Panel data examples

ncgraetz

September 2020

1 Mixed effects

OLS:

$$Y_{ij} = \beta_0 + \beta_1(Age) + \epsilon_{ij} \tag{1}$$

Random intercepts:

$$Y_{ij} = \beta_0 + \beta_1(Age) + u_i + \epsilon_{ij} \tag{2}$$

$$u_i = N(0, \sigma_u^2) \tag{3}$$

$$Y_{ij} = \beta_0 + \beta_1(Age) + \boldsymbol{F}(i) + \epsilon_{ij} \tag{4}$$

Independent random slops and intercepts:

$$Y_{ij} = \beta_0 + \beta_1(Age) + u_i + t_j + \epsilon_{ij} \tag{5}$$

$$u_i = N(0, \sigma_u^2) \tag{6}$$

$$t_j = N(0, \sigma_t^2) \tag{7}$$

Correlated random slops and intercepts:

$$Y_{ij} = \beta_0 + \beta_1(Age) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{8}$$

$$\boldsymbol{Z_{ij}} = N(0, \boldsymbol{G}) \tag{9}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{10}$$

Correlated random slops and intercepts, Age*Program from class:

$$Y_{ij} = \beta_0 + \beta_1(Age) + \beta_2(Program) + \beta_3(Age * Program) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{11}$$

$$\boldsymbol{Z_{ij}} = N(0, \boldsymbol{G}) \tag{12}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{13}$$

Interaction model

$$Y_{ij} = \beta_0 + \beta_1(Age) + \beta_2(Program) + \beta_3(Age * Program) \tag{14}$$

Interaction model with random slopes/intercepts

$$Y_{ij} = \beta_0 + \beta_1(Age) + \beta_2(Program) + \beta_3(Age * Program) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{15}$$

$$\boldsymbol{Z_{ij}} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G}\right) \tag{16}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{17}$$

$$\epsilon_{ij} = N(0, \sigma_\epsilon^2) \tag{18}$$

Interaction model with random slopes/intercepts

$$Y_{ij} = \beta_0 + \beta_1(Age) + \beta_2(Program) + \beta_3(Age * Program) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{19}$$

$$\boldsymbol{Z_{ij}} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G}\right) \tag{20}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{21}$$

$$\epsilon_{ij} = N(0, \sigma_\epsilon^2) \tag{22}$$

Interaction model with random slopes/intercepts

$$Y_{ij} = \beta_0 + \beta_1 (Age) + \beta_2 (Program) + \beta_3 (Age * Program) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{23}$$

$$\boldsymbol{Z_{ij}} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G} \right) \tag{24}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{25}$$

$$\epsilon_{ij} = N(0, \sigma_\epsilon^2) \tag{26}$$

Interaction model with random slopes/intercepts

$$Y_{ij} = \beta_0 + \beta_1 (Age) + \beta_2 (Program) + \beta_3 (Age * Program) + \boldsymbol{Z_{ij}} + \epsilon_{ij} \tag{27}$$

$$\boldsymbol{Z_{ij}} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{G} \right) \tag{28}$$

$$\boldsymbol{G} = \begin{bmatrix} \sigma_u^2 & \sigma_{ut}^2 \\ \sigma_{ut}^2 & \sigma_t^2 \end{bmatrix} \tag{29}$$

$$\epsilon_{ij} = N(0, \sigma_\epsilon^2) \tag{30}$$

Assumptions/decisions

- The functional form of the structural specification, or *fixed effects* (likelihood function, link function, variables)
- Distribution(s) of the stochastic specification, or *random effects*, $\boldsymbol{Z_{ij}}$
- What is the distribution of $\boldsymbol{Z_{ij}}$? What is the shape of $\boldsymbol{G}$? What else could it be? What is your data generating process?

Why are these mixed effects models (or “hierarchical models”) very often Bayesian?

- **Practically**: With computers, way easier and more efficient to fit complex model specifications. Often virtually impossible to estimate in frequentist framework (MLE) - for example, autocorrelated models (spatial, temporal) with large N.

- **Conceptually**: At a very high level, there are trade offs involving considerations such as sample size. The Bayesian approach will give you the most different answer when samples are small (<30 in strata). This is because of what people often refer to as “drawing strength” across levels. Within-group estimates in small samples are pulled towards the global mean (essentially becomes a weighted average based on the within-group N and the global N).

Model A) "unconditional growth model" (Week 4, slide 22)

$$Y_{ij} = \beta_0 + \beta_1(T_{ij}) + \alpha_i + \gamma_i(T_{ij}) + \epsilon_{ij} \tag{31}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\gamma} \\ \sigma^2_{\alpha\gamma} & \sigma^2_\gamma \end{bmatrix} \right) \tag{32}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{33}$$

Yellow highlights your "fixed effects", or what we are treating as fixed unknown constants (β_0 is the global intercept or average initial value, and β_1 is the global time effect or average time slope). The other parameters are your "random effects", or what we are treating as unknown random variables. Because these are random variables, they require specifying another equation to define their distributions. Here we are saying they are multivariate normal, so their joint distribution is defined by the variance parameters highlighted in green: the variance for the intercept random effect σ^2_α, the variance for the time random effect (σ^2_γ), and their covariance ($\sigma^2_{\alpha\gamma}$). In the R package `nlme`, running `summary()` on a fitted model object gives you these three terms as standard deviations, so you could square them to get the variances. Most software for mixed effect models, including `nlme`, automatically return summary statistics for the fixed effects (yellow terms above), fitted variance terms defining your random effects (green terms above), and the residual variance (orange term above). This is because you know the mean of all your random effects is 0. If you want the actual random effects themselves for every ij, you can do `random.effects()` on your fitted model object.

