SMART
School Mental Health Assessment Research & Training Center

HARBORVIEW
INJURY PREVENTION & RESEARCH CENTER
UNIVERSITY of WASHINGTON

SMARTSTATS: FOUNDATION OF MIXED MODELS

KEITH HULLENAAR, PHD
QUANTITATIVE METHODS ENJOYER
UNIVERSITY OF WASHINGTON
SMART CENTER & HIPRC
PSYCHIATRY AND BEHAVIORAL SCIENCE

SMART School Mental Health Assessment Research & Training Center

HARBORVIEW INJURY PREVENTION & RESEARCH CENTER

UNIVERSITY of WASHINGTON

LAND ACKNOWLEDGEMENT

We live and work on the **unceded** ancestral lands of the **Coast Salish people**, the land which touches the shared waters of all tribes and bands within the **Duwamish**, **Puyallup**, **Suquamish**, **Tulalip** and **Muckleshoot** nations, and pay our respects to elders past and present.

We work to create inclusive and respectful partnerships that **honor Indigenous cultures, histories, identities, and sociopolitical realities**, that dismantle ongoing legacies of settler colonialism, and that recognize the hundreds of Indigenous Nations who continue to resist, live, and uphold their sacred relations across their lands.

WELCOME TO SMARTSTATS

Our **mission** is to make quantitative methodologies freely accessible to **all who want to learn**.

Keith Hullenaar, PhD
SMARTstats Founder
Postdoc

Bethlehem Kebede, BS
SMART Center
Research Analyst

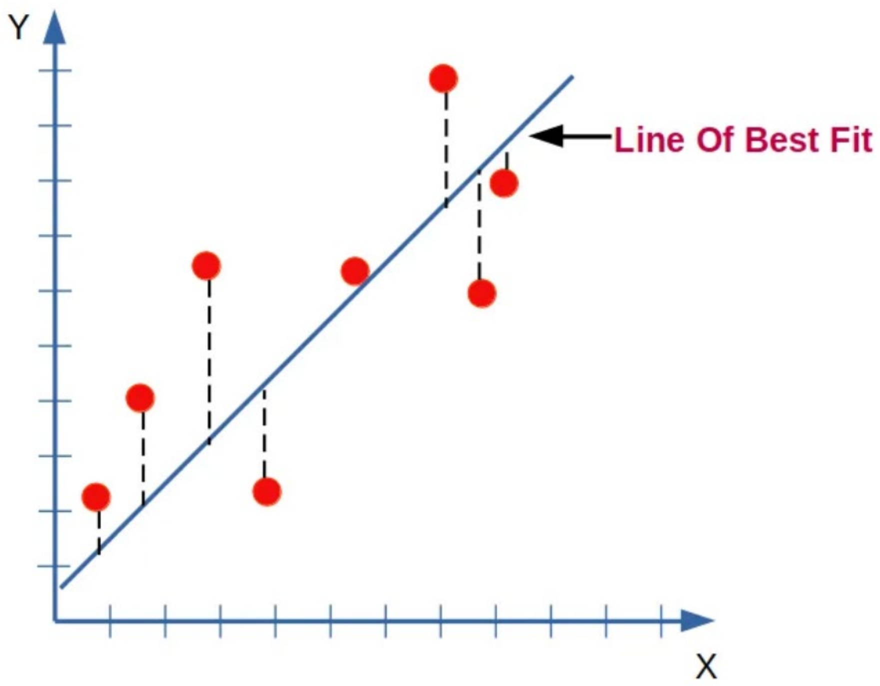
Casey Ehde, BA
SMART Center
Research Coordinator

Mahima Joshi, MPH
SMART Center
Research Scientist I

AGENDA

- What is a mixed model?
- Key concepts and notations
- Workshop and applied example

THE OLS WORKHORSE



The OLS Equation

$$Y = \beta_0 + \beta_X + \varepsilon$$

Y = dependent variable

X = intercept

β_1 = slope

ε = error term

One assumption:

Errors are **INDEPENDENT & IDENTICALLY DISTRIBUTED**

Our **standard errors, p-values, confidence intervals** can be biased if violated

DOES THIS ASSUMPTION HOLD IN REAL LIFE?

