Multilevel models

E0430

Week 2

HLM?

MLR?

Random effects model?

MLM?

HLR?

Mixed effects model?

HLM = MLM

MLM = mixed effects model

mixed effects model = random effects model

MLR = multiple linear regression

HLR = hierarchical linear regression

HL/ML regression

- **Multiple linear regression (MLR)**
  - Using several explanatory variables (predictors) to predict an outcome

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Health = β_0 + AgeX_1 + FemaleX_2 + DepressionX_3 + ε

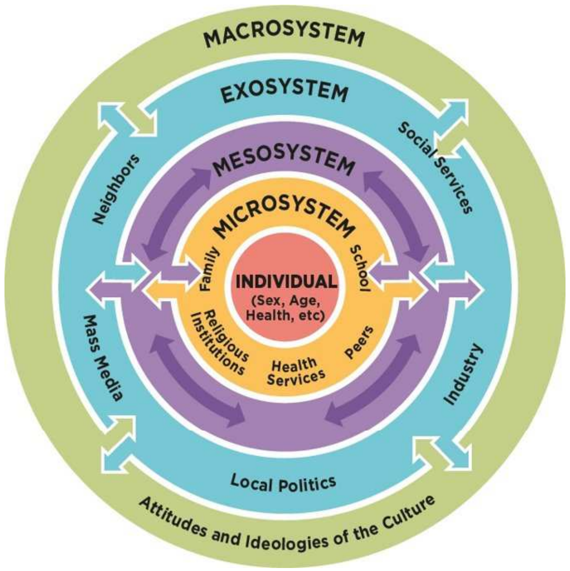
- **Hierarchical linear regression (HLR)**
  - Type of MLR for model comparison
  - Predictors separated into sets

Model 1: Health = β_0 + AgeX_1 + FemaleX_2 (R^2)

Model 2: Health = β_0 + AgeX_1 + FemaleX_2 + DepressionX_3 (R^2, ΔR^2 *(Model 2 – Model 1 R^2)*)

What is multilevel modeling?

- Modeling the **nesting/clustering** of individuals in higher-order structures = hierarchical structure
- Clusters = classroom, school, household, region, country, timepoint…



Necessity of multilevel models

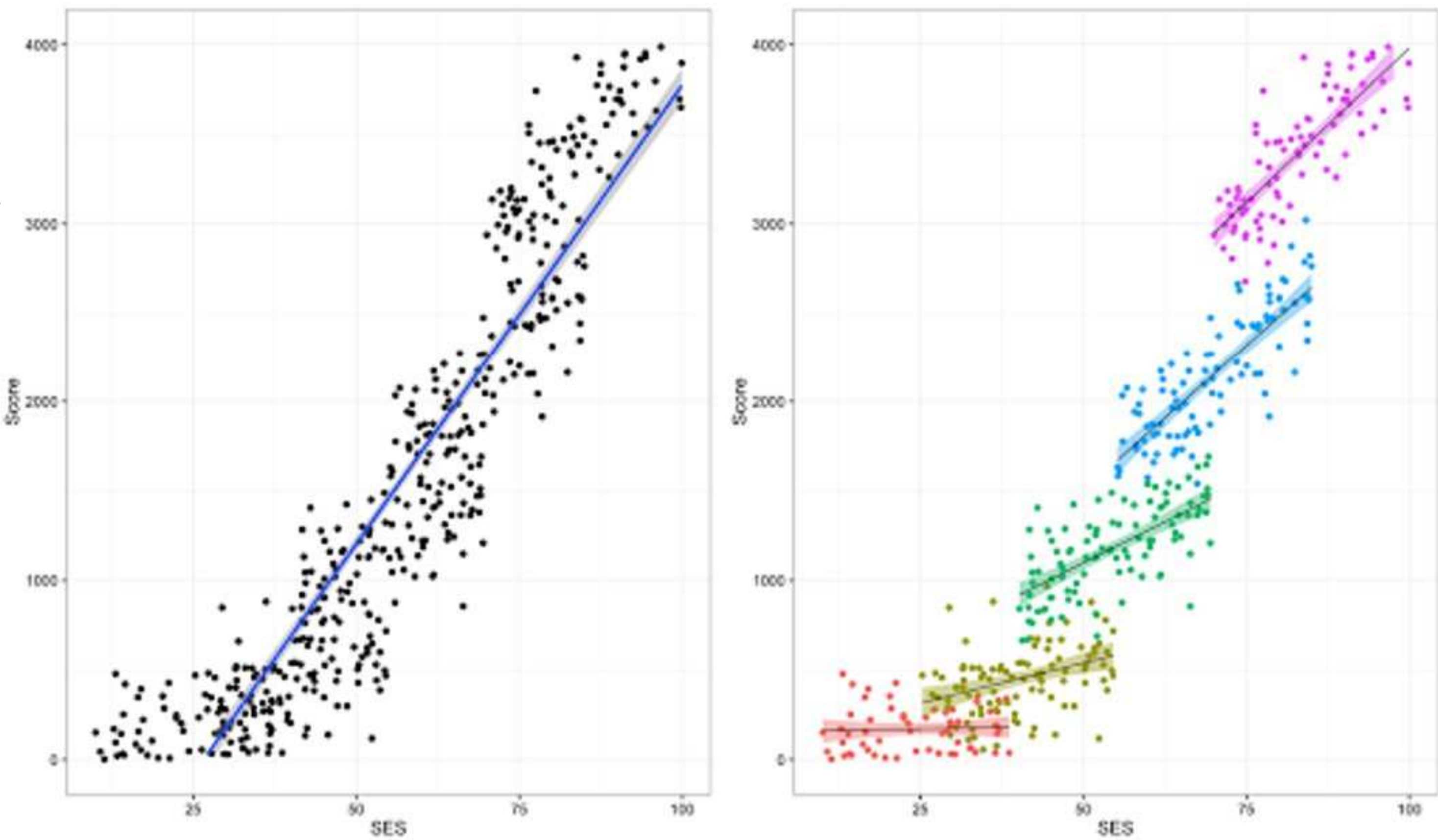
- **Ignoring** the nesting of individuals **violates** basic regression assumption – independence of errors
- Example:

