Multilevel Modeling (Part 2)

Random-Intercept and Random-Slope Modeling with Covariates

Desmond D. Wallace

Department of Political Science
The University of Iowa
Iowa City, IA

April 13, 2018

Including Level-1 Covariates

- Predicting the outcome from an intercept that varies between groups and individual-level independent variables.
- The 2-level model takes the following form:
  - Level-1 Model: $y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij}$
  - Level-2 Models
    - Intercept: $\beta_{0j} = \gamma_{00} + U_{0j}$ where

      γ_{00} – Average (general) intercept holding across all groups (fixed effect)

      U_{0j} – Group-specific effect on the intercept (random effect)
    - Slope: $\beta_{1j} = \gamma_{10}$ where

      γ_{10} – Amount of increase (decrease) in dependent variable for a one-unit change in x_{ij} (fixed effect)
  - Full Specification: $y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + \varepsilon_{ij}$

Assumptions

1. U_{0j} and ε_{ij} are mutually independent with mean 0, given the values of x_{ij}
2. U_{0j} is randomly drawn from a population distribution with mean 0 and variance τ_0^2
3. Population variance of level-1 residuals, σ^2, is constant across groups
4. U_{0j} are interpretable as group-level residuals, or group effects left unexplained by x_{ij}
5. **Unexplained variability at multiple levels is essence of multilevel modeling**

Variance of y_{ij}

- Variance of y_{ij}, conditioned on the values of x_{ij} is again the sum of the level-two and level-one variances
  - $Var(y_{ij} \mid X_{ij}) = Var(U_{0j}) + Var(\varepsilon_{ij}) = \tau_0^2 + \sigma^2$
- Covariance between two observations from the same group (ij and $i'j$) is equal to the variance of the contribution U_{0j} that is shared by these observations
  - $Cov(y_{ij}, y_{i'j} \mid x_{ij}, x_{i'j}) = Var(U_{0j}) = \tau_0^2$

Residual Intraclass Correlation Coefficient

- A part of the covariance or correlation between two observations from the same group is *explained* by values of the independent variable(s)
- The rest of the covariance or correlation is *unexplained*
- $\rho(y \mid X) = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$
- Represents correlation between the y-values of two randomly drawn individuals in a randomly drawn group, controlling for x
- If $\rho(y \mid X) = 0$, OLS is appropriate; If $\rho(y \mid X) > 0$, then multilevel model is better

Example: Math Achievement

- Goal: Examine the influence students' socioeconomic status (SES) has on math achievement scores while controlling for students' minority and gender identification.
- Model Specification
  - Level-1 Model (student-level):
    $\text{mathach}_{ij} = \beta_{0j} + \beta_{1j}\text{SES}_{ij} + \beta_{2j}\text{minority}_{ij} + \beta_{3j}\text{female}_{ij} + \varepsilon_{ij}$
  - Level-2 Models (school-level):
    - $\beta_{0j} = \gamma_{00} + U_{0j}$
    - $\beta_{1j} = \gamma_{10}$
    - $\beta_{2j} = \gamma_{20}$
    - $\beta_{3j} = \gamma_{30}$
  - Full Model:
    $\text{mathach}_{ij} = \gamma_{00} + \gamma_{10}\text{SES}_{ij} + \gamma_{20}\text{minority}_{ij} + \gamma_{30}\text{female}_{ij} + U_{0j} + \varepsilon_{ij}$

Example: Math Achievement

- Parameter Interpretation
  - γ_{00}: Overall mean of student's math achievement scores
  - γ_{10}: Effect SES has on math achievement
  - γ_{20}: Difference in math achievement between minorities and non-minorities
  - γ_{30}: Difference in math achievement between females and males
  - U_{0j}: Unique effect of school j on mean math achievement score

Random Slopes

- Belief the relationship between independent and dependent variables differs across groups
- The 2-level model takes the following form:
  - Level-1 Model: $y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij}$
  - Level-2 Models
    - Intercept: $\beta_{0j} = \gamma_{00} + U_{0j}$ where

      γ_{00} – Average (general) intercept holding across all groups (fixed effect)

      U_{0j} – Group-specific effect on the intercept (random effect)
    - Slope: $\beta_{1j} = \gamma_{10} + U_{1j}$ where

      γ_{10} – Average relationship of x_{ij} and y_{ij} across groups (fixed effect)

      U_{1j} – Group-specific variation of the relationship between x_{ij} and y_{ij} (random effect)
  - Full Specification: $y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{0j} + U_{1j} x_{ij} + \varepsilon_{ij}$

Assumptions

- All residuals (U_{0j}, U_{1j}, and ε_{ij}) have mean 0, given the values of the independent variable(s)
- The pair of random effects (U_{0j}, U_{1j}) are independent and identically distributed ($i.i.d$)
- (U_{0j}, U_{1j}) are independent of (ε_{ij})
- ε_{ij} is $i.i.d$

