

Multilevel Linear Modeling

PSY 5102: Advanced Statistics for Psychological and Behavioral Research 2

Goals

- What is multilevel modeling?
- Examples of multilevel data structures
- Current applications
- Why multilevel modeling?
- What types of studies use multilevel modeling?

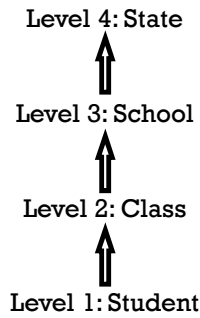
Hierarchical Data

- Data structures are often hierarchical
 - This is referred to as a **Nested Data Structure**
- Examples:
 - Children nested within classrooms within schools
 - Data points nested within people
 - Employees nested within organizations
 - Patients nested within hospitals
 - Patients nested within doctors nested within clinics
 - Families within neighborhoods within communities

What Are Nested Data?

- ◉ Sub-units are grouped (or “nested”) within larger units
 - Many statistical techniques assume that observations are independent...but this assumption is violated with nested data
- ◉ Often the data are observations of individuals nested within groups
 - This is important because individuals within groups are often more similar to one another than to individuals in the other groups
 - We can empirically verify this
- ◉ Sometimes data are multiple observations nested within an individual

Multilevel Data Structures



Basic Terminology

- ◉ A **Level** consists of individual units (e.g., students, classrooms, schools, observations)
- ◉ **Level-1 variables**
 - These are the variables that are nested within groups
 - Example: individual students
- ◉ **Level-2 variables**
 - Typically these are group-level variables
 - Example: classes of students
- ◉ There can also be Level-3 or higher variables
 - There must be at least two levels for it to be a multilevel design
- ◉ Note that growth models have time periods at Level-1 and individuals at Level-2

Why Multilevel Modeling?

- ◉ Nested data are very common in psychology
- ◉ Analysis of nested data poses a unit of analysis problem
 - Should we analyze the individual or the group?
 - Unfortunately, we are often unable to choose one over the other
- ◉ Traditional linear models offer a simple view of a complex world by generally assuming the same effects across groups
- ◉ If effects do differ across groups, we can explain these differences with multilevel modeling

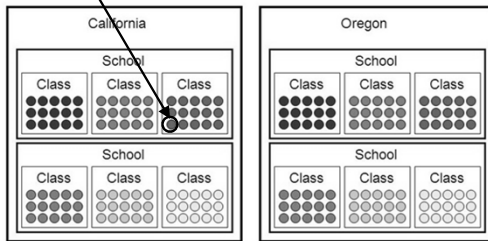
Basic Terminology

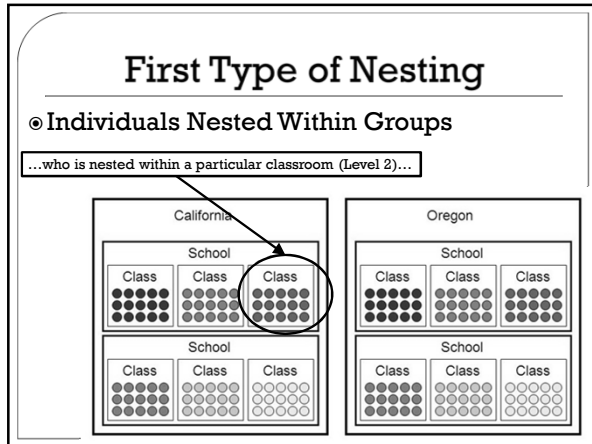
- ◉ There are multiple terms that are used to refer to Multilevel Modeling
 - Hierarchical Linear Modeling (HLM) is both a statistical technique and a software package
 - Multilevel Random-Coefficient Models
 - Random-Coefficient Regression Models

