Efficient Analysis of Three-Level Cross-Classified Linear Models with Ignorable Missing Data

Yongyun Shin
Virginia Commonwealth University

COMPSTAT
Aug. 24, 2010

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305D090022 to NORC.

Summary

1. Desired Model
   - Data Structure
   - Hierarchical Cross-classified Linear Model (HCLM)

2. Motivation for Research
   - Missing Data at Multiple Levels
   - Efficient Analysis of HCLM

3. Estimation and MI with Ignorable Missing Data

4. Discussion/Future Research

Desired Model: Data Structure

(a) child (row) j, school (column) k, n_{jk} occasions (cell size)

j\k	1	2	3	4	K=5
1	$(1,1,n_{11})$		$(1,3,n_{13})$	$(1,4,n_{14})$	
2		$(2,2,n_{22})$			
3	$(3,1,n_{31})$	$(3,2,n_{32})$			
4	$(4,1,n_{41})$				$(4,5,n_{45})$
5			$(5,3,n_{53})$		
J=6				$(6,4,n_{64})$	

Desired Model: HCM

$$R_{ijk} = W_{ijk}^T\gamma_1 + Y_{rj}^T\gamma_2 + Y_{ck}^T\gamma_3$$
$$+ C_{ijk}^T\gamma_4 + C_{rj}^T\gamma_5 + C_{ck}^T\gamma_6 + D_{rijk}^T u_{rj} + D_{cijk}^T u_{ck} + e_{ijk}$$

for occasion $i = 1, \cdots, n_{jk}$, child $j = 1, \cdots, J$, school $k = 1, \cdots, K$

where

R_{ijk} is a response variable;

covariates are (W_{ijk}, C_{ijk}) at time level or level 1,

(Y_{rj}, C_{rj}) at child or row level,

(Y_{ck}, C_{ck}) at school or column level;

D_{rijk} having child-specific $u_{rj} \sim N(0, \tau_r)$;

D_{cijk} having school-specific $u_{ck} \sim N(0, \tau_c)$;

$e_{ijk} \sim \mathrm{N}(0, \sigma^2)$ is a level-1 random error.

Desired Model: HCM

$$R_{ijk} = W_{ijk}^T \gamma_1 + Y_{rj}^T \gamma_2 + Y_{ck}^T \gamma_3 + C_{ijk}^T \gamma_4 + C_{rj}^T \gamma_5 + C_{ck}^T \gamma_6 + D_{rijk}^T u_{rj} + D_{cijk}^T u_{ck} + e_{ijk}.$$

-Estimation Challenging

(Raudenbush '93; Goldstein and Rasbash '94; Hill and Goldstein '98)

-Bayesian methods to ease computational burden

(Clayton and Rasbash '99; Browne et al. '01)

Motivation: Efficient Analysis

$$R_{ijk} = W_{ijk}^T \gamma_1 + Y_{rj}^T \gamma_2 + Y_{ck}^T \gamma_3 + C_{ijk}^T \gamma_4 + C_{rj}^T \gamma_5 + C_{ck}^T \gamma_6 + D_{rijk}^T u_{rj} + D_{cijk}^T u_{ck} + e_{ijk}$$

where

$Y_{1ijk} = \begin{bmatrix} R_{ijk} \\ W_{ijk} \end{bmatrix}, Y_{rj}, Y_{ck}$ are subject to missingness;

C_{ijk}, C_{rj}, C_{ck} are completely observed;

D_{rijk} having child-specific $u_{rj} \sim N(0, \tau_r)$;

D_{cijk} having school-specific $u_{ck} \sim N(0, \tau_c)$;

$e_{ijk} \sim \mathrm{N}(0, \sigma^2)$ is a level-1 random error.