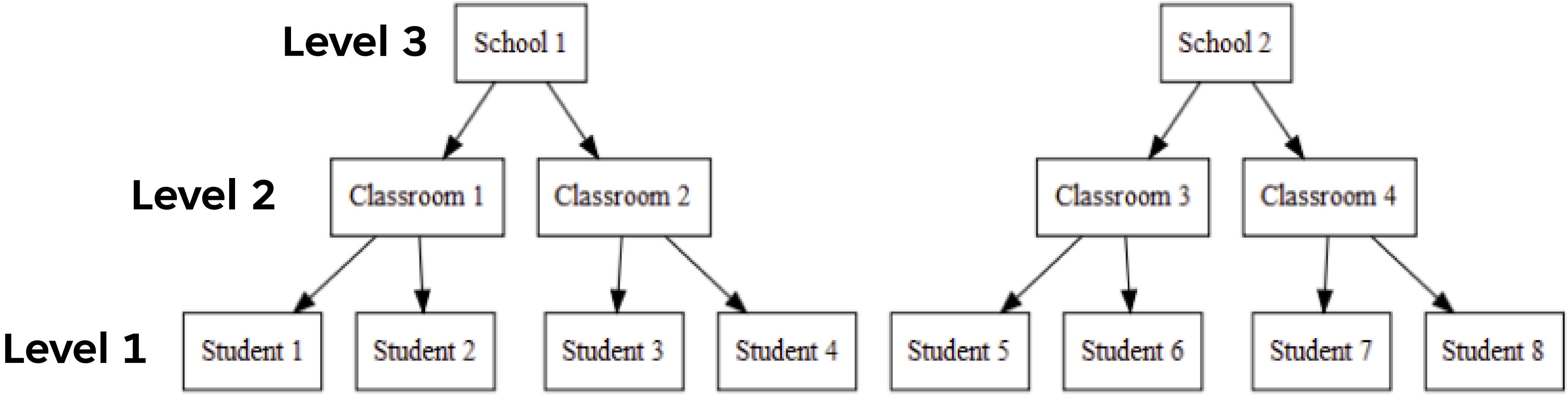
Sometimes but **definitely not always:]**

WHAT IS A MIXED/MULTILEVEL/HIERARCHICAL MODEL?

A **mixed (multilevel/hierarchical)** model estimates associations using **data structured at multiple levels** of observations

- Referred to as **clustered, hierarchical, multi-level, or nested data**

THE MULTI-LEVEL DATA STRUCTURE



3 Level Model

WHAT DOES A MIXED MODEL ESTIMATE?

