Introduction to Multilevel Modeling Using HLM 6

By ATS Statistical Consulting Group

Multilevel data structure

- Students nested within schools
- Children nested within families
- Respondents nested within interviewers
- Repeated measures nested within individuals – longitudinal data, growth curve modeling

In the example of student nested within schools:

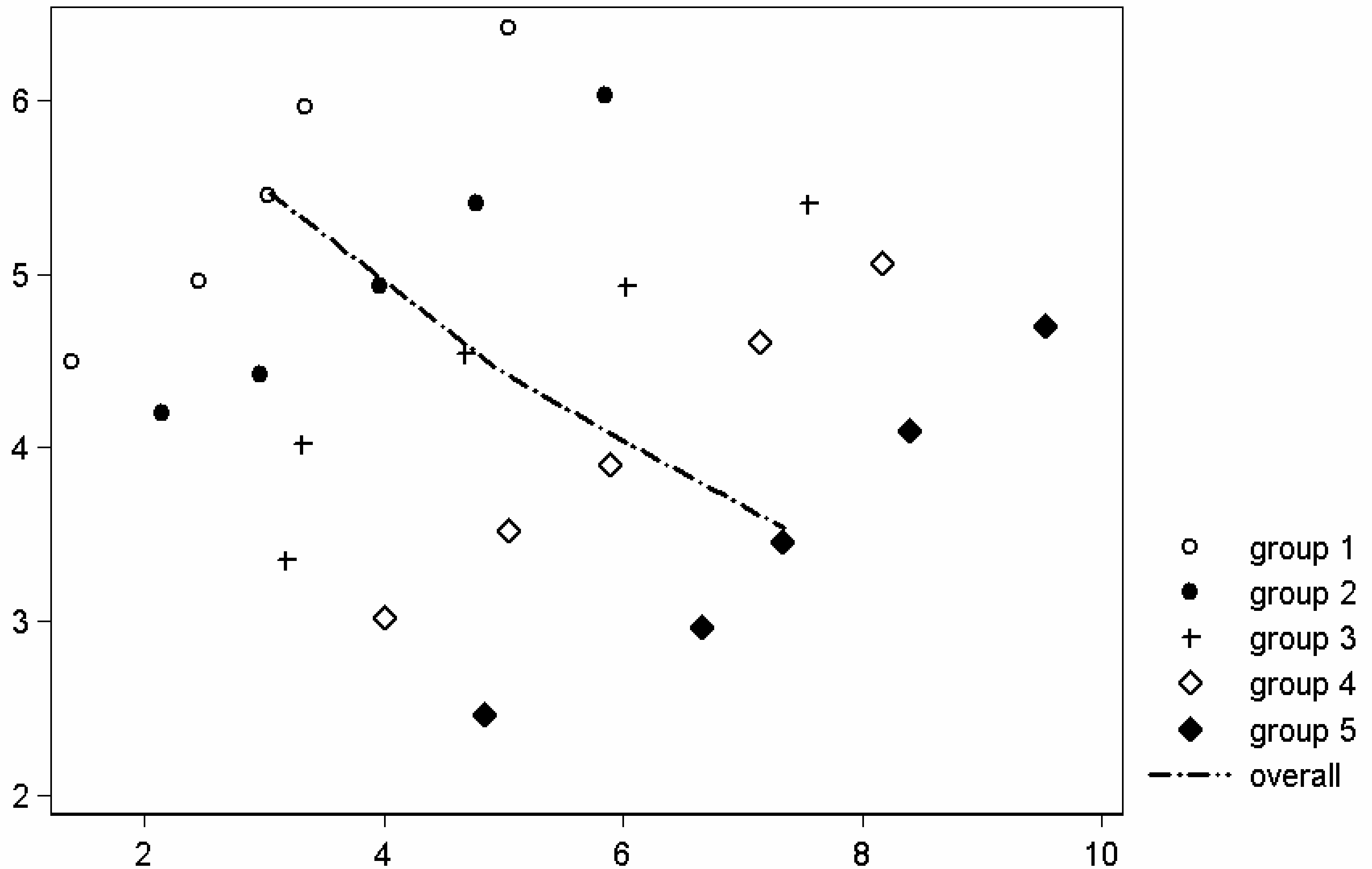
- Level-1 variables, such as student's gender and age
- Level-2 variables, such as school type and size