Estimation with Ignorable Missing Data

- Key Idea : 3 steps

1. Express joint model of $(Y_{1ijk}, Y_{rj}, Y_{ck})$ given $(C_{ijk}, C_{rj}, C_{ck}, D_{rijk}, D_{cijk})$;

2. Estimate the joint model efficiently
   - computation not burdened by cross-classification
   - via the EM;

3. Multiply impute based on the joint model.

Step 1: Express Joint Model

$$\begin{bmatrix} Y_{1ijk} \\ Y_{rj} \\ Y_{ck} \end{bmatrix} = \begin{bmatrix} X_{1ijk} & 0 & 0 \\ 0 & X_{rj} & 0 \\ 0 & 0 & X_{ck} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_r \\ \beta_c \end{bmatrix} + \begin{bmatrix} Z_{rijk} & 0 \\ 0 & I_{p_{r2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1j} \\ r_{2j} \end{bmatrix} + \begin{bmatrix} Z_{cijk} & 0 \\ 0 & 0 \\ 0 & I_{p_{c2}} \end{bmatrix} \begin{bmatrix} c_{1k} \\ c_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ijk} \\ 0 \\ 0 \end{bmatrix}$$

where

$$X_{1ijk} = I_{p_1} \otimes [C_{ijk}^T C_{rj}^T C_{ck}^T], \; X_{rj} = I_{p_{r2}} \otimes [C_{rj}^T C_{ck}^T], \; X_{ck} = I_{p_{c2}} \otimes C_{ck}^T,$$

$$Z_{rijk} = diag\{D_{rijk}^T, I_{p_1 - 1}\}, \quad Z_{cijk} = diag\{D_{cijk}^T, I_{p_1 - 1}\},$$

$$r_j \sim N(0, \pi_r), \quad c_k \sim N(0, \pi_c), \quad \varepsilon_{ijk} \sim \mathrm{N}(0, \pi_1)$$

for

$$r_j = \begin{bmatrix} r_{1j} \\ r_{2j} \end{bmatrix}_{p_{r1}+p_{r2}}, \pi_r = \begin{bmatrix} \pi_{r11} & \pi_{r12} \\ \pi_{r21} & \pi_{r22} \end{bmatrix}, c_k = \begin{bmatrix} c_{1k} \\ c_{2k} \end{bmatrix}_{p_{c1}+p_{c2}}, \pi_c = \begin{bmatrix} \pi_{c11} & \pi_{c12} \\ \pi_{c21} & \pi_{c22} \end{bmatrix}.$$

Estimate $\theta = (\beta_1, \beta_r, \beta_c, \pi_r, \pi_c, \pi_1)$.

Step 2: Estimate Joint Model

(a)

j\k	1	2	3	4	K=5
1	$(1,1,n_{11})$		$(1,3,n_{13})$	$(1,4,n_{14})$	
2		$(2,2,n_{22})$			
3	$(3,1,n_{31})$	$(3,2,n_{32})$			
4	$(4,1,n_{41})$				$(4,5,n_{45})$
5			$(5,3,n_{53})$		
J=6				$(6,4,n_{64})$	

(b)

			Summed cell sizes	# schools attended
$(1,1,n_{11})$	$(1,3,n_{13})$	$(1,4,n_{14})$	n_1	$K_1 = 3$
$(2,2,n_{22})$			n_2	$K_2 = 1$
$(3,1,n_{31})$	$(3,2,n_{32})$		n_3	$K_3 = 2$
$(4,1,n_{41})$	$(4,5,n_{45})$		n_4	$K_4 = 2$
$(5,3,n_{53})$			n_5	$K_5 = 1$
$(6,4,n_{64})$			n_6	$K_6 = 1$

Estimate $\theta_r = (\beta_r, \pi_r, \beta_1, \pi_1)$

(c)

	$(1,1,n_{11})$	$(2,2,n_{22})$	$(1,3,n_{13})$	$(1,4,n_{14})$	$(4,5,n_{45})$
	$(3,1,n_{31})$	$(3,2,n_{32})$	$(5,3,n_{53})$	$(6,4,n_{64})$	
	$(4,1,n_{41})$				
Summed cell sizes	N_1	N_2	N_3	N_4	N_5
# students In school	$J_1 = 3$	$J_2 = 2$	$J_3 = 2$	$J_4 = 2$	$J_5 = 1$

