecs}\,\boldsymbol{\Sigma}_i \ .$$

Using the result (Browne, 1974), on vector operations, *viz*,

$$\mathsf{vec}\,(\mathbf{CAC}') = (\mathbf{C} \otimes \mathbf{C})\mathsf{vec}\mathbf{A} \ ,$$

it follows that

$$\begin{aligned} \mathsf{vec}\mathbf{V}_i \quad &= \quad \mathbf{X}_{(3)i} \otimes \mathbf{X}_{(3)i} \mathsf{vec}\,\boldsymbol{\Phi}_{(3)} + \sum_{j=1}^{n_i} (\mathbf{Z}_{(2)ij} \otimes \mathbf{Z}_{(2)ij}) \mathsf{vec}\,\boldsymbol{\Phi}_{(2)} + \\ &\quad + \sum_{l=1}^{n_i^*} (\mathbf{U}_{(1)il} \otimes \mathbf{U}_{(1)il}) \mathsf{vec}\,\boldsymbol{\Phi}_{(1)} \ . \end{aligned}$$

There exists a unique matrix (see Browne, 1974, or McCulloch, 1982) $\mathbf{G}_p$: $p^2 \times \frac{1}{2}p(p+1)$ such that

$$\mathsf{vec}\,\mathbf{A} = \mathbf{G}_p \mathsf{vecs}\,\mathbf{A}$$

with $\mathbf{A}$ a symmetric $p \times p$ matrix. There is also a non-unique matrix $\mathbf{H}_p : \frac{1}{2}p(p+1) \times p^2$ such that

$$\text{vecs}\,\mathbf{A} = \mathbf{H}_p \text{vec}\,\mathbf{A}$$

The vector vecs $\boldsymbol{\Sigma}_i$, consisting of the non-duplicated elements of $\boldsymbol{\Sigma}_i$, can then be written as

$$\text{Vecs}\,\boldsymbol{\Sigma}_i = \mathbf{X}_i^* \boldsymbol{\tau}$$

where

$$\mathbf{X}_i^{*\prime} = \mathbf{H}_{n_i^*} \begin{bmatrix} (\mathbf{X}_{(3)i} \otimes \mathbf{X}_{(3)i})\mathbf{G}_q \\ \left(\sum_{j=1}^{n_i} \mathbf{Z}_{(2)ij} \otimes \mathbf{Z}_{(2)ij}\right) \mathbf{G}_m \\ \left(\sum_{l=1}^{n_i^*} \mathbf{U}_{(1)il} \otimes \mathbf{U}_{(1)il}\right) \mathbf{G}_r \end{bmatrix} \tag{2.26}$$

and

$$\boldsymbol{\tau} = \begin{bmatrix} \text{vecs}\,\boldsymbol{\Phi}_{(3)} \\ \text{vecs}\,\boldsymbol{\Phi}_{(2)} \\ \text{vecs}\,\boldsymbol{\Phi}_{(1)} \end{bmatrix} . \tag{2.27}$$

Now consider the quadratic form

$$\mathbf{Q}_T = \sum_{i=1}^{N} \left\{ [\mathbf{y}_i^* - \mathbf{X}_i^* \boldsymbol{\tau}]' \, \mathbf{W}_i^{-1} \, [\mathbf{y}_i^* - \mathbf{X}_i^* \boldsymbol{\tau}] \right\}$$

where $\mathbf{W}_i$ is a consistent estimator of the covariance of $\mathbf{y}_i^*$.

It has, for example, been shown by Browne (1974) and Goldstein (1989) that, if

$$\mathbf{W}_i^{-1} = \frac{1}{2}\mathbf{G}'_{n_i} \left(\mathbf{V}_i^{-1} \otimes \mathbf{V}_i^{-1}\right) \mathbf{G}_{n_i} \ , \tag{2.28}$$

then $\mathbf{W}_i$ is a consistent estimator of the covariance of $\mathbf{y}_i^*$.

Minimization of $\mathbf{Q}_T$ with respect to $\boldsymbol{\tau}$ yields

$$\hat{\boldsymbol{\tau}} = \left[\sum_{i=1}^{N} \mathbf{X}_i^{*\prime}\mathbf{W}_i^{-1}\mathbf{X}_i^*\right]^{-1} \left[\sum_{i=1}^{N} \mathbf{X}_i^{*\prime}\mathbf{W}_i^{-1}\mathbf{y}_i^*\right] \tag{2.29}$$

In order to ensure computational efficiency, the components of $\hat{\boldsymbol{\tau}}$ are further simplified as discussed in du Toit (1995).

IGLS estimators

The IGLS estimators $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ and $\hat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ can be obtained as follows:

(i) Set $\mathbf{V} = \mathbf{I}$
(ii) Calculate $\boldsymbol{\beta}_k$ (*cf.* (2.24))
(iii) Calculate $\mathbf{y}_i^* = \mathsf{vecs}\,(\mathbf{y}_i - \mathbf{X}_{(3)i}\boldsymbol{\beta}_k)(\mathbf{y}_i - \mathbf{X}_{(3)i}\boldsymbol{\beta}_k)'$ (*cf.* (2.25))
(iv) Calculate $\boldsymbol{\tau}_k$ by initially setting $\boldsymbol{\Phi}_{(1)} = \mathbf{I}$, $\boldsymbol{\Phi}_{(2)} = \mathbf{I}$, and $\boldsymbol{\Phi}_{(3)} = \mathbf{I}$
(v) Update $\mathbf{V}_i$ (*cf.* (2.28))

Repeat steps (ii) to (v) until convergence is obtained. The algorithm described above is known as Iterative Generalized Least Squares (IGLS).

2.2.5 Statistical inference

In the following section (*Standard errors*), results are given which are required for the calculation of the standard errors of the estimated parameters.

Next, in *Contrasts*, we discuss hypotheses of the form $c_1\beta_1 + c_2\beta_2 + \cdots + c_q\beta_q = k$ about the elements of the fixed parameter vector $\boldsymbol{\beta}$.

We continue with the calculation of empirical Bayes estimates (p. 30) and conclude with discussing likelihood ratio tests (p. 31).

Standard errors

From (2.24) it follows that the covariance matrix of $\hat{\boldsymbol{\beta}}$ is given by

$$\mathsf{Cov}\,(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}') = \left[\mathbf{X}^{*\prime}_{(f)}\boldsymbol{\Sigma}^{-1}\mathbf{X}^{*}_{(f)}\right]^{-1} . \tag{2.30}$$

In practice, $\boldsymbol{\Sigma}_i$ is unknown and is replaced by maximum likelihood estimator $\boldsymbol{\Sigma}_i(\hat{\boldsymbol{\tau}}) = \mathbf{V}_i$. Hence, a consistent estimate of the covariance matrix of is given by

$$\mathsf{Cov}\,(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}') = \left[\mathbf{X}^{*\prime}_{(f)}\mathbf{V}^{-1}\mathbf{X}^{*}_{(f)}\right]^{-1} . \tag{2.31}$$

Similarly, it can be shown that a consistent estimator of the covariance matrix of $\hat{\boldsymbol{\tau}}$ (*cf.* (2.29)) is given by

$$\mathsf{Cov}\,(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\tau}}') = \left[\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{W}_i^{-1}\mathbf{X}_i^{*}\right]^{-1} . \tag{2.32}$$

The diagonal elements of the covariance matrix (2.31) may be used to obtain large-sample estimates of the standard errors for the fixed parameter estimates. For large samples, $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\tau}}$ have approximate multivariate normal distributions. See, for example, Malinvaud (1970) for general results on the distribution of least squares estimators.

Contrasts

The construction of contrasts or linear functions of the parameters is a useful statistical analysis tool and enables the researcher to perform hypothesis testing concerning the equality of subsets of parameters. In this section a summary of the results required for contrast testing is given.

A $p \times q$ contrast matrix $\mathbf{C}$, where p denotes the number of contrasts, can be used to formulate a complex hypothesis about several elements of $\boldsymbol{\beta}$. The hypothesis is written in the form $\mathbf{C}\boldsymbol{\beta} = \mathbf{k}$, where $\mathbf{k}$ is a known $p \times 1$ vector.

Consider as an example the case where $q = 3$ and the following hypothesis is to be tested:

$$\begin{aligned} \beta_1 - \beta_2 &= 0 \\ \beta_3 - \beta_2 &= 0 \ . \end{aligned}$$

The null hypothesis can be formulated as

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \ .$$

For large samples, it can be shown that the vector variate $\mathbf{C}\hat{\boldsymbol{\beta}}$ will be approximately distributed as $N(\mathbf{C}\boldsymbol{\beta}, \mathbf{C}(\mathbf{X}'_{(f)}\hat{\mathbf{V}}^{-1}\mathbf{X}_{(f)})^{-1}\mathbf{C}')$.

Therefore, if H_0 is true,

$$\mathbf{M} = (\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{k})' \left\{ \mathbf{C}(\mathbf{X}'_{(f)}\hat{\mathbf{V}}^{-1}\mathbf{X}_{(f)})^{-1}\mathbf{C}' \right\} (\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{k}) \tag{2.33}$$

has an approximate χ^2 distribution with p degrees of freedom.

Let $\mathbf{c}'$ denote the i-th row of $\mathbf{C}$ and $\chi^2_{q,\alpha}$ the critical value of the χ^2 distribution with q degrees of freedom. A set of $100(1-\alpha)$ % simultaneous confidence intervals for the p elements of $\mathbf{C}\beta$ is given by the p intervals

$$\mathbf{c}'_i\hat{\boldsymbol{\beta}} \pm \left\{ \mathbf{c}'_i(\mathbf{X}'_{(j)}\hat{\mathbf{V}}^{-1}\mathbf{X}_{(j)})^{-1}\mathbf{c}_i\chi^2_{q,\alpha} \right\}^{0.5} \ , \quad p < q \ . \tag{2.34}$$

The null hypothesis $H_0 : \hat{\beta}_j = 0, \ j = 1, 2, \ldots, q$ is tested by using the test statistic

$$z = \frac{\hat{\beta}_k}{\text{S.E.}\,(\hat{\beta}_k)}$$

which, for large samples, has an approximate $N(0,1)$ distribution if H_0 is true.

Empirical Bayes estimates

The residuals $\hat{\mathbf{v}}_i$, $\hat{\mathbf{u}}_i$, and $\hat{\mathbf{e}}_i$ (see Section 2.2.3) may be estimated as follows.

Let

$$\tilde{\mathbf{y}}_i = \mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta} \ ,$$

Under the assumption of multivariate normality it follows that $\tilde{\mathbf{y}}_i$ is approximately distributed as $N(\mathbf{0}, \boldsymbol{\Sigma}_i)$ and $\mathbf{v}_i$ is approximately distributed as $N(\mathbf{0}, \boldsymbol{\Phi}_{(3)})$, and hence the joint distribution of $\tilde{\mathbf{y}}_i$ and $\mathbf{v}_i$ is

$$\begin{pmatrix} \tilde{\mathbf{y}}_i \\ \mathbf{v}_i \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} , \begin{matrix} \boldsymbol{\Sigma} & \mathbf{X}_{(3)i}\boldsymbol{\Phi}_{(3)} \\ \boldsymbol{\Phi}_{(3)}\mathbf{X}'_{(3)i} & \boldsymbol{\Phi}_{(3)} \end{matrix}\right) \ .$$

From standard results on conditional distributions (see for example Morrison, 1991) it follows that

$$\begin{aligned} \mathsf{E}(\mathbf{v}_i|\tilde{\mathbf{y}}_i) &= \mathbf{0} + (\mathbf{X}_{(3)i}\boldsymbol{\Phi}_{(3)})'\boldsymbol{\Sigma}_i^{-1}(\tilde{\mathbf{y}}_i - \mathbf{0}) \\ &= (\mathbf{X}_{(3)i}\boldsymbol{\Phi}_{(3)})'\boldsymbol{\Sigma}_i^{-1}\tilde{\mathbf{y}}_i \\ &= \boldsymbol{\Phi}_{(3)}\mathbf{X}'_{(3)i}\boldsymbol{\Sigma}_i^{-1}\tilde{\mathbf{y}}_i \ , \end{aligned}$$

Thus, the empirical Bayes estimate of $\mathbf{v}_i$ is

$$\hat{\mathbf{v}}_i = \hat{\boldsymbol{\Phi}}_{(3)}\mathbf{X}'_{(3)i}\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_{(f)i}\hat{\boldsymbol{\beta}}) \ . \tag{2.35}$$

Similarly,

$$\hat{\mathbf{u}}_{ij} = \hat{\boldsymbol{\Phi}}_{(2)}\mathbf{X}'_{(2)ij}\mathbf{V}_i^{-1}(\mathbf{y}_{ij} - \mathbf{X}_{(f)ij}\hat{\boldsymbol{\beta}}) \ . \tag{2.36}$$

Likelihood ratio tests

Finally, likelihood ratio tests are considered. Tests of a null hypothesis against a restricted alternative hypothesis can be constructed, provided that two conditions are met. Firstly, the models under H_0 and H_1 should be estimable and secondly, the parameter space $\mathbf{\Omega}_0$ for H_0 must be a subset of the parameter space $\mathbf{\Omega}$ of H_1.

Use is made of the likelihood ratio test statistic

$$\lambda = \frac{L_0(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\tau}})}{L_1(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\tau}})} , \tag{2.37}$$

where L_0 and L_1 denote the likelihood functions under H_0 and H_1, respectively. For N large (see for example Anderson, 1984), $-2\ln\lambda$ has an approximate $\chi^2(v)$ distribution where the number of degrees of freedom v is the difference in the number of parameters estimated under H_1 and the number of parameters estimated under H_0.

Example:

Consider the null hypothesis

$$H_0\text{: Cov}\,(\mathbf{y}_{ij}, \mathbf{y}'_{ij}) = \mathbf{X}_{(2)ij}\mathbf{\Phi}_{(2)}\mathbf{X}'_{(2)ij} + \mathbf{\Phi}_{(1)}\mathbf{I}_{n_{ij}}$$

as opposed to the alternative hypothesis

$$H_1\text{: Cov}\,(\mathbf{y}_{ij}, \mathbf{y}'_{ij}) = \mathbf{X}_{(3)ij}\mathbf{\Phi}_{(3)}\mathbf{X}'_{(3)ij} + \mathbf{X}_{(2)ij}\mathbf{\Phi}_{(2)}\mathbf{X}'_{(2)ij} + \mathbf{\Phi}_{(1)}\mathbf{I}_{n_{ij}} .$$

Let

$$\lambda = \frac{L_0(\hat{\mathbf{\Phi}}_{(1)}, \hat{\mathbf{\Phi}}_{(2)})}{L_1(\hat{\mathbf{\Phi}}_{(1)}, \hat{\mathbf{\Phi}}_{(2)}, \hat{\mathbf{\Phi}}_{(3)})} .$$

For N large, $-2\ln\lambda = -2(\ln L_0 - \ln L_1)$ has an approximate $\chi^2(v)$ distribution with the number of degrees of freedom, $v = \frac{1}{2}q(q+1)$, which is the number of non-duplicated elements of $\mathbf{\Phi}_{(3)}$. Note that $\ln L$ is the log-likelihood function

$$\ln L = -\frac{1}{2}\sum_{i=1}^{N}\left\{n_i \ln(2\pi) + \ln|\mathbf{\Sigma}_i| + \mathrm{tr}\mathbf{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta})(\mathbf{y}_i - \mathbf{X}_{(f)i}\boldsymbol{\beta})'\right\}$$

with $\mathbf{X}_i^*\boldsymbol{\beta}$ and $\mathbf{\Sigma}_i$, respectively, the expected value and covariance of $\mathbf{y}_i$.

2.2.6 Multilevel logit models

Introduction

Survey data usually consist of a mixture of biographical, geographical, and response variables and frequently have a hierarchical structure. Quite often, these response variables are categorical in nature. In the last few decades a wide variety of methods for the analysis of categorical data have been proposed. Many of these are generalizations of continuous data analysis methods (see for example Bishop, Fienberg, & Holland, 1975, and Agresti, 1990). In order to accommodate the structure of hierarchical data, a multilevel modeling approach may be used.

A level-3 logit model

To introduce a logit model with a categorical response variable, consider the following frequency table, where the subscripts ijk refer to subpopulation k within the i-th level-3 and j-th level-2 combination.

For illustrative purposes, it is assumed that a maximum of six subpopulations are formed according to the gender and age of respondents, as shown below for the i-th level-3 and j-th level-2 unit.

Gender	Age	Number of responses			Total
		Negative	Don't know	Yes	
Male	18 to 29 years	$f_{ij1,1}$	$f_{ij1,2}$	$f_{ij1,3}$	$f_{ij1\cdot}$
Male	30 to 49 years	$f_{ij2,1}$	$f_{ij2,2}$	$f_{ij2,3}$	$f_{ij2\cdot}$
Male	50 years and older	$f_{ij3,1}$	$f_{ij3,2}$	$f_{ij3,3}$	$f_{ij3\cdot}$
Female	18 to 29 years	$f_{ij4,1}$	$f_{ij4,2}$	$f_{ij4,3}$	$f_{ij4\cdot}$
Female	30 to 49 years	$f_{ij5,1}$	$f_{ij5,2}$	$f_{ij5,3}$	$f_{ij5\cdot}$
Female	50 years and older	$f_{ij6,1}$	$f_{ij6,2}$	$f_{ij6,3}$	$f_{ij6\cdot}$

The frequencies $f_{ijk,1}$, $f_{ijk,2}$, and $f_{ijk,3}$ denote the number of negative, don't know, and positive responses, respectively, while $f_{ijk\cdot}$ is the total number of responses for the k-th subpopulation, $k = 1, 2, \ldots, 6$. The number of response categories is denoted by c which, in this case, is equal to three.

The vector of responses is formed by using the third category, *i.e.*, the number of positive responses, as reference category. For each of the six subpopulations described above, two elements of the vector of responses are formed as follows:

$$\mathbf{y}_{ij} = \left[\begin{array}{ccc} \ln\frac{f_{ij1,1}}{f_{ij1,3}} \ln\frac{f_{ij1,2}}{f_{ij1,3}} & \ldots & \ln\frac{f_{ij6,1}}{f_{ij6,3}} \ln\frac{f_{ij6,2}}{f_{ij6,3}} \end{array} \right]. \tag{2.38}$$

For each subpopulation k within the i-th level-3 and j-th level-2 unit, the probability π_{ijkl} of the l-th response $(l = 1, 2)$ occurring is estimated by $p_{ijkl} = \frac{f_{ijk,l}}{f_{ijk\cdot}}$. Hence, a typical element of $\mathbf{y}_{ij}$ is $\ln\frac{p_{ijkl}}{p_{ijkc}}$.