About 100 married adults fill out a survey with questions about their household, marriage etc. The researchers then test the hypothesis where they predict marital satisfaction by household income in an OLS regression. They ignore the fact that many of the respondents are in fact married to each other!

Test scores and SES of students

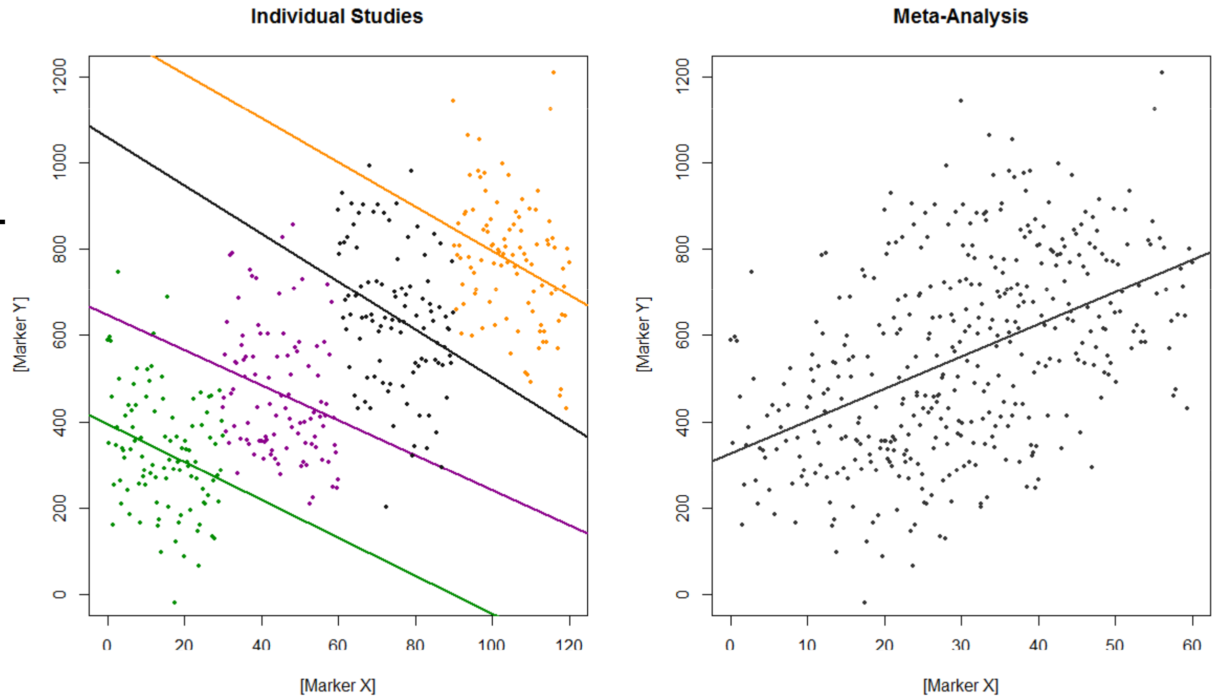
- Left panel –
  simple OLS regression

- Right panel
  each colour is a different school



https://www.r-bloggers.com/2016/10/multilevel-modeling-of-educational-data-using-r-part-1/

- Simpson's paradox – trend is different (opposite sign or non-existent) in groups vs in the whole sample



https://stats.stackexchange.com/questions/168732/simpsons-paradox-random-effects

MLM – the basics

- OLS Regression

$$Y_i = \beta_0 + \beta_1 X_i + r_i$$

Y_i = predicted outcome of *i*th person

β_0 = intercept when predictor β1 is zero

β_1 = slope of predictor (expected change with 1 unit increase in X)

r_i = unique effect associated with person *I* (random error)