Estimate $\theta_c = (\beta_c, \pi_c)$

Step 2: Estimate Joint Model

$$\text{Notations:}$$

$$Y_{1jk} = \begin{bmatrix} Y_{11jk} \\ \vdots \\ Y_{1n_{jk}jk} \end{bmatrix}, X_{1jk} = \begin{bmatrix} X_{11jk} \\ \vdots \\ X_{1n_{jk}jk} \end{bmatrix}, Z_{rjk} = \begin{bmatrix} Z_{r1jk} \\ \vdots \\ Z_{rn_{jk}jk} \end{bmatrix}, Z_{cjk} = \begin{bmatrix} Z_{c1jk} \\ \vdots \\ Z_{cn_{jk}jk} \end{bmatrix}, \varepsilon_{jk} = \begin{bmatrix} \varepsilon_{1jk} \\ \vdots \\ \varepsilon_{n_{jk}jk} \end{bmatrix},$$

$$d_{1jk} = Y_{1jk} - X_{1jk}\beta_1, \quad d_{rj} = Y_{rj} - X_{rj}\beta_r, \quad d_{ck} = Y_{ck} - X_{ck}\beta_c.$$

$$Y_r = \begin{bmatrix} Y_{r1} \\ \vdots \\ Y_{rJ} \end{bmatrix}, \quad X_r = \begin{bmatrix} X_{r1} \\ \vdots \\ X_{rJ} \end{bmatrix}, \quad d_r = Y_r - X_r\beta_r, \quad r_1 = \begin{bmatrix} r_{11} \\ \vdots \\ r_{1J} \end{bmatrix}, \quad r_2 = \begin{bmatrix} r_{21} \\ \vdots \\ r_{2J} \end{bmatrix},$$

$$Y_c = \begin{bmatrix} Y_{c1} \\ \vdots \\ Y_{cK} \end{bmatrix}, \quad X_c = \begin{bmatrix} X_{c1} \\ \vdots \\ X_{cK} \end{bmatrix}, \quad d_c = Y_c - X_c\beta_c, \quad c_1 = \begin{bmatrix} c_{11} \\ \vdots \\ c_{1K} \end{bmatrix}, \quad c_2 = \begin{bmatrix} c_{21} \\ \vdots \\ c_{2K} \end{bmatrix}.$$

Step 2: Estimate Joint Model

(b)

			Summed cell sizes	# schools attended
$(1,1,n_{11})$	$(1,3,n_{13})$	$(1,4,n_{14})$	n_1	$K_1=3$
$(2,2,n_{22})$			n_2	$K_2=1$
$(3,1,n_{31})$	$(3,2,n_{32})$		n_3	$K_3=2$
$(4,1,n_{41})$	$(4,5,n_{45})$		n_4	$K_4=2$
$(5,3,n_{53})$			n_5	$K_5=1$
$(6,4,n_{64})$			n_6	$K_6=1$

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{K_1 \times K}$$

$$C_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}_{K_2 \times K}$$

$$\vdots$$

$$C_6 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{K_6 \times K}$$

$$d_{cK_j} = (C_j \otimes I_{p_{c2}})d_c, \quad c_{1K_j} = (C_j \otimes I_{p_{c1}})c_1, \quad c_{2K_j} = (C_j \otimes I_{p_{c2}})c_2.$$

Step 2: Estimate Joint Model

$$\begin{bmatrix} d_{1j} \\ d_{rj} \\ d_{cK_j} \end{bmatrix} = \begin{bmatrix} Z_{rj} & 0 \\ 0 & I_{p_{r2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1j} \\ r_{2j} \end{bmatrix} + \begin{bmatrix} \oplus_{k=1}^{K_j} Z_{cjk} & 0 \\ 0 & 0 \\ 0 & I_{K_j \times p_{c2}} \end{bmatrix} \begin{bmatrix} c_{1K_j} \\ c_{2K_j} \end{bmatrix} + \begin{bmatrix} \varepsilon_j \\ 0 \\ 0 \end{bmatrix}$$

where

$$d_{1j} = \begin{bmatrix} d_{1j1} \\ \vdots \\ d_{1jK_j} \end{bmatrix}, \quad Z_{rj} = \begin{bmatrix} Z_{rj1} \\ \vdots \\ Z_{rjK_j} \end{bmatrix}, \quad \varepsilon_j \sim N(0, I_{n_j} \otimes \pi_1),$$

$$\begin{bmatrix} c_{1K_j} \\ c_{2K_j} \end{bmatrix} \sim N\left(0, \begin{bmatrix} I_{K_j} \otimes \pi_{c11} & I_{K_j} \otimes \pi_{c12} \\ I_{K_j} \otimes \pi_{c21} & I_{K_j} \otimes \pi_{c22} \end{bmatrix} \right)$$

for $j = 1, \cdots, J$.