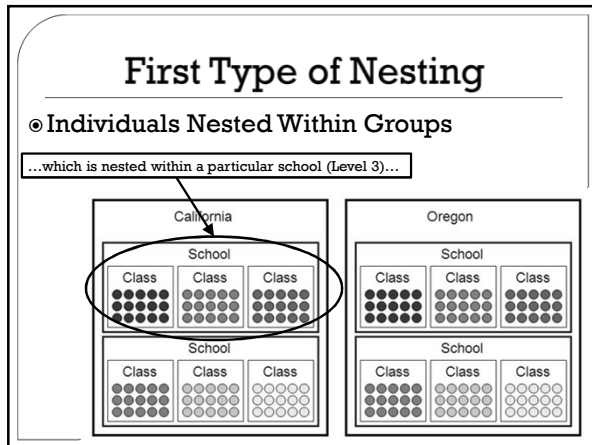
First Type of Nesting

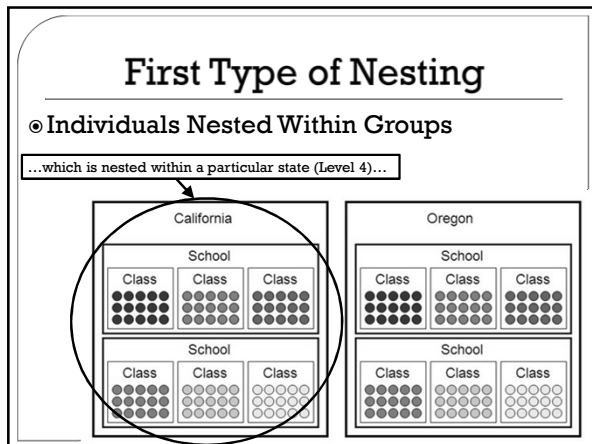
- ◉ Individuals Nested Within Groups

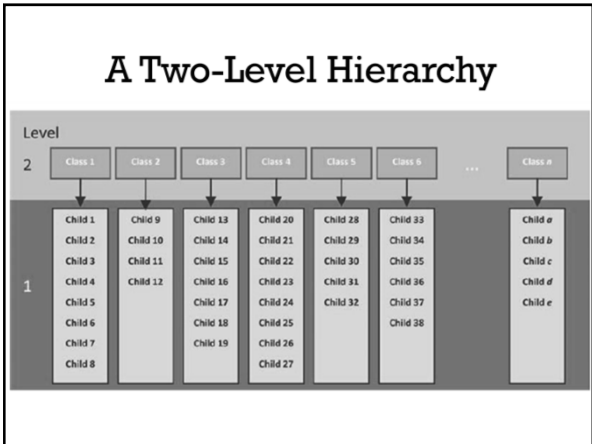
Here is one child (Level 1)...

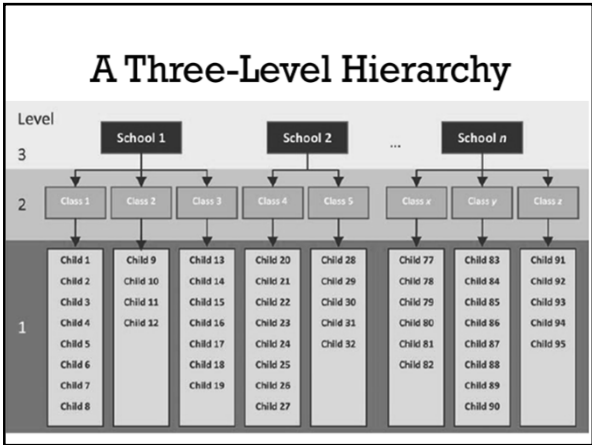












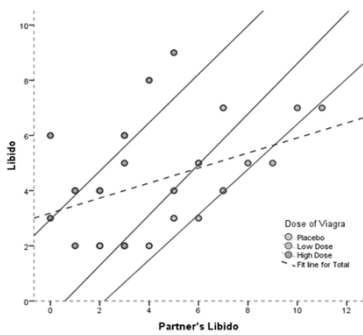
Benefits of Multilevel Models

- It allows you to model the variability in intercepts and regression slopes
- It allows you to account for the associations between cases
 - This allows you to deal with the violation of the independence assumption that accompanies nested data
- Multilevel models can cope with “missing” data
 - It is quite common for nested data to be “uneven” (e.g., some classes may have more children than others, some participants may complete more daily measures than others)

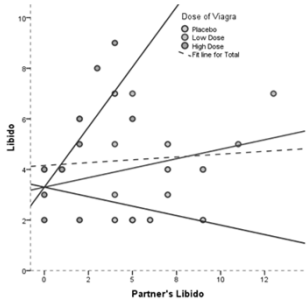
Fixed vs. Random Coefficients

- Intercepts and slopes can be fixed or random
 - In simple linear regression they are fixed
- Fixed Coefficients
 - Intercepts and/or slopes are assumed to be the same across different groups
- Random Coefficients
 - Intercepts and/or slopes are allowed to vary across different groups

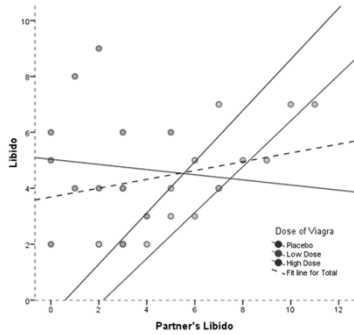
Fixed Slope, Random Intercept



Random Slope, Fixed Intercept



Random Slope, Random Intercept



Representing These Models

Random Intercepts and Fixed Slopes

$$Y_{ij} = b_{0j} + b_1 X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

Fixed Intercepts and Random Slopes

$$Y_{ij} = b_{0i} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{1j} = b_1 + u_{1j}$$

