Research Design - - Topic 17
Hierarchical Linear Modeling (Persons within Groups)

General Rationale, Purpose, and Applications

Random Coefficients Model
- Level 1 and Level 2 equations
- An example
- Ordinary least Squares (OLS) and Maximum Likelihood (ML) Estimates
- Major output and tests of significance

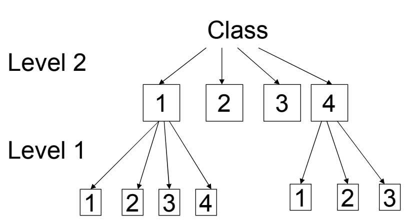
Intercepts- and Slopes-as-outcomes model
- Level 1 and level 2 equations
- Major output and tests of significance

General Overview

HLM can be considered as a variant of multiple regression. There is one dependent (i.e., **outcome)** variable and one or more independent variables (i.e., predictors). It assumes further that the basic data can be considered to be made up of individual observations at Level 1 that are meaningfully grouped at a higher level (e.g., Level 2). One rationale for using this procedure is that the Level 2 groupings reflect meaningful inter-subject variation that should be considered in the estimation process.

The next slide portrays this as it applies to HLM Persons within Groups.

Format | Variables



Class size, class attitudes, etc…

Motivation, Anxiety, Achievement, etc..

Note: Class is a Level 2 variable (that may or may not have some defined feature(s)), while "persons" is a Level 1 variable nested in Class. Models can be made much more complex but we will focus on this basic one.

HLM is a relatively new procedure introduced in the 1990's. It has various names such as:
- Multilevel linear models (Sociology)
- Mixed effects models and random effects models (Biometrics)
- Random-coefficient regression models (Econometrics)
- Covariance components models (Statistics)
- Multilevel models (Tabachnick & Fidell, 2006)

The analysis can be performed on many computer programs. The following provide parameter estimates using Maximum Likelihood (Full) or Restricted Maximum Likelihood (REML) methods, and often give slightly different answers.
- HLM
- SAS Proc Mixed
- SPSS Mixed Models

There are generally two types of models:
- Random Coefficients models
- Intercepts- and Slopes-as-Outcomes models

Applications

HLM is essentially multiple regression with **a twist**, with two basic applications:

1. **Persons Within Groups.** Analyses involving independent groups of participants. For example, we could investigate the effects of motivation (and possibly other variables) on student's grades in French taking into consideration the class in which students are enrolled (possibly including some measure of the class as a whole).
2. **Measures Within Persons.** Analyses based on repeated measures of participants. That is, we could investigate grades in French over 4 terms (and also possibly the effects of other measures each term (Level 1) or for each student (Level 2).

Note. Each of these could be considered from the point of view of MRC analysis. **The twist** is that in HLM we make use of maximum likelihood to estimate the mean and variance of intercepts and slopes (and conduct tests of significance on their values) rather than use ordinary least squares to compute the semi-partial R^2 and F-ratios as in MRC analysis.

5

Application to Persons within Groups

The purpose is to test the adequacy of a regression model where a number of groups are involved. Thus, to study the "effects" of motivation of students in various classes on French grades, we would write an equation that predicts grades as a function of motivation in each class. This is referred to as a **Level 1 equation** because the focus is on the individual. At the same time we could write a regression equation where we define the slopes and intercepts for each class as a function of their deviation from the mean slope and intercept (and for more complex models as a function of other characteristics of the class). These are referred to as **Level 2 equations** because the focus is on the class.

The next slide shows examples of these equations in the simplest case, the random coefficients model. In this model, class is considered a grouping variable, but there is no interest in determining whether some definable characteristic of the class influences the results.

6

Random Coefficients Model

Level 1 model $y_{ij} = \beta_{0j} + \beta_{1j}C_{ij} + \gamma_{ij}$

where: β_{0j} is the intercept for group j
β_{1j} is the unstandardized regression coefficient (slope) for group j
C_{ij} is a code for the group for individual j
γ_{ij} is the error in prediction.

