Lecture 9
Three levels Logistic Random Intercept Model

Did the Guatemalan immunization campaign work?

- Data are available from the National Survey of Maternal and Child Health conducted in Guatemala in 1987
- A nationally representative sample of 5160 women aged between 15 and 44 were interviewed
- The questionnaire included questions determining the immunization status of children who were born in the previous 5 years and alive at the time of the interview

Did the Guatemalan immunization campaign work?

- Beginning 1986, the Guatemalan government undertook a series of campaign to immunize the population against major childhood diseases
- An important explanatory variable is whether the child was at least 2 years old at the time of the interview, in which case the child was old enough to be immunized during the 1986 campaign
- If this variable is associated with immunization, there is some indication that the government campaign work

Data structure

- Level 1 (child)
  - Immun: dummy variable for child being immunized (y)
  - Kid2p: child at least 2 years old at the time of the interview (x2)
- Level 2 (mother)
  - Mom: identifier for the mother
  - Ethinicity (dummy variables with latino as a reference category)
    - indNospa: mother is indigenous, not Spanish speaking (x3)
    - indSpa: mother is indigenous, Spanish speaking (x4)

Data structure

- Level 2 (mother)
  - Mother’s education (dummy variables with no education as a reference category)
    - momEdPri: mother has primary education (x5)
    - momEdSec: mother has secondary education (x6)
  - Husband’s education (dummy variables with no education as a reference category)
    - husEdpri: husband has primary education (x7)
    - husSecpri: husband has secondary education (x8)
    - husEdK: husband education is not known (x9)

Data structure

- Level 3 (community)
  - Cluster: identifier for communities (k)
  - Rural: dummy variable for community being rural (x10)
  - pcInd81: percentage of population that was indigenous in 1981 (x11)

A three-level logistic random-intercept model

$$\operatorname{logit} P(y_{ijk} = 1 \mid x_{ijk}, \varsigma^{(2)}_{jk}, \varsigma^{(3)}_{k}) = \beta_1 + \varsigma^{(2)}_{jk} + \varsigma^{(3)}_{k} +$$

$$+\beta_2 x_{2ijk} + \sum_{p=3}^{9} \beta_p x_{pjk} + \beta_{10} x_{10k} + \beta_{11} x_{11k}$$

$$\varsigma^{(2)}_{jk} \sim N(0, \tau_2^2)$$

$$\varsigma^{(3)}_{k} \sim N(0, \tau_3^2)$$

Latent variable formulation

$$y^*_{ijk} = \beta_1 + \varsigma^{(2)}_{jk} + \varsigma^{(3)}_{k} +$$

$$+\beta_2 x_{2ijk} + \sum_{p=3}^{9} \beta_p x_{pjk} + \beta_{10} x_{10k} + \beta_{11} x_{11k} + \varepsilon_{ijk}$$

$$\varsigma^{(2)}_{jk} \sim N(0, \tau_2^2)$$

$$\varsigma^{(3)}_{k} \sim N(0, \tau_3^2)$$

$$y_{ijk} = 1 \Leftrightarrow y^*_{ijk} > 0$$

$$\Pr(\varepsilon_{ijk} < h) = \exp(h)/(1 + \exp(h))$$

$$E[\varepsilon_{ijk}] = 0, \operatorname{var}(\varepsilon_{ijk}) = \pi^2/3$$

Intraclass correlation for latent responses

Correlation across mothers within the same community

$$\rho(comm) = cor(y^*_{ijk}, y^*_{i'j'k} \mid x_{ijk}, x_{i'j'k}) = \frac{\tau_3^2}{\tau_2^2 + \tau_3^2 + \pi^2/3}$$

Correlation across children for the same mother and within the same community

$$\rho(mother,comm) = cor(y^*_{ijk}, y^*_{i'jk} \mid x_{ijk}, x_{i'k}) = \frac{\tau_2^2 + \tau_3^2}{\tau_2^2 + \tau_3^2 + \pi^2/3}$$

$$\rho(mother,comm) > \rho(comm)$$

Children of a given moth are more similar than children Within the same community but with different mothers