From this it follows that

$$\mathbf{y}_i = f(\mathbf{p}_{ij}) \tag{2.39}$$

where

$$\mathbf{p}'_{ij} = \left[\begin{array}{ccc} p_{ij1,1}\; p_{ij1,2} & \ldots & p_{ij6,1}\; p_{ij6,2} \end{array} \right].$$

It is assumed that the expected value of the response vector $\mathbf{y}_{ij}$ can be expressed in the form

$$\mathsf{E}(\mathbf{y}_{ij}) = \mathbf{X}_{(f)ij}\boldsymbol{\beta} \tag{2.40}$$

where the elements of $\mathbf{X}_{(f)ij}$ depend on whether provision is made for the inclusion of an intercept or constant effect and on the way in which subpopulations are formed. The elements of the vector are fixed, but unknown, parameters to be estimated.

Denote the covariance matrix of $\mathbf{y}_{ij}$ by $\boldsymbol{\Sigma}_{ij}$. An approximate expression for $\boldsymbol{\Sigma}_i$ is obtained through use of the following first order Taylor expansion of $\mathbf{y}_{ij}$ evaluated in the neighborhood of $\boldsymbol{\pi}_{ij}$ (see for example Cramer, 1946)

$$\mathbf{y}_{ij} \simeq f(\boldsymbol{\pi}_{ij}) + \mathbf{J}_{ij}(\mathbf{p}_{ij} - \boldsymbol{\pi}_{ij}) \tag{2.41}$$

where

$$\mathbf{J}_{ij} : \; s \times cs = \frac{\partial \mathbf{y}_{ij}}{\partial \mathbf{y}'_{ij}} | \mathbf{p}_{ij} = \boldsymbol{\pi}_{ij} \; ,$$

c is the number of categories of the response variable and s denotes the total number of subpopulations in the (i, j)-th unit.

The covariance of $\mathbf{y}_{ij}$, to the first order of approximation, can be written as

$$\boldsymbol{\Sigma}_{ij} = \mathbf{J}_{ij} \, \mathsf{Cov}(\mathbf{p}_{ij}, \mathbf{p}'_{ij})\mathbf{J}'_{ij} = \mathbf{J}_{ij}\mathbf{V}_{ij}\mathbf{J}'_{ij} \; . \tag{2.42}$$

It can be shown that, if the c-th category of the response variable is used as the reference category, $\boldsymbol{\Sigma}_{ij}$ has, to the first order of approximation, the following form

$$\boldsymbol{\Sigma}_{ij} = \begin{bmatrix} \frac{1}{f_{ij1\cdot}}\boldsymbol{\Psi}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{1}{f_{ij2\cdot}}\boldsymbol{\Psi}_{22} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{1}{f_{ijs\cdot}}\boldsymbol{\Psi}_{ss} \end{bmatrix} \tag{2.43}$$

For a c category response variable

$$\boldsymbol{\Psi}_{kk} = \mathbf{D}\boldsymbol{\pi}_{ijk} + \frac{\mathbf{jj}'}{\pi_{ijk,c}} , \tag{2.44}$$

where $\pi_{ijk,c}$ is the probability that respondents in the k-th subpopulation from the i-th level-3 and j-th level-2 unit will select the c-th category of the response variable.

In order to accommodate the level-1 error structure, the model can be written as

$$\mathbf{y}_{ij} = \mathbf{X}_{(f)ij}\boldsymbol{\beta} + \mathbf{X}_{(1)ij}\mathbf{e}_{ij} . \tag{2.45}$$

It is assumed that $\mathbf{e}_{i1}$, $\mathbf{e}_{i2}$, ..., $\mathbf{e}_{in_i}$ are *i.i.d.* with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(1)}$. Assuming that the number of response categories is c, it follows that the matrix $\mathbf{X}_{(1)ij}$ denoting the random parameter weight matrix on level 1 of the model can be written as the symmetric matrix

$$\mathbf{X}_{(1)\mathrm{ij}} = \mathbf{D}_{\mathrm{ij}} = \begin{bmatrix} \frac{1}{\sqrt{f_{ij1\cdot}}}\mathbf{I}_{c-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{f_{ij2\cdot}}}\mathbf{I}_{c-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{f_{ij3\cdot}}}\mathbf{I}_{c-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{f_{ij4\cdot}}}\mathbf{I}_{c-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{f_{ij5\cdot}}}\mathbf{I}_{c-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{f_{ij6\cdot}}}\mathbf{I}_{c-1} \end{bmatrix} . \tag{2.46}$$

The following constraints are imposed on the level-1 covariance matrix $\boldsymbol{\Phi}_{(1)}$:

$$\boldsymbol{\Phi}_{(1)} \begin{bmatrix} \boldsymbol{\Phi}_{(1)_{11}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{\Phi}_{(1)_{22}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{\Phi}_{(1)_{33}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{\Phi}_{(1)_{44}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{\Phi}_{(1)_{55}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{\Phi}_{(1)_{66}} \end{bmatrix} . \tag{2.47}$$

The model given in (2.45) is further extended to include a random component on level 3 of the model as follows

$$\begin{aligned} \mathbf{y}_{ij} &= \mathbf{X}_{(f)ij}\left[\boldsymbol{\beta} + \mathbf{S}_{(3)ij}\mathbf{v}_i + \mathbf{S}_{(2)ij}\mathbf{u}_{ij}\right] + \mathbf{X}_{(1)ij}\mathbf{e}_{ij} \\ &= \mathbf{X}_{(f)ij}\boldsymbol{\beta} + \mathbf{X}_{(3)ij}\mathbf{v}_i + \mathbf{X}_{(2)ij}\mathbf{u}_{ij} + \mathbf{X}_{(1)ij}\mathbf{e}_{ij} \ . \end{aligned} \tag{2.48}$$

The vector $\mathbf{v}_i$ represents the coefficients of variables allowed to be random on level 3 and it is assumed that $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_N$ are independently and identically distributed with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Phi}_{(3)}$. It is assumed that $\mathbf{v}_i$, $\mathbf{u}_{ij}$, and $\mathbf{e}_{ijk}$ are independent.

$\mathbf{S}_{(2)}$ is a $t \times m$ ($m \leq t$) matrix formed by the selection of columns from the identity matrix of order t. These columns correspond to those elements of $\mathbf{b}_{ij}$ that are allowed to be random on level 2. If, for example, $t = 4$ and only the first and fourth coefficients are allowed to vary randomly on level 2, then

$$\mathbf{S}_{(2)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} .$$

The matrix $\mathbf{S}_{(3)}$ is defined in a similar way.

Let

$$\mathbf{y}'_i = \begin{bmatrix} \mathbf{y}'_{i1} & \cdots & \mathbf{y}'_{ij} & \cdots & \mathbf{y}'_{in_i} \end{bmatrix}$$

and denote the covariance matrix of $\mathbf{y}_i$ by $\mathbf{\Sigma}_i$. It follows that

$$\mathbf{\Sigma}_i = \mathbf{X}_{(3)ij}\mathbf{\Phi}_{(3)}\mathbf{X}'_{(3)ij} + \mathbf{\Lambda}_i \ , \tag{2.49}$$

where

$$\mathbf{\Lambda}_i = \begin{bmatrix} \mathbf{\Lambda}_{i1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{i2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Lambda}_{in_j} \end{bmatrix}$$

with

$$\mathbf{\Lambda}_{ij} = \mathbf{X}_{(2)ij}\mathbf{\Phi}_{(2)}\mathbf{X}'_{(2)ij} + \mathbf{X}_{(1)ij}\mathbf{\Phi}_{(1)}\mathbf{X}'_{(1)ij} \; .$$

Estimation considerations

In this section it is shown that parameter estimation of a level-3 logit model is achieved by fitting a general level-3 model with random coefficients on level 3 and level 2 of the hierarchy.

Let

$$\mathbf{y}_{ijk} = \begin{bmatrix} \mathbf{y}_{ijk,1} \\ \mathbf{y}_{ijk,2} \\ \vdots \\ \mathbf{y}_{ijk,c-1} \end{bmatrix} \tag{2.50}$$

denote a set of $c-1$ responses from subpopulation k, $k = 1, 2, \dots, s$, from the j-th level-2 and i-th level-3 unit.

It follows from (2.48) that

$$\begin{aligned} y_{ijk,1} &= \mathbf{x}'_{(f)ijk,1}\boldsymbol{\beta} + \mathbf{x}'_{(3)ijk,1}\mathbf{v}_i + \mathbf{x}'_{(2)ijk,1}\mathbf{u}_{ij} + \mathbf{x}'_{(1)ijk,1}\mathbf{e}_{ijk} \qquad (2.51) \\ y_{ijk,2} &= \mathbf{x}'_{(f)ijk,2}\boldsymbol{\beta} + \mathbf{x}'_{(3)ijk,2}\mathbf{v}_i + \mathbf{x}'_{(2)ijk,2}\mathbf{u}_{ij} + \mathbf{x}'_{(1)ijk,2}\mathbf{e}_{ijk} \\ \vdots \quad & \quad \vdots \\ y_{ijk,c-1} &= \mathbf{x}'_{(f)ijk,c-1}\boldsymbol{\beta} + \mathbf{x}'_{(3)ijk,c-1}\mathbf{v}_i + \mathbf{x}'_{(2)ijk,c-1}\mathbf{u}_{ij} + \mathbf{x}'_{(1)ijk,c-1}\mathbf{e}_{ijk} \end{aligned}$$

Under the assumption of a multinomial level-1 error structure (see the previous Section), it follows that the elements of $\mathbf{e}_{ijk}$ are correlated.

A multilevel analysis of logit models requires that

(i) $\mathbf{e}_{ijk}$, $k = 1, 2, \ldots, s_j$ are independently distributed and that
(ii) $\mathrm{Cov}(\mathbf{e}_{ijk}, \mathbf{e}'_{ijk}) = \mathbf{\Phi}_{(1)_{kk}}$ (*cf.* (2.43) and (2.46))

which implies that the $c-1$ elements of $\mathbf{e}_{ijk}$ are correlated, but that $\mathbf{e}_{ijk}$ and $\mathbf{e}_{ijm}$, $k \neq m$, are uncorrelated.

Let $\mathbf{X}_{(f)ijk}$, $\mathbf{X}_{(3)ijk}$, $\mathbf{X}_{(2)ijk}$, and $\mathbf{X}_{(1)ijk}$ have typical rows

$$\mathbf{x}'_{(f)ijk,l},\ \mathbf{x}'_{(3)ijk,l},\ \mathbf{x}'_{(2)ijk,l}\ \text{and}\ \mathbf{x}'_{(1)ijk,l}, \quad l = 1, 2, \ldots, c-1,$$

respectively. From (2.51) it follows that

$$\begin{aligned} y_{ijk} \quad = \quad & \mathbf{X}_{(f)ijk}\boldsymbol{\beta} + \mathbf{X}_{(3)ijk}\mathbf{v}_i + \mathbf{X}_{(2)ijk}\mathbf{u}_{ij} + \mathbf{X}_{(1)ijk}\mathbf{e}_{ijk} \\ & k = 1, 2, \ldots, s\ . \end{aligned} \tag{2.52}$$

The set of regression equations given by (2.52) can be written as

$$\mathbf{y}_{ij} = \mathbf{X}_{(f)ij}\boldsymbol{\beta} + \mathbf{X}_{(3)ij}\mathbf{v}_i + \mathbf{X}^*_{(2)ij}\mathbf{u}^*_{ij}$$

where

$$\mathbf{X}^*_{(2)ij} = \begin{bmatrix} \mathbf{X}_{(2)ij1} & \mathbf{X}_{(1)ij1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{X}_{(2)ij2} & \mathbf{0} & \mathbf{X}_{(1)ij2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}_{(2)ijs} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{X}_{(1)ijs} \end{bmatrix}, \tag{2.53}$$

and

$$\mathbf{u}^*_{ij} = \begin{bmatrix} \mathbf{u}_{ij} \\ \mathbf{e}_{ij1} \\ \mathbf{e}_{ij2} \\ \vdots \\ \mathbf{e}_{ijs} \end{bmatrix},$$

where s is the maximum number of subpopulations.

Note that

$$\mathbf{X}^*_{(2)ij} = \begin{bmatrix} \mathbf{X}_{(2)ij} & \mathbf{X}_{(1)ij} \end{bmatrix},$$

where (*cf.* (2.46))

$$\mathbf{X}_{(1)ij} = \mathbf{D}_{ij}.$$

Suppose that for a given (i, j)-combination only $s_j < s$ subpopulations are present. In this case $\mathbf{X}_{(2)ij}$ has $(c-1) \times s_j$ rows, but the number of columns remains $m + (c-1) \times s$, which is the dimension of the random coefficient vector $\mathbf{u}^*_{ij}$.

The covariance of $\mathbf{u}^*_{ij}$ can then be written as

$$\mathbf{\Phi}^*_{(2)} = \mathsf{Cov}(\mathbf{u}^*_{ij}, \mathbf{u}^{*\prime}_{ij}) = \begin{bmatrix} \mathbf{\Phi}_{(2)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_{(1)_{11}} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{\Phi}_{(1)_{ss}} \end{bmatrix}. \quad (2.54)$$

2.3 Multilevel Input Files

Although users of the Windows version of the LISREL program have the option to construct the required command files through the dialog-box interface, we will focus on the use of the multilevel syntax to write a proper input file directly in a text editor, which is the way the program works on each platform that it is available for.[2].

Some analysis specifications cannot even be produced through the dialog screens, in which case the editing of the syntax file becomes necessary.

2.3.1 Data file

All the multilevel examples that are included with the program and discussed starting on page 69, use a PRELIS system file or PSF file as starting point. For example, the very first example contains the following command in the input file:

```
SY=K:\LISREL83\MLEVELEX\MOUSE.PSF;
```

The data are in a system file (indicated with the command SY=) with the name MOUSE.PSF in a certain directory on the system.

Here is an example of a simple PRELIS file that creates such a system file (file EX2.PR2):

```
EXAMPLE 2: ATTITUDES OF MORALITY AND EQUALITY
DA NI=8 NO=200 MI=0 TR=PA
LA
HUMRGHTS EQUALCON RACEPROB EQUALVAL EUTHANAS CRIMEPUN CONSCOBJ GUILT
RA FI=DATA.EX2
OU RA=DATA.PSF
```

After running PRELIS with this input file, a PRELIS system file DATA.PSF is created in the default directory.

We recommend that the user follows the same path and concentrates first on the data step using all the features that PRELIS offers. Once the data are in shape, create a PSF file and concentrate on the data analysis.

---

[2] SSI publishes a separate guide that includes the Windows dialog-box interface for multilevel analysis: *Interactive LISREL*

2.3.2 Syntax file

This section provides an overview of the syntax conventions for multilevel analysis.

Syntax overview

The basic structure of the analysis input file is as given in Table 2.1, and the required commands are indicated. Turn to the page given in the last column for a detailed description of the particular command.

Table 2.1 Basic Structure of the Multilevel Input File

Command	Description	Required?	Page
OPTIONS	[list of options] ;	Yes	42
IDn =	[name of variable identifying level-n units] ;	Yes	46
RANDOMn =	[names of variables random on level n of the model] ;	Yes	47
RESPONSE =	[name(s) of response variables(s)] ;	Yes	48
FIXED =	[names of variables included as fixed effects in the model] ;	Yes	48
COVnPAT =	[pattern for level-n random coefficient covariance matrix] ;	No	49
COVnVAL =	[starting values for level-n random coefficient covariance matrix] ;	No	54
FIXVAL =	[starting values for fixed effect parameters] ;	No	55
CONTRAST =	[name of contrast file] ;	No	56
MISSING_DAT =	[integer value] ;	No	58
MISSING_DEP =	[integer value] ;	No	59
SUBPOP =	[names of variables to be used to construct subpopulations] ;	No	60
TITLE =	[a descriptive title for the analysis] ;	No	61

Guidelines for constructing or changing the input file

When the input file is constructed or edited through a text editor, the following guidelines should be kept in mind:

- All commands start with a command name and conclude with a semi-colon.
- Note that the maximum length of a command in the input file is 80 characters.
- There is no specific required order in which commands have to be given, with the exception that the OPTIONS command must always be the first command in the input file.
- Blank lines between commands are allowed.
- The input file may be given in either upper or lower case, as long as use of a preference is consistent throughout the file.
- A keyword cannot be a mixture of upper and lower case letters.

OPTIONS

Required command

Each input file for a multilevel analysis starts with an OPTIONS command. The options are used to control the estimation procedure and the amount of output to be supplied at convergence of the iterative procedure.

The list of options may be given in any order between the command name and the semi-colon signaling the conclusion of this command. If the default values of the options are deemed sufficient, the corresponding options may also be omitted.

The options that may be used with the OPTIONS command and their default values are summarized below.

OLS

OLS estimates of the fixed effects are calculated as a first step of the iterative procedure unless otherwise specified. The OLS keyword is used to indicate whether the OLS estimates are to be calculated during the first iteration. Valid values are YES and NONE.

If NONE is specified, no OLS estimates will be calculated during the first iteration. The NONE value, when used in combination with the optional FIXVAL command (see page 55), allows the user to provide a set of initial values for the fixed coefficients in the model. The default value for this keyword is YES, indicating calculation of OLS estimates during iteration 1.

CONVERGENCE

A test for convergence is made at the end of each iteration. The difference in the numerical value between successive values of the estimated parameters is compared to a convergence criterion of the form 10^{-x}, where x is an integer with possible values $1, 2, \ldots$. If the difference between successive values is smaller than the convergence criterion, convergence is said to have been reached. The default convergence criterion is 0.001 (or 10^{-2}). In order to use a different value, for example, 0.0001 as convergence criterion, the specification CONVERGE = 0.0001 must be included on the OPTIONS command.

MAXITER

The keyword MAXITER is used to indicate the maximum number of iterations to be performed. The default number of iterations is 10, which should be sufficient for convergence to be reached in most cases. If, however, a more stringent convergence criterion is used or previous experience with a particular data set indicates slow convergence, this keyword may be used to increase the maximum number of iterations. If, on the other hand, the user wishes to obtain only the OLS estimates calculated in the first iteration, MAXITER should be set equal to 1.