Level 2 models $(1)\ \beta_{0j} = \gamma_{00} + \mu_{0j}$

where:

$\gamma_{00} =$ mean of the intercepts

$\mu_{0j} = \beta_{0j} - \gamma_{00} =$ deviation of each intercept from the mean intercept

$(2)\ \beta_{1j} = \gamma_{10} + \mu_{1j}$

where:

$\gamma_{10} =$ mean of the slopes

$\mu_{1j} = \beta_{1j} - \gamma_{10} =$ deviation of each slope from the mean slope

7

Substituting the two Level 2 values into the Level 1 equation yields the full equation:

$$y_{ij} = \gamma_{00} + \gamma_{10}C_{ij} + \mu_{0j} + \mu_{1j}C_{ij} + \gamma_{ij}$$

Given this equation we can solve for the following parameters and their standard errors of estimate:

Mean Intercept γ_{00}

Mean Slope γ_{10}

Variance of the intercepts - - based on μ_{0j}

Variance of the slopes - - based on μ_{1j}

Error - - based on γ_{ij}

8

Generalizing to more than one Level 1 Variable, HLM is a very versatile system. For example, in the random-coefficients model you can have more than one level 1 variable. Thus, for two predictors, C and D the Level 1 equation is:

$$y_{ij} = \beta_{0j} + \beta_{1j}C_{ij} + \beta_{2j}D_{ij} + \gamma_{ij}$$

and the Level 2 equations are:

$\beta_{0j} = \gamma_{00} + \mu_{0j}$ Intercept for each group

$\beta_{1j} = \gamma_{10} + \mu_{1j}$ Slope against C for group j

$\beta_{2j} = \gamma_{20} + \mu_{2j}$ Slope against D for group j

And the full equation is:

$$\therefore y_{ij} = \overset{1}{\gamma_{00}} + \overset{2}{\mu_{0j}} + \overset{3}{\gamma_{10}} C_{ij} + \overset{4}{\mu_{1j}} C_{ij} + \overset{5}{\gamma_{20}} D_{ij} + \overset{6}{\mu_{2j}} D_{ij} + \gamma_{ij}$$

These values give rise to the following parameter estimates

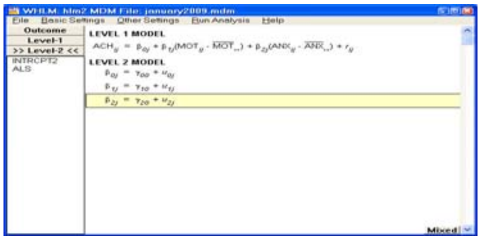
1. γ_{00} - mean intercept
2. μ_{0j} - variance of intercepts
3. γ_{10} - mean slope against variable C
4. μ_{1j} - variance of slopes against variable C
5. γ_{20} - mean slope against variable D
6. μ_{2j} - variance of slopes against variable D

Consider the following simple example dealing with the prediction of French grades with scores on measures of motivation and anxiety in 5 classes of students. This example is too small for such an analysis, but it serves to demonstrate the process. The sample data file is presented on the next slide.

Variables:
Class. Note, there are different numbers of Ss in the classes
Ach (Y). Grades in French, the outcome variable
Mot (C). Scores on a motivation test
Anx (D). Scores on a language anxiety test.
ALS (W). Score for each class on attitudes toward the class