schid	minority	female	ses	mathach	size	schtype	meanses
1224	0	0	-.878	10.557	842	0	-.428
1224	0	0	-.938	.868	842	0	-.428
1224	0	0	-.548	8.296	842	0	-.428
1224	1	0	.142	-1.688	842	0	-.428
1224	0	1	.972	2.059	842	0	-.428
1224	0	0	.372	6.714	842	0	-.428
1224	1	0	-1.658	.122	842	0	-.428
1224	0	1	-1.068	2.381	842	0	-.428
1224	0	1	-.248	16.336	842	0	-.428
1224	0	1	-1.398	6.134	842	0	-.428
1224	0	1	.752	23.584	842	0	-.428
1224	0	1	.012	14.053	842	0	-.428
1224	0	1	-.418	2.183	842	0	-.428
1288	0	1	-.788	7.857	1855	0	.128
1288	1	0	-.328	10.171	1855	0	.128
1288	0	0	.472	15.699	1855	0	.128
1288	0	1	.352	22.919	1855	0	.128
1288	1	1	-1.468	10.664	1855	0	.128
1288	0	1	.202	13.543	1855	0	.128
1288	0	0	-.518	18.207	1855	0	.128
1288	1	0	-.158	5.552	1855	0	.128
1288	0	0	.042	7.416	1855	0	.128
1288	0	0	.682	18.792	1855	0	.128
1288	0	0	1.262	1.575	1855	0	.128
1288	0	1	.152	3.534	1855	0	.128
1288	0	1	-.678	20.173	1855	0	.128
1288	0	1	.332	10.772	1855	0	.128

How would we analyze such multilevel data?

- OLS regression
- OLS regression with robust standard error
- Aggregation
- Disaggregation
- Ecological fallacy – interpreting analyses on aggregated data at the individual level

Ecological Fallacy



See figure 3.1, on page 14 from *Multilevel Analysis* by Snijders and Bosker

Hierarchical linear model

- Random Intercept model

$$Y_{ij} = \beta_{0j} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$

- Written in mixed model format:

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

  - i is for individuals and j is for schools
  - β_{0j} is the mean of Y_{ij} for school j
  - γ_{00} is the average of all the β_{0j}'s, therefore the grand
  - r_{ij} and u_{0j} are normally distributed
  - r_{ij} and u_{0j} are independent of each other
  - Parameters to be estimated include regression coefficients and variance components: γ_{00}, var(r_{ij}) and var(u_{0j})

Hierarchical linear model

- Random Intercept and random slope model

$$Y_{ij} = \beta_{0j} + \beta_{1j}X + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

- Written in mixed model format:

$$Y_{ij} = \gamma_{00} + \gamma_{10}X + u_{0j} + u_{1j}X + r_{ij}$$

  - β_{0j} is the mean of Y_{ij} for school j when X is zero
  - β_{1j} is the slope of X for school j (or the effect of X for school j)
  - r_{ij}, u_{0j} and u_{1j} are normally distributed
  - u_{0j} and u_{1j} are assumed to be correlated
  - cross-level error terms are assumed to be independent
  - parameters: γ_{00}, γ_{10}, $var(u_{0j})$, $var(u_{1j})$, $cov(u_{0j}, u_{1j})$ and $var(r_{ij})$

Hierarchical linear model

- Random Intercept and random slope model
- Level-2 variable(s) to predict intercept and/or slope

$Y_{ij} = \beta_{0j} + \beta_{1j}X + r_{ij}$

$\beta_{0j} = \gamma_{00} + \gamma_{01}W + u_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}W + u_{1j}$

- Written in mixed model format:

$Y_{ij} = \gamma_{00} + \gamma_{01}W + \gamma_{10}X + \gamma_{11}W*X + u_{0j} + u_{1j}*X + r_{ij}$

  - β_{0j} is the mean of Y_{ij} for school j when X is zero
  - β_{1j} is the slope of X for school j (or the effect of X for school j)
  - γ_{00} is the average intercept
  - γ_{11} is the coefficient for the cross-level interaction term
  - r_{ij}, u_{0j} and u_{1j} are normally distributed
  - u_{0j} and u_{1j} are assumed to be correlated
  - Cross-level error terms are assumed to be independent
  - parameters to be estimated: γ_{00}, γ_{01}, γ_{10}, γ_{11}, var(u_{0j}), var(u_{1j}), cov(u_{0j}, u_{1j}) and var(r_{ij})

