Introduction to Multilevel Models (MLMs) for Clustered Data

- Topics:
  - Features of clustered data
  - Features of clustered models
  - What can MLM do for you?
  - What to expect in this course

MLMs for Clustered* Data

- **Clustering = Nesting = Grouping = Hierarchies*
  - Key idea: More than one dimension of sampling simultaneously
  - **Micro** units are nested within one or more types of **macro** units; each type of unit is considered a "**level**" → **multilevel**
  - We will analyze data from two-level and three-level sampling designs with sources of nesting and/or crossing using **MLMs**

- MLMs are same as other terms you may have heard of:
  - **General Linear Mixed-Effects Models** (if you are from statistics)
    - *Mixed Effects* = Fixed and Random effects
  - **Random Coefficients Models** (also if you are from statistics)
    - Random coefficients = Random effects = latent variables/factors
  - **Hierarchical Linear Models** (if you are from education)
    - Not the same as *hierarchical regression*

Examples of *Nested* Designs

- Examples of **two-level** sampling designs:
  - Students (level 1) nested in classrooms/teachers (level 2)
  - Patients (level 1) nested in doctors (level 2)
  - Citizens (level 1) nested in countries (level 2)
  - Occasions (level 1) nested in persons (level 2)

- Examples of **three-level** sampling designs:
  - Students (level 1) nested in classrooms/teachers (level 2) nested in schools (level 3)
  - Patients (level 1) nested in doctors (level 2) nested in hospitals (level 3)
  - Citizens (level 1) nested in survey years (level 2) nested in countries (level 3)
  - Occasions (level 1) nested in persons (level 2) nested in families (level 3)

Examples of *Crossed* Designs

- Examples of **two-level cross-classified** sampling designs:
  - Students (level 1) nested in both schools (level 2) and neighborhoods (crossed level 2, so two level 2s)
  - Repeated measures design: Responses (level 1) nested in both subjects (level 2) and items (crossed level 2)
  - Reliability assessment: Ratings (level 1) nested in both raters (level 2) and targets (crossed level 2)
  - Students who change classrooms over time: occasions (level 1) nested in both students (level 2) and classrooms (crossed level 2)

- Example of **three-level cross-classified** sampling designs:
  - Time-specific ratings (level 1) nested in both children (level 2) and raters (crossed level 2); raters are nested within sites (level 3)
  - Responses (level 1) nested in both students (level 2) and items (crossed level 2); students are nested within schools (level 3)

Requirements for MLMs

- Multiple outcomes from same sampling unit (cluster)
  - Outcome must be measured at level 1 or else it's not level 1
  - 2 per cluster is the minimum (e.g., dyadic data), but 3+ are needed to distinguish "real" cluster differences in effects of person-level predictors from error
- More data is better (with diminishing returns)
  - More people → more power for person-level predictors
    - Clustered data will always require more people than independent data to achieve the same amount of person-level power
  - More clusters → better estimates of size of cluster differences; better prediction of those cluster differences
  - More items/stimuli/whatever → more power to show effects of differences between items/stimuli/whatever

Data Requirements for Our Models

- A useful outcome variable:
  - Has an interval scale* (for now)
    - A one-unit difference means the same thing across all scale points
    - **Other kinds of outcomes will be analyzed using generalized versions of multilevel models instead, but estimation will be more challenging*
    - **MLMs can be fit latent variables or observed variables*
  - Has scores with the same meaning over observations
    - Includes meaning of construct
    - Includes how items relate to the scale
    - Implies measurement invariance across clusters
- FANCY MODELS CANNOT SAVE BADLY MEASURED VARIABLES OR CONFOUNDED RESEARCH DESIGNS.

Levels of Analysis in Two-Level Nested Data

- Between-Cluster (BC) Variation:
  - **Level-2** – "**INTER**-cluster differences" – cluster characteristics
- Within-Cluster (WC) Variation:
  - **Level-1** – "**INTRA**-cluster differences" – person characteristics
- Any level-1 variable could have both BC and WC variation
  - BC = some clusters have more/less than other clusters
  - WC = some people have more/less than their cluster's average
  - Univariate MLMs handle this differently for level-1 predictors versus level-1 outcomes, but multivariate MLMs treat both the same way → aka "Multilevel Structural Equation Models"
- **So how do MLMs "handle" multiple levels of sampling?**