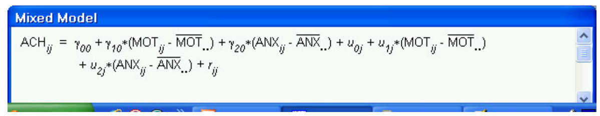
Class	Ach	Mot	Anx	ALS
1	70	5.5	4.5	2.5
1	78	5.6	4.0	2.5
1	69	4.7	4.0	2.5
1	65	4.2	5.2	2.5
1	72	5.6	5.1	2.5
2	76	4.9	4.7	3.7
2	74	5.6	4.8	3.7
2	79	5.7	4.2	3.7
2	81	6.0	4.1	3.7
2	84	6.1	3.9	3.7
2	86	5.9	4.7	3.7
2	79	6.9	4.7	3.7
2	83	6.9	4.3	3.7
3	72	6.8	3.8	5.2
3	71	5.8	3.7	5.2
3	69	4.5	3.4	5.2
3	76	5.9	5.6	5.2
3	77	6.7	2.6	5.2
3	75	6.4	3.4	5.2
4	76	5.4	3.5	4.1
4	73	4.7	3.2	4.1
4	67	5.7	5.4	4.1
4	62	4.8	4.8	4.1
4	69	5.9	4.3	4.1
4	75	6.1	4.1	4.1
5	56	3.6	2.5	4.9
5	54	2.9	2.6	4.9
5	75	4.7	1.2	4.9
5	63	5.1	3.2	4.9
5	61	3.7	3.5	4.9
5	77	4.8	2.5	4.9
5	80	5.3	2.4	4.9
5	74	5.1	2.7	4.9

Following is the model as constructed on HLM6. Instructions for running HLM are available in Gardner (2007).



The Level 1 model predicts ACH as a function of MOT and ANX (each of which are grand-mean centered). The Level 2 model defines each intercept and slope as the mean value plus the deviation from the mean. The uij are estimated; though they could be set to 0 by clicking them off.

HLM uses the unique approach (i.e., Model I or SPSS TYPE 3) by estimating each parameter given the other parameters in the model. The complete model estimated is the aggregate of the Level 1 and Level 2 models and is as follows:



The following example is concerned with a model that treats Achievement as a function of motivation, anxiety and Class. In this model, motivation and anxiety are Level 1 variables and Class is a Level 2 variable with no associated covariate.

The primary function of HLM is to estimate the parameters and conduct tests of significance. It does not routinely output the intercepts and slopes but, if requested, it will produce the Ordinary Least Squares (OLS) and maximum likelihood (REML) estimates. The OLS values are simply the values you would obtain if you were to perform a regression analysis separately on each group. The results for this data set are as follows:

	Ordinary Least Squares			Maximum Likelihood (REML)		
Class	Intercept	Slopes		Intercept	Slopes	
		Mot	Anx		Mot	Anx
1	74.34	5.32	-2.99	73.98	5.02	-2.80
2	81.55	2.31	-4.65	78.93	3.09	-2.18
3	71.75	2.57	.63	71.91	3.12	- .00
4	72.32	5.25	-5.98	71.93	6.99	-4.39
5	70.81	9.15	-4.48	69.45	8.11	-4.85
Means	74.15	4.92	-3.49	73.24	5.26	-2.84

Note that the values are different for the OLS and ML solutions. The OLS values are statistics calculated on the data. The ML values are estimates of the parameters that are most likely, given the nature of the data in this sample.

There are 5 major sections of output of interest. The following three are not presented here, while the two on the next slide are shown and discussed.

1. Tau Matrix: Matrix of variances and covariances of the intercepts and slopes of the level 1 variables
2. Tau as Correlations:
3. Reliability estimates – reliability of the level 1 estimates over the groups, defined as:

$$R = \frac{\text{true variance}}{\text{true + error variance}}$$

A low reliability doesn't invalidate the analysis. Instead it indicates that there will be shrinkage from the OLS estimates to the maximum likelihood estimates. The higher the reliability, the more similar the two sets of estimates.

4. Final Estimation of fixed effects

These are tests of significance of the mean intercept and mean slope against 0. They make use of single sample t-tests where:

$$t = \frac{\text{estimated mean intercept or slope} - 0}{\text{estimated standard error of intercept or slope}}$$

Note that the coefficients are the actual means of the maximum likelihood estimates, but the standard errors are maximum likelihood estimates, and not simply the standard error of the actual estimates.