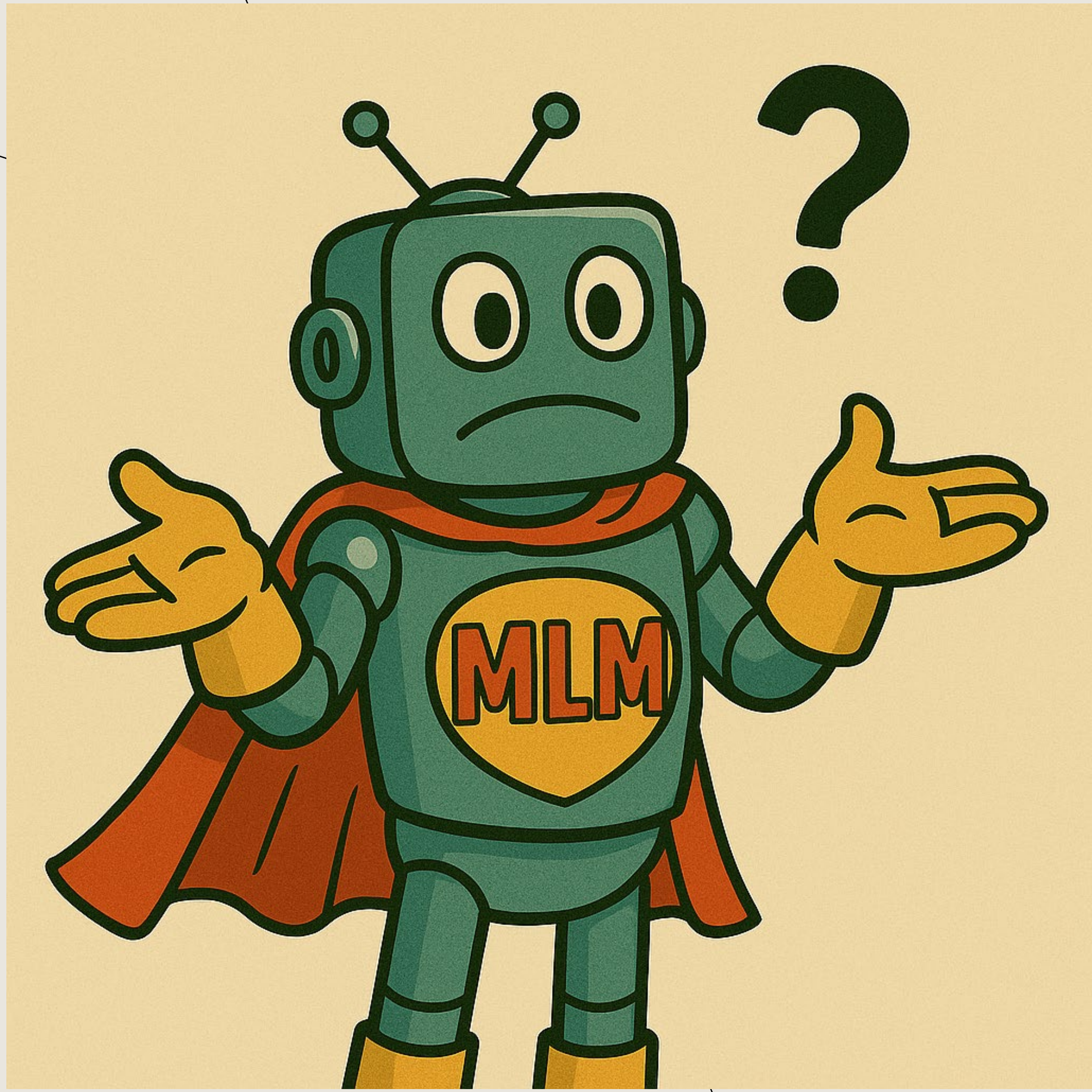
A **mixed (multilevel/hierarchical)** model estimates **fixed** and **random** effects

- **Fixed effect->**estimate for the **population averages;**
- **Random effect->**estimates **group-specific deviations and their variability**

Why mixed models really help:

- Estimates variation **across and within** groups
- Accounts for correlations within **higher-level units**
- **Can** produce **more accurate** standard errors (compared to OLS)

WHY IS THIS ALL WORTH IT?



POWER OF USING GROUP-LEVEL INFORMATION

How we **pool** the group-level information matters…

1) *No pooling:* Use an OLS for each group->noisy, unstable means
2) *Pool everything*: Use an OLS for all->biased if group means differ
3) *Partial pooling (mixed models)*: Group and total sample means inform the estimates through **a balanced approach**

Mixed models balance **group** and **total sample** info through **"shrinkage"**

Small or noisy groups "shrunk" **toward total sample** estimates

Large or stable groups remain **close to their own** estimates

Degree of group similarity/difference also matters (**ICC**)

INTRODUCING ICC

Intraclass correlation coefficient (Ranges from 0 to 1)

- Proportion of total variance explained by group differences
- Between-group var / (Between-group var+Within-group var)
- Correlation among individuals within a group
- **Higher ICC->Mixed models more impactful**
- **But even a low ICC can be impactful (with large clusters)**

How **ICC** affects levels of “shrinkage”

- **High ICC-**> shrinkage to the mean **is smaller**
- **Low ICC**-> shrinkage toward the mean **is larger**

MIXED MODEL CONSIDERATIONS AND ASSUMPTIONS

Define your levels (e.g., people within groups, time within people)

Choose random effects carefully (model problems can increase)

Groups need enough data and variance (fewer groups are harder)

Key assumptions

- Similar to **OLS**
- Random effects **independent** of residual errors
- **Exchangeability:** Any pair within a group has the same covariance/correlation
  - *Within groups:*
    - **Random-intercept-only models imply exchangeable within-group correlation** (same correlation for any pair within a group).
    - **With random slopes (or longitudinal data), correlation can depend on predictors/time** (not necessarily exchangeable).
  - *Between groups:* Random group effects drawn from same distribution