Estimate $\theta_r = (\beta_r, \pi_r, \beta_1, \pi_1)$.

Step 2: Estimate Joint Model

(c)

$(1,1,n_{11})$	$(2,2,n_{22})$	$(1,3,n_{13})$	$(1,4,n_{14})$	$(4,5,n_{45})$
$(3,1,n_{31})$	$(3,2,n_{32})$	$(5,3,n_{53})$	$(6,4,n_{64})$	
$(4,1,n_{41})$				
Summed cell sizes: N_1	N_2	N_3	N_4	N_5
# students In school: $J_1 = 3$	$J_2 = 2$	$J_3 = 2$	$J_4 = 2$	$J_5 = 1$

$$d_{rJ_k} = (R_k \otimes I_{p_{r2}})d_r,$$
$$r_{1J_k} = (R_k \otimes I_{p_{r1}})r_1,$$
$$r_{2J_k} = (R_k \otimes I_{p_{r2}})r_2.$$

$$R_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}_{J_1 \times J}$$

$$R_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}_{J_2 \times J}$$

$$\vdots$$

$$R_5 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}_{J_6 \times J}$$

Step 2: Estimate Joint Model

$$\begin{bmatrix} d_{1k} \\ d_{rJ_k} \\ d_{ck} \end{bmatrix} = \begin{bmatrix} \oplus_{j=1}^{J_k} Z_{rjk} & 0 \\ 0 & I_{J_k \times p_{r2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1J_k} \\ r_{2J_k} \end{bmatrix} + \begin{bmatrix} Z_{ck} & 0 \\ 0 & 0 \\ 0 & I_{p_{c2}} \end{bmatrix} \begin{bmatrix} c_{1k} \\ c_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_k \\ 0 \\ 0 \end{bmatrix}$$

where

$$d_{1k} = \begin{bmatrix} d_{11k} \\ \vdots \\ d_{1J_k k} \end{bmatrix}, \quad Z_{ck} = \begin{bmatrix} Z_{c1k} \\ \vdots \\ Z_{cJ_k k} \end{bmatrix}, \quad \varepsilon_k \sim N(0, I_{N_k} \otimes \pi_1),$$

$$\begin{bmatrix} r_{1J_k} \\ r_{2J_k} \end{bmatrix} \sim N\left(0, \begin{bmatrix} I_{J_k} \otimes \pi_{r11} & I_{J_k} \otimes \pi_{r12} \\ I_{J_k} \otimes \pi_{r21} & I_{J_k} \otimes \pi_{r22} \end{bmatrix} \right)$$

for $k = 1, \cdots, K.$

Estimate $\theta_c = (\beta_c, \pi_c).$

Step 2: Estimate Joint Model

- Observed value indicator matrices like C_j, R_k :

$$diag\{O_{1j}, O_{rj}, O_{cK_j}\}\begin{bmatrix} d_{1j} \\ d_{rj} \\ d_{cK_j} \end{bmatrix} = \begin{bmatrix} d_{1jobs} \\ d_{rjobs} \\ d_{cK_jobs} \end{bmatrix}, Z_{rjobs} = O_{1j} Z_{rj}, Z_{cK_jobs} = O_{1j} \oplus_{k=1}^{K_j} Z_{cjk};$$

$$diag\{O_{1k}, O_{rJ_k}, O_{ck}\}\begin{bmatrix} d_{1k} \\ d_{rJ_k} \\ d_{ck} \end{bmatrix} = \begin{bmatrix} d_{1kobs} \\ d_{rJ_kobs} \\ d_{ckobs} \end{bmatrix}, Z_{rJ_kobs} = O_{1k} \oplus_{j=1}^{J_k} Z_{rjk}, Z_{ckobs} = O_{1k} Z_{ck};$$

(Shin and Raudenbush '07).