Random Intercepts and Random Slopes

$$Y_{ij} = b_{0ij} + b_{1ij} X_{ij} + \epsilon_{ij}$$

$$b_{0ij} = b_0 + u_{0ij}$$

$$b_{1ij} = b_1 + u_{1ij}$$

Representing These Models

Random Intercepts and Fixed Slopes

$$Y_{ij} = b_{0j} + b_1 X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

Fixed Intercepts and Random Slopes

$$Y_{ij} = b_{0i} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{1j} = b_1 + u_{1j}$$

Random Intercepts and Random Slopes

$$Y_{ij} = b_{0ij} + b_{1ij} X_{ij} + \epsilon_{ij}$$

$$b_{0ij} = b_0 + u_{0ij}$$

$$b_{1ij} = b_1 + u_{1ij}$$

Representing These Models

Random Intercepts and Fixed Slopes

$$Y_{ij} = b_{0j} + b_1 X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

Fixed Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{1j} = b_1 + u_{1j}$$

Random Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

$$b_{1j} = b_1 + u_{1j}$$

Representing These Models

Random Intercepts and Fixed Slopes

$$Y_{ij} = b_{0j} + b_1 X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

Fixed Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{1j} = b_1 + u_{1j}$$

Random Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

$$b_{1j} = b_1 + u_{1j}$$

Representing These Models

Random Intercepts and Fixed Slopes

$$Y_{ij} = b_{0j} + b_1 X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

Fixed Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{1j} = b_1 + u_{1j}$$

Random Intercepts and Random Slopes

$$Y_{ij} = b_{0j} + b_{1j} X_{ij} + \epsilon_{ij}$$

$$b_{0j} = b_0 + u_{0j}$$

$$b_{1j} = b_1 + u_{1j}$$

Fixed and Random Classifications

Random classification

Generalization of a level (e.g., schools)

Random effects come from a *distribution*

All schools contribute to between-school variance

Fixed classification

Discrete categories of a variable (e.g., sex)

Not sample from a population

Specific categories only contribute to their respective means

Second Type of Nesting

- ◉ Repeated measures nested within individuals
 - The focus of these analyses is either change or growth over time

Growth Models

- ◉ Growth models look at the rate of change of a variable over time
 - Examples
 - Depression over 8 weeks of treatment
 - Back pain over 10 weeks of physiotherapy
 - Profits over months of the year
 - Radioactive decay

**Repeated Measures
Nested Within Individuals**















Level 2: Individual

↑

Level 1: Repeated measures for one variable
over time

**Repeated Measures
Nested Within Individuals**

Track mood of participants over the course of 7 days

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Participant 1							
Participant 2							
...							
...							

Third Type of Nesting

◉ Repeated measures nested within individuals

- Focus is not on change
- Focus is on associations between variables within an individual over time

Repeated Measures Within Individuals (Not Change)















Level 2: Individual

↑

Level 1: Repeated measures for more than one variable over time

Repeated Measures Within Individuals (Not Change)

Track mood and sleep quality of participants over the course of 7 days

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Participant 1							
Participant 2							
...							
...							

Fourth Type of Nesting

○ One-with-Many Design

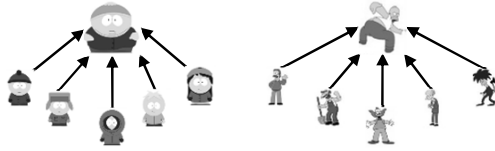
Level 2: Target (i.e., the participant)

↑

Level 1: Perceivers (e.g., friends or family members) rate the target on one or more dimensions

Fourth Type of Nesting

One-with-Many Design



Centering variables

- Generally two types of centering
 - Grand mean centering – subtract the mean for the entire sample from each observation in the sample
 - Group mean centering – subtract the mean for each group from each member of the group
- These two types of centering will allow you to answer slightly different questions
 - Grand mean centering helps you understand how an observation compares to the total average
 - Group mean centering helps you understand how an observation compares to the group average

Summary

- Data can be hierarchical and this nested structure is important
 - Most of the tests that you have learned about are not appropriate for hierarchical data
- Multilevel models are a variation of regression that allows you to estimate the variability in the slopes and intercepts within entities
 - Slopes and intercepts can be random variables (allowed to vary) rather than fixed (assumed to be equal in different situations)
- Growth curves model trends in the data over time
 - These trends can also have variable intercepts and slopes