WE HAVE TO LOOK AT NOTATION NOW…

Level 1

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk} t_{ijk} + R_{ijk}$$

Level 2

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} TX_{jk} + U_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k} TX_{jk} + U_{1jk}$$

Level 3

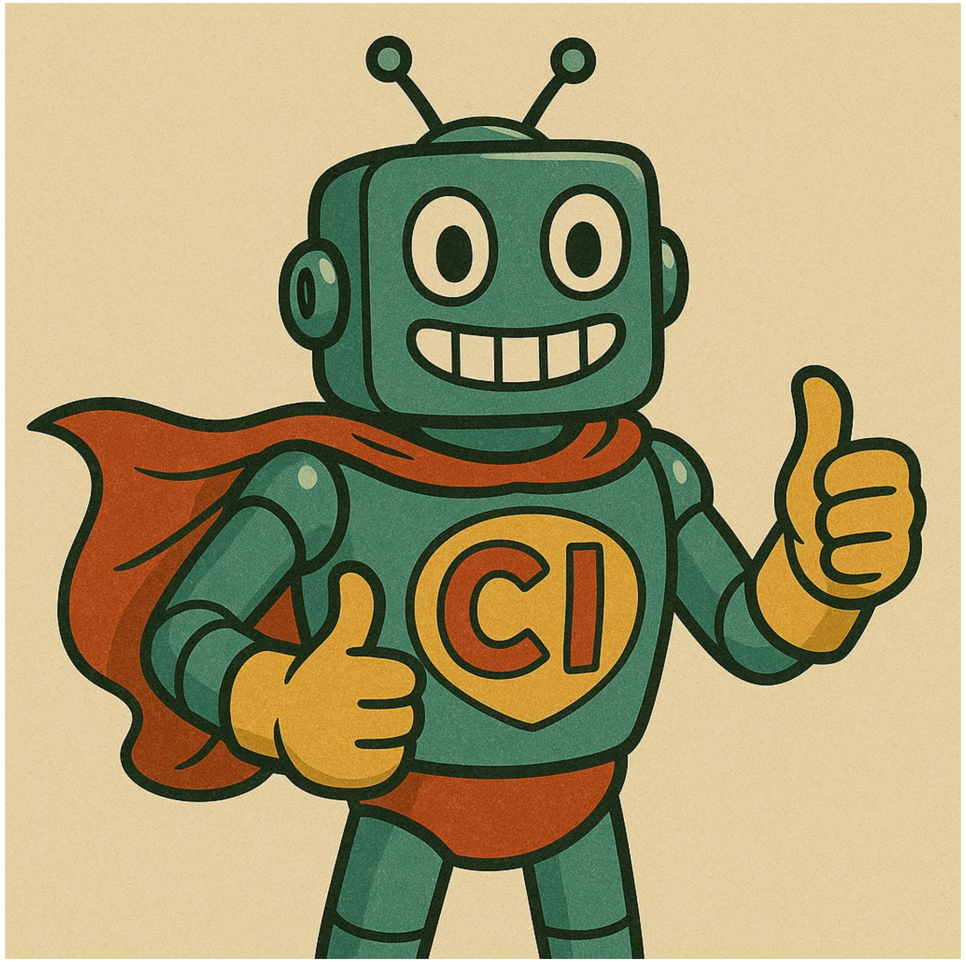
$$\gamma_{00k} = \delta_{000} + V_{0k}$$

$$\gamma_{10k} = \delta_{100} + V_{1k}$$

$$\gamma_{01k} = \delta_{010} + V_{2k}$$

$$\gamma_{11k} = \delta_{110} + V_{3k}$$

BUT IT WILL BE FINE!