Step 2: Estimate Joint Model

- Observed model for child j:

$$\begin{bmatrix} d_{1jobs} \\ d_{rjobs} \end{bmatrix} \mid d_{cK_j obs}, \theta \sim N\left(\begin{bmatrix} Z_{cK_j obs} \oplus_{k=1}^{K_j} c_{1k}^{*} \\ 0 \end{bmatrix}, \begin{bmatrix} V_{1jobs} & Z_{rjobs}\pi_{r12}O_{rj}^{T} \\ O_{rj}\pi_{r21}Z_{rjobs}^{T} & \pi_{r22j} \end{bmatrix} \right)$$

where $\quad c_{1k} \mid d_{ckobs} \sim N(c_{1k}^{*}, \pi_{c11k}^{*}), \quad \pi_{r22j} = O_{rj}\pi_{r22}O_{rj}^{T}$

$$V_{1jobs} = Z_{rjobs}\pi_{r11}Z_{rjobs}^{T} + Z_{cK_j obs} \oplus_{k=1}^{K_j} \pi_{c11k}^{*} Z_{cK_j obs}^{T} + \oplus_{k=1}^{K_j} \oplus_{i=1}^{n_{jk}} \pi_{1ijk}$$

for $\pi_{1ijk} = O_{1ijk}\pi_{1}O_{1ijk}^{T}$

- Observed model for school k:

$$\begin{bmatrix} d_{1kobs} \\ d_{ckobs} \end{bmatrix} \mid d_{rJ_k obs}, \theta \sim N\left(\begin{bmatrix} Z_{rJ_k obs} \oplus_{j=1}^{J_k} r_{1j}^{*} \\ 0 \end{bmatrix}, \begin{bmatrix} V_{1kobs} & Z_{ckobs}\pi_{c12}O_{ck}^{T} \\ O_{ck}\pi_{c21}Z_{ckobs}^{T} & \pi_{c22k} \end{bmatrix} \right)$$

for $\quad r_{1j} \mid d_{rjobs} \sim N(r_{1j}^{*}, \pi_{r11j}^{*}), \quad \pi_{c22k} = O_{ck}\pi_{c22}O_{ck}^{T}$

$$V_{1kobs} = Z_{rJ_k obs} \oplus_{j=1}^{J_k} \pi_{r11j}^{*} Z_{rJ_k obs}^{T} + Z_{ckobs}\pi_{c11}Z_{ckobs}^{T} + \oplus_{j=1}^{J_k} \oplus_{i=1}^{n_{jk}} \pi_{1ijk}.$$

Step 2: Estimate Joint Model

- E - step:

$$\varepsilon_j \mid d_{1jobs}, d_{rjobs}, d_{cK_jobs}, \theta \sim N(\tilde{\varepsilon}_j, V_{\varepsilon j}),$$

$$r_j | d_{1jobs}, d_{rjobs}, d_{cK_jobs}, \theta \sim N(\tilde{r}_j, V_{rj}),$$

$$c_k | d_{1kobs}, d_{rJ_kobs}, d_{ckobs}, \theta \sim N(\tilde{c}_k, V_{ck}).$$

- M - step: $L(\theta) = \prod_j f(\varepsilon_j \mid \mathbf{r}, \mathbf{c}, \theta) \prod_j f(r_j \mid \theta) \prod_k f(c_k \mid \theta)$

$$\hat{\pi}_1 = \sum_{j=1}^{J} \sum_{k=1}^{K_j} \sum_{i=1}^{n_{jk}} \varepsilon_{ijk} \varepsilon_{ijk}^T / N; \quad \hat{\pi}_r = \sum_{j=1}^{J} r_j r_j^T / J; \quad \hat{\pi}_c = \sum_{k=1}^{K} c_k c_k^T / K;$$