OUTPUT

The OUTPUT keyword determines the amount of output required. Valid specifications are:

STANDARD	The default output only, as described below
BAYES	The default output and empirical Bayes estimates

RESIDUAL	The default output and residuals
ALL	The default output, residuals and empirical Bayes estimates.

OUTPUT=STANDARD (Default output)

The following information is written to the default output file:

1. Input specifications as supplied by the user in the input file.
2. A summary of the hierarchical structure of the raw data.
3. Details of the iterative procedure. For each iteration, these details include the estimates, their standard errors, z-values, and probabilities of exceeding those limits.
4. The covariance and correlation matrices of the random parameters on the different levels of the model are also given.
5. The value of $-2\ln L$ (likelihood function) at each iteration
6. The computation time for completion of the iterative procedure and writing of required results to the output file.

See also the section on the format of the default output file (p. 61) for a discussion of the default output file.

OUTPUT=BAYES (Empirical Bayes estimates)

If OUTPUT=BAYES is specified, 1 to 6 listed above are written to the output file. Two additional output files are created (in the case of a level-3 model).

The empirical Bayes estimates on levels 2 and 3 of the model are calculated and, along with their variance and relevant variable codes, are written to the files *.BA2 and *.BA3, where these filenames refer to the second and third level of the hierarchy, respectively.

See also the section on the format of the empirical Bayes estimate output file(s) (p. 67) for a discussion of the EB estimate output file obtained for the mice data.

OUTPUT=RESIDUAL (Residuals)

If OUTPUT=BAYES is specified, items 1 to 6 listed above are written to the output file. An additional file, *.RES, is created, containing the residuals as at convergence. The following information is given:

- the residuals $(y_{ijk} - \mathbf{x}'_{(f)ijk} * \hat{\boldsymbol{\beta}})$,
- expected value $(\tilde{y}_{ijk})$, and
- observed value (y_{ijk}) for each observation in the raw data set.

See also the section on the format of the residual output file (p. 68) for a discussion of the residual file obtained for the analysis of the mice data.

OUTPUT=ALL (All output)

All of the above files are created.

Examples of OPTIONS commands

(a) OPTIONS ;

By using this form of the OPTIONS command, the procedure to be used is IGLS and the convergence criterion is 0.001. A maximum number of 10 iterations will be carried out and partial output will be written to either the user specified output file or the default output file *.OUT. OLS estimates are calculated during the first iteration.

(b) OPTIONS MAXITER=5 OUTPUT=ALL;

This command specifies the use of a convergence criterion of 0.001 with the use of OLS estimation in the first iteration. A maximum of 5 iterations is allowed and the creation of all additional output files is requested.

(c) OPTIONS OLS=NONE MAXITER=20 OUTPUT=ALL CONVERGE=0.0001;

Use of this command will exclude the calculation of the OLS estimates during the first iteration. Subsequent iterations will be performed using the IGLS procedure. The convergence criterion is 10^{-4} and the maximum number of iterations 20, indicating that the iterative procedure will terminate according to these two criteria. If convergence is not reached after 20 iterations, the procedure will be terminated. In this case, lack of convergence will be noted in the default output file. Otherwise, the procedure will stop when convergence according to the specified criterion occurs before the maximum number of iterations. All additional output files as described under the OUTPUT keyword above will be created on the hard disk.

IDn

Required command

The ID commands are used to indicate the variable(s) identifying the units on the different levels of the hierarchy.

If the model specified by the user is a level-2 model, the commands ID1 and ID2 are required. Likewise, if a level-3 model is to be considered, the ID1, ID2, and ID3 commands are required in the input file. The exceptions to this rule are in the case of a multivariate model and in the case of a model with no random component on level 1 of the model, where the ID1 command may be omitted.

Variables listed in the ID commands must be included in the PRELIS system file (*.PSF file). The spelling and case in which they are given need to correspond to that given in the PSF file. The syntax of this command is:

IDn = [variable name identifiying level-n units] ;

Examples of IDn commands

(a) If the raw data file contains information on the test scores, age, and gender of pupils belonging to classes within schools, and the variables school, class, pupil, age, gender, and result are contained in the PSF file, the following ID commands may be used to identify the levels of the hierarchical structure:

```
ID3=school ;
ID2=class ;
ID1=pupil;
```

(b) If the variables iden1, iden2, iden3, Y1, Y2, and X1 to X4 are in the PSF file, IDEN1 to IDEN3 may be used as shown below to identify the levels of the hierarchy.

```
ID3=iden3;
ID2=iden2;
ID1=iden1;
```

RANDOMn

Required command

One random command of the form

RANDOMn = [list of variable names] ;

is allowed for each level of the hierarchy. The RANDOM command is used to identify those coefficients that are allowed to vary randomly over a given level of the hierarchy.

Variables listed here must be included in the PSF file. The spelling and case in which they are given need to correspond to that given in the file.

Example of RANDOM commands

RANDOM3= X1:X4 ;

RANDOM2 = X2 X1;

RANDOM1 = X3;

From this hypothetical example, it can be seen that:

1. The random variables may be listed in any order.
2. Any or all of the possible predictors may be used at any level of the model.

As in the case of the ID commands, the RANDOM1 command may be omitted in the case of a multivariate model or if a model with no random component on level 1 of the hierarchy is to be fitted. Thus, in the case of a multivariate model, the following set of commands may be used:

```
ID3=iden3;
ID2=iden2;
RANDOM3= X1:X4 ;
RANDOM2= X3:X4 ;
```

It is possible to place constraints on elements of the random coefficient covariance matrices. Information on the constraints permitted and on the provision of initial values for elements of these matrices will be discussed when the COVnPAT and COVnVAL commands, which are optional, are considered.

RESPONSE

Required command

The RESPONSE command is of the form:

RESPONSE = [response variable(s)] ;

This command contains information on the response variable(s) to be used in the analysis.

In the case of a multivariate model, more than one response variable may be listed in the RESPONSE command. Spelling, etc., of the names of the response variables must once again be the same as that used in the PSF file and may be entered in any order.

Example of RESPONSE commands

(a) If the hypothetical example discussed previously is to be analyzed as a general level-3 model with Y1 as response variable, the relevant RESPONSE command will be:

RESPONSE = Y1;

(b) If, however, the same data is to be analyzed with a multivariate model with both Y1 and Y2 as response variables, the relevant RESPONSE command is either

RESPONSE = Y1 Y2 ;

or

RESPONSE = Y2 Y1 ;

FIXED

Required command

The FIXED command is used to identify the fixed effects for the model to be analyzed.

The syntax is:

FIXED = [list of variables names to be included as fixed effects] ;

The fixed effects may be all of the predictor variables contained in the raw data file or any subset of these predictors and may be specified in any order. Spelling of the names, however, must correspond to the spelling used in the PSF file.

If a covariate is included in the analysis, this should be reflected in the FIXED command. The format in which the covariate should be entered is:

FIXED = [var1] [. . .] [var$n$] [covariate*var1] [. . .] [covariate*var$n$];

Note that the covariate can be used in combination with any of the predictors listed in the FIXED command.

Examples of FIXED commands

(a) For the educational example, any one of the following FIXED commands is permissible:

FIXED = AGE ;

FIXED = GENDER ;

FIXED = GENDER AGE ;

(b) For the second of our previous examples, any of the following FIXED commands may be used:

FIXED X1:X2 ;

FIXED X1 X4 X3 X2 ;

or any other similar command.

(c) If the variable X3 is to be included as covariate in the first command in (b), the appropriate FIXED command is:

FIXED X1 X2 X3*X1 X3*X2;

Initial estimates for the fixed effects may be provided by the user. This is done through use of the optional FIXVAL command (discussed on p. 55).

COVnPAT

Optional command

The COVnPAT commands are used to place constraints on the covariance matrices of random coefficients on the different levels of the model.

The syntax for a COVnPAT command is:

COVnPAT= [options] ;

One COVnPAT command is allowed for each level of the hierarchy. If, for instance, a level-3 model with random components on all three levels of the hierarchy is to be fitted, up to three COVnPAT commands may be included in the input file.

Options that may be used with this command are:

COVnPAT = DIAG;

In this case the covariance matrix of random parameters on level n of the model will be constrained to be a diagonal matrix of the form

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \phi_{(n)11} & 0 & \dots & 0 \\ 0 & \phi_{(n)22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{(n)pp} \end{bmatrix},$$

where p is the number of random coefficients on level n of the hierarchy.

COVnPAT = TOEPLITZ;

The covariance matrix on level n will, in this case, be constrained to be of the form of a so-called Toeplitz matrix, *i.e.*

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix}.$$

This option may also be invoked by using only the first four letters of the required structure, *i.e.*, TOEP.

COVnPAT = INTRA;

When the INTRA option is used, the covariance matrix of random parameters on level n is constrained to have an intra-class structure, *i.e.*

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \alpha & \beta & \beta & \dots & \beta \\ \beta & \alpha & \beta & \dots & \beta \\ \beta & \beta & \alpha & \dots & \beta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \beta & \dots & \alpha \end{bmatrix} .$$

This command may also be used with the abbreviation INTR as option.

COVnPAT = MA1;

In order to constrain the covariance matrix on level n to be similar to that of a time series process of order MA1, the MA1 option may be used. The form of the covariance matrix will then be

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \gamma & \beta & 0 & \dots & 0 \\ \beta & \gamma & \beta & \dots & 0 \\ 0 & \beta & \gamma & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \gamma \end{bmatrix} .$$

COVnPAT = LOGIT;

In the case of categorical data analysis it may be necessary to constrain the elements of the covariance matrix on level n to be of the following form

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \mathbf{\Psi}_{11} & 0 & \dots & 0 \\ 0 & \mathbf{\Psi}_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{\Psi}_{ss} \end{bmatrix} ,$$

where, for example in the case of 4 response categories, the submatrix $\mathbf{\Psi}_{ii}$ is of the form

$$\boldsymbol{\Psi}_{ii} = \begin{bmatrix} \frac{1}{\pi_{i1}} + \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i4}} \\ \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i2}} + \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i4}} \\ \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i4}} & \frac{1}{\pi_{i3}} + \frac{1}{\pi_{i4}} \end{bmatrix},$$

For each subpopulation i, π_{ij} denotes the probability of the j-th response ($j = 1, 2, 3, 4$) occurring. The subpopulations are defined using the optional SUBPOP command.

COVnPAT = [as specified by user] ;

If it is deemed necessary to constrain the elements of the covariance matrix to be of a form other than those discussed above, the user may specify this required structure through use of the COVnPAT command. This can be done by entering a lower-triangular matrix with the required structure, using the COVnPAT command. If, for example, the covariance matrix for the RANDOMn command

RANDOMn = X1 X2 X3 X4 ;

is to be constrained, this can be done by keeping in mind that the lower triangular elements of the covariance matrix are numbered row-wise as shown below.

```
1
2  3
4  5  6
7  8  9  10;
```

The elements to be fixed are then replaced with '0.' If, for example, the matrix is constrained to be diagonal, the command to be used is:

```
COVnPAT = 1
          0  3
          0  0  6
          0  0  0  10;
```

The structure as specified indicates that there are four parameters to be estimated (*i.e.*, numbers 1, 3, 6, and 10, corresponding to the variances) and five fixed parameters (corresponding to the covariances), indicated with 0. The values to which the fixed parameters are to be set equal to can be supplied using the COVnVAL command . If the COVnVAL command is omitted, the fixed parameters will be constrained to be equal to zero, as the initial structure of all covariance matrices are assumed to be diagonal at the start of the iterative procedure.

In the case of an MA1 process, for example, the command will be:

```
COVnPAT = 1
          2  1
          0  2  1
          0  0  2  1;
```

From this structure, it follows that there are only two parameters to be estimated (numbers 1 and 2) while all other parameters are constrained to be equal to zero, unless otherwise specified using the COVnVAL command.

It is permissible to constrain diagonal elements of the level-n covariance matrix to be equal to 0 through the use of the COVnVAL command.

The following commands, for example, are permissible

```
COVnPAT = 1
          2  0
          3  2  0
          0  0  2  0;

COVnPAT = 0
          2  0
          3  2  0
          0  0  2  0;
```

Note that zero values indicate that the corresponding elements remain fixed at the values specified in the COVnVAL commands.

In conclusion, the user should note that:

(a) No line of input may exceed 80 characters. Thus, if a large COVnPAT matrix is entered, care should be taken that no row of this matrix exceeds this limit. If a row of the matrix is too long, it may simply be continued on the next line of the input file.

(b) If elements of the covariance matrix to be estimated are constrained to be equal in value, the number assigned to those elements must be the same.

(c) As with all other commands in the input file, the command should end with a semi-colon that may be placed directly after the last element of the matrix as specified or on the next line of the input file.

(d) The matrix specified by the user must have the same number of elements as implied by the RANDOMn command. That is, if there are p variables listed in the RANDOMn command, a total number of $\frac{1}{2}p(p+1)$ elements must be entered by the user.

(e) In order to assign initial values to elements of the covariance matrix at level n or to set fixed elements of the matrix to user specified values, the COVnPAT command should be used in conjunction with the COVnVAL command.

COVnVAL

Optional command

COVnVAL commands may be used to provide either initial values for elements of the covariance matrix on level n of the model or to provide values for elements fixed through the use of options of the COVnPAT command.

The syntax of these commands is:

COVnVAL = [as specified by user] ;

One COVnVAL command is allowed for each level of the hierarchy. If, for instance, a level-3 model with random coefficients on all three levels of the hierarchy is to be fitted, up to three COVnVAL commands may be included in the input file.

The values to be used for the elements of the covariance matrix must be entered in the form of a lower-triangular matrix. The number of values entered must be the same as the number of elements implied by the relevant RANDOMn command. If there are p variables listed in the RANDOMn command, $\frac{1}{2}p(p+1)$ values must be entered by the user. If a large number of values are entered, a row of the lower-triangular matrix may be continued on the next line of the input file if the number of characters in that row of the matrix exceeds 80 characters. The command must end with a semicolon, which may be entered on the last line of the values given or on the next line of the input file.

Examples of COVnVAL commands

(a) Continuing with the example used to illustrate the use of the COVnPAT command to obtain a user specified covariance structure, the following command illustrates how the user may provide values for the elements of the covariance matrix (n).

```
COVnVAL = 1.00
          0.32  0.85
          0.63  0.62  0.78
          0.19  0.00  0.25  0.99;
```

If an accompanying COVnPAT command is not used, these values will function as starting values for the level-n covariance matrix. When good starting values for the elements of this covariance matrix are known, the use of the command as shown above together with the use of the option OLS=NONE on the OPTIONS command will likely decrease the number of iterations required to obtain convergence.

(b) When the command

COVnPAT = DIAG;

is used together with the command given in (a), the values specified on the diagonal of the lower-triangular matrix in (a) will be used as initial values for these parameters which are to be estimated. The off-diagonal elements of the covariance matrix will then be constrained to be equal to the values of off-diagonal elements of the matrix in (a).

FIXVAL

Optional command

It is also possible to provide initial values for the fixed parameters in the model to be analyzed. This may be achieved through the use of the FIXVAL command, which allows the user to provide starting values for these parameters. The syntax of this command is:

FIXVAL = [as specified by user] ;

The number of values entered by the user using this command must be equal to the number of fixed parameters to be estimated. There is no specific format in which the values have to be entered. The following three input styles

```
FIXVAL = 0.151 0.355 0.654;
```

```
FIXVAL = 0.151
         0.355
         0.654;
```

and

```
FIXVAL = 0.151
         0.355
         0.654
;
```

are all permissible. If the first of these commands is used and the number of characters in the user specified string exceeds 80 characters, the next line of the input file should be used.

The use of the FIXVAL command and the OLS=NONE option of the OPTIONS command may be particularly effective when good starting values of these parameters are available.

CONTRAST

Optional command

The CONTRAST command is used to specify the path to an optional additional input file containing information on any contrast(s) between the fixed effects in the model to be tested. The syntax of this command is similar to that of the PRELIS DA command in that the contrast file may be located in another directory or even subdirectory. Thus, both

```
CONTRAST = MLEVEL.CTR ;
```

and

```
CONTRAST= C:\MLEVEL\EXAMPLES\MLEVEL.CTR ;
```

are valid examples of the CONTRAST command. Information contained in the contrast file must adhere to certain specifications, as illustrated with the following examples.

Examples

Suppose that there are six fixed effects in a particular model: INTERCEPT, GENDER, MATHS, READING, SCIENCE, and WRITING.

If, for example, we wish to test

$$H_0 : \quad \begin{aligned} &\beta_{\text{READING}} - \beta_{\text{WRITING}} = 0 \\ &\beta_{\text{MATHS}} - \beta_{\text{SCIENCE}} = 0 \ , \end{aligned}$$

we can test this by specifying

$$H_0 : \quad \mathbf{C}\boldsymbol{\beta} = \mathbf{0} \ ,$$

where

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$$

and

$$\boldsymbol{\beta}' = \begin{bmatrix} \beta_{\text{INTERCEPT}} & \beta_{\text{GENDER}} & \beta_{\text{MATHS}} & \beta_{\text{READING}} & \beta_{\text{SCIENCE}} & \beta_{\text{WRITING}} \end{bmatrix}$$

Note that each row of $\mathbf{C}$ has six elements, corresponding to the six fixed effects. Since the fourth element in the first row equals 1, this denotes β_{READING} while the sixth element denotes β_{WRITING}.

The contrast file will have the following form

```
2
0   0   0   1   0  -1
0   0   1   0  -1   0
```

The first row indicates the number of contrasts and the second and third rows the actual contrasts to be tested.

If the contrast file is specified as

```
1
0   0   0   1   0  -1
1
0   0   0   0  -1   0,
```

two separate contrast tests are performed as opposed to a simultaneous test for two contrasts.

MISSING_DAT

Optional command

There are two optional commands, which may be used when missing data is present in the raw data file. These commands are the MISSING_DAT and MISSING_DEP commands.

The MISSING_DAT command allows the user to specify an integer value, which will represent a missing value on any of the variables used in the analysis. The syntax of this command is

MISSING_DAT = [integer value] ;

Any positive or negative integer may be used. Only one value is allowed in this command. All records with data values equal to the code specified in this command will subsequently be removed from the analysis.