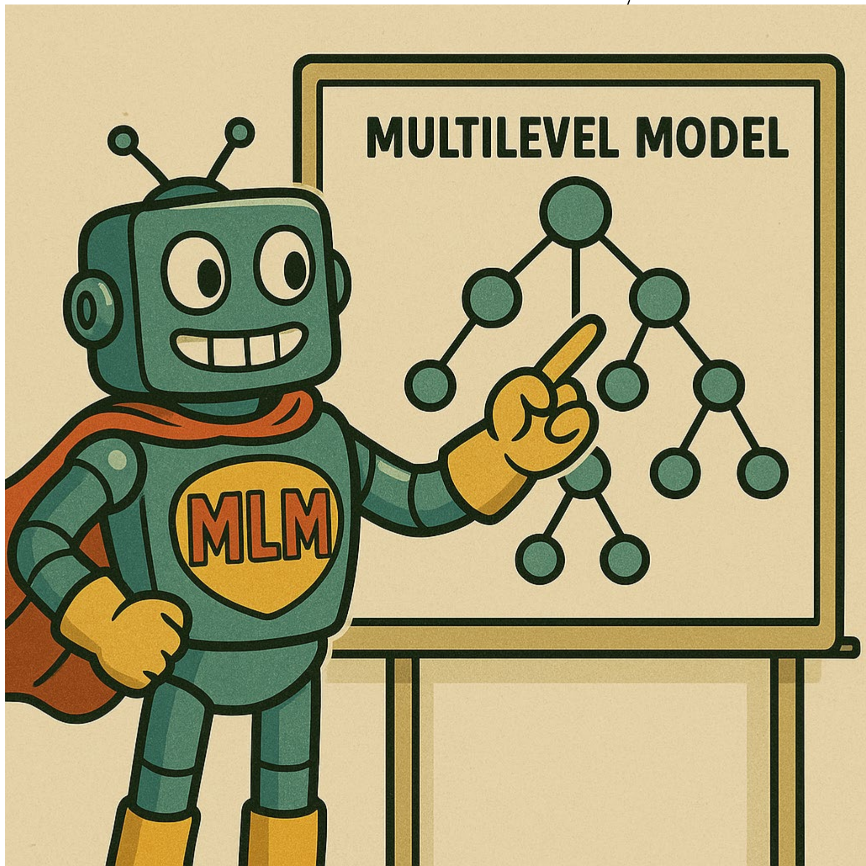
SIMPLE OLS MODEL

OLS: $y_i = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} x_i + \varepsilon_i$

$\boldsymbol{\beta_0}$: intercept (same for all)

$\boldsymbol{\beta_1}$: slope (same for all)

Assumes errors are independent and identically distributed



2-LEVEL WITH RANDOM INTERCEPT (NULL EMPTY MODEL)

Level 1 (individual *i* in group *j*):

$y_{ij} = \boldsymbol{\beta_{0j}} + \varepsilon_{ij}$

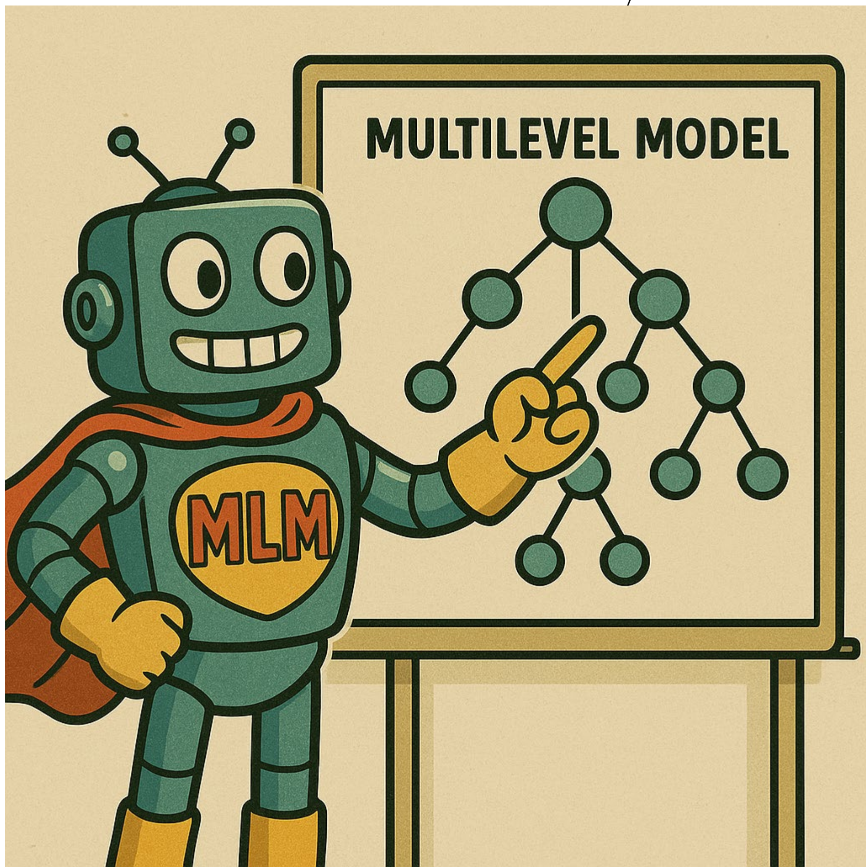
Level 2 (group *j*):