$$\hat{\beta}_1 = \beta_1 + \left(\sum_{j=1}^{J} \sum_{k=1}^{K_j} \sum_{i=1}^{n_{jk}} X_{1ijk}^T \pi_1^{-1} X_{1ijk} \right)^{-1} \sum_{j=1}^{J} \sum_{k=1}^{K_j} \sum_{i=1}^{n_{jk}} X_{1ijk}^T \pi_1^{-1} \varepsilon_{ijk};$$

$$\hat{\beta}_r = \beta_r + \left(\sum_{j=1}^{J} X_{rj}^T \operatorname{var}(r_{2j} \mid r_{1j})^{-1} X_{rj} \right)^{-1} \sum_{j=1}^{J} X_{rj}^T \operatorname{var}(r_{2j} \mid r_{1j})^{-1} (r_{2j} - \pi_{r21} \pi_{r11}^{-1} r_{1j});$$

$$\hat{\beta}_c = \beta_c + \left(\sum_{k=1}^{K} X_{ck}^T \operatorname{var}(c_{2k} \mid c_{1k})^{-1} X_{ck} \right)^{-1} \sum_{k=1}^{K} X_{ck}^T \operatorname{var}(c_{2k} \mid c_{1k})^{-1} (c_{2k} - \pi_{c21} \pi_{c11}^{-1} c_{1k});$$

Step 3: Proper MI Based on Joint Model

1. For diagonal π_{ii}, off-diagonal π_{ij} in (π_1, π_r, π_c) s.t.

$$\mathrm{var}[\log(\hat{\pi}_{ii})] = v_{ii},\ \mathrm{var}\left[\log\frac{1+\hat{\rho}_{ij}}{1-\hat{\rho}_{ij}}\right] = v_{ij} \text{ for } \hat{\rho}_{\mathrm{ij}} = \frac{\hat{\pi}_{ij}}{\sqrt{\hat{\pi}_{ii}\hat{\pi}_{jj}}},$$

Generate $\tilde{\theta} = (\tilde{\beta}, \tilde{\pi}_1, \tilde{\pi}_r, \tilde{\pi}_c)$ from $N[\hat{\beta}, \mathrm{var}(\hat{\beta})]$, $N(\log\pi_{ii}, v_{ii})$, $N\left(\log\frac{1+\rho_{ij}}{1-\rho_{ij}}, v_{ij}\right)$.

2. Let $Y = (Y_{obs}, Y_{mis})$.

$$f(r,\varepsilon,c \mid Y_{obs},\tilde{\theta}) = f(c \mid Y_{obs},\tilde{\theta})\, f(r,\varepsilon \mid c, Y_{obs},\tilde{\theta})$$
$$= \prod_{k=1}^{K} f(c_k \mid Y_{1kobs}, Y_{rJ_k obs}, Y_{ckobs}, \tilde{\theta}) \prod_{j=1}^{J} f(r_j, \varepsilon_j \mid c_{1K_j}, Y_{1jobs}, Y_{rjobs}, \tilde{\theta});$$

$i)$ $c_k^{mi} \sim f(c_k \mid Y_{1kobs}, Y_{rJ_k obs}, Y_{ckobs}, \tilde{\theta}) \Rightarrow c^{mi} = (c_1^{mi}, \cdots, c_K^{mi});$

$ii)$ $r_j, \varepsilon_j \sim f(r_j, \varepsilon_j \mid c^{mi}, Y_{1jobs}, Y_{rjobs}, \tilde{\theta}) \Rightarrow r^{mi} = (r_1^{mi}, \cdots, r_J^{mi}), \varepsilon^{mi} = (\varepsilon_1^{mi}, \cdots, \varepsilon_J^{mi});$

$iii)$ Construct Y_{mis} given $c^{mi}, r^{mi}, \varepsilon^{mi}, \tilde{\theta}$.

Discussion/Future Research

Discussion :

1. Efficient estimation method for HCLM with missing data;
2. Computation not burdened by cross-classification;
3. Requires knowledge of complete-data HCLM;
4. Appeals to a broad range of scientists.

Future Research :

1. random interaction effects;
2. more nesting $\Rightarrow$ additional level(s);
3. Applications : e.g./
   - ECLS data

     7 occasions, 21409 students, 3022 schools.