Examples of MISSING_DAT commands

Valid examples of the use of this command includes the following:

```
MISSING_DAT = 99 ;
MISSING_DAT = -998 ;
MISSING_DAT = 0 ;
```

Note that this command may also be used in conjunction with the MISSING_DEP command.

MISSING_DEP

Optional command

The MISSING_DEP command may be used to specify a code assigned to missing values on the response variables only. The syntax of this command is

MISSING_DEP = [integer value] ;

The same rules for values, which may be used with the MISSING_DAT command, applies to the MISSING_DEP command. The consequence of using the MISSING_DEP command is that only records with response variable values equal to the code assigned through the MISSING_DEP command will be removed from the analysis.

The MISSING_DEP command is recommended for use in the case of multivariate analysis. If only one of the response variables to be used in the multivariate analysis has a missing response, only that particular response will be considered missing while the remaining responses will still be used.

Example of a MISSING_DEP command

Consider the observations

Response variables				*Predictor variables*			
4.0	5.3	1.7	99	1	10	14.5	999
3.2	4.4	99	7.7	3	12	13.7	53.2

and the command

```
MISSING_DEP=99 ;
```

If the code 99 is identified as the code for missing data values on the dependent variables, this will imply that the analysis of this record will use the first three response values and disregard the fourth one in the case of the first observation. The third response variable will, however, be removed where the second observation is concerned.

If the code 999 is specified as the code for missing data values on all the variables included in the analysis, however, the whole first record as given above will be deleted from the data set to be analyzed. The second observation will be retained with the exception of the third response variable value.

This is accomplished by using both the MISSING_DEP and MISSING_DAT commands:

```
MISSING_DEP = 99 ;
MISSING_DAT = 999;
```

Note that if only the MISSING_DEP command is used for the two observations given above, the value of 999 for the last predictor variable on the first observation will be considered valid data and will be used as such in the analysis.

SUBPOP

Optional command

When categorical data are to be analyzed, subpopulations may be created through use of the SUBPOP command.

The syntax of this command is:

SUBPOP = [names of variables used to create subpopulations] ;

Use of, for example, two variables with two response categories each will lead to the creation of four subpopulations, *i.e.*, a two-way table for these two variables. Counts of responses to a particular question for these subpopulations may then be used as the response variable in a categorical data analysis.

Example of a SUBPOP command

Consider the two variables GENDER and AGE. If GENDER has two possible outcomes, for example, 1 = Male and 2 = Female and AGE has three outcomes, for example, 1 = less than 20 years old, 2 = 21–40 years old, and 3 = 41+ years old, the use of the SUBPOP command

SUBPOP = GENDER AGE;

will induce the creation of six subpopulations for the combination (GENDER;AGE), namely:

```
(GENDER=1;AGE=1)   (GENDER=1;AGE=2)   (GENDER=1;AGE=3)
(GENDER=2;AGE=1)   (GENDER=2;AGE=2)   (GENDER=2;AGE=3)
```

TITLE

Optional command

The TITLE command allows the user to provide a description of the analysis to be performed. This command, like all commands excluding the OPTIONS command, can be placed anywhere in the input file. The maximum permissible length of this command is 70 characters. The syntax of this command is

TITLE = [title as provided by the user] ;

2.4 Multilevel Output Files

In this section, the list output and the optional additional output files are discussed.

Each successful analysis generates list output that is displayed on the screen for review and saved to a file with the extension OUT. The following is a reference overview of the different output sections that a multilevel analysis may generate. It uses a default output file from one of the examples (*Analysis of 2-level repeated measures data*, starting on p. 69), annotated with brief descriptions of the various parts.

There are three optional output files that may be saved by the program after each analysis. They are discussed in the sections *Empirical Bayes estimates: the output files *.BA2 and *.BA3* (see p. 67), and *Residuals: the output file *.RES* (see p. 68).

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. For a detailed analysis of this data, please see the example *Analysis of 2-level repeated measures data*, starting on p. 69.

The annotated output file for a level-2 analysis follows.

Analysis requested

The first section of the output file contains an echo of the input file used for analysis. It is concluded with a message indicating that no problems occurred during the processing of the input file. If an error had occurred, for instance if the OPTIONS command was missing or one of the keywords was misspelled, an appropriate error message will be displayed here.

```
The following lines were read from file K:\LISREL83\MLEVELEX\MOUSE4.PR2

 OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=ALL ;
 TITLE=Mouse data: variance decomposition;
 SY=K:\LISREL83\MLEVELEX\MOUSE.PSF;
 ID1=iden1;
 ID2=iden2;
 RESPONSE=weight;
 FIXED=constant time timesq;
 RANDOM1=constant;
 RANDOM2=constant time timesq;
--------------------------------------------------------------------------
  NO ERROR DIAGNOSTICS GENERATED AT MODEL SPECIFICATIONS STAGE
--------------------------------------------------------------------------

                         +--------------------+
                         |  DATE : 01/16/1999 |
                         +--------------------+
```

Data summary

The data summary section contains information on the number of units at different levels of the hierarchy. From the data summary given here, it can be concluded that there were eight or nine separate measurements for each level-2 unit, the level-2 units being the 82 mice.

```
                             +--------------+
                             | DATA SUMMARY |
                             +--------------+

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
```

Fixed part of model

Estimates of the elements of the fixed coefficient vector at each iteration are given here. The standard errors, z-values, and probabilities of exceeding those limits are also reported.

Note that, before convergence is attained, the standard errors, z-values, and probabilities of exceeding are only an indication of the relevant values at convergence. In this case, all the fixed effects are statistically significant.

```
Mouse data: variance decomposition

ITERATION NUMBER     5

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

----------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
----------------------------------------------------------------------------
 constant                   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
```

Log-likelihood value

The log-likelihood value reported here is the value of $-2\ln L$ (likelihood function) evaluated at the parameter values during the relevant iteration. Comparison of these values over iterations gives an indication of how stable the iterative procedure is. For a model that provides a good description of the data, this statistic should decrease from one iteration to the next.

It may also be used to evaluate the goodness of fit of different models for the same data. The difference between the $-2\ln L$ values for models based on the same data has a chi-square distribution. The degrees of freedom equal the difference in the number of parameters estimated in these models.

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

-2 LOG-LIKELIHOOD =     3400.93248789791
```

Random part of model

This section contains a list of estimates of all the free elements of the covariance matrices of the random parameters at the different levels of the hierarchy. The standard errors, z-values, and the probabilities of exceeding those limits are also given. Note that, before convergence is attained, the standard errors, z-values, and probabilities of exceeding are only an indication of the relevant values at convergence.

The coefficient CONSTANT, denoting the intercept term, has the largest variation over the level-2 units. From the output it is seen that all variances and covariances are significant.

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


------------------------------------------------------------------------------
 LEVEL 2                         TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
------------------------------------------------------------------------------
 constant/constant               11.72906     2.70296      4.33933   0.00001
 time    /constant               -5.86552     1.58228     -3.70699   0.00021
 time    /time                    6.59372     1.23130      5.35509   0.00000
 timesq  /constant                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|
------------------------------------------------------------------------------
 constant/constant                3.07863     0.20465     15.04328   0.00000
```

Random coefficient covariance and correlation matrices

In this section, the covariance and correlation matrices of the random parameters in the model are given. These matrices include the elements that may have been considered fixed during the iterative procedure.

In certain situations, there may be differences between elements listed in the section of random parameter matrices and those given here. This will be the case if a covariance matrix was not positive definite. The matrix reported will be adjusted to be positive definite. This is accomplished by replacing the negative eigenvalues with a small positive number by the corresponding covariance matrix. If there are large differences between the elements reported in these two sections, this may indicate a misspecification of the model or the selection of an inappropriate covariance pattern.

Problems of this nature will also be reflected in the correlation matrix for the same level of the model as correlations close to either 1 or -1 in value. For this reason, the correlation matrices are provided as part of the output.

```
       LEVEL 2 COVARIANCE MATRIX

              constant        time      timesq

constant      11.72906
time          -5.86552     6.59372
timesq         0.38927    -0.55909     0.05814

       LEVEL 2 CORRELATION MATRIX

            constant      time    timesq

constant      1.0000
time         -0.6670    1.0000
timesq        0.4714   -0.9030    1.0000

       LEVEL 1 COVARIANCE MATRIX

             constant

constant      3.07863

       LEVEL 1 CORRELATION MATRIX

            constant

constant      1.0000
```

Technical details

This section is the last part of the output and gives the number of iterations performed as well as the total CPU time used for the iterative procedure.

```
CONVERGENCE REACHED IN  5 ITERATIONS


                  +-------------------------------------+
                  |  CPU TIME (SECONDS) :       0.37    |
                  +-------------------------------------+

                      +------------------+
                      |  END OF ANALYSIS |
                      +------------------+
```

Empirical Bayes estimates: the output files *.BA2 and *.BA3

The model fitted in this example was a level-2 model. A description of the data and model are given on page 61.

For a two level model no *.BA3 file is created. The structure of a BA3 file for a 3-level model is similar to that of the file MOUSE4.BA2.

The contents of the output file MOUSE4.BA2 for mice 7 and 8 are as follows:

```
         7         1     1.5979         3.3066      constant
         7         2    -1.0797        0.73621      time
         7         3    0.25660        0.73356E-02  timesq
         8         1    -1.7937         3.3066      constant
         8         2     3.2809        0.73621      time
         8         3   -0.22528        0.73356E-02  timesq
```

The first column of the output file MOUSE4.BA2 indicates the level-2 unit, in this case a specific mouse. There are three fixed effects in the model fitted to the data and this is reflected in the second column which gives the number of the fixed effect for each mouse.

The third column contains the deviations of the empirical Bayes estimates from the estimated population parameters. For mouse 7, it follows from the results for the fixed part of the model in the default output file that

$$\hat{\beta}' = \begin{bmatrix} 4.16213 & 6.90560 & -0.29629 \end{bmatrix}.$$

The deviations from these parameters for mouse 7, for example, are 1.5979, −1.0797, and 0.25660, respectively. Hence, the empirical Bayes estimates of the vector of fixed parameters for mouse 7 is given by

$$\begin{aligned} 4.16213 + 1.5979 &= 5.76003 \\ 6.90560 - 1.0797 &= 5.8259 \\ -0.29629 + 0.25660 &= -0.03969 \end{aligned}$$

The fourth column of the MOUSE4.BA2 file gives the variances of the empirical Bayes estimates while the last column contains the name of the relevant fixed effect.

Residuals: the output file *.RES

This file contains one line of information for each of the observations from the raw data file that was used in the analysis. The first 4 lines of this file for the mouse data analysis are:

```
1    1    1   15.000    10.771     4.2286
2    1    2   17.000    16.788    0.21183
3    1    3   23.000    22.212    0.78768
4    1    4   24.000    27.044    -3.0439
```

The first three columns of this file consist of information identifying the observations. In the first column, the residuals are numbered sequentially, in the second column the level-2 ID number identifying the 82 mice and in the third column the time points at which measurements were made (*i.e.*, the level-1 ID number) are given.

The fourth column contains the observed values of the response variables for each observation and thus represents the vector $\boldsymbol{y}$. This is followed by the expected values which represent the vector $\boldsymbol{X}_{(f)} * \hat{\boldsymbol{\beta}}$. Finally, the residuals are given, which are calculated as $\boldsymbol{y} - \boldsymbol{X}_{(f)} * \hat{\boldsymbol{\beta}}$, being the difference between the observed and estimated values of the response variable.

2.5 Multilevel Examples

This section presents four examples using multilevel analysis. An analysis of 2-level repeated measures data (see p. 69), a multivariate analysis of educational data (see p. 102), an analysis of air traffic control data (see p. 91), and an analysis of CPC Survey data (see p. 114). For a brief review of the various parts of the default output for multilevel analysis, see p. 61.

2.5.1 Analysis of 2-level repeated measures data

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

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.PSF, and the variable names to be used are shown below.

IDEN1	IDEN2	WEIGHT	CONSTANT	TIME	TIMESQ	GENDER
1.00	1.00	15.00	1.00	1.00	1.00	1.00
1.00	2.00	17.00	1.00	2.00	4.00	1.00
1.00	3.00	23.00	1.00	3.00	9.00	1.00
1.00	4.00	24.00	1.00	4.00	16.00	1.00
1.00	5.00	26.00	1.00	5.00	25.00	1.00
1.00	6.00	31.00	1.00	6.00	36.00	1.00
1.00	7.00	37.00	1.00	7.00	49.00	1.00
1.00	8.00	42.00	1.00	8.00	64.00	1.00
1.00	9.00	46.00	1.00	9.00	81.00	1.00
2.00	1.00	11.00	1.00	1.00	1.00	1.00
2.00	2.00	14.00	1.00	2.00	4.00	1.00

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 will be used to estimate the intercept term in the model. It may also be used to estimate the amount in which the intercept varies over the different levels of the hierarchy, as illustrated in the following section.

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 denotes 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:

$$y_{ij} = \text{CONSTANT}_{ij}\beta_0 + \text{CONSTANT}_{ij}u_{ij} + \text{CONSTANT}_{ij}e_{ij}$$

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.

Leave all the options default, or simply cut and paste them from another input file. Type a title command for this analysis. In this example we use *Mouse data: Variance decomposition*. Next, specify the location of the PSF file.

Select the variable IDEN2 as the level-2 identification variable and IDEN1 as the level-1 identification variable. Select the variable WEIGHT as response variable and the variable CONSTANT as fixed variable.

This variable CONSTANT, representing the intercept term, is then also specified as the level-1 and level-2 random variable.

The complete syntax file MOUSE1.PR2 should look like this:

```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=STANDARD ;
TITLE=Mouse data: Variance decomposition;
SY=K:\LISREL83\MLEVELEX\MOUSE.PSF;
ID1=iden1;
ID2=iden2;
```

```
RESPONSE=weight;
FIXED=constant;
RANDOM1=constant;
RANDOM2=constant;
```

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 (see the section on syntax on p. 41 for more detailed information).

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 given below.

```
                           +--------------+
                           | DATA SUMMARY |
                           +--------------+

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

ITERATION NUMBER     3

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

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

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

-2 LOG-LIKELIHOOD =     5425.49001592990

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

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

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

       LEVEL 2 COVARIANCE MATRIX

              constant

constant      11.32910

       LEVEL 2 CORRELATION MATRIX
```

```
          constant

constant     1.0000

       LEVEL 1 COVARIANCE MATRIX

              constant

constant     130.32083

       LEVEL 1 CORRELATION MATRIX

          constant

constant     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.

Fom 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.

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

$$y_{ij} = \mathsf{CONSTANT}_{ij}\beta_0 + \mathsf{TIME}_{ij}\beta_1 + \mathsf{CONSTANT}_{ij}u_{ij} + \mathsf{CONSTANT}_{ij}e_{ij}$$

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

The only change to the input file MOUSE2.PR2 is in the FIXED command, which now becomes:

```
FIXED=CONSTANT TIME;
```

The 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|
-------------------------------------------------------------------------------
 constant                  9.09586       0.60387      15.06258       0.00000
 time                      4.09218       0.06258      65.39108       0.00000


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

-2 LOG-LIKELIHOOD =      4137.57876020825


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


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

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


        LEVEL 2 COVARIANCE MATRIX

               constant
```

```
constant      20.69397

       LEVEL 2 CORRELATION MATRIX

            constant

constant      1.0000

       LEVEL 1 COVARIANCE MATRIX

              constant

constant      16.46288

       LEVEL 1 CORRELATION MATRIX

            constant

constant      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 weight measurement units is expected for each unit increase in TIME, that is a weekly interval between measurements.
- The log-likelihood value for this model is 4137.5788, compared to the value of 5425.4900 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 schools 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 56 percent.
- From the reduction in the level-1 variance component it can be 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 (MOUSE3.PR2):

```
RANDOM2 = CONSTANT TIME ;
```

After running PRELIS on this input file to fit the model to the data, we find the output for this model (here abbreviated) in the MOUSE3.OUT file:

```
ITERATION NUMBER     4

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


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


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

-2 LOG-LIKELIHOOD =     3873.66054782870


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


-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             12.85320       2.80949       4.57492   0.00000
 time    /constant             -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|
-----------------------------------------------------------------------------
 constant/constant              8.90088       0.54469      16.34126   0.00000

       LEVEL 2 COVARIANCE MATRIX
```

```
              constant        time

constant      12.85320
time          -1.64430     1.07389


       LEVEL 2 CORRELATION MATRIX

             constant      time

constant      1.0000
time         -0.4426    1.0000


       LEVEL 1 COVARIANCE MATRIX

              constant

constant      8.90088


       LEVEL 1 CORRELATION MATRIX

             constant

constant      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 CONSTANT 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-1 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 variables CONSTANT and TIME and may be written as:

$$\mathsf{Var}\left(\mathsf{CONSTANT}_{ij} + \mathsf{TIME}_{ij}\Phi_{(2)\mathsf{TIME,TIME}}\right)$$

$$= \Phi_{(2)\mathsf{CONSTANT,CONSTANT}} + 2\Phi_{(2)\mathsf{TIME,CONSTANT}}(\mathsf{TIME}_{ij}) +$$

$$+\Phi_{(2)\mathsf{TIME},\mathsf{TIME}}(\mathsf{TIME}_{ij})^2$$
$$= 2.85320 + 2(-1.64430)(\mathsf{TIME}_{ij}) + 1.07389(\mathsf{TIME}_{ij})^2$$

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 Figure 2.1.

- 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.

Figure 2.1 Between-mice variation as a function of the time points

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 non-linear 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:

$$y_{ij} = \text{CONS TIME TIMESQ}_{ij}\beta + \text{CONS TIME TIMESQ}_{ij}u_{ij} + \text{CONS}_{ij}e_{ij}$$

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 = CONSTANT  TIME  TIMESQ ;
RANDOM2 = CONSTANT  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=BAYES is added to the OPTIONS command.