$\boldsymbol{\beta_{0j} = \gamma_{00} + u_{0j}}$

$\boldsymbol{\beta_{0j}}$: intercept (different by groups)

$\boldsymbol{\gamma_{00}}$: grand mean (mean for all)

$\boldsymbol{u_{0j}}$: group-specific deviation



2-LEVEL WITH RANDOM INTERCEPT (NULL EMPTY MODEL)

Level 1 (individual *i* in group *j):*

$y_{ij} = \boldsymbol{\beta_{0j}} + \varepsilon_{ij}$

Level 2 (group *j*):

$\boldsymbol{\beta_{0j}} = \boldsymbol{\gamma_{00}} + \mathbf{u_{0j}}$

$\boldsymbol{\beta_{0j}}$: intercept (different by groups)

$\boldsymbol{\gamma_{00}}$: grand mean (mean for all)

$\mathbf{u_{0j}}$: group-specific deviation

$$\mathbf{ICC} = \frac{\mathrm{Var}(u_{0j})}{\mathrm{Var}(u_{0j}) + \mathrm{Var}(\varepsilon_{ij})}$$

2-LEVEL MODEL WITH RANDOM INTERCEPT AND FIXED SLOPE

Level 1 (individual *i* in group *j*):

$y_{ij} = \boldsymbol{\beta}_{\mathbf{0}j} + \beta_1 x_{ij} + \varepsilon_{ij}$

Level 2 (group *j*):

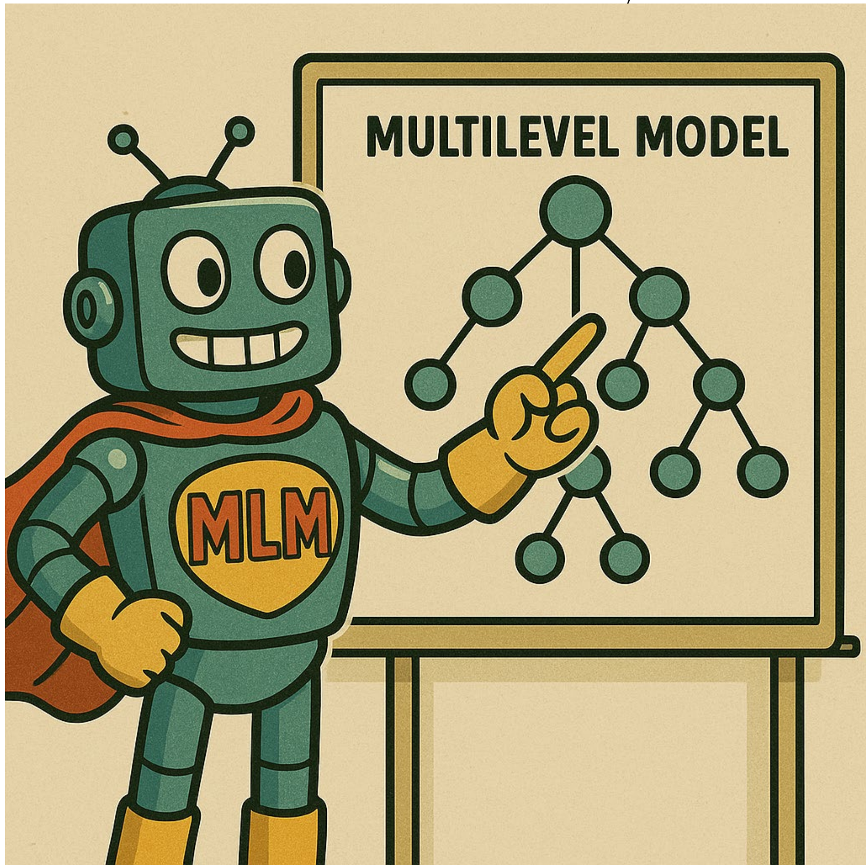
$\boldsymbol{\beta}_{\mathbf{0j}} = \boldsymbol{\gamma}_{\mathbf{00}} + \mathbf{u}_{\mathbf{0j}}$