5. Final Estimation of variance components

In this model, the variance of the intercepts and the variance of the slopes are tested for significance. These are referred to as random effects because groups are considered random samples from the population. These tests are evaluated by chi-square.

The major output for this analysis

The outcome variable is ACH

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	73.237619	1.915799	38.228	4	0.000
For MOT slope, B1					
INTRCPT2, G10	5.264860	1.480674	3.556	4	0.036
For ANX slope, B2					
INTRCPT2, G20	-2.843210	1.440980	-1.973	4	0.117

Interpretation: The mean intercept (73.24) differs significantly from 0, t(4)=38.23. The mean slope against motivation (5.26) differs significantly from 0, t(4)=3.56. The mean slope for anxiety (-2.84) does not differ significantly from 0, t(4)=-1.97.

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	3.93828	15.51003	4	7.88873	0.095
MOT slope, U1	2.55837	6.54524	4	9.66466	0.046
ANX slope, U2	2.35723	5.55656	4	7.36967	0.116
level-1, R	3.67928	13.53710			

Interpretation: The variance of the intercepts (15.51) does not differ significantly from 0, $\chi^2(4)=7.89$. The variance of the slopes against motivation (6.55) is significantly greater than 0, $\chi^2(4)=9.66$. The variance of the slopes against anxiety (5.56) does not differ significantly from 0, $\chi^2(4)=7.37$.

Conclusions from the Results

These results indicate that in a model that is concerned with the effects of motivation and anxiety on achievement in French, there is a clear effect of motivation but not anxiety. Moreover, although this effect is consistent overall, it is also the case that the effects of motivation tend to vary from class to class.

21

Some Terminology

Centering: Used to locate the value of the intercept, it is appropriate only to the predictors (Level 1 and 2). There are 3 options:

Grand Mean Centering. The grand mean of the predictor is subtracted from each predictor.

Group Mean Centering. For level 1 predictors, the mean for the group is subtracted from each predictor.

No Centering.

The type of centering has an effect on the estimate of the intercepts but not the slopes. That is the intercepts refer to the value of the outcome variable at the mean of all the subjects (grand mean centering) or the means of the individual groups (group mean centering) on the predictors.

22

Iterations: Maximum likelihood is an iterative procedure in which estimates are continually made until they produce values that are maximally likely given the nature of the sample data.

Fixed Effects: Involves tests of the mean intercepts and slopes vs 0. That is, is there an effect over all.

Random Effects: Asks whether the intercepts and the means differ among themselves over the groups. If not significant, it indicates whatever fixed effect that was obtained is consistent over classes. If desired, these estimates can be fixed at 0, by clicking off (i.e., making it lighter in the set-up window on slide 13) the appropriate μ_{0j}, μ_{1j}, μ_{2j}, etc. This will of course change the model, and the estimates, tests of significance, etc…

23

Intercepts- and Slopes-as-Outcome Variables

This model can be essentially the same as the preceding one, but in this case, there is some definable characteristic of the grouping variable, and the interest is in whether and how this characteristic influences the results. This example uses the same data but includes the level 2 variable, ALS.

That is, the model concerns the effects of motivation and anxiety on achievement in French as moderated by attitudes toward the class as revealed in the mean score on an attitudes toward the class measure.

24

Equations

Level 1 $$y_{ij} = \beta_{0j} + \beta_{1j}C_{ij} + \beta_{2j}D_{ij} + \gamma_{ij}$$

Level 2 $$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + \mu_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}W_j + \mu_{2j}$$

Full Equation:

$$y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}C_{ij} + \gamma_{11}C_{ij}W_j + \gamma_{20}D_{ij} + \gamma_{21}D_{ij}W_j + \mu_{0j} + \mu_{1j}C_{ij} + \mu_{2j}D_{ij} + \gamma_{ij}$$