Comparing the assumptions for hierarchical linear models with OLS models

OLS Assumptions

- Linearity: function form is linear
- Normality: residuals are normally distributed
- Homoscedasticity: residual variance is constant
- Independence: observations are independent of each other

HLM assumptions

- Linearity: function forms are linear at each level
- Normality: level-1 residuals are normally distributed and level-2 random effects u's have a multivariate normal distribution
- Homoscedasticity: level-1 residual variance is constant
- Independence: level-1 residuals and level-2 residuals are uncorrelated
- Independence: observations at highest level are independent of each other

Estimation Methods: REML vs. ML

- Reading: Section 4.6 Parameter Estimation from Snijder and Bosker
- REML and ML produce similar regression coefficients
- REML and ML differ in terms of estimating the variance components
- If the number of level-2 units is small , then ML variance estimates will be smaller than REML, leading to artificially short confidence interval and biased significant tests.
- REML is the default estimation method for HLM
- Likelihood ratio test for nested models
  - When fixed effects are the same, model has fewer random effects , then both REML or ML may be used
  - When one model has fewer fixed effects and possibly fewer random effects, then ML may be used

Issues with Centering

- Reading: Section 5.2 *The effects of centering* from Kreft and De Leeuw
- In OLS centering is to change the interpretation of the intercept
- Centering in HLM is not a simple issue
- Grand-mean centering

  "The raw score model and the grand mean centered model are *equivalent linear models.*"
- Group-mean centering
  Most of the times, the group mean centered model and the raw score model are neither equivalent in the fixed part nor in the random part.
- Combining substantive and statistical reasons in choosing
  - raw score
  - group-centering with reintroducing the means
  - group-centering without reintroducing the means

An Example

- The dataset is a subsample from the 1982 High School and Beyond Survey and is used extensively in *Hierarchical Linear Models* by Raudenbush and Bryk.
- It consists of 7185 students nested in 160 schools.
- The outcome variable of interest is the student-level math achievement score, **mathach**.
- Predictor variables
  - Level-1 (student level) predictor variables:
    - **ses**: social-economic-status of a student
    - **female** 0 = male and 1 = female
  - Level-2 (school level) predictor variables:
    - **meanses**: mean ses at school level, aggregated from student level
    - **schtype**: type of school: 0 = public and 1 = private, there are 90 public and 70 private schools
    - **size**: size of a school

Model Building

- Reading: Section 6.4 Model specification from Snijder and Bosker

- Unconditional model:

$$mathach_{ij} = \beta_{0j} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$

- Random intercept model with level-2 predictor(s):

$$mathach_{ij} = \beta_{0j} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(meanses) + u_{0j}$$

- Random intercept and random slope model:

$$mathach_{ij} = \beta_{0j} + \beta_{1j}(ses) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

- Full model:

$$mathach_{ij} = \beta_{0j} + \beta_{1j}(group\_mean\_centered\_ses) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(schtype) + \gamma_{02}(meanses) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(schtype) + \gamma_{12}(meanses) + u_{1j}$$

Model 1: Unconditional Means Model

mathachij = β0j + rij β0j = γ00 + u0j

γ00 = 12.636972

var(rij) = 39.14831 var(u0j) = 8.61431

Rho = var(u0j)/(var(u0j) + var(rij))

= 8.61431/(8.61431+39.14831) = .18035673

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	12.636972	0.243628	51.870	159	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	2.93501	8.61431	159	1660.23259	0.000
level-1,	R	6.25686	39.14831			

Final model

$$\text{mathach}_{ij} = \beta_{0j} + \beta_{1j}(\text{group\_mean\_centered\_ses}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{schtype}) + \gamma_{02}(\text{meanses}) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{schtype}) + \gamma_{12}(\text{meanses}) + u_{1j}$$

Tau

INTRCPT1,B0	2.37996	0.19058
SES,B1	0.19058	0.14892

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	12.096006	0.173699	69.638	157	0.000
SCHTYPE, G01	1.226384	0.308484	3.976	157	0.000
MEANSES, G02	5.333056	0.334600	15.939	157	0.000
For SES slope, B1					
INTRCPT2, G10	2.937981	0.147620	19.902	157	0.000
SCHTYPE, G11	-1.640954	0.237401	-6.912	157	0.000
MEANSES, G12	1.034427	0.332785	3.108	157	0.003