$\boldsymbol{\beta}_{\mathbf{0j}}$: intercept (different by groups)

$\boldsymbol{\gamma}_{\mathbf{00}}$: grand mean (mean for all)

$\mathbf{u}_{\mathbf{0j}}$: group-specific deviation

β_1: fixed slope (same for all)



2-LEVEL MODEL WITH RANDOM INTERCEPT AND RANDOM SLOPE

Level 1 (individual *i*): $y_{ij} = \boldsymbol{\beta_{0j}} + \boldsymbol{\beta_{1j}} x_{ij} + \varepsilon_{ij}$

Level 2 (group *j*):

$\boldsymbol{\beta_{0j}} = \boldsymbol{\gamma_{00}} + \mathbf{u_{0j}}$

$\boldsymbol{\beta_{1j}} = \boldsymbol{\gamma_{10}} + \mathbf{u_{1j}}$

$\boldsymbol{\beta_{0j}}$: intercept (different by groups)

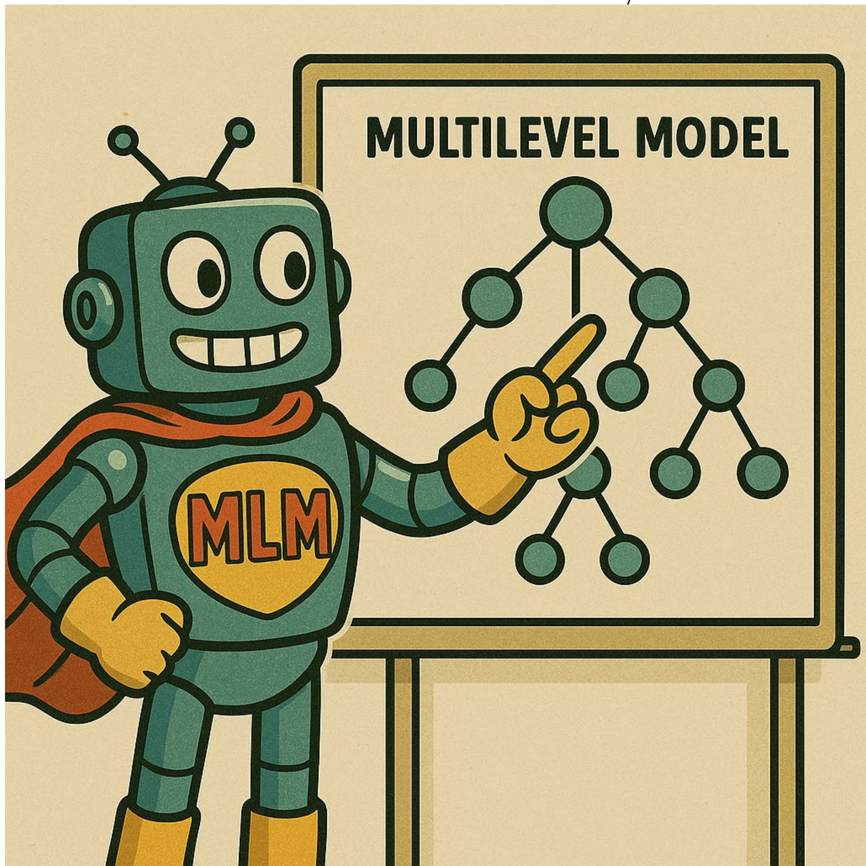
$\boldsymbol{\gamma_{00}}$: grand mean (mean for all)

$\mathbf{u_{0j}}$: group-specific deviation

$\boldsymbol{\beta_{1}}$: slope (different by groups)

$\boldsymbol{\gamma_{10}}$: overall slope (slope for all)