Variances

- Random Effects
  - $Var(U_{0j}) = \tau_{00} = \tau_0^2$
  - $Var(U_{1j}) = \tau_{11} = \tau_1^2$
  - $Cov(U_{0j}, U_{1j}) = \tau_{01}$
- y_{ij}
  - $Var(y_{ij} \mid X_{ij}) = \tau_0^2 + 2\tau_{01}x_{ij} + \tau_1^2 x_{ij}^2 + \sigma^2$
  - $Cov(y_{ij}, y_{i'j} \mid x_{ij}, x_{i'j}) = \tau_0^2 + \tau_{01}(x_{ij} + x_{i'j}) + \tau_1^2 x_{ij} x_{i'j}$
  - Residual variance is minimal for $x_{ij} = \frac{-\tau_{01}}{\tau_1^2}$
    - If within the range of possible x_{ij} values, variance will first decrease, then increase
    - If smaller than all x_{ij} values, variance will increase as a function of x
    - If larger than all x_{ij} values, variance will decrease as a function of x

Example: Math Achievement

- Goal: Examine the influence students' socioeconomic status (SES) has on math achievement scores while controlling for students' minority and gender identification, while accounting for the effect of SES varying across schools.
- Model Specification
  - Level-1 Model (student-level): $\text{mathach}_{ij} = \beta_{0j} + \beta_{1j}\text{SES}_{ij} + \beta_{2j}\text{minority}_{ij} + \beta_{3j}\text{female}_{ij} + \varepsilon_{ij}$
  - Level-2 Models (school-level):
    - $\beta_{0j} = \gamma_{00} + U_{0j}$
    - $\beta_{1j} = \gamma_{10} + U_{1j}$
    - $\beta_{2j} = \gamma_{20}$
    - $\beta_{3j} = \gamma_{30}$
  - Full Model: $\text{mathach}_{ij} = \gamma_{00} + \gamma_{10}\text{SES}_{ij} + \gamma_{20}\text{minority}_{ij} + \gamma_{30}\text{female}_{ij} + U_{0j} + U_{1j} + \varepsilon_{ij}$

Example: Math Achievement

- Parameter Interpretation
  - γ_{00}: Overall mean of student's math achievement scores
  - γ_{10}: Average effect SES has on math achievement
  - γ_{20}: Difference in math achievement between minorities and non-minorities
  - γ_{30}: Difference in math achievement between females and males
  - U_{0j}: Unique effect of school j on mean math achievement score
  - U_{1j}: Unique effect of school j on SES effect on math achievement score

Explaining Random Intercept and Random Slope Variation

- So far, coefficients have been the sum of an average and random effect.
- One could further explain this random variability via inclusion of group-level variables (Z)
- Example (Single group-level variable):
  - Random Intercept: $\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + U_{0j}$
  - Random Slope: $\beta_{1j} = \gamma_{10} + \gamma_{11} z_j + U_{1j}$
- Including a group-level variable in the random intercept equation leads to a main effect of z_j
- Including a group-level variable in the random slope equation leads to an interaction effect of $z_j x_{ij}$ (Cross-level Interaction)
- Just as with level-1 variables, can feature multiple level-2 variables

Example: Math Achievement

- Goal: Examine the influence students' socioeconomic status (SES) has on math achievement scores.
- Model Specification
  - Level-1 Model (student-level):
    $\text{mathach}_{ij} = \beta_{0j} + \beta_{1j}\text{SES}_{ij} + \beta_{2j}\text{minority}_{ij} + \beta_{3j}\text{female}_{ij} + \varepsilon_{ij}$
  - Level-2 Models (school-level):
    - $\beta_{0j} = \gamma_{00} + \gamma_{01}\text{size}_j + \gamma_{02}\text{sector}_j + U_{0j}$
    - $\beta_{1j} = \gamma_{10} + \gamma_{11}\text{size}_j + U_{1j}$
    - $\beta_{2j} = \gamma_{20}$
    - $\beta_{3j} = \gamma_{30}$
  - Full Model: $\text{mathach}_{ij} = \gamma_{00} + \gamma_{01}\text{size}_j + \gamma_{02}\text{sector}_j + \gamma_{10}\text{SES}_{ij} + \gamma_{11}\text{size}_j * \text{SES}_{ij} + \gamma_{20}\text{minority}_{ij} + \gamma_{30}\text{female}_{ij} + U_{0j} + U_{1j} + \varepsilon_{ij}$

Example: Math Achievement

- Parameter Interpretation
  - γ_{00}: Overall mean of student's math achievement scores
  - γ_{01}: Effect school size has on overall mean of student's math achievement scores when SES $= 0$
  - γ_{02}: Difference in overall mean of student's math achievement scores for schools in sectors coded as 1 compared to schools in sectors coded as 0.
  - γ_{10}: Average effect SES has on math achievement when size $= 0$
  - γ_{11}: Average effect SES has on math achievement depends on school size
  - γ_{20}: Difference in math achievement between minorities and non-minorities
  - γ_{30}: Difference in math achievement between females and males
  - U_{0j}: Unique effect of school j on mean math achievement score
  - U_{1j}: Unique effect of school j on SES effect on math achievement score

Email: desmond-wallace@uiowa.edu
Any Questions?