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	1.54271	2.37996	157	605.29503	0.000
SES slope,	U1	0.38590	0.14892	157	162.30867	0.369
level-1,	R	6.05831	36.70313			

Final Model (continued)

$mathach_{ij} = \beta_{0j} + \beta_{1j}(group\_mean\_centered\_ses) + r_{ij}$

$\beta_{0j} = \gamma_{00} + \gamma_{01}(schtype) + \gamma_{02}(meanses) + u_{0j}$

$\beta_{1j} = \gamma_{10} + \gamma_{11}(schtype) + \gamma_{12}(meanses) + u_{1j}$

γ_{00} = 12.096: the intercept for public schools with meanses =0 (average ses)

γ_{01} = 1.226: the change in intercept from a public school to a private school

- the intercept for private school with meanses = 0 is 12.096+1.226 = 13.322

γ_{02} = 5.333: the change in intercept for a one-unit change in meanses

- the intercept for public school with meanses = 1 is 12.096 + 5.333 = 17.429

γ_{10} = 2.94: the slope of gcses for public schools with meanses = 0.

- the effect of gcses for public schools with meanses = 0 is 2.94

γ_{11} = -1.641: the change in slope from a public school to a private school

- the effect of gcses for private schools with meanses = 0 is 2.94 – 1.641 = 1.299

γ_{12} = 1.034: the change in slope for a one-unit change in meanses

- the effect of gcses for public schools with meanses = 0 is 2.94
- the effect of gcses for public schools with meanses = 1 is 2.94 + 1.034 = 3.974

```
For       INTRCPT1, B0
     INTRCPT2, G00          12.096006   0.173699    69.638       157    0.000
      SCHTYPE, G01           1.226384   0.308484     3.976       157    0.000
      MEANSES, G02           5.333056   0.334600    15.939       157    0.000
For      SES slope, B1
     INTRCPT2, G10           2.937981   0.147620    19.902       157    0.000
      SCHTYPE, G11          -1.640954   0.237401    -6.912       157    0.000
      MEANSES, G12           1.034427   0.332785     3.108       157    0.003
```

What's new in HLM 6

The following paragraph is based on:

http://www.ssicentral.com/hlm/new.html

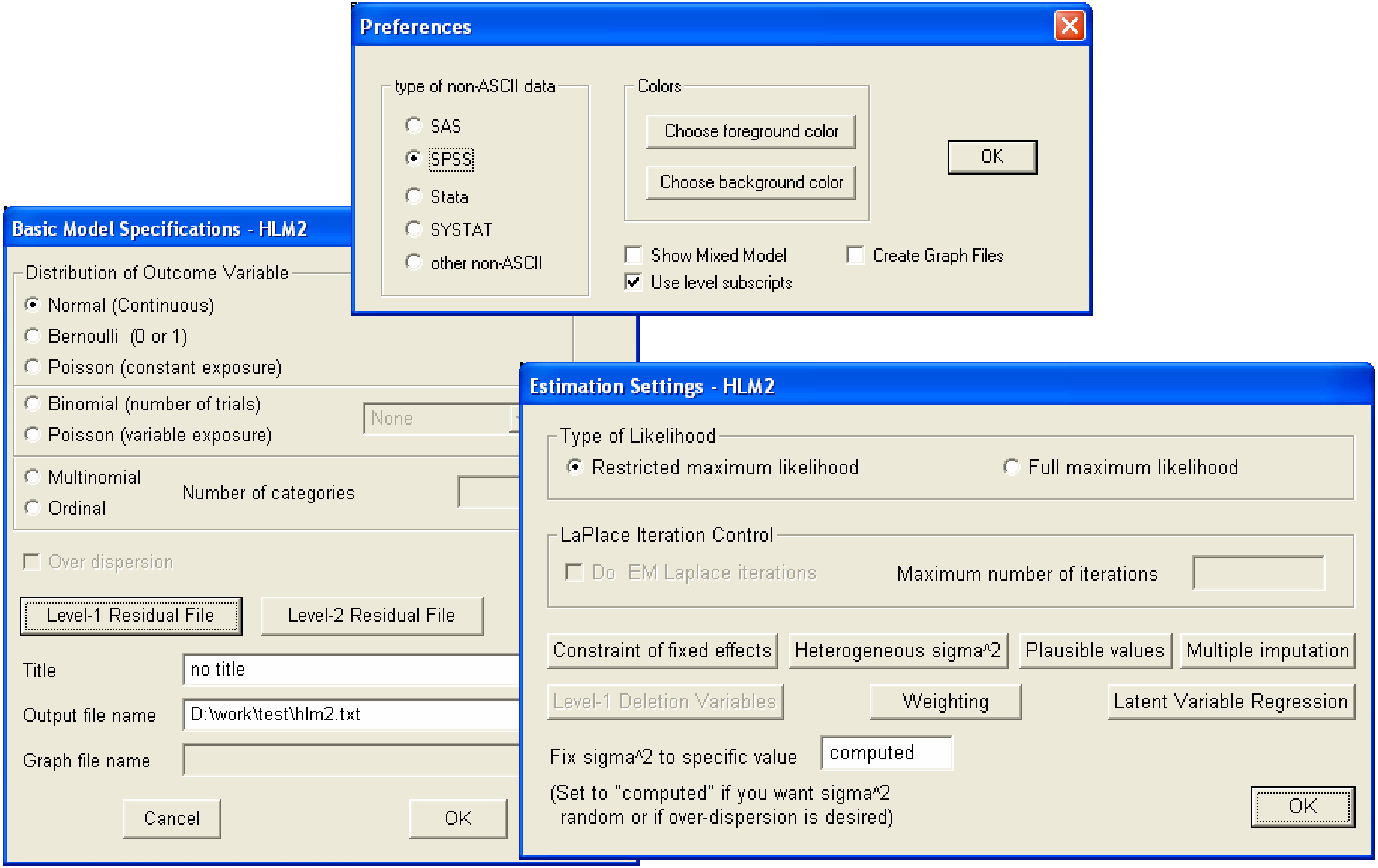
HLM 6 greatly broadens the range of hierarchical models that can be estimated. It also offers greater convenience of use than previous versions. Here is a quick overview of key new features and options:

