Introduction to this Course: Multilevel Models (MLMs) for Clustered Data

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

What To Expect This Semester

- I believe that **everyone is capable** and **can significantly benefit**** from learning more types of quantitative methods (*especially* multilevel models)!
- **Philosophy:** Focus on accessibility + mastery learning
  - No anxiety-prone tasks (e.g., hand calculations, memorizing formulas)
  - No need to resort to AI for help! That's the job of me and the TAs!
- **Materials:** Unit = lecture + example(s); 8 units planned
  - **Lecture** slides present concepts—the (wordy) **what** and the **why**
  - **Example** documents: reinforce the concepts and demonstrate the **how using software**—STATA or R (SAS also provided); Mplus too
  - All available at the course website (hosted outside of ICON)

** ***Benefits*** *include but are not limited to: Better research, more authorship opportunities, and actual money*

What To Expect This Semester

- I will NOT:
  - Use infrequent high-stakes testing to assess your level of learning
  - Assume you know any models beyond the GLM to begin with!
    - But we will also use generalized models (for non-normal outcomes)
    - I try to make connections to other modeling frameworks (i.e., MLM≈M-SEM)
- I WILL:
  - Use **formative assessments (FA)** to help you figure out what you need to review (6 planned; 12 points for **completing them at all**)
  - Require **online homework (HW) assignments** that give you real-data practice (6 planned; 88 points for **completing them accurately**)
    - Complete demo "HW 0" for 2 points extra credit
    - All canned data except for HW5, for which you can choose another canned data activity OR to do individual analysis using your own data
  - Link research designs, data, questions, and models explicitly
    - If we don't cover the exact combination you need in class, just ask me—odds are good I have an example from elsewhere that I can send you!

About the Course Requirements

- **Everything** is take-home, open-note, and untimed
- Late* work will be accepted (−2 for HW; −1 for FA)
  - **Extensions granted if requested at least 2 weeks in advance*
  - HW due dates **may be pushed later** (to ensure approximately 1 week after covering the material before it's due), but never sooner

- **Formative assessments:** Big-picture questions to provide a structured review (will go over answers at the next class)
  - Also likely to include case studies and correcting AI-generated text!

- **Homework assignments:** Practice doing data analysis
  - Based directly on examples given (no googling required)
  - You will (usually) have a unique dataset (with a common story)
  - **Computation** sections: Instant feedback, infinite attempts
  - **Results** (interpretation) sections: Delayed feedback, single attempt (but repetition of concepts and vocabulary across the semester)

Our Other Responsibilities

- My job (besides providing materials and assignments):
  - **Answer questions** via email, in individual meetings, or in group-based zoom office hours—you can each work on homework during office hours and get (near) immediate assistance (and then keep working)

- Your job (in descending order of timely importance):
  - **Ask questions**—preferably in class, but any time is better than none
  - **Frequently review** the class material, focusing on mastering the vocabulary (words and symbols), logic, and procedural skills
  - 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
    - Please email me a screenshot of your code+error so I can respond easily
  - **Do the readings** for a broader perspective and additional examples (best after lecture; readings are for the whole unit, not just that day)
  - **Practice** using the software to implement the techniques you are learning **on data you care about**—this will help you so much more!

More About Your Experience in this Class

- **Attendance**: Strongly recommended but not required
  - **You choose** (for any reason): In-person or zoom
  - **Masks** are still welcome for in-person attendees
  - **Please do not attend in-person if you were exposed to Covid!** (or monkey pox, or whatever new pandemic is next)
  - You won't miss out: I will post **YouTube-hosted recordings** (audio + screenshare only) for each class at the course website
  - **Ask questions aloud or in the zoom chat window (+DM)** (even if you are attending class in person)

- **Changes** will be sent via email by 9 am on class days
  - I will change to zoom-only if I am exposed to Covid!
  - I will change to zoom-only for dangerous weather
  - Nothing is more important than our health and safety…

Class-Sponsored Statistical Software

- To help address the needs of different Iowa degree programs, I will show examples using **STATA, R, and SAS** software
  - **STATA** (aka, Stata) = "Software for Statistics and Data Science"
  - **R** = free implementation of what was initially the "S" language
  - **SAS** = "Statistical Analysis System" (won't be featured in handouts, though)
- **Why not SPSS**? Because it doesn't have as much room to grow (and thus it isn't used in any other EMS advanced classes)
  - As in SPSS, drop-down windows can also generate syntax in STATA and in SAS "enterprise" (which I don't use, and you won't need to)
  - SPSS could be used for some—but not all—of our content
- **My story:** After SPSS, I became a heavy-duty **SAS enthusiast** who:
  - Picked up enough STATA initially to teach workshops using it, and I am learning it better now that I teach it in my classes
  - Is (begrudgingly) learning enough (base) R to add it to my classes, so it's possible you learned how to do the same things differently but correctly
  - So if you have **STATA or R tips**, <u>please</u> share them with me!

Which Program: SAS, STATA, or R?