Example

- Predicting Happiness by the number of pets
- Happiness measured on a 1-10 scale
- Number of pets (0-5)

$$Happiness_i = 6.7 + 0.24(Pets)_i$$

MLM

$$\text{Level 1: } Y_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j}$$

Y_{ij} = outcome for *i*th level 1 unit nested within *j*th level-2 unit

β_{0j} = level-1 intercept for the *j*th level-2 unit

r_{ij} = residual/unexplained variance at level 1

γ_{00} = level-2 intercept – overall grand mean

u_{0j} = residual/unexplained variance at level 2

Mixed model

Level 1: $Y_{ij} = \beta_{0j} + r_{ij}$
Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$

- Substituting for β_{0j} gives us

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

- Sometimes called "Mixed Model"
- This simple model is equivalent to one-way ANOVA with random effects

Example

- Match achievement of students in school (50 – 120)
- Students clustered in classrooms
  - Students = Level 1
  - Classroom = Level 2

$Math_{ij} = 84.35 + u_{0j} + r_{ij}$

84.35 = average math achievement of all classrooms (grand mean)

u_{0j} = variation due to differences among classrooms (difference from grand mean)

r_{ij} = variation due to individual

MLM with predictors

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$

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

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

Y_{ij} = outcome for *i*th level-1 unit nested within *j*th level-2 unit
β_{0j} = level-1 intercept for the *j*th level-2 unit
β_{1j} = slope of predictor (expected change with 1 unit increase in X)
r_{ij} = residual/unexplained variance at level 1

γ_{00} = level-2 intercept
γ_{01} = regression coefficient associated with W relative to level-1 intercept
W = value on the level-2 predictor for jth level-2 unit
u_{0j} = random effects of the jth level-2 unit (intercepts are random)

β_{1j} = slope for the jth level-2 unit
γ_{11} = effect of W for jth level-2 unit on level-1 slope
u_{1j} = random effect of j on the slope

Mixed model with predictors

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$

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

$= Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \upsilon_{0j} + \beta_{1j}X_{ij} + r_{ij} =$

$\mathbf{Y_{ij} = \gamma_{00} + \beta_{1j}X_{ij} + \gamma_{01}W_j + r_{ij} + \upsilon_{0j}}$

$$\mathbf{Y_{ij} = \gamma_{00} + \beta_{1j}X_{ij} + \gamma_{01}W_j + r_{ij} + u_{0j}}$$

Level 1 predictor = self-efficacy (1-5)

Level 2 predictor = classroom size (5-27)

$$Math_{ij} = \gamma_{00} + \beta_{1j}SelfEff_{ij} + \gamma_{01}ClassSiz_j + r_{ij} + u_{0j}$$

$$Math_{ij} = 80.16 + 0.81SelfEff_{ij} - 0.14ClassSiz_j + r_{ij} + u_{0j}$$

Variance partitioning

- We see that individual math achievement is affected by individual factors (self efficacy) as well as classroom factors (classroom size)

- In this way, there is variance in math achievement at the individual level (Level 1) and at the classroom level (Level 2)

How much of variance is at higher level?

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

γ_{00} = grand-mean
$u_{0j} + r_{ij}$ = variance Y_{ij}

Level-1 variance = σ^2
Level-2 variance = τ_{00}
In this case, $r_{ij} = \sigma^2$ and $u_{0j} = \tau_{00}$

$$\rho = \frac{\tau_{00}}{(\sigma^2 + \tau_{00})}$$

Intraclass coefficient

ρ = ICC

- Intraclass correlation = correlation between two observations within a cluster
- Tells how much of variance is due to level-2 grouping
- Ranges from 0 – 1
- Important first step in assessing suitability of MLM
- If negligible, one can go with OLS

What is substantial ICC?

- Not a single answer
- Some authors (Nezlek, 2008) argue that one should always use multilevel modeling when dealing with hierarchical data structure
- With low ICC (e.g., ICC = 0.03), the OLS and the MLM results will be almost identical
- Another issue when opting for MLM might be the number of level-2 units = if it is too few

Unconditional model

$$\text{Level 1: } Y_{ij} = \beta_{0j} + r_{ij}$$
$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j}$$

- Unconditional model = no predictors
- Also called unconstrained or null model
- One-way ANOVA testing for variability in the outcome by level-2 group
- Model used for ICC and for R^2

Exercise Data – popular2.sav

Simulated data for 2000 pupils in 100 schools (class)

Outcome: pupil popularity (1-10 sociometic rating)

Predictors:

Sex (girl = 1)

Extroversion (1-10, teacher rating)

Teacher experience (years)