The complete input file (MOUSE4.PR2) is now:

```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=ALL ;
TITLE=Mouse data: Modeling non-linear growth ;
SY=K:\LISREL83\MLEVELEX\MOUSE.PSF;
ID1=iden1;
ID2=iden2;
RESPONSE=weight;
FIXED=constant time timesq gender gender*time gender*timesq;
RANDOM1=constant;
RANDOM2=constant time timesq;
```

Convergence of the iterative procedure was reached in five iterations and the following output obtained.

```
ITERATION NUMBER     5

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

----------------------------------------------------------------------
 COEFFICIENTS           BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
----------------------------------------------------------------------
 constant               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   |
                      +-----------------------+

-2 LOG-LIKELIHOOD =     3400.93248789791

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

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             11.72906      2.70296      4.33933   0.00001
 time    /constant             -5.86552      1.58228     -3.70699   0.00021
 time    /time                  6.59372      1.23130      5.35509   0.00000
 timesq  /constant              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|
-----------------------------------------------------------------------------
 constant/constant              3.07863      0.20465     15.04328   0.00000

       LEVEL 2 COVARIANCE MATRIX

              constant        time      timesq

constant      11.72906
time          -5.86552     6.59372
timesq         0.38927    -0.55909     0.05814

       LEVEL 2 CORRELATION MATRIX

            constant      time    timesq

constant      1.0000
time         -0.6670    1.0000
timesq        0.4714   -0.9030    1.0000

       LEVEL 1 COVARIANCE MATRIX

             constant
```

```
constant      3.07863

      LEVEL 1 CORRELATION MATRIX

           constant

constant      1.0000
```

- The fixed effects are all highly significant. There is an expected decrease of 0.2963 grams for every 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

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

- 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.

1	1	5.9047	3.3066	CONSTANT
1	2	-3.6104	0.73621	TIME
1	3	0.36730	0.73356E-02	TIMESQ
2	1	0.16302	3.3066	CONSTANT
2	2	-1.1345	0.73621	TIME
2	3	0.13707	0.73356E-02	TIMESQ
3	1	-3.1320	3.3066	CONSTANT
3	2	2.3674	0.73621	TIME
3	3	-0.13993	0.73356E-02	TIMESQ
4	1	0.16759	3.3066	CONSTANT
4	2	-0.18088	0.73621	TIME
4	3	0.65171E-01	0.73356E-02	TIMESQ
5	1	-5.5456	3.3066	CONSTANT
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 estimate
1	CONSTANT	$4.1621 + 5.9047 = 10.0668$
	TIME	$6.9056 - 3.6104 = 3.2952$
	TIMESQ	$-0.2963 + 0.3673 = 0.0710$
2	CONSTANT	$4.1621 + 0.1630 = 4.3251$
	TIME	$6.9056 - 1.1345 = 5.7711$
	TIMESQ	$-0.2963 + 0.1371 = 0.1592$
3	CONSTANT	$4.1621 - 3.1320 = 1.0301$
	TIME	$6.9056 + 2.3674 = 9.2730$
	TIMESQ	$-0.29630.1399 = 0.4362$
4	CONSTANT	$4.1621 + 0.1676 = 4.3297$
	TIME	$6.9056 - 0.1809 = 6.7247$
	TIMESQ	$-0.2963 + 0.0652 = 0.2311$
5	CONSTANT	$4.1621 + 5.5456 = 9.7077$
	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:

$$\text{WEIGHT}_{12} = 10.0668 + 2.00(3.2952) + 4.00(0.0710) = 16.9412 \text{ grams}$$
$$\text{WEIGHT}_{22} = 4.3251 + 2.00(5.7711) + 4.00(0.1592) = 15.2305 \text{ grams}$$
$$\text{WEIGHT}_{32} = 1.0301 + 2.00(9.2730) + 4.00(0.4362) = 17.8313 \text{ grams}$$
$$\text{WEIGHT}_{42} = 4.3297 + 2.00(6.7247) + 4.00(0.2311) = 16.8547 \text{ grams}$$
$$\text{WEIGHT}_{52} = 9.7077 + 2.00(7.8961) + 4.00(0.2104) = 24.6583 \text{ grams}$$

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:

1	1	1	15.000	10.771	4.2286
2	1	2	17.000	16.788	0.21183
3	1	3	23.000	22.212	0.78768
4	1	4	24.000	27.044	-3.0439
5	1	5	26.000	31.283	-5.2829
6	1	6	31.000	34.929	-3.9293
7	1	7	37.000	37.983	-0.98314
8	1	8	42.000	40.444	1.5556
9	1	9	46.000	42.313	3.6869
10	2	1	11.000	10.771	0.22856
11	2	2	14.000	16.788	-2.7882
12	2	3	20.000	22.212	-2.2123
13	2	4	24.000	27.044	-3.0439
14	2	5	29.000	31.283	-2.2829
15	2	6	35.000	34.929	0.70698E-01
16	2	7	36.000	37.983	-1.9831
17	2	8	41.000	40.444	0.55561
18	2	9	43.000	42.313	0.68693
19	3	1	11.000	10.771	0.22856
20	3	2	16.000	16.788	-0.78817
21	3	3	24.000	22.212	1.7877
22	3	4	30.000	27.044	2.9561
23	3	5	39.000	31.283	7.7171
24	3	6	39.000	34.929	4.0707
25	3	7	48.000	37.983	10.017
26	3	8	47.000	40.444	6.5556

27	3	9	48.000	42.313	5.6869
28	4	1	13.000	10.771	2.2286
29	4	2	16.000	16.788	-0.78817
30	4	3	21.000	22.212	-1.2123
31	4	4	26.000	27.044	-1.0439
32	4	5	31.000	31.283	-0.28289
33	4	6	37.000	34.929	2.0707
34	4	7	44.000	37.983	6.0169
35	4	8	44.000	40.444	3.5556
36	4	9	44.000	42.313	1.6869
37	5	1	5.0000	10.771	-5.7714
38	5	2	13.000	16.788	-3.7882
39	5	3	20.000	22.212	-2.2123
40	5	4	26.000	27.044	-1.0439
41	5	5	32.000	31.283	0.71711
42	5	6	39.000	34.929	4.0707
43	5	7	45.000	37.983	7.0169
44	5	8	48.000	40.444	7.5556
45	5	9	52.000	42.313	9.6869

The largest residual for the male mice is for observation number 25 which represents the fifth measurement for mouse 3, where a residual of 10.017 is encountered. Plots of the residuals against the observation number are given in Figure 2.2. For a more detailed discussion of the analysis of residuals in a multilevel context, the user is referred to Goldstein (1987, pp. 21–26).

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 = CONSTANT  TIME  TIMESQ  GENDER GENDER*TIME  GENDER*TIMESQ;
```

Edit the previous file accordingly and save it as a new input file (MOUSE5.PR2). 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.

Figure 2.2 Plot of residuals for the first 45 records in the data set

```
ITERATION NUMBER     4

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

-----------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 constant                   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    |
                      +------------------------+

-2 LOG-LIKELIHOOD =     3389.53831478844
```

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

LEVEL 2	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
constant/constant	11.07339	2.60146	4.25660	0.00002
time /constant	-5.16047	1.46703	-3.51763	0.00044
time /time	5.83687	1.11334	5.24265	0.00000
timesq /constant	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|
constant/constant	3.07844	0.20464	15.04338	0.00000

LEVEL 2 COVARIANCE MATRIX

	constant	time	timesq
constant	11.07339		
time	-5.16047	5.83687	
timesq	0.34138	-0.50761	0.05464

LEVEL 2 CORRELATION MATRIX

	constant	time	timesq
constant	1.0000		
time	-0.6419	1.0000	
timesq	0.4389	-0.8989	1.0000

LEVEL 1 COVARIANCE MATRIX

	constant
constant	3.07844

LEVEL 1 CORRELATION MATRIX

	constant
constant	1.0000

❑ The coefficients of GENDER*TIME and GENDER*TIME are significant at a 5 percent level, but the coefficient for GENDER*CONSTANT 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 $-2\ln L$ (likelihood function) 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 denoted by the variables CONSTANT, there are significant differences between the growth patterns of male and female mice prior to physical maturity.

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

```
RANDOM1 = CONSTANT TIME;
```

This change in the level-1 covariance structure implies that the total variation at this level of the model can now be written as:

$$\text{Var}\left(\text{CONSTANT}_{ij} + \text{TIME}_{ij}\Phi_{(1)\text{TIME,TIME}}\right)$$
$$= \Phi_{(1)\text{CONSTANT,CONSTANT}} + 2\Phi_{(1)\text{TIME,CONSTANT}}(\text{TIME}_{ij}) + \Phi_{(1)\text{TIME,TIME}}(\text{TIME}_{ij})^2$$

The input file (MOUSE6.PR2) should look like:

```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=25 OUTPUT=ALL ;
TITLE=Mouse data: variance decomposition;
SY=K:\LISREL83\MLEVELEX\MOUSE.PSF;
ID1=iden1;
ID2=iden2;
```

```
RESPONSE=weight;
FIXED=constant time;
RANDOM1=constant time;
RANDOM2=constant time;
```

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

```
ITERATION NUMBER    18

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

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

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

-2 LOG-LIKELIHOOD =     3822.83149832868

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

-----------------------------------------------------------------------------
 LEVEL 2                         TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              13.35035      2.77510      4.81076   0.00000
 time    /constant              -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|
-----------------------------------------------------------------------------
 constant/constant              13.73386      2.13539      6.43156   0.00000
 time    /constant              -2.91233      0.54471     -5.34654   0.00000
 time    /time                   0.83668      0.12967      6.45252   0.00000
```

```
      LEVEL 2 COVARIANCE MATRIX

             constant       time

constant     13.35035
time         -1.61832    1.17651


      LEVEL 2 CORRELATION MATRIX

           constant     time

constant     1.0000
time        -0.4083   1.0000


      LEVEL 1 COVARIANCE MATRIX

             constant       time

constant     13.73386
time         -2.91233    0.83668


      LEVEL 1 CORRELATION MATRIX

           constant     time

constant     1.0000
time        -0.8591   1.0000
```

From the random part of the output it can be seen that there is a marked increase in the variance of the CONSTANT 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. 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 .

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 non-linear 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 covariate. From Figure 2.1, however, it seems as if other non-linear functions for the modeling of the growth of the mice can also be considered.

2.5.2 Analysis of air traffic control data

The data used in this example are described by Kanfer & Ackerman (1989). The data consists of information for 141 U.S. Air Force enlisted personnel. The personnel carried out a computerized air traffic controller task developed by Kanfer and Ackerman.

The subjects were instructed to accept planes into their hold pattern and land them safely and efficiently on one of four runways, varying in length and compass directions, according to rules governing plane movements and landing requirements. For each subject, the success of a series of between three and six 10-minute trials was recorded. The measurement employed was the number of correct landings per trial.

The Armed Services Vocational Battery (ASVB) was also administered to each subject. A global measure of cognitive ability, obtained from the sum of scores on the 10 subscales, is included in the data.

The data for this example can be found in the KANFER.PSF file. The variable labels and first few records of this data file are shown below.

CONTROL	TIME	MEASURE	ABILITY	CONSTANT	TIMESQ
1.00	1.00	24.00	142.16	1.00	1.00
1.00	2.00	27.00	142.16	1.00	4.00
1.00	3.00	30.00	142.16	1.00	9.00
1.00	4.00	32.00	142.16	1.00	16.00
1.00	5.00	38.00	142.16	1.00	25.00

1.00	6.00	41.00	142.16	1.00	36.00
2.00	1.00	2.00	-7.63	1.00	1.00
2.00	2.00	3.00	-7.63	1.00	4.00
2.00	3.00	9.00	-7.63	1.00	9.00
2.00	4.00	13.00	-7.63	1.00	16.00
2.00	5.00	13.00	-7.63	1.00	25.00
2.00	6.00	14.00	-7.63	1.00	36.00
3.00	1.00	12.00	-67.43	1.00	1.00
3.00	2.00	18.00	-67.43	1.00	4.00
3.00	3.00	24.00	-67.43	1.00	9.00

The variables in the data set are:

CONTROL	The identifying number of the air traffic controller
TIME	The number of the trial (between 1 and 6)
MEASURE	The number of successful trial for the trial
ABILITY	The cognitive ability score (combined ASVB score)
CONSTANT	The intercept term, with value 1 throughout
TIMESQ	TIME*TIME, a quadratic term

Using this data, three models will be fitted:

- The first model, a variance decomposition model, will investigate the variation in the number of correct trials over subjects and also over measurements for each subject.
- In the next model, a non-linear growth model will be considered.
- Finally, the cognitive ability measure, a controller-related variable, will be introduced into the non-linear growth model.

Variance decomposition model for Air Force data

To start the input file, we select the defaults for output options, maximum number of iterations, and convergence criterion, then provide an optional title for the analysis.

The variable CONTROL, identifying the air traffic controller, is used as level-2 identification, as up to six measurements are available for each controller. The variable TIME, indicating the number of the trial, is used as level-1, or measurement, identification.

The number of successful landings per trial is represented by the variable MEASURE, which we select as the response variable for this particular model. We include the variable CONSTANT, representing the intercept term, as fixed effect in a similar way.

Finaly, the intercept term, as represented by the variable CONSTANT is selected as a random variable at both levels of the model.

The resulting input file (KANFER1.PR2) is as follows:

```
OPTIONS ;
TITLE=Kanfer and Ackerman data: Variance decomposition;
SY=K:\LISREL83\MLEVELEX\KANFER.PSF;
ID1=time;
ID2=control;
RESPONSE=measure;
FIXED=constant;
RANDOM1=constant;
RANDOM2=constant;
```

Partial output is given and discussed below.

```
                             +--------------+
                             | DATA SUMMARY |
                             +--------------+

NUMBER OF LEVEL 2 UNITS :       141
NUMBER OF LEVEL 1 UNITS :       840

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

N2   :        9       10       11       12       13       14       15       16
N1   :        6        6        6        6        6        6        6        6

N2   :       17       18       19       20       21       22       23       24
N1   :        6        6        6        6        6        6        6        6

N2   :       25       26       27       28       29       30       31       32
N1   :        6        6        6        6        6        6        6        6

N2   :       33       34       35       36       37       38       39       40
N1   :        6        6        6        6        6        6        6        6

N2   :       41       42       43       44       45       46       47       48
N1   :        6        6        6        6        6        6        6        6

N2   :       49       50       51       52       53       54       55       56
```

```
N1  :     6     6     6     6     6     6     6     6

N2  :    57    58    59    60    61    62    63    64
N1  :     6     6     6     6     6     6     6     6

N2  :    65    66    67    68    69    70    71    72
N1  :     6     6     6     6     6     6     6     6

N2  :    73    74    75    76    77    78    79    80
N1  :     6     6     6     6     6     6     6     6

N2  :    81    82    83    84    85    86    87    88
N1  :     6     6     6     6     6     6     6     6

N2  :    89    90    91    92    93    94    95    96
N1  :     6     6     6     6     6     6     6     6

N2  :    97    98    99   100   101   102   103   104
N1  :     6     6     6     6     6     6     6     6

N2  :   105   106   107   108   109   110   111   112
N1  :     6     6     6     6     6     6     6     6

N2  :   113   114   115   116   117   118   119   120
N1  :     6     6     6     6     6     6     6     6

N2  :   121   122   123   124   125   126   127   128
N1  :     6     6     6     6     6     6     6     6

N2  :   129   130   131   132   133   134   135   136
N1  :     6     6     6     6     6     6     6     6

N2  :   137   138   139   140   141
N1  :     6     5     5     5     3
________________________________________________________________

ITERATION NUMBER     4

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

-----------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 constant                  26.15108       0.66707     39.20293       0.00000

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

```
                      +-----------------------+

-2 LOG-LIKELIHOOD =     6380.91436790376

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

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             47.22948      7.51714      6.28290   0.00000

-----------------------------------------------------------------------------
 LEVEL 1                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             92.17651      4.93041     18.69552   0.00000

       LEVEL 2 COVARIANCE MATRIX

              constant

constant      47.22948

       LEVEL 2 CORRELATION MATRIX

            constant

constant      1.0000

       LEVEL 1 COVARIANCE MATRIX

              constant

constant      92.17651

       LEVEL 1 CORRELATION MATRIX

            constant

constant      1.0000

_______________________________________________________________________

CONVERGENCE REACHED IN  4 ITERATIONS
```

- From the *Data Summary*, we see that we have complete data for all but the last four air traffic controllers.
- The mean number of successful trials, as obtained from the output for the fixed part of the model, is 26.15108. From exploratory analysis of this data done prior to analysis, it is known that the number of successful trials ranged between 0 and 57.
- From the output for the random part of the model, we see that the variation over measurements is approximately half of the variation over the controllers (47.2298 versus 92.1765), indicating significant differences in the intercepts between controllers.
- A value of 6380.914 was obtained for $-2\ln L$.

In the next model, we will consider a non-linear growth model and compare results with the results obtained here.

Non-linear model for air traffic data

In the example above, a basic model for the analysis of the air traffic data of Kanfer & Ackerman was considered.

In data of this nature, it can be expected that the number of successful landings per trial may be influenced by previous experience. In order to take this into account, we now introduce the order of the measurements into the model. The variable TIME indicates the number of the trial, and TIMESQ the quadratic term equal to TIME*TIME.

To fit such a model, the FIXED and RANDOM2 commands used previously have to be changed to:

```
FIXED = constant time timesq;
RANDOM2 = constant time timesq;
```

These changes can be made by directly editing the KANFER1.PR2 file used in the previous analysis, and saving it as KANFER2.PR2:

```
OPTIONS ;
TITLE=Kanfer and Ackerman data: Non-linear model;
SY=K:\LISREL83\MLEVELEX\KANFER.PSF;
```

```
ID1=time;
ID2=control;
RESPONSE=measure;
FIXED=constant time timesq;
RANDOM1=constant;
RANDOM2=constant time timesq;
```

Partial output for this model follows.

```
ITERATION NUMBER     4

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

-----------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 constant                   1.67806       0.87343       1.92123       0.05470
 time                      11.48966       0.50397      22.79838       0.00000
 timesq                    -1.03583       0.06141     -16.86761       0.00000

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

-2 LOG-LIKELIHOOD =     5095.46585430938

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

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             77.02126     12.98193      5.93296   0.00000
 time    /constant            -24.73085      6.50474     -3.80197   0.00014
 time    /time                 22.66462      4.35831      5.20033   0.00000
 timesq  /constant              2.26771      0.76607      2.96017   0.00307
 timesq  /time                 -2.38073      0.52370     -4.54595   0.00001
 timesq  /timesq                0.27282      0.06566      4.15473   0.00003

-----------------------------------------------------------------------------
 LEVEL 1                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
```

```
--------------------------------------------------------------------------
 constant/constant             9.49299      0.65567     14.47823   0.00000


       LEVEL 2 COVARIANCE MATRIX

              constant        time      timesq

constant      77.02126
time         -24.73085    22.66462
timesq         2.26771    -2.38073     0.27282


       LEVEL 2 CORRELATION MATRIX

            constant      time    timesq

constant      1.0000
time         -0.5919    1.0000
timesq        0.4947   -0.9574    1.0000


       LEVEL 1 COVARIANCE MATRIX

             constant

constant      9.49299


       LEVEL 1 CORRELATION MATRIX

            constant

constant      1.0000
```

```
CONVERGENCE REACHED IN  4 ITERATIONS
```

From the output above, we see that:

- Both the TIME and TIMESQ variables, introduced as fixed effects in this model, have coefficents which are highly significant. From the large coefficient for TIME (11.48966) we see that there is a positive increase of 11 successful landings expected between successive trials.
- From the random part of the output, it follows that all the estimates of the random coefficients on level 2 of the model are significant.