- **I am assuming you know how to use at least one of these!**
  - Each is available (with VPN) in the free U Iowa Virtual Desktop
  - More programs = more "technical skills" for your CV; easier collaboration with colleagues (who only know one program)
  - For intro handouts and videos, please see materials posted 2/7 here

- **To consider** when choosing which program to focus on:
  - Future use: R can be freely installed on your own machine; SAS has a free web-based SAS OnDemand; STATA install = $$$
  - **STATA** is popular in fields that use **large, weighted survey data** (e.g., sociology, political science, public health, EPLS at Iowa)
  - **R** will be used exclusively in classes by Drs. Aloe, LeBeau, or Templin, and it has become increasingly mainstream, **but**:
    - R packages are only as good as their authors (so little quality control)
    - Syntax and capabilities are idiosyncratic to the packages (grrrrrr)
    - I have found incorrect or impossible results, even for "good" packages

Class-Sponsored Statistical Software

- **STATA and R** will be the primary packages emphasized
- PilesOfVariance.com currently has examples of every model in my textbook using SAS, SPSS, STATA, and Mplus (and some using R)
  - Some textbook examples may be used in class, along with unique examples
  - **STATA MIXED and R (various packages):** can used for homework and I will provide syntax and output for all unique class examples
    - For models in which the answers diverse across programs, two versions of the same homework will be offered (one in STATA, and one in R)
  - **SPSS MIXED:** syntax is very similar to that of SAS MIXED, and it has largely the same functionality (so you can use it for some of the course homework), but I will not be adding it to any unique course examples
  - **Mplus:** syntax is very different, and it does not have REML estimation, and thus you cannot use it for most of the homework in this class
    - We will use Mplus to do multivariate multilevel models (aka, multilevel "structural equation modeling"), such as in multilevel mediation

This Semester's Topics

- Review of single-level linear general and generalized models on your own as needed (especially interaction terms!)
  - See PSQF 6243 (for general) and PSQF 6270 (for generalized)
- 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 (aka, explanatory item response theory models)
- Multivariate MLMs (*aka*, multilevel structural equation models):
  - For path analysis and mediation; symmetry in predictors and outcomes
- Readings also provided for special topics as time permits:
  - Power analysis, three-level models, location–scale models

Multilevel Models (MLMs) for Clustered* Data

- ***Clustering = Nesting = Grouping = Hierarchies**
  - Key idea: Outcomes with >1 dimension of sampling simultaneously ("**micro**" units are nested in one or more types of "**macro**" units)
  - Each sampling dimension is considered its own "**level**" → **MLM**
  - MLMs can be used to predict outcomes from two-level (or more-level) sampling designs that result in nested and/or crossed observations

- The term "Multilevel Model" (MLM) has many synonyms:
  - **General Linear Mixed-Effects Models** (Fixed + Random = Mixed)
  - **Random Coefficients Models** (Random effects = latent variables)
  - **Hierarchical Linear Models** (HLM, but not = hierarchical regression)
  - Most MLM software is "univariate" → predict 1 outcome at a time
  - Multivariate MLMs can be estimated as "multilevel structural equation models" to predict 2+ outcomes at once (+ address missing predictors)

Examples of **Nested** Designs

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

- Examples of **three-level** sampling designs:
  - Students (level 1) nested in classes/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)

Examples of **Crossed** Designs

- Examples of **two-level cross-classified** sampling designs:
  - Two kinds of nesting: Students (level 1) nested in both schools (level-2) and neighborhoods (crossed at level 2)
  - Repeated measures: Responses (level 1) nested in both subjects (level 2) and items (crossed at level 2)
  - Reliability assessment: Ratings (level 1) nested in both raters (level 2) and targets (crossed at level 2)
  - Students who change classes over time: occasions (level 1) nested in both students (level 2) and classes (crossed at level 2)

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

Labels 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)

Note: Ordinary Least Squares is only for GLM

- **General*ized* Linear Models:** **any conditional outcome distribution**, **fixed** effects through **link functions**, no random effects (only one dimension)
- **General Linear Mixed Models:** conditionally normal outcome distribution, **fixed and random effects** (identity link, but **multiple dimensions** of sampling)
- **General*ized* Linear Mixed Models:** **any conditional outcome distribution**, **fixed and random effects** through **link functions** (**multiple dimensions**)
  - Same concepts as for general*ized* or mixed separately, but with more complexity in estimation

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

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 variable measured per person** could have both **L2 between and L1 within** variation!
  - BC = some clusters are higher/lower on average than other clusters
  - WC = some people are higher/lower than the rest of their cluster
  - Btw, univariate MLMs must address this differently for level-1 predictors vs. level-1 outcomes, but multivariate MLMs treat both the same way
- **So how do MLMs "handle" multiple levels of sampling?**

The Two Sides of *Any* Model

- **Model for the Means:**
  - **Fixed Effects**, the "structural" part (= latent variables means)
  - What you are used to **caring about for testing hypotheses**
  - How the expected outcome for a given observation varies as a function of their values for the predictor variables

- **Model for the Variance:**
  - **Random Effects and Residuals**, the "stochastic" or "error" part
    - Btw, random effect variances = latent variable variances
  - 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*** *(GLMs; "regression")*

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 Mixed, 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 “cluster” (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 independent observations 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, *so long as effects are disaggregated across levels of analysis!*
  - Standard errors (and thus *p*-values) will 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 Options for Cluster Dependency

Represent Cluster Differences via Fixed Effects

- Include (#clusters−1) binary predictors for cluster membership in the **model for the means** → **so cluster is NOT a model "level"**
  - Main effects control for cluster mean differences only; interactions with person predictors are also needed to control for cluster slope differences
- Useful if #clusters < 10ish or you care about specific clusters, but then you cannot include cluster predictors → saturated mean diffs

Represent Cluster Differences via Random Effects

- Include a random intercept variance across clusters in the **model for the variance** → **then cluster IS a model "level"**
  - A random intercept controls for cluster mean differences only; a random slope variance is needed for cluster differences in person predictor slopes
- Better if #clusters > 10ish or you want to **predict** cluster differences

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:

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

Only rows missing data are excluded

100 uses 4 cases

101 uses 3 cases

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 slope of a continuous variable?
  - the slope of a categorical variable?
  - a variance component ("pile of variance")?