The Two Sides of Any Model

- **Model for the Means:**
  - *Aka* **Fixed Effects**, Structural Part of Model
  - What you are used to **caring about for testing hypotheses**
  - How the expected outcome for a given observation varies as a function of values on predictor variables

- **Model for the Variance:**
  - *Aka* **Random Effects and Residuals**, Stochastic Part of Model
  - What you are used to **making assumptions about** instead
  - How residuals are distributed and related across observations (persons, clusters, items, etc.) → these relationships are called "dependency" and ***this is the primary way that multilevel models differ from general linear models*** *(e.g., regression)*

Dimensions for Organizing Models

- Outcome type: General (normal) vs. General*ized* (not normal)
- Dimensions of sampling: One (so one variance term per outcome) vs. **Multiple** (so multiple variance terms per outcome) → **OUR WORLD**

- **General Linear Models:** conditionally normal outcome distribution, **fixed effects** (identity link; only one dimension of sampling)
- **General*ized* Linear Models: any conditional outcome distribution**, **fixed** effects through **link functions**, no random effects (one dimension)

Note: Least Squares is only for GLM

- **General Linear Mixed Models:** conditionally normal outcome distribution, **fixed and random effects** (identity link, but multiple sampling dimensions)
- **General*ized* Linear Mixed Models: any conditional outcome distribution**, **fixed and random effects** through **link functions** (multiple dimensions)
  - Same concepts, but with more complexity in estimation

- "Linear" means fixed effects predict the *link-transformed* conditional mean of DV in a linear combination of (effect*predictor) + (effect*predictor)…

What can MLM do for you?

1. **Model dependency across observations**
   - Longitudinal, clustered, and/or cross-classified data? No problem!
   - Tailor your model of sources of correlation to your data

2. **Include categorical or continuous predictors at any level**
   - Time-varying, person-level, cluster-level predictors for each variance
   - Explore reasons for dependency, don't just control for dependency

3. **Does not require same data structure for each unit**
   - Unbalanced or missing data? No problem!

4. **You already know how (or you will soon)!**
   - Use SPSS Mixed, **SAS Mixed**, **Stata**, Mplus, R, HLM, MlwiN…
   - What's an intercept? What's a slope? What's a pile of variance?

1. Model Dependency

- Sources of dependency depend on the sources of **variation** created by your sampling design: residuals for outcomes from the same unit are likely to be related, which violates the GLM "independence" assumption

- **"Levels" for dependency** ="levels of random effects"
  - Sampling dimensions can be **nested**
    - e.g., time within person, person within cluster, school within district
  - If you can't figure out the direction of your nesting structure, odds are good you have a **crossed sampling design** instead
    - e.g., persons crossed with items, raters crossed with targets
  - To have a "level", there must be random outcome variation due to sampling that **remains** after including the model's fixed effects
    - e.g., treatment vs. control does not create another level of "group" (but it would if you had multiple treatment and multiple control groups)

Dependency comes from…

- Mean differences across sampling units (e.g., clusters)
  - Creates *constant* dependency (covariance) over persons
  - Will be represented by a random intercept in our models

- Cluster differences in effects of person predictors
  - Creates non-constant dependency, the size of which depends on the value of the predictor for each person
  - Will be represented by random slopes in our models

- Side note: MLMs can be extended to test for and allow heterogeneity of variance of other kinds
  - e.g., within-classroom variance differs across classrooms
  - These are called "location-scale mixed-effects models"
    - See Don Hedeker's work: https://hedeker-sites.uchicago.edu/

Why care about dependency?

- In other words, what happens if we have the wrong model for the variance (assume independence instead)?

- **Validity of the tests of the predictors** depends on having the "most right" model for the variance → *all cluster dependency is accounted for via random effects*
  - Estimates will usually be ok → come from model for the means
  - Standard errors (and thus *p*-values) can be compromised

- The sources of variation that exist in your outcome will dictate **what kinds of predictors** will be useful
  - Between-Cluster variation needs Between-Cluster predictors
  - Within-Cluster variation needs Within-Cluster predictors
  - Between-Item variation needs Between-Item predictors…

2. Include categorical or continuous predictors at any level of analysis

- "ANOVA" test differences among discrete groups
- "Regression" tests slopes for continuous predictors
- What if a predictor is assessed repeatedly but can't be characterized by discrete "conditions"?
  - RM ANOVA or Regression won't work → you need MLM