- All new graphical displays of data.
- Greater expanded graphics for fitted models.
- Model equations displayed in hierarchical or mixed-model format with or without subscripts - easy to save for publication. Distribution assumptions and link functions are presented in detail.
- Slightly different and easier way for specifying random effects.
- Cross-classified random effects models for linear models and non-linear link functions with convenient Windows interface.
- High-order Laplace approximation with EM algorithm for stable convergence and accurate estimation in two-level hierarchical generalized linear models (HGLM).
- Multinomial and ordinal models for three-level data. Also see the types of models.
- New flexible and accurate sample design weighting for two- and three-level HLMs and HGLMs.
- Easier automated input from a wide variety of software packages, including the current versions of SAS, SPSS, and STATA.
- Residual files can be saved directly as SPSS (*.sav) or STATA (*.dta) files.
- Analyses are based on MDM files, replacing the older less flexible SSM format.

Getting ready for using HLM software for multilevel data analysis

- Creating MDM file
  - separate level-1 and level2 files for HLM2, or a single file
  - original file can be in different format, such as SPSS, Stata and SAS
  - linking variable can be either numeric or character
  - variables in the analyses have to be numeric
  - mdm file: binary file used for analyses and graphics
  - mdmt file: template file in text format for creating mdm file
  - hlm2mdm.sts: text file containing the summary statistics
- Data management
  - HLM does not have data management capability
  - One has to use other stat package(s) to clean the data and to create variables, such as dummy variables and within-level interaction terms
  - HLM handles cross-level interactions nicely

Choosing preferences and other settings



Demo on using HLM

- Input Data and Creating the "MDM" file
  - from a single SPSS file
- data-based graphs
  - box-plot
  - scatter plot
- Model Building
  - unconditional means model
  - regression with means-as-outcomes
  - random-coefficient model
  - intercepts and slopes-as-outcomes model
- Hypothesis Testing, Model Fit
  - Multivariate hypothesis tests on fixed effects
  - Multivariate Tests of variance-covariance components specification
  - Model-based graphs
- Other Issues
  - Modeling Heterogeneity of Level-1 Variances
  - Models Without a Level-1 Intercept
  - Constraints on Fixed Effects

References

- Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling by Tom Snijders and Roel Bosker
- Introduction to Multilevel Modeling by Ita Kreft and Jan de Leeuw
- Multilevel Analysis: Techniques and Applications by Joop Hox
- Hierarchical Linear Models, Second Edition by Stephen Raudenbush and Anthony Bryk
- HLM 6 - Hierarchical Linear and Nonlinear Modeling by Raudenbush et al.