Step 2: Estimate Joint Model

$$\begin{bmatrix} Y_{1jk} \\ Y_{rj} \\ Y_{ck} \end{bmatrix} = \begin{bmatrix} X_{1jk} & 0 & 0 \\ 0 & X_{rj} & 0 \\ 0 & 0 & X_{ck} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_r \\ \beta_c \end{bmatrix} + \begin{bmatrix} Z_{rjk} & 0 \\ 0 & I_{p_{r2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{1j} \\ r_{2j} \end{bmatrix} + \begin{bmatrix} Z_{cjk} & 0 \\ 0 & 0 \\ 0 & I_{p_{c2}} \end{bmatrix} \begin{bmatrix} c_{1k} \\ c_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{jk} \\ 0 \\ 0 \end{bmatrix}$$

for $i = 1, \cdots, n_{jk}, \quad j = 1, \cdots, J, \quad k = 1, \cdots, K_j$ to estimate θ_r, or

$i = 1, \cdots, n_{jk}, \quad j = 1, \cdots, J_k, \quad k = 1, \cdots, K$ to estimate θ_c

where

$$Y_{1jk} = \begin{bmatrix} Y_{11jk} \\ \vdots \\ Y_{1n_{jk}jk} \end{bmatrix}, X_{1jk} = \begin{bmatrix} X_{11jk} \\ \vdots \\ X_{1n_{jk}jk} \end{bmatrix}, Z_{rjk} = \begin{bmatrix} Z_{r1jk} \\ \vdots \\ Z_{rn_{jk}jk} \end{bmatrix}, Z_{cjk} = \begin{bmatrix} Z_{c1jk} \\ \vdots \\ Z_{cn_{jk}jk} \end{bmatrix}, \varepsilon_{jk} = \begin{bmatrix} \varepsilon_{1jk} \\ \vdots \\ \varepsilon_{n_{jk}jk} \end{bmatrix}.$$

Let $d_{1jk} = Y_{1jk} - X_{1jk}\beta_1, \quad d_{rj} = Y_{rj} - X_{rj}\beta_r, \quad d_{ck} = Y_{ck} - X_{ck}\beta_c.$

Step 2: Estimate Joint Model

- Observed value indicator matrices O like C_j, R_k : (Shin and Raudenbush '07)

$$e.g./\,Y_{1ijk} = \begin{bmatrix} R_{ijk} \\ W_{ijk} \end{bmatrix} \text{s.t.}\, O_{1ijk} = \begin{cases} I_2 & \text{for both observed,} \\ [1\,0] & \text{for } R_{ijk} \text{ observed}, W_{ijk} \text{ missing,} \\ [0\,1] & \text{for } R_{ijk} \text{ missing}, W_{ijk} \text{ observed.} \end{cases}$$

$$O_{1j} = \oplus_{k=1}^{K_j} \oplus_{i=1}^{n_{jk}} O_{1ijk} \Rightarrow Y_{1jobs} = O_{1j} Y_{1j},\; X_{1jobs} = O_{1j} X_{1j},\; d_{1jobs} = O_{1j} d_{1j},$$

$$Z_{rjobs} = O_{1j} Z_{rj},\; Z_{cK_j obs} = O_{1j} \oplus_{k=1}^{K_j} Z_{cjk};$$

$$O_{1k} = \oplus_{j=1}^{J_k} \oplus_{i=1}^{n_{jk}} O_{1ijk} \Rightarrow Y_{1kobs} = O_{1k} Y_{1k}, X_{1kobs} = O_{1k} X_{1k}, d_{1kobs} = O_{1k} d_{1k},$$

$$Z_{rJ_k obs} = O_{1k} \oplus_{j=1}^{J_k} Z_{rjk}, Z_{ckobs} = O_{1k} Z_{ck},$$

Likewise,

$$d_{rjobs} = O_{rj} d_{rj}, d_{ckobs} = O_{ck} d_{ck};$$

$$O_{rJ_k} = \oplus_{j=1}^{J_k} O_{rj} \Rightarrow d_{rJ_k obs} = O_{rJ_k} d_{rJ_k},$$

$$O_{cK_j} = \oplus_{k=1}^{K_j} O_{ck} \Rightarrow d_{cK_j obs} = O_{cK_j} d_{cK_j},$$