- Some things don't change over time → time-invariant
- Some things do change over time → time-varying
- Some things are measured at higher levels
- Interactions are possible at same level or across levels

3. Does not require same data structure per person (by accident or by design)

RM ANOVA: uses **multivariate** (wide) data structure:

ID	Sex	T1	T2	T3	T4
100	0	5	6	8	12
101	1	4	7	.	11

People missing any data are excluded (data from ID 101 are not included at all)

MLM: uses **stacked** (long) data structure:

Only rows missing data are excluded

100 uses 4 cases

101 uses 3 cases

ID	Sex	Time	Y
100	0	1	5
100	0	2	6
100	0	3	8
100	0	4	12
101	1	1	4
101	1	2	7
101	1	3	.
101	1	4	11

Time can also be **unbalanced** across people such that each person can have his or her own measurement schedule: Time "0.9" "1.4" "3.5" "4.2"…

4. You already know how!

- If you can do GLM, you can do MLM
  (*and if you can do generalized linear models,*
  you can do generalized multilevel models, too)

- How do you interpret an estimate for…
  - the intercept?
  - the effect of a continuous variable?
  - the effect of a categorical variable?
  - a variance component (“pile of variance”)?

Course Requirements

- Readings were selected to expand on lecture materials and provide future reference → somewhat optional

- Homework assignments: NOT OPTIONAL and GRADED
  - Data analysis and interpretation using online homework system
  - 6 planned with due dates of Mondays by 11:59 PM
  - 3 points deducted for late submissions (86 points possible)

- Formative assessments: NOT OPTIONAL, NOT GRADED*
  - Provide structured review of topics during off-homework weeks
  - 7 planned with due dates of Mondays by 11:50 PM
  - 1 point deducted for late or lazy submissions (14 points possible)

Our Responsibilities

- My job:
  - Provide custom lecture materials and examples that are accurate, comprehensive, and with the necessary scaffolding for your use
  - Answer questions via email, in individual meetings, or in group-based office hours—you are ALL invited to work on homework during office hours and get immediate assistance if you want it

- Your job:
  - **Ask questions**—preferably in class, but any time is better than none
  - **Review** the class material **frequently**, focusing on mastering the vocabulary (words and symbols), logic, and procedural skills
  - **Practice** using the software to implement the techniques you are learning **on data you care about**—this will help you so much more!
  - **Read**: S&B is for reference, but others will be more directly relevant
  - Don't wait until the last minute to start homework, and don't be afraid to **ask for help** if you get stuck on one thing for more than 15 minutes

Class-Sponsored Statistical Software

- To help address the needs of different programs, I will show examples using both **SAS and STATA**
- I am a heavy-duty SAS user who picked up enough STATA to teach multilevel modeling workshops using it
  - So if you have STATA tips, please share them with me!

- Things to consider when choosing which one to focus on:
  - More programs = more entries in your "technical skills" part of CV
  - Both programs are available through the Uiowa Virtual Desktop, but **only SAS is available from off campus, too**
  - What program will be used in your quant classes to follow? What do the other members of your research lab use?
  - Btw, I can also help you in SPSS, Mplus, and a little bit of R

SAS vs. STATA: My Opinion

Activity	Winner	Commentary
Working with raw files or multiple datasets	SAS, hands down	As of STATA 15, only one dataset can be open at once—problematic for messy data management
Within-dataset manipulations	Tie, but STATA for some tasks	STATA wins for group-mean-centering, stacking, and unstacking (stay tuned)
Data analysis	Tie, but SAS for some tasks	I've had estimation problems in STATA for certain MLM variants (unstructured R, crossed random)
Post-estimation (get predictions or simple effects)	STATA, hands down	STATA has simple yet powerful options for doing these tasks in bulk that SAS doesn't have
Automating data tasks (i.e., loops)	Tie	Both programs have ways to do this, but I only know how in SAS…

This Semester's Topics

- Review of single-level linear models and interactions
- Univariate MLMs:
  - General MLMs for two-level nested data
  - General MLMs for two-level cross-classified data
  - General*ized* MLMs for two-level nested data
  - General*ized* MLMs for two-level cross-classified data
- Multivariate MLMs:
  - For testing differences in effect size across outcomes
  - For path analysis and mediation
- Three-level MLMs (type to be decided)
- Other topics as time permits