25

Estimates

1. γ_{00} - mean intercept
2. μ_{0j} - variance of intercepts
3. γ_{01} - regression of the class intercepts on W
4. γ_{10} - mean slope for variable C
5. γ_{20} - mean slope for variable D
6. μ_{1j} - variance of slopes for variable C
7. μ_{2j} - variance of slopes for variable D
8. γ_{11} - regression of the class slopes for C on W
9. γ_{21} - regression of the class slopes for D on W
10. γ_{ij} - error

26

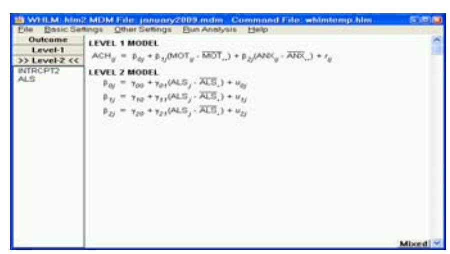
An example of the Intercepts- and Slopes-as-Outcomes Model

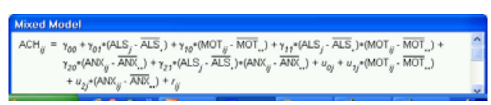
In this case, Achievement is the Outcome variable and motivation and anxiety are the Level 1 predictor variables as before, but in this case, Attitudes toward the Learning Situation is added as a Level 2 variable based on the class mean scores on this measure. In this example, all three variables, Mot, Anx, and ALS, are grand mean centered.

This is a different model from before and as would be anticipated the estimates and their tests of significance differ from the previous analysis.

The changes are illustrated on the next slides.

27





28

	Ordinary Least Squares			Maximum Likelihood (REML)		
Class	Intercept	Slopes		Intercept	Slopes	
		Mot	Anx		Mot	Anx
1	74.34	5.32	-2.99	77.21	5.66	-5.88
2	81.55	2.31	-4.65	78.04	3.15	- .91
3	71.75	2.57	.63	71.90	3.08	.21
4	72.32	5.25	-5.98	71.87	6.59	-4.80
5	70.81	9.15	-4.48	69.04	8.00	-4.93
Means	74.15	4.92	-3.49	73.61	5.30	-3.26

Note that the OLS values are the same as in the last table, but that the maximum likelihood values are different because of the addition of ALS as a Level 2 variable to the model.

29

The outcome variable is ACH

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	73.610586	1.653743	44.512	3	0.000
ALS, G01	-2.858148	1.794497	-1.593	3	0.208
For MOT slope, B1					
INTRCPT2, G10	5.295367	1.580919	3.350	3	0.071
ALS, G11	0.001096	1.685158	0.001	3	1.000
For ANX slope, B2					
INTRCPT2, G20	-3.260301	1.728956	-1.886	3	0.150
ALS, G21	1.256815	1.844128	0.682	3	0.544

Note. In this analysis, only the mean intercept differs significantly from 0. When the moderating effect of Attitudes toward the Learning Situation is introduced into the model, the mean slope for motivation is no longer significant.

30

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	3.01472	9.08855	3	6.71931	0.080
MOT slope, U1	2.81268	7.91118	3	9.08379	0.028
ANX slope, U2	2.96511	8.79186	3	5.25581	0.152
level-1, R	3.79303	14.38711			

Note. As before, the variance of the slopes for motivation is significantly greater than 0, indicating that the slopes vary over the classes. This agrees with the results from the first analysis. Thus, although the mean slope for motivation is not significantly different from 0, there is variation in these slopes across the classes.

31

References

Raudenbush, S.W., & Bryk, A.S.(2002). Hierarchical Linear Models: Applications and Data Analysis Methods (second edition). Thousand Oaks, CA: Sage.

Gardner, R. C. (2006). Hierarchical Linear Modeling: Persons within Groups. Unpublished manuscript. Department of Psychology, University of Western Ontario. http://publish.uwo.ca/~gardner/DataAnalysisDotCalm.

32