Estimation with Missing Data: Step 2

(a)

j\k	1	2	3	4	K=5
1	$(1,1,n_{11})$		$(1,3,n_{13})$	$(1,4,n_{14})$	
2		$(2,2,n_{22})$			
3	$(3,1,n_{31})$	$(3,2,n_{32})$			
4	$(4,1,n_{41})$				$(4,5,n_{45})$
5			$(5,3,n_{53})$		
J=6				$(6,4,n_{64})$	

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{K_1 \times K}$$

$$C_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}_{K_2 \times K}$$

$$\vdots$$

$$C_6 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}_{K_6 \times K}$$

$$R_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}_{J_1 \times J}$$

$$R_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}_{J_2 \times J}$$

$$\vdots$$

$$R_5 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}_{J_6 \times J}$$

Step 2: Estimate Joint Model

- Observed log-likelihood $l(\theta)_{obs}$

$$
\begin{aligned}
2l(\theta)_{obs} \propto \sum_{j=1}^{J} \Big[& \sum_{k=1}^{K_j} \sum_{i=1}^{n_{jk}} \log|\pi_{1ijk}^{-1}| + \log|\pi_{r11j}^{*-1}| + \log|\Delta_j^{-1}| + \sum_{k=1}^{K_j} \log|\pi_{c11k}^{*-1}| \\
& + \log|\Gamma_j^{-1}| + \log|\pi_{r22j}^{-1}| - e_{1jobs}^{*T} V_{jobs}^{-11} e_{1jobs}^{*} - d_{rjobs}^{T} \pi_{r22j}^{-1} d_{rjobs} \Big] \\
& + \sum_{k=1}^{K} \Big[\log|\pi_{c22k}^{-1}| - d_{ckobs}^{T} \pi_{c22k}^{-1} d_{ckobs} \Big]
\end{aligned}
$$

for

$$\Delta_j = Z_{rjobs}^{T} \oplus_{k=1}^{K_j} \oplus_{i=1}^{n_{jk}} \pi_{1ijk}^{-1} Z_{rjobs} + \pi_{r11j}^{*-1};$$

$$\psi_j = Z_{rjobs} \pi_{r11j}^{*} Z_{rjobs}^{T} + \oplus_{k=1}^{K_j} \oplus_{i=1}^{n_{jk}} \pi_{1ijk}^{-1};$$

$$\Gamma_j = Z_{cjobs} \psi_j^{-1} Z_{cjobs}^{T} + \oplus_{k=1}^{K_j} \pi_{c11k}^{*-1};$$

$$e_{1jobs}^{*} = d_{1jobs} - Z_{rjobs} r_{1j}^{*} - Z_{cjobs} \oplus_{k=1}^{K_j} c_{1k}^{*};$$

$$V_{jobs}^{-11} = \psi_j^{-1} - \psi_j^{-1} Z_{cjobs} \Gamma_j^{-1} Z_{cjobs}^{T} \psi_j^{-1}.$$

Step 2: Estimate Joint Model

- More Notations:

$$Y_1 = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1J} \end{bmatrix}, \quad X_1 = \begin{bmatrix} X_{11} \\ \vdots \\ X_{1J} \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ Y_r \\ Y_c \end{bmatrix}, \quad X = diag\{X_1, X_r, X_c\},$$

$$O = diag\{O_1, O_r, O_c\} \text{ for } O_1 = \oplus_{j=1}^{J} O_{1j}.$$

$$\Rightarrow Y_{obs} = OY \sim N(X_{obs}\beta, V) \text{ for } X_{obs} = OX$$

- Variances: $\varphi = \text{distinct elements in } (\pi_{\text{r}}, \pi_{\text{c}}, \pi_1)$

$$\text{var}(\hat{\beta}) = \left(X_{obs}^T V^{-1} X_{obs}\right)^{-1};$$

$$\text{var}(\hat{\varphi}) = \left[\frac{1}{2}\left(\frac{dvecV}{d\varphi}\right)^T V^{-1} \otimes V^{-1} \frac{dvecV}{d\varphi}\right].$$