We can add unemployment status (U) to the fixed effects above to get Model B, and write Model B in both the composite form above by simply adding one more yellow term (β_2):

$$Y_{ij} = \beta_0 + \beta_1(T_{ij}) + \beta_2(U_{ij}) + \alpha_i + \gamma_i(T_{ij}) + \epsilon_{ij} \tag{34}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\gamma} \\ \sigma^2_{\alpha\gamma} & \sigma^2_\gamma \end{bmatrix} \right) \tag{35}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{36}$$

Or we can write this same Model B using a level-1/level-2 specification. Notice how we are just rearranging terms from the composite specification into multiple lines (Week

4, slide 23):

$$Y_{ij} = \beta_{0i} + \beta_{1i}(T_{ij}) + \beta_{2i}(U_{ij}) + \epsilon_{ij} \tag{37}$$

$$\beta_{0i} = \beta_0 + \alpha_i \tag{38}$$

$$\beta_{1i} = \beta_1 + \gamma_i \tag{39}$$

$$\beta_{2i} = \beta_2 \tag{40}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\gamma} \\ \sigma^2_{\alpha\gamma} & \sigma^2_\gamma \end{bmatrix} \right) \tag{41}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{42}$$

And we can fit Model A and Model B using `nlme` in R:

```
modelA <- lme(cesd~months, data=unemployment, random=~months|id, method="ML")
```

```
modelB <- lme(cesd~months+unemp, data=unemployment, random=~months|id,
method="ML")
```

In Model C, we interact time and unemployment in our fixed effects:

$$Y_{ij} = \beta_0 + \beta_1(T_{ij}) + \beta_2(U_{ij}) + \beta_3(T_{ij} * U_{ij}) + \alpha_i + \gamma_i(T_{ij}) + \epsilon_{ij} \tag{43}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\gamma} \\ \sigma^2_{\alpha\gamma} & \sigma^2_\gamma \end{bmatrix} \right) \tag{44}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{45}$$

Model C as level-1/level-2:

$$Y_{ij} = \beta_{0i} + \beta_{1i}(T_{ij}) + \beta_{2i}(U_{ij}) + \beta_{3i}(T_{ij} * U_{ij}) + \epsilon_{ij} \tag{46}$$

$$\beta_{0i} = \beta_0 + \alpha_i \tag{47}$$

$$\beta_{1i} = \beta_1 + \gamma_i \tag{48}$$

$$\beta_{2i} = \beta_2 \tag{49}$$

$$\beta_{3i} = \beta_3 \tag{50}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\gamma} \\ \sigma^2_{\alpha\gamma} & \sigma^2_\gamma \end{bmatrix} \right) \tag{51}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{52}$$

Fit Model C in R:

```
modelC <- lme(cesd~months*unemp, data=unemployment, random=~months|id,
method="ML")
```

Notice above, we assume the main effect and the interactive effect of unemployment are fixed and do not vary by individual. Finally in Model D, we introduce a random effect to both of these terms, but remove the fixed main effect of time:

$$Y_{ij} = \beta_0 + \beta_1(U_{ij}) + \beta_2(T_{ij} * U_{ij}) + \alpha_i + \zeta_i(U_{ij}) + \delta_i(T_{ij} * U_{ij}) + \epsilon_{ij} \tag{53}$$

$$\begin{bmatrix} \alpha_i \\ \zeta_i \\ \delta_i \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\alpha & \sigma^2_{\alpha\zeta} & \sigma^2_{\alpha\delta} \\ \sigma^2_{\zeta\alpha} & \sigma^2_\zeta & \sigma^2_{\zeta\delta} \\ \sigma^2_{\delta\alpha} & \sigma^2_{\delta\zeta} & \sigma^2_\delta \end{bmatrix} \right) \tag{54}$$

$$\epsilon_{ij} = N(0, \sigma^2_\epsilon) \tag{55}$$

Model D as level-1/level-2 (again, notice how we are just rearranging terms from the composite model):

$$\begin{aligned}
Y_{ij} &= \beta_{0i} + \beta_{1i}(U_{ij}) + \beta_{2i}(T_{ij} * U_{ij}) && (56)\\
\beta_{0i} &= \beta_0 + \alpha_i && (57)\\
\beta_{1i} &= \beta_1 + \zeta_i && (58)\\
\beta_{2i} &= \beta_2 + \delta_i && (59)\\
\begin{bmatrix} \alpha_i \\ \zeta_i \\ \delta_i \end{bmatrix} &= N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\alpha} & \sigma^2_{\alpha\zeta} & \sigma^2_{\alpha\delta} \\ \sigma^2_{\zeta\alpha} & \sigma^2_{\zeta} & \sigma^2_{\zeta\delta} \\ \sigma^2_{\delta\alpha} & \sigma^2_{\delta\zeta} & \sigma^2_{\delta} \end{bmatrix} \right) && (60)\\
\epsilon_{ij} &= N(0, \sigma^2_{\epsilon}) && (61)
\end{aligned}$$

Fit Model D in R:

```
modelD <- lme(cesd unemp+unemp:months,
random= unemp+unemp:months|id,
data=unemployment,
control = list(opt = "optim", tolerance=1e-6))
```

Models A-D here are fit in R in the script I uploaded for the TA session, where we also visualized the outputs of the fixed effects (example_code.R under Module 5). These correspond to the results table from Week 4, slide 24. Notice how here, we only report the variance components from the random effects, the diagonal of the variance-covariance matrix above, rather than the covariance terms. We could also report these, but it can quickly get cluttered if those are not interesting for your research question.