This is also true for all the interaction terms on this level of the model.

- The estimate of the coefficient of the level-1 variance is now 9.49299, compared to the coefficient for the similar term in the previous model of 92.17651, where TIME and TIMESQ were not allowed to vary randomly at the controller level. Variation over measurements has been drastically reduced in the extended model.
- For this model, a $-2\ln L$ value of 5095.466 was obtained, compared to the previous model where $-2\ln L$ was 6380.914. In the previous model, 3 parameters were estimated, while for the extended model 10 parameters were estimated. The inclusion of TIME and TIMESQ significantly improved the fit of the model.

In the two models considered thus far, the measure of cognitive ability available for each controller has not been considered. In the third and final analysis of this data, we will include this measure in the model considered in this section.

Including additional variables in the air traffic data analysis

In the previous example a non-linear model was considered.

Recall that the data set also includes a measure of cognitive ability (composite ASVB score) for each of the 141 air traffic controllers, denoted by ABILITY in the file KANFER.PSF. We now include this variable as a fixed effect in the model.

To add ABILITY to the model, the input file used in the previous example may be edited. After it has been added to the fixed variables field, the input file for this analysis is as follows (KANFER3.PR2).

```
OPTIONS ;
TITLE=Kanfer and Ackerman data: Non-linear model with covariate;
SY=K:\LISREL83\MLEVELEX\KANFER.PSF;
ID1=time;
ID2=control;
RESPONSE=measure;
FIXED=constant time timesq ability;
RANDOM1=constant;
RANDOM2=constant time timesq;
```

Partial output for this analysis is shown below.

```
ITERATION NUMBER     4

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

-----------------------------------------------------------------------------
 COEFFICIENTS             BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 constant                  1.67596       0.78914       2.12379       0.03369
 time                     11.49203       0.50394      22.80458       0.00000
 timesq                   -1.03639       0.06140     -16.87917       0.00000
 ability                   0.03703       0.00519       7.12990       0.00000

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

-2 LOG-LIKELIHOOD =      5052.79195898966

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

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant             57.26989     10.66634      5.36921   0.00000
 time    /constant            -22.58400      5.99981     -3.76412   0.00017
 time    /time                 22.66330      4.35771      5.20074   0.00000
 timesq  /constant              2.09446      0.70778      2.95921   0.00308
 timesq  /time                 -2.38017      0.52358     -4.54592   0.00001
 timesq  /timesq                0.27273      0.06565      4.15457   0.00003

-----------------------------------------------------------------------------
 LEVEL 1                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              9.49057      0.65551     14.47819   0.00000

       LEVEL 2 COVARIANCE MATRIX

              constant        time      timesq
```

```
constant      57.26989
time         -22.58400    22.66330
timesq         2.09446    -2.38017     0.27273


       LEVEL 2 CORRELATION MATRIX

            constant      time    timesq

constant      1.0000
time         -0.6269    1.0000
timesq        0.5300   -0.9574    1.0000


       LEVEL 1 COVARIANCE MATRIX

             constant

constant      9.49057


       LEVEL 1 CORRELATION MATRIX

            constant

constant      1.0000

------------------------------------------------------------------------

CONVERGENCE REACHED IN  4 ITERATIONS
```

From the output above, we see that:

- The effect of ability on the number of successful landings is small in size and positive, indicating a small, and highly significant increase in the expected number of successful landings per trial. The largest effect is that of the order of trials, indicating that a higher number of succesful landings are expected in a following trial.
- From the large coefficient for TIME (11.4920), we see that there is a positive increase of 11 successful landings expected between successive trials.
- From the random part of the output, it follows that all the estimates of the random coefficients on level 2 of the model are significant. This is also true for all the interaction terms at this level of the

model. Note that the coefficient for the intercept at this level has been reduced to 57.26989, compared to 77.02126 in the previous model, where the cognitive ability measure was not included. The introduction of this measure into the model leads to a reduction in the variation of mean outcome scores over controllers.

- The estimate of the coefficient of the level-1 variance is now 9.4057, which is very similar to the estimate obtained for the model without ABILITY. Inclusion of this variable did not contribute to the reduction in the variation over measurements nested within each controller.
- For this model, a $-2\ln L$ value of 5052.792 was obtained, compared to the previous model where $-2\ln L$ was 5095.466. As only one extra parameter was estimated in this model, we conclude that the inclusion of ABILITY in the non-linear model significantly improved the fit of the model.

2.5.3 A multivariate analysis of educational data

The data set used in this section forms part of the data library of the Multilevel Project at the University of London and comes from the Junior School Project (Mortimer, *et al.*, 1988).

Mathematics and language tests were administered in three consecutive years to more than 1000 students from 50 primary schools, which were randomly selected from primary schools maintained by the Inner London Education Authority.

The following variables are available in the data file JSP.PSF:

The school number (SCHOOL) is used as the level-3 identification variable, with Student ID (STUDENT) as the level-2 identification variable. The level-1 units are the scores obtained by a student for the mathematics and language tests, as represented by MATH1 to MATH3 and ENG1 to ENG3.

The aim of this analysis is to examine the variation in test scores over pupils. It would also be interesting to determine the extent to which schools vary with respect to the response variable(s). One of the main benefits of analyzing different responses simultaneously in one multivariate analysis is that the way in which measurements relate to the explanatory variables can be directly explored. The gender, Ravens test score, and

SCHOOL	School code (1 to 50)
CLASS	Class code (1 to 4)
STUDENT	Student ID (1 to 1402)
GENDER	Gender (boy=1; girl=0)
RAVENS	Ravens test score in year 1(score 4–36)
MATH1	Score on mathematics test in year 1 (score 1–40)
MATH2	Score on mathematics test in year 2 (score 1–40)
MATH3	Score on mathematics test in year 3 (score 1–40)
ENG1	Score on language test in year 1 (score 0–98)
ENG2	Score on language test in year 2 (score 0–98)
ENG3	Score on language test in year 3 (score 0–98)
CONSTANT	Intercept, value=1 throughout

an intercept term (GENDER, RAVENS, and CONSTANT, respectively) will be used as explanatory variables in the second of the models described in this section.

As the residual covariance matrices for the second and third levels of the hierarchy are also obtained from this analysis, differences between coefficients of explanatory variables for different responses can be studied.

Finally, each respondent does not have to be measured on each response, as is the case for the data set we consider here, where scores for all three years are not available for all students. This type of model is one in which the MISSING_DEP and MISSING_DAT commands can be used to good advantage, as will be shown. In the case of missing responses, a multilevel multivariate analysis of the responses that are available for respondents can be used to provide information in the estimation of those that are missing.

The first five observations in the JSP.PSF file are shown below.

SCHOOL	STUDENT	GENDER	RAVENS	MATH1	MATH2	MATH3	ENG1	ENG2	ENG3	CONSTANT
1.00	1.00	0.00	23.00	23.00	24.00	23.00	72.00	80.00	39.00	1.00
1.00	2.00	1.00	15.00	14.00	11.00	-9.00	7.00	17.00	-9.00	1.00
1.00	3.00	1.00	22.00	36.00	32.00	39.00	88.00	89.00	83.00	1.00
1.00	4.00	1.00	14.00	24.00	26.00	32.00	12.00	25.00	12.00	1.00
1.00	5.00	0.00	19.00	22.00	23.00	-9.00	67.00	78.00	-9.00	1.00
1.00	6.00	0.00	16.00	19.00	23.00	11.00	52.00	76.00	19.00	1.00
1.00	7.00	1.00	17.00	22.00	22.00	26.00	37.00	68.00	31.00	1.00

1.00	8.00	0.00	21.00	18.00	29.00	28.00	57.00	86.00	40.00	1.00
1.00	9.00	1.00	30.00	30.00	31.00	-9.00	42.00	59.00	-9.00	1.00
1.00	10.00	0.00	25.00	29.00	29.00	-9.00	46.00	79.00	-9.00	1.00
1.00	11.00	0.00	32.00	31.00	28.00	32.00	69.00	84.00	50.00	1.00
1.00	12.00	0.00	15.00	18.00	26.00	-9.00	54.00	74.00	-9.00	1.00
1.00	13.00	0.00	25.00	23.00	-9.00	27.00	63.00	-9.00	39.00	1.00
1.00	14.00	0.00	29.00	39.00	35.00	36.00	83.00	88.00	80.00	1.00
1.00	15.00	0.00	34.00	24.00	30.00	33.00	37.00	44.00	37.00	1.00

One line of information is given for each student. Note that, for the second and fourth students, no mathematics score was available in the third year of the study (MATH3). A missing data code of -9 was assigned to all missing values on both explanatory and response variables in the data set.

In order to perform a multilevel analysis, we need one line of information for each level-1 unit, in this case each of the six test scores. The data manipulation required in creating this revised data file format will be performed automatically by the program in the case of a multivariate model. Six dummy variables are created for each of the explanatory variables used. For the explanatory variable GENDER, for example, the dummy variables GENDER1, GENDER2, . . ., GENDER6 will denote the gender effect for each of the response variables.

Two models will be fitted and discussed:

- A variance decomposition model
- Adding explanatory variables to the model

A variance decomposition model

As a first step, we provide a title for our analysis. No specifications (other than the defaults) are needed for the OPTIONS command, and thus the default convergence criterion and maximum number of iterations will be used. Only the default output file will be produced.

In the case of a multivariate model, only levels 2 and 3 of the hierarchy have to be identified, because the actual measurements for each level-2 unit serve as level-1 units. In this case, we select SCHOOL as the level-3 identification variable. The student ID (STUDENT) is selected as the level-2 identification variable.

We choose the six variables MATH1 to MATH3 and ENG1 to ENG3 as response variables.

In this model, we wish to start our analysis of the data with a look at the differences in intercepts over students and schools. Note that in this analysis, no RANDOM commands should be included. These commands are only required when running a multivariate model when more variables than just the intercept is required at higher levels of the hierarchy. If no RANDOM commands are included, the inclusion of an intercept term on higher levels is automatically assumed by the program and dummy variables for the explanatory predictors given in the FIXED command are generated.

The resulting input file (JSP1.PR2) is given below.

Note, that the optional MISSING_DEP command is used to identify -9 as the missing data code for all response variables.

```
OPTIONS ;
TITLE=Multivariate Analysis of Education Data;
SY=K:\LISREL83\MLEVELEX\JSP.PSF;
ID2=student;
ID3=school;
RESPONSE=math1 math2 math3 eng1 eng2 eng3;
FIXED=constant;
MISSING_DEP=-9;
```

The output for this model is written to the default output file JSP1.OUT. Partial output is given below.

```
                           +--------------+
                           | DATA SUMMARY |
                           +--------------+

NUMBER OF LEVEL 3 UNITS :       49
NUMBER OF LEVEL 2 UNITS :     1192
NUMBER OF LEVEL 1 UNITS :     6472

ID3 :        1        2        3        4        5        6        7        8
N2  :       34       13       21       24       29       24       15       31
N1  :      184       72       96      144      166      120       78      174

ID3 :        9       10       11       12       13       14       15       16
```

```
N2  :      22      14      12      28      25      15      24      19
N1  :     130      48      70     154     138      86     106     106

ID3 :      17      18      19      20      21      22      23      24
N2  :       7      20      17      15      32      16      20      21
N1  :      40     110      88      84     184      92     116     126

ID3 :      25      26      27      28      29      30      31      32
N2  :      18      29      27      18      25      37      36      26
N1  :      94     156     146      96     124     182     214     148

ID3 :      33      34      35      36      37      38      39      40
N2  :      46      34      19      32      22      14      16      13
N1  :     262     184     106     178     118      78      90      68

ID3 :      41      42      43      44      45      46      47      48
N2  :      13      47       6      17      19      37      72      25
N1  :      70     262      32      88     102     204     412     144

ID3 :      49
N2  :      46
N1  :     202
```

[Output omitted]

```
Multivariate Analysis of Education Data

ITERATION NUMBER     8

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

-----------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 constan1                  24.90370      0.33546      74.23792       0.00000
 constan2                  24.87234      0.40108      62.01311       0.00000
 constan3                  30.04909      0.37761      79.57736       0.00000
 constan4                  47.15338      1.32158      35.67945       0.00000
 constan5                  64.96594      1.22017      53.24352       0.00000
 constan6                  40.71988      1.38296      29.44399       0.00000

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

```
-2 LOG-LIKELIHOOD =     45991.2578958991

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

------------------------------------------------------------------------------
 LEVEL 3                          TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
------------------------------------------------------------------------------
 math1    /math1                  3.31028       1.10053      3.00791   0.00263
 math2    /math1                  2.29203       1.09364      2.09579   0.03610
 math2    /math2                  5.23047       1.57789      3.31486   0.00092
 math3    /math1                  2.36877       1.02957      2.30075   0.02141
 math3    /math2                  3.13485       1.25421      2.49945   0.01244
 math3    /math3                  4.79640       1.39721      3.43285   0.00060
 eng1     /math1                  9.95162       3.74223      2.65927   0.00783
 eng1     /math2                  9.49602       4.22124      2.24958   0.02448
 eng1     /math3                  9.95900       4.00265      2.48810   0.01284
 eng1     /eng1                  59.32680      17.15964      3.45735   0.00055
 eng2     /math1                  9.93290       3.47650      2.85716   0.00427
 eng2     /math2                 11.41833       4.08446      2.79555   0.00518
 eng2     /math3                 11.59906       3.87949      2.98984   0.00279
 eng2     /eng1                  42.49228      14.15024      3.00294   0.00267
 eng2     /eng2                  53.08595      14.64628      3.62453   0.00029
 eng3     /math1                 10.16830       3.82941      2.65532   0.00792
 eng3     /math2                 10.71620       4.43202      2.41790   0.01561
 eng3     /math3                 13.71315       4.43882      3.08937   0.00201
 eng3     /eng1                  45.00410      15.55552      2.89313   0.00381
 eng3     /eng2                  51.61419      15.11775      3.41415   0.00064
 eng3     /eng3                  71.08796      18.81475      3.77831   0.00016

------------------------------------------------------------------------------
 LEVEL 2                          TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
------------------------------------------------------------------------------
 math1    /math1                 47.17671       1.99173     23.68631   0.00000
 math2    /math1                 38.63798       1.91420     20.18492   0.00000
 math2    /math2                 55.45831       2.35660     23.53314   0.00000
 math3    /math1                 31.21276       1.65633     18.84448   0.00000
 math3    /math2                 36.54802       1.84653     19.79278   0.00000
 math3    /math3                 41.33874       1.86018     22.22293   0.00000
 eng1     /math1                109.03175       5.79022     18.83032   0.00000
 eng1     /math2                109.44780       6.15532     17.78100   0.00000
 eng1     /math3                 88.00030       5.36198     16.41191   0.00000
 eng1     /eng1                 549.41232      23.15761     23.72492   0.00000
 eng2     /math1                 88.73168       4.92512     18.01613   0.00000
 eng2     /math2                 94.99391       5.32980     17.82315   0.00000
 eng2     /math3                 79.52971       4.69229     16.94901   0.00000
 eng2     /eng1                 388.36174      18.26625     21.26117   0.00000
 eng2     /eng2                 409.27297      17.36258     23.57213   0.00000
 eng3     /math1                 94.37662       5.19417     18.16973   0.00000
```

eng3	/math2	99.80683	5.60600	17.80356	0.00000
eng3	/math3	86.38268	4.95170	17.44505	0.00000
eng3	/eng1	382.76743	18.73683	20.42861	0.00000
eng3	/eng2	317.45309	15.98496	19.85949	0.00000
eng3	/eng3	425.49404	18.97401	22.42510	0.00000

LEVEL 3 COVARIANCE MATRIX

	math1	math2	math3	eng1	eng2
math1	3.31028				
math2	2.29203	5.23047			
math3	2.36877	3.13485	4.79640		
eng1	9.95162	9.49602	9.95900	59.32680	
eng2	9.93290	11.41833	11.59906	42.49228	53.08595
eng3	10.16830	10.71620	13.71315	45.00410	51.61419

	eng3
eng3	71.08796

LEVEL 3 CORRELATION MATRIX

	math1	math2	math3	eng1	eng2	eng3
math1	1.0000					
math2	0.5508	1.0000				
math3	0.5945	0.6259	1.0000			
eng1	0.7101	0.5391	0.5904	1.0000		
eng2	0.7493	0.6852	0.7269	0.7572	1.0000	
eng3	0.6629	0.5557	0.7426	0.6930	0.8402	1.0000

LEVEL 2 COVARIANCE MATRIX

	math1	math2	math3	eng1	eng2
math1	47.17671				
math2	38.63798	55.45831			
math3	31.21276	36.54802	41.33874		
eng1	109.03175	109.44780	88.00030	549.41232	
eng2	88.73168	94.99391	79.52971	388.36174	409.27297
eng3	94.37662	99.80683	86.38268	382.76743	317.45309

	eng3
eng3	425.49404

```
      LEVEL 2 CORRELATION MATRIX

               math1     math2     math3     eng1      eng2      eng3

math1          1.0000
math2          0.7554    1.0000
math3          0.7068    0.7633    1.0000
eng1           0.6772    0.6270    0.5839    1.0000
eng2           0.6386    0.6305    0.6114    0.8190    1.0000
eng3           0.6661    0.6497    0.6513    0.7917    0.7607    1.0000
```

```
CONVERGENCE REACHED IN  8 ITERATIONS
```

- From the *Data Summary*, we see that data from 1192 students from 49 schools were used in the analysis. The number of level-1 units (*i.e.*, measurements) per school ranged from 32 in the case of school number 43 to 412 for school number 47.
- From the output for the fixed part of the model, it can be seen that all six fixed effects are highly significant, indicating significant differences in the six measurements over the schools.
- There is significant variation in the mean effects of the six response variables over both schools and students. The variation over schools (level 3) is higher than over students. At the student level, the largest variation is in the mean effects for the language tests, ranging between 59.326 (for measurements from the first year), to 71.088 (for measurements from the third year of the study). The same tendency is observed for the mean effects of language test scores over schools, with the highest variation recorded for the language test score from the third year of the study.
- The $-2\ln L$ value for this analysis at convergence after 7 iterations was 45991.2579.

Adding explanatory variables to the model

Using the model discussed in the previous section as point of departure, we now proceed to add fixed effects to the model. The variables GENDER and RAVENS, indicating the gender of a student and the student's score on the Ravens test in the first year of the study, respectively, are added to the FIXED command.

This is how the FIXED command should look:

```
FIXED = CONSTANT GENDER RAVENS;
```

To make this change, the input file previously used (JSP1.PR2) can be edited directly, and then saved as the new input file JSP2.PR2:

```
OPTIONS ;
TITLE=Multivariate Analysis of Education Data;
SY=K:\LISREL83\MLEVELEX\JSP.PSF;
ID2=student;
ID3=school;
RESPONSE=math1 math2 math3 eng1 eng2 eng3;
FIXED=constant gender ravens;
MISSING_DEP=-9;
```

For this model the following output was obtained.