$\mathbf{u_{1j}}$: group-specific deviation



2-LEVEL MODEL WITH RANDOM INTERCEPT AND RANDOM SLOPE

Level 1 (individual *i*): $y_{ij} = \boldsymbol{\beta_{0j}} + \boldsymbol{\beta_{1j}} x_{ij} + \varepsilon_{ij}$

Level 2 (group *j*):

$\boldsymbol{\beta_{0j}} = \boldsymbol{\gamma_{00}} + \mathbf{u_{0j}}$

$\boldsymbol{\beta_{1j}} = \boldsymbol{\gamma_{10}} + \mathbf{u_{1j}}$

Cov($\mathbf{u_{0j}}, \mathbf{u_{1j}}$) [typically]

$\boldsymbol{\beta_{0j}}$: intercept (different by groups)

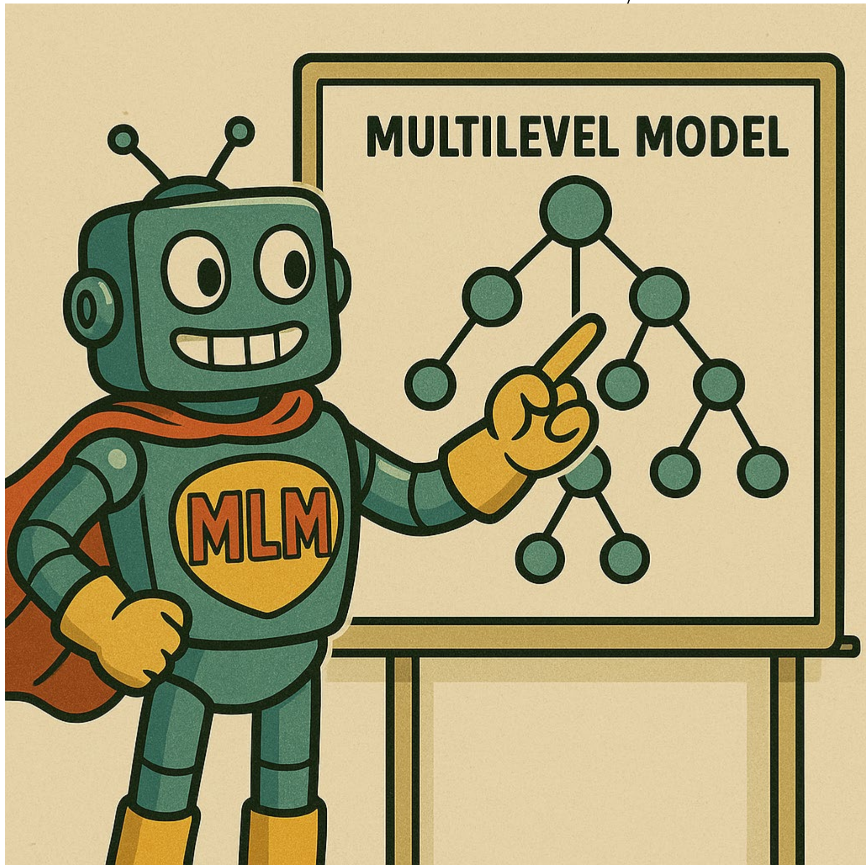
$\boldsymbol{\gamma_{00}}$: grand mean (mean for all)

$\mathbf{u_{0j}}$: group-specific deviation

$\boldsymbol{\beta_{1}}$: slope (different by groups)

$\boldsymbol{\gamma_{10}}$: overall slope (slope for all)

$\mathbf{u_{1j}}$: group-specific deviation



CENTERING VARIABLES AND INTERPRETATION

If X (ie, primary exposure) is not centered, intercepts **can be meaningless**

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij}$$

Grand mean centering

$$\boldsymbol{X_{ij} - \overline{X}}$$

Intercept = expected outcome at the *sample mean* of X

Group mean centering

$$\boldsymbol{X_{ij} - \overline{X_j}}$$

Intercept = expected outcome at the *group mean* of X

SELECTING A MODEL USING TESTS

Do we need a more complicated model?

We can test that!

Likelihood ratio test (Testing simple model against complex model)

- Significant LRT (p < .05) indicates that you should keep the complex model

AIC and BIC

- Lower values indicate a “better fitting” model