```
ITERATION NUMBER     7

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

COEFFICIENTS	BETA-HAT	STD.ERR.	Z-VALUE	PR > |Z|
constan1	7.68926	0.80046	9.60607	0.00000
constan2	6.48252	0.90269	7.18136	0.00000
constan3	14.84541	0.84586	17.55061	0.00000
constan4	2.64745	2.90410	0.91163	0.36197
constan5	28.95062	2.57309	11.25128	0.00000
constan6	-2.58204	2.74794	-0.93963	0.34741
gender1	-0.48513	0.33815	-1.43466	0.15138
gender2	-0.79986	0.37167	-2.15205	0.03139
gender3	-0.45790	0.34277	-1.33588	0.18159
gender4	-10.57862	1.20875	-8.75172	0.00000
gender5	-9.21809	1.06819	-8.62960	0.00000
gender6	-7.02657	1.10494	-6.35926	0.00000
ravens1	0.69639	0.02943	23.66040	0.00000

```
ravens2              0.74945      0.03248     23.07353      0.00000
ravens3              0.61505      0.03029     20.30821      0.00000
ravens4              1.97944      0.10546     18.77035      0.00000
ravens5              1.61390      0.09328     17.30225      0.00000
ravens6              1.86293      0.09754     19.09957      0.00000

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

-2 LOG-LIKELIHOOD =     45321.6134304450

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

-----------------------------------------------------------------------------
LEVEL 3                         TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
math1   /math1                  2.35970       0.77231      3.05538   0.00225
math2   /math1                  1.47571       0.79018      1.86756   0.06182
math2   /math2                  4.59531       1.29317      3.55353   0.00038
math3   /math1                  1.45284       0.73222      1.98414   0.04724
math3   /math2                  2.34726       0.97705      2.40240   0.01629
math3   /math3                  3.94231       1.11612      3.53215   0.00041
eng1    /math1                  5.89266       2.50131      2.35583   0.01848
eng1    /math2                  5.10887       3.00832      1.69825   0.08946
eng1    /math3                  5.75844       2.83649      2.03013   0.04234
eng1    /eng1                  40.46361      11.99950      3.37211   0.00075
eng2    /math1                  5.42374       2.23795      2.42353   0.01537
eng2    /math2                  6.60379       2.82875      2.33453   0.01957
eng2    /math3                  7.26912       2.69805      2.69422   0.00706
eng2    /eng1                  22.52145       9.10418      2.47375   0.01337
eng2    /eng2                  33.03809       9.66632      3.41786   0.00063
eng3    /math1                  5.14461       2.49143      2.06492   0.03893
eng3    /math2                  5.67697       3.12932      1.81413   0.06966
eng3    /math3                  9.01133       3.15762      2.85384   0.00432
eng3    /eng1                  22.93497      10.12809      2.26449   0.02354
eng3    /eng2                  30.00480       9.77962      3.06810   0.00215
eng3    /eng3                  47.71171      12.97810      3.67632   0.00024

-----------------------------------------------------------------------------
LEVEL 2                         TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
math1   /math1                 31.84368       1.34833     23.61706   0.00000
math2   /math1                 22.14572       1.23554     17.92388   0.00000
math2   /math2                 37.70830       1.61030     23.41693   0.00000
math3   /math1                 17.66778       1.10211     16.03093   0.00000
math3   /math2                 21.92750       1.24318     17.63828   0.00000
math3   /math3                 29.34575       1.34391     21.83617   0.00000
```

eng1	/math1	65.59655	3.92248	16.72322	0.00000
eng1	/math2	62.26680	4.16530	14.94895	0.00000
eng1	/math3	49.65248	3.75896	13.20909	0.00000
eng1	/eng1	406.75188	17.17856	23.67788	0.00000
eng2	/math1	53.71682	3.40837	15.76025	0.00000
eng2	/math2	56.85501	3.69230	15.39826	0.00000
eng2	/math3	48.64973	3.37484	14.41540	0.00000
eng2	/eng1	272.02299	13.42728	20.25897	0.00000
eng2	/eng2	314.33951	13.37559	23.50099	0.00000
eng3	/math1	53.74137	3.50517	15.33203	0.00000
eng3	/math2	55.52744	3.78879	14.65570	0.00000
eng3	/math3	50.44039	3.44613	14.63681	0.00000
eng3	/eng1	256.17082	13.47180	19.01534	0.00000
eng3	/eng2	214.35205	11.70207	18.31744	0.00000
eng3	/eng3	310.90414	14.08313	22.07635	0.00000

LEVEL 3 COVARIANCE MATRIX

	math1	math2	math3	eng1	eng2
math1	2.35970				
math2	1.47571	4.59531			
math3	1.45284	2.34726	3.94231		
eng1	5.89266	5.10887	5.75844	40.46361	
eng2	5.42374	6.60379	7.26912	22.52145	33.03809
eng3	5.14461	5.67697	9.01133	22.93497	30.00480

	eng3
eng3	47.71171

LEVEL 3 CORRELATION MATRIX

	math1	math2	math3	eng1	eng2	eng3
math1	1.0000					
math2	0.4481	1.0000				
math3	0.4763	0.5515	1.0000			
eng1	0.6030	0.3747	0.4559	1.0000		
eng2	0.6143	0.5360	0.6369	0.6160	1.0000	
eng3	0.4849	0.3834	0.6571	0.5220	0.7557	1.0000

LEVEL 2 COVARIANCE MATRIX

math1	math2	math3	eng1	eng2

```
math1         31.84368
math2         22.14572     37.70830
math3         17.66778     21.92750     29.34575
eng1          65.59655     62.26680     49.65248    406.75188
eng2          53.71682     56.85501     48.64973    272.02299    314.33951
eng3          53.74137     55.52744     50.44039    256.17082    214.35205

                  eng3

eng3         310.90414


       LEVEL 2 CORRELATION MATRIX

               math1      math2      math3       eng1       eng2       eng3

math1         1.0000
math2         0.6391     1.0000
math3         0.5780     0.6592     1.0000
eng1          0.5764     0.5028     0.4545     1.0000
eng2          0.5369     0.5222     0.5065     0.7607     1.0000
eng3          0.5401     0.5128     0.5281     0.7204     0.6857     1.0000
-----------------------------------------------------------------------

CONVERGENCE REACHED IN  7 ITERATIONS
```

Results for the fixed part of the model show that:

- The expected score of girls (GENDER=0) for all test scores is higher than the expected score for boys (GENDER=1). For boys (GENDER=1), the coefficients GENDER4 to GENDER6 are negative and highly significant. In the case of the mathematics test scores (MATH1 to MATH3 as represented by GENDER1 to GENDER3) the effects are smaller. Keep in mind that the range of scores differed between the mathematics and language tests.
- The effects of the RAVENS test are positive and highly significant for all six response variables, with the largest effects for the mathematics tests. An increase of one unit in the RAVENS test score implies an expected increase in the third year language test of 0.61 and an expected increase of 1.86 for the expected third year mathematics test score.

Results for the random part of the model are consistent with the results for the previous model fitted, with larger variation for the three language tests and, in general, more variation over schools than over students.

The $-2\ln L$ recorded for this model is 45321.61343. When compared to the $-2\ln L$ of 45991.2579 obtained previously, a marked decrease is observed. In the first model, 48 parameters were estimated, compared to 60 parameters for the model discussed here. The introduction of the GENDER and RAVENS variables have contributed significantly to the explanation of variance in the response variables.

2.5.4 Analysis of CPC survey data

In this section, data from the March 1995 Current Population Survey are used. The data set is a subset of data obtained from the Data Library at the Department of Statistics at UCLA. A small number of demographic variables for two occupation groups was extracted, and all analyses are based on unweighted data.

Only respondents between the ages of twenty-one and sixty-five, who held full time positions in 1994 and had an annual income of US$1 or more were considered. The two groups we will focus on here are defined as follows:

Educational sector	Respondents with professional specialty in the educational sector
Construction sector	Operators, fabricators, and laborers in the construction sector

The variable GROUP in the PRELIS system file INCOME.PSF represents the groups, with GROUP=0 for the respondents in the construction sector and GROUP=1 for respondents in the educational sector. A 3–D bar chart showing the sample sizes of the two groups,[3] is shown in Figure 2.3.

Other demographic variables and their codes are:

[3] As obtained with the Graphs option in the Windows version

Figure 2.3 Bar Chart of the GROUP Variable in INCOME.PSF

GENDER	0 = female; 1 = male
AGE	Age in single years
MARITAL	1 = married; 0 = other
HOURS	Hours worked during last week at all jobs
CITIZEN	1 for native Americans, 0 for all foreign born respondents
INCOME	The natural logarithm of the personal income during 1994
DEGREE	1 for respondents with master's degrees, professional school degree, or doctoral degree; 0 otherwise

Respondents were from 9 regions of the USA, and the state of residence was also given. The variables REGION and STATE represent this information. The full description of the regions and states within regions is presented in Table 2.2. On the respondent level, the variable PERSON is a respondent identity number.

Finally, the variable CONSTANT denotes the intercept term and has a value of 1 for all respondents. The variable INCOME will be used as response variable in all analyses.

The models considered here are:

Table 2.2 Region and State codes

Region	State Code	State Name
New England region	11	Maine
(REGION = 1)	12	New Hampshire
	13	Vermont
	14	Massachusetts
	15	Rhode Island
	16	Connecticut
Middle Atlantic region	21	New York
(REGION=2)	22	New Jersey
	23	Pennsylvania
East North Central region	31	Ohio
(REGION=3)	32	Indiana
	33	Illinois
	34	Michigan
	35	Wisconsin
West North Central region	41	Minnesota
(REGION=4)	42	Iowa
	43	Missouri
	44	North Dakota
	45	South Dakota
	46	Nebraska
	47	Kansas
South Atlantic region	51	Delaware
(REGION=5)	52	Maryland
	53	District of Columbia
	54	Virginia
	55	West Virginia
	56	North Carolina
	57	South Carolina
	58	Georgia
	59	Florida
East South Central region	61	Kentucky
(REGION=6)	62	Tennessee
	63	Alabama
	64	Mississippi
West South Central region	71	Arkansas
(REGION=7)	72	Louisiana
	73	Oklahoma
	74	Texas
Mountain region	81	Montana
(REGION=8)	82	Idaho
	83	Wyoming
	84	Colorado
	85	New Mexico
	86	Arizona
	87	Utah
	88	Nevada
Pacific region	91	Washington
(REGION=9)	92	Oregon
	93	California
	94	Alaska
	95	Hawaii

- A 3-level model for the combined group, using INCOME.PSF
- A similar model for the education sector only, using a subset of the data
- A similar model for the construction sector only, using a subset of the data

3-level model for subset of CPC survey data

The data set used, as described in the previous section, is contained in the PSF file INCOME.PSF. The variable labels (first two lines) and the first fifteen data records of this file are shown below.

```
  REGION    STATE      AGE   GENDER  MARITAL    HOURS
 CITIZEN   PERSON CONSTANT   DEGREE    GROUP   INCOME
  1 11 30  1  1 40  1  14751 1  0  0  9.6291
  1 11 41  1  1 40  1  14768 1  0  0 10.6041
  1 11 46  0  1 36  1  14781 1  0  0 10.1897
  1 11 30  1  1 40  1  14813 1  0  0 10.9618
  1 11 46  1  1 10  0  14825 1  0  0  9.8009
  1 11 63  1  1 25  1  14830 1  0  0 10.2071
  1 11 38  1  0 75  1  14833 1  0  0 10.0587
  1 11 33  1  1 40  1  14861 1  0  0 10.2400
  1 11 25  1  0 60  1  14890 1  0  0 10.2324
  1 11 43  1  0 40  1  14949 1  0  0 10.0648
  1 11 34  1  1  6  1  14984 1  0  0 10.8208
  1 11 46  1  1  0  1  14999 1  0  0  9.6803
  1 11 41  1  1 39  1  15009 1  0  0  9.9998
  1 11 61  1  1 32  1  15012 1  0  0  9.5468
  1 11 51  1  1 25  1  15068 1  0  0 10.4913
```

Again, we start creating the input file by accepting the defaults for the maximum number of iterations, the convergence criterion, and the output options, then provide the title for the analysis (optional).

As all respondents are nested within state of residence, and states are in turn nested within the nine regions, we select REGION as the variable for the level-3 identification variable field. The variables STATE and PERSON are selected as level-2 and level-1 identification variables, respectively.

Next, we select INCOME, representing the natural logarithm of personal income, as the response variable for this analysis. The variables AGE,

GENDER, MARITAL, HOURS, CITIZEN, CONSTANT, DEGREE, and GROUP are all entered into the model as fixed effects.

Finally, the variable CONSTANT, representing the intercept term, is identified as a random effect on all levels of the hierarchy.

As a result, the input file (INCOME.PR2) looks like this:

```
OPTIONS ;
TITLE=Analysis of CPC data: combined group;
SY=K:\LISREL83\MLEVELEX\INCOME.PSF;
ID1=person;
ID2=state;
ID3=region;
RESPONSE=income;
FIXED=age gender marital hours citizen constant degree group;
RANDOM1=constant;
RANDOM2=constant;
RANDOM3=constant;
```

The data summary and output for the final iteration are given below.

```
                              +--------------+
                              | DATA SUMMARY |
                              +--------------+

NUMBER OF LEVEL 3 UNITS :          9
NUMBER OF LEVEL 2 UNITS :         51
NUMBER OF LEVEL 1 UNITS :       6062

ID3  :       1       2       3       4       5       6       7       8
N2   :       6       3       5       7       9       4       4       8
N1   :     545     862     785     521    1095     291     598     704

ID3  :       9
N2   :       5
N1   :     661
```

[Output omitted]

```
ITERATION NUMBER      3
```

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


-----------------------------------------------------------------------------
 COEFFICIENTS             BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 age                      0.01636       0.00101      16.17417       0.00000
 gender                   0.23710       0.02853       8.31169       0.00000
 marital                  0.08456       0.02243       3.76975       0.00016
 hours                    0.01344       0.00065      20.63542       0.00000
 citizen                  0.28652       0.03449       8.30714       0.00000
 constant                 8.19488       0.06867     119.34530       0.00000
 degree                   0.41226       0.02846      14.48697       0.00000
 group                    0.19798       0.03135       6.31519       0.00000


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

-2 LOG-LIKELIHOOD =      14222.6158872357


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


-----------------------------------------------------------------------------
 LEVEL 3                        TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.00783      0.00472      1.65774   0.09737

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.00522      0.00250      2.08936   0.03668

-----------------------------------------------------------------------------
 LEVEL 1                        TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.60688      0.01107     54.83990   0.00000


       LEVEL 3 COVARIANCE MATRIX

            constant

constant     0.00783
```

```
      LEVEL 3 CORRELATION MATRIX

           constant

constant      1.0000

      LEVEL 2 COVARIANCE MATRIX

           constant

constant     0.00522

      LEVEL 2 CORRELATION MATRIX

           constant

constant      1.0000

      LEVEL 1 COVARIANCE MATRIX

           constant

constant      0.6069

      LEVEL 1 CORRELATION MATRIX

           constant

constant      1.0000
______________________________________________________________________

CONVERGENCE REACHED IN  3 ITERATIONS
```

From the output given above, we see that:

- The nine regions had between 291 and 1095 respondents, nested within states. The smallest number of level-2 units within a level-3 unit was 3, for the middle Atlantic region which included only New York, New Jersey, and Pennsylvania.
- All the fixed effects were highly significant. The coefficient for the variable CONSTANT, representing the mean income, was 8.19488. Since the response variable is the natural logarithm of a respondent's annual income, this number translates to a mean income of

$$\exp(8.19488 + 21(0.01636) + 40(0.01344)) = \$8,743$$

for a respondent from the construction sector who is 21 years of age, working 40 hours per week, unmarried, without a higher degree, and not a USA citizen. Although the size of the coefficients is quite small, it should be kept in mind that the natural logarithm of income is used as response variable. The relatively large positive coefficients for GENDER (0.23710), CITIZEN (0.28652), and DEGREE (0.41226) indicate that males, citizens of the USA, and respondents with a high education level tend to earn more when other variables are held constant. A comparison of two respondents with different demographic profiles as given below illustrates this point.

Respondent 1	**Respondent 2**
AGE=30	AGE=30
HOURS=40	HOURS=40
GROUP=1	GROUP=1
MARITAL=0	MARITAL=0
GENDER=0	GENDER=1
CITIZEN=0	CITIZEN=1
DEGREE=0	DEGREE=1

The first respondent's expected income is calculated as

Expected income =

$$\exp[8.19488 + 30(0.01636) + 40(0.01344) + 0.19798] =$$
$$\exp[9.42126] = \$12,348$$

while the expected income of the second respondent is

Expected income =

$$\exp[8.19488 + 30(0.01636) + 40(0.01344) + 0.19798 + 0.23710+$$
$$0.28652 + 0.41226] = \exp[10.35714] = \$31,481$$

- Income varies most over the respondents (level-1 units), and least over the nine regions (level-3) units as we can see from the variances at these levels, given as 0.60688 and 0.00783, respectively.

In order to take a closer look at the relationships within the construction and educational sectors, two separate data sets will be created for these groups and similar models fitted in the next two examples. In the next section, a model for respondents from the education sector will be considered.

Three-level model for the educational sector

In the previous example, a 3-level model for the combined education and construction sector respondents from the 1995 CPC survey data was considered. In order to study effects for the educational sector only, a subset of the data in the file INCOME.PSF is used.

We select respondents belonging to the educational sector by using the PRELIS SC (select cases) command and select only those cases with GROUP=1.

```
CREATE A SUBSET OF THE FULL DATASET
SY=L:\LISREL83\MLEVELEX\INCOME.PSF
SC GROUP=1
OU XM
```

The new dataset is saved as EDUC.PSF.

The input file for the analysis is exactly the same as in the previous example, with one exception – the variable GROUP is not included as a fixed effect, as this variable now has the value 1 for all respondents in the data set EDUC.PSF.

The input file for this analysis (EDUC.PR2) is shown below.

```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=STANDARD ;
TITLE=Analysis of CPC data: education sector group;
SY=L:\LISREL83\MLEVELEX\EDUC.PSF;
ID1=person;
ID2=state;
ID3=region;
RESPONSE=income;
FIXED=age gender marital hours citizen constant degree;
RANDOM1=constant;
RANDOM2=constant;
RANDOM3=constant;
```

Partial output for this analysis follows.

```
ITERATION NUMBER     4

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

-----------------------------------------------------------------------------
 COEFFICIENTS             BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
-----------------------------------------------------------------------------
 age                       0.02001      0.00129      15.49123       0.00000
 gender                    0.20943      0.02759       7.58961       0.00000
 marital                  -0.01506      0.02851      -0.52812       0.59741
 hours                     0.01458      0.00079      18.55257       0.00000
 citizen                   0.17746      0.05042       3.51950       0.00043
 constant                  8.38120      0.07850     106.76220       0.00000
 degree                    0.39622      0.02693      14.71524       0.00000

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

-2 LOG-LIKELIHOOD =     6991.19474301050

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

-----------------------------------------------------------------------------
 LEVEL 3                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.00502      0.00445      1.12840   0.25915

-----------------------------------------------------------------------------
 LEVEL 2                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.01283      0.00498      2.57744   0.00995

-----------------------------------------------------------------------------
 LEVEL 1                        TAU-HAT      STD.ERR.     Z-VALUE   PR > |Z|
-----------------------------------------------------------------------------
 constant/constant              0.50458      0.01267     39.83364   0.00000

       LEVEL 3 COVARIANCE MATRIX
```

```
            constant

constant     0.00502

       LEVEL 3 CORRELATION MATRIX

            constant

constant      1.0000

       LEVEL 2 COVARIANCE MATRIX

            constant

constant     0.01283

       LEVEL 2 CORRELATION MATRIX

            constant

constant      1.0000

       LEVEL 1 COVARIANCE MATRIX

            constant

constant      0.5046

       LEVEL 1 CORRELATION MATRIX

            constant

constant      1.0000
```

```
CONVERGENCE REACHED IN  4 ITERATIONS
```

For the education sector, the following results are obtained.

- All fixed effects are highly significant and positive, with the exception of the coefficient for marital status (MARITAL). For this group, the coefficient for the intercept (CONSTANT) is 8.38120. From the results of the previous analysis, the intercept for the group of respondents with GROUP=1 was $8.19488 + 0.19798 = 8.39286$, with all other variables held constant. In general, the same trends are observed for the combined and educational sector only groups: larger positive coefficients are obtained for the variables GENDER, CITIZEN, and DEGREE. Using the same two respondent profiles as in the previous example, with the exception of the GROUP variable which was not included in this analysis, we calculate the expected incomes of the two repondents.

Respondent 1	**Respondent 2**
AGE=30	AGE=30
HOURS=40	HOURS=40
GROUP=1	GROUP=1
MARITAL=0	MARITAL=0
GENDER=0	GENDER=1
CITIZEN=0	CITIZEN=1
DEGREE=0	DEGREE=1

The first respondent's expected income is calculated as

Expected income =

$$\exp[8.38120+30(0.02001)+40(0.01458)] = \exp[9.5647] = \$14,252$$

while the expected income of the second respondent is

Expected income =

$$\exp[8.38120+30(0.02001)+40(0.01458)+0.20943+0.17746+0.39622] =$$

$$\exp[10.34781] = \$31,118$$

The difference between the expected income of these respondents is slightly smaller when only the educational sector is considered.

- For this sector, the mean income varies little over the nine regions. The variation at level 3 of the model is smaller than for the combined model (0.00783 versus 0.00502) and is not significant at any commonly used level of significance. The conclusion may be reached that most of the variation previously observed at a region level (level 3) was due to differences between the two sectors. Variation at levels 1 and 2 remained significant.

In the last example, we will consider a similar model for the construction sector only.

Three-level model for the construction sector

In the previous two examples, a model for the combined education and construction sectors and a model for the educational sector only were fitted to the 1995 CPC survey data . As a final example, we consider a separate model for those respondents active in the construction sector during 1994.

In order to study effects for the construction sector only, a subset of the data in the file INCOME.PSF is used. As before, we select respondents belonging to the construction sector by running a small PRELIS input file using the select cases (SC) command:

```
CREATE A SUBSET OF THE FULL DATASET
SY=L:\LISREL83\MLEVELEX\INCOME.PSF
SC GROUP=0
OU XM
```

The resulting file CONS.PSF contains only those cases with GROUP=0.

The input file is exactly the same as in the previous example, with one exception – the variable GROUP is not included as a fixed effect, as this variable now has the value 0 for all respondents in the data set CONS.PSF.

The input file for this analysis is shown below.

```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=STANDARD ;
TITLE=Analysis of CPC data: construction sector only;
SY=L:\LISREL83\MLEVELEX\CONS.PSF;
ID1=person;
ID2=state;
ID3=region;
RESPONSE=income;
FIXED=age gender marital hours citizen constant degree;
RANDOM1=constant;
RANDOM2=constant;
RANDOM3=constant;
```

Partial output for this analysis is given below.

```
ITERATION NUMBER     6

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

----------------------------------------------------------------------------
 COEFFICIENTS            BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
----------------------------------------------------------------------------
 age                      0.01208       0.00157       7.69715       0.00000
 gender                   0.39847       0.09723       4.09819       0.00004
 marital                  0.20337       0.03505       5.80180       0.00000
 hours                    0.01183       0.00110      10.73244       0.00000
 citizen                  0.32688       0.04805       6.80230       0.00000
 constant                 8.15061       0.13001      62.69199       0.00000
 degree                   0.21725       0.22626       0.96018       0.33697

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

-2 LOG-LIKELIHOOD =     7087.74890333658

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

----------------------------------------------------------------------------
 LEVEL 3                       TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
----------------------------------------------------------------------------
 constant/constant             0.01203      0.00747      1.61013   0.10737

----------------------------------------------------------------------------
 LEVEL 2                       TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
----------------------------------------------------------------------------
 constant/constant             0.00486      0.00408      1.18984   0.23411

----------------------------------------------------------------------------
 LEVEL 1                       TAU-HAT      STD.ERR.      Z-VALUE   PR > |Z|
----------------------------------------------------------------------------
 constant/constant             0.70303      0.01880     37.39738   0.00000

       LEVEL 3 COVARIANCE MATRIX
```

```
          constant

constant    0.01203

       LEVEL 3 CORRELATION MATRIX

          constant

constant     1.0000

       LEVEL 2 COVARIANCE MATRIX

          constant

constant    0.00486

       LEVEL 2 CORRELATION MATRIX

          constant

constant     1.0000

       LEVEL 1 COVARIANCE MATRIX

          constant

constant     0.7030

       LEVEL 1 CORRELATION MATRIX

          constant

constant     1.0000
```

```
CONVERGENCE REACHED IN  6 ITERATIONS
```

The following conclusions may be reached from the output given above:

- When the fixed effects for this model is compared to those obtained for the education sector, the coefficients for AGE and HOURS are smaller. The age of a respondent in the construction sector and the number of hours worked will result in a smaller expected increase in annual personal income. In contrast with the education sector, where the effect of marital status (MARITAL) was not significant, a respondent in the construction sector is likely to earn more when the respondent is married, with all other variables held constant.

- The coefficient for CITIZEN is approximately twice that of a respondent from the education sector (0.32688 versus 0.17746). The mean income, with all other variables held fixed at 0, is 8.15061 (as natural logarithm). With all other variables held constant, this translates into a $\$1,339$ difference in baseline income between citizens and non-citizens in the construction sector. The baseline expected income for a US citizen working in the construction sector can be calculated as

  Expected baseline =

  $$\exp(8.15061 + 0.32688) = \$4,805$$

  For a US citizen in the education sector, the expected baseline income is calculated as

  Expected baseline =

  $$\exp(8.38120 + 0.17746) = \$5,211$$

- Again, the largest coefficients obtained are for GENDER, CITIZEN, and DEGREE. Where the coefficient for GENDER was 0.20943 in the education sector, the coefficient for the construction sector is approximately twice that, at 0.39847. From the output above, it is seen that the coefficient for DEGREE is not significant. A closer examination of the data, using PRELIS data screening features, reveals that only 14 respondents have a high level of education (masters, professional, or PhD degree). If the same model is fitted without the degree predictor, all other estimated parameter values basically remain unchanged.
- Turning to the random effects, we see that the only significant variation is over respondents. At both state and division level, the variation is not significant. From this, combined with the results of the analysis for the education sector (see p. 2.5.4), we conclude that

the significant variation at level 2 seen in the combined model (see p. 2.5.4), is probably due to the differences between these two groups and the significant variation at this level for the education sector.

References

Anderson, T.W., & Rubin, H. (1956)
Statistical inference in factor analysis.
In *Proceedings of the Third Berkeley Symposium*, Volume V. Berkeley: University of California Press.

Bollen, K. A. (1995)
Structural equation models that are nonlinear in latent variables: A least squares estimator.
In P.M. Marsden (Ed.): *Sociological Methodology 1995*. Cambridge, MA: Blackwell.

Bollen, K.A. (1996)
An alternative two stage least squares (2SLS) estimator for latent variable equations.
Psychometrika, **61**, 109–121

Bollen, K.A. & Paxton, P. (1998)
Two-stage least squares estimation of interaction effects.
Pp. 125–151 in G.A. Marcoulides & R.E. Schumacker (Eds): *Interaction and nonlinear effects in structural equation models*. Mahwah, N.J: Lawrence Erlbaum Associates, Publishers.

Browne, M.W. (1974)
Generalized least squares estimators in the analysis of covariance structures.
South African Statistical Journal, **8**, 1–24.

Browne, M.W. (1977)
Generalized least squares estimators in the analysis of covariance structures.
Pp. 205–226 in D.J. Aigner and A.S. Goldberger (Eds.): *Latent variables in socio-economic models*. Amsterdam: North-Holland.

Browne, M.W. (1984)

Asymptotically distribution-free methods for the analysis of covariance structures.
British Journal of Mathematical and Statistical Psychology, **37**, 62-83.

Browne, M.W. (1987)
Robustness of statistical inference in factor analysis and related models.
Biometrika, **74**, 375–384.

Browne, M. W. & Cudeck, R. (1993)
Alternative ways of assessing model fit.
In K. A. Bollen & J. S. Long (Editors): *Testing Structural Equation Models*.
Sage Publications.

Bryk, A.S. & Raudenbush, S.W. (1992)
Hierarchical Linear Models.
Sage.

Cudeck, R. (1989)
Analysis of correlation matrices using covariance structure models.
Psychological Bulletin, **105**, 317–327.

Finn, J.D. (1974)
A general model for multivariate analysis.
New York: Holt, Reinhart and Winston.

French, J.V. (1951)
The description of aptitude and achievement tests in terms of rotated factors.
Psychometric Monographs, **5**.

Goldberger, A.S. (1964)
Econometric theory.
New York: Wiley.

Goldstein, H. (1987)
Multilevel models in educational and social research.
C. Griffin and Co., London, Oxford University Press, New York.

Goldstein, H. (1995)
Multilevel Statistical Models.
Edward Arnold.

Guilford, J.P. (1956)
The structure of intellect.
Psychological Bulletin, **53**, 267–293.

Hendrickson, A.E. & White, P.O. (1964)
Promax: A quick method for rotation to oblique simple structure.
British Journal of Mathematical and Statistical Psychology, **17**, 65-70.

Holzinger, K., & Swineford, F. (1939)
A study in factor analysis: The stability of a bifactor solution.
Supplementary Educational Monograph no. 48. Chicago: University of Chicago Press.

Hu, L., Bentler, P.M., & Kano, Y. (1992)
Can test statistics in covariance structure analysis be trusted?
Psychological Bulletin, Vol. **112**, 351–362.

Hägglund, G. (1982)
Factor analysis by instrumental variable methods.
Psychometrika, **47**, 209–222.

Jöreskog, K.G. (1967)
Some contributions to maximum likelihood factor analysis.
Psychometrika, **32**, 443–482.

Jöreskog, K.G. (1969)
A general approach to confirmatory maximum likelihood factor analysis.
Psychometrika, **34**, 183–202.

Jöreskog, K.G. (1977)
Factor analysis by least-squares and maximum-likelihood methods.
Pp. 125–153 in K. Enslein, A. Ralston, & H.S. Wilf (Eds.):
Statistical methods for digital computers. New York: Wiley.

Jöreskog, K.G. (1979)
Basic ideas of factor and component analysis.
In K.G. Jöreskog & D. Sörbom: *Advances in factor analysis and structural equation models.* Cambridge, Mass.: Abt Books, 5–20.

Jöreskog, K.G. (1981)
Analysis of covariance structures.
Scandinavian Journal of Statistics, **8**, 65–92.

Jöreskog, K.G. (1990)
New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares.
Quality and Quantity, **24**, 387–404.

Jöreskog, K.G. (1993)
Testing structural equation models.
In K.A. Bollen & J.S. Long (Eds), *Testing Structural Equation Models*.
Sage Publications.

Jöreskog, K.G. (1994).
On the estimation of polychoric correlations and their asymptotic covariance matrix.
Psychometrika, **59**, 381–389.

Jöreskog, K.G., & Sörbom, D. (1996a)
PRELIS 2 *User's Reference Guide.*
Chicago: Scientific Software International.

Jöreskog, K.G., & Sörbom, D. (1996b)
LISREL 8 *User's Reference Guide.*
Chicago: Scientific Software International.

Jöreskog, K.G., & Sörbom, D. (1996c)
LISREL 8*: Structural Equation Modeling with the SIMPLIS Command Language.*
Chicago: Scientific Software International.

Jöreskog, K.G., & Yang, F. (1996).
Nonlinear structural equation models: The Kenny–Judd model with interaction effects.
Pp. 57–88 in G.A. Marcoulides & R.E. Schumacker (Eds):
Advanced structural equation modeling: Issues and techniques.
Lawrence Erlbaum Associates, Publishers.

Kaiser, H.F. (1958)
The varimax criterion for analytical rotation in factor analysis.
Psychometrika, **23**, 187–200.

Kanfer, R. & Ackerman, P. L. (1989)
Motivation and cognitive abilities: an integrative/aptitude-treatment interaction approach to skill acquisition.
Journal of Applied Psychology, Monograph, **74**, 657–690.

Kenny, D.A., & Judd, C.M. (1984).
Estimating the nonlinear and interactive effects of latent variables.
Psychological Bulletin, **96**, 201–210.

Klein, L.R. (1950)
Economic fluctuations in the United States 1921–1941.
Cowles Commission Monograph No. 11. New York: Wiley.

Kreft, I. & de Leeuw, J. (1998)
Introducing multilevel modeling.
Sage.

Lawley, D.N., & Maxwell, A.E. (1971)
Factor analysis as a statistical method, (2nd edition).
London: Butterworths.

Longford, N.T. (1987)
A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested effects.
Biometrika, **741**, 4, 817–827.

Magnus, J.R. & Neudecker, H. (1988)
Matrix differential calculus with applications in statistics and econometrics.
New York, Wiley.

Mardia, K.V., Kent, J.T., & Bibby, J.M. (1980)
Multivariate analysis.
New York: Academic Press.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988)
School Matters, the Junior Years.
Wells, Open Books.

Prosser, R., Rasbash, J. & Goldstein, H. (1991)
ML3 software for Three-level Analysis, Users' Guide for V.2.
Institute of Education, University of London.

Rasbash, J. (1993)
ML3E Version 2.3 Manual Supplement.
University of London: Institute of Education..

Reyment, R. & Jöreskog K.G. (1993)
Applied factor analysis in the natural sciences.
Cambridge: Cambridge University Press.

Satorra, A. (1987)
Alternative test criteria in covariance structure analysis: A unified approach.
Psychometrika, **54**, 131–151.

Satorra, A. (1993)
Multi-sample analysis of moment structures: Asymptotic validity of inferences based on second-order moments.
Pp. 283–298 in K. Haagen, D.J. Bartholomew, & M. Deistler (Editors): Statistical Modelling and Latent Variables. Elsevier Science Publishers.

Satorra, A., & Bentler, P.M. (1988)
Scaling corrections for chi-square statistics in covariance structure analysis.
Proceedings of the Business and Economic Statistics Section of the American Statistical Association, 1988, 308–313.

Schumacker, R.E. & Marcoulides, G.A. (Eds) (1998)
Interaction and nonlinear effects in structural equation models.
Mahwah, N.J: Lawrence Erlbaum Associates, Publishers.

Sörbom, D. (1974)
A general method for studying differences in factor means and factor structures between groups.
British Journal of Mathematical and Statistical Psychology, Vol. **27** 229–239.

Steiger, J.H. (1990)
Structural model evaluation and modification: An interval estimation approach.
Multivariate Behavioral Research, **25**, 173–180.

Theil, H. (1971)
Principles of econometrics.
New York: Wiley.

Thurstone, L.L. (1938)
Primary mental abilities.
PsychometricMonographs, **1**.

Tintner, G. (1952)
Econometrics.
New York: Wiley.

Yuan, K-H., & Bentler, P.M. (1997)
Mean and covariance structure analysis: Theoretical and practical improvements.
Journal of the American Statistical Association, Vol. **92**, 767–774.