Three-stage formulation

$$level-1$$

$$\log itP(y_{ijk} = 1 \mid \eta_{1jk}, x_{2ijk}) = \eta_{1jk} + \beta_2 x_{2ijk}$$

$$level-2$$

$$\eta_{1jk} = \pi_{11k} + \pi_{12} w_{2jk} + ... + \pi_{18} w_{8jk} + \varsigma^{(2)}_{jk}$$

$$level-3$$

$$\pi_{11k} = \gamma_{111} + \gamma_{112} v_{2k} + \gamma_{113} v_{3k} + \varsigma^{(3)}_k$$

Table 7.2: Maximum likelihood estimates for three-level random-intercept logistic model

	Log odds $=\beta$		Odds ratios $=\exp(\beta)$	
	Est	(SE)	OR	(95% CI)
Fixed part				
β_1 [_cons]	−1.03	(0.41)		
β_2 [kid2p]	1.71	(0.22)	5.55	(3.64, 8.46)
β_3 [indNoSpa]	−0.30	(0.48)	0.74	(0.29, 1.89)
β_4 [indSpa]	−0.16	(0.36)	0.85	(0.42, 1.72)
β_5 [momEdPri]	0.38	(0.22)	1.47	(0.96, 2.25)
β_6 [momEdSec]	0.36	(0.47)	1.44	(0.57, 3.63)
β_7 [husEdPri]	0.50	(0.23)	1.65	(1.05, 2.57)
β_8 [husEdSec]	0.44	(0.40)	1.55	(0.70, 3.42)
β_9 [husEdDK]	−0.01	(0.35)	0.99	(0.50, 1.97)
β_{10} [rural]	−0.91	(0.30)	0.41	(0.23, 0.74)
β_{11} [pcInd81]	−1.15	(0.49)	0.32	(0.12, 0.83)
Random part				
$\psi^{(2)}$	5.19	(1.19)		
$\psi^{(3)}$	1.03	(0.32)		
Log likelihood		−1328.50		

Results

- There is evidence of an effect of the government campaign on immunization (OR=5.55)
- The correlation among children within the same community the correlation is 0.11
- The correlation among children of the same mother is 0.65
- The effects of all other covariates is not statistically significant

Introducing a random coefficient at level 3: does the effect of the campaign varies across communities?

$$\log it P(y_{ijk} = 1 \mid x_{ijk}, \zeta^{(2)}_{jk}, \zeta^{(3)}_{1k}, \zeta^{(3)}_{2k},) = \beta_1 + \beta_2 x_{2ijk} +$$

$$+\beta_{10} x_{10k} + \beta_{11} x_{11k} + \zeta^{(2)}_{jk} + \zeta^{(3)}_{1k} + \zeta^{(3)}_{2k} x_{2ijk}$$

$$\log it P(y_{ijk} = 1 \mid x_{ijk}, \zeta^{(2)}_{jk}, \zeta^{(3)}_{1k}, \zeta^{(3)}_{2k},) = (\beta_1 + \zeta^{(2)}_{jk} + \zeta^{(3)}_{1k}) +$$

$$+(\beta_2 + \zeta^{(3)}_{2k}) x_{2ijk} + \beta_{10} x_{10k} + \beta_{11} x_{11k}$$

Table 7.3: Maximum likelihood estimates for three-level random-intercept and random-coefficient logistic models

	Random Intercept		Random Coefficient	
	Est	(95% CI)	Est	(95% CI)
Fixed part: odds ratios				
$\exp(\beta_2)$ [kid2p]	5.37	(3.53, 8.17)	6.73	(3.79, 11.96)
$\exp(\beta_{10})$ [rural]	0.35	(0.20, 0.60)	0.33	(0.18, 0.59)
$\exp(\beta_{11})$ [pcInd81]	0.19	(0.09, 0.38)	0.18	(0.08, 0.37)
Random part				
$\psi^{(2)}$	5.21		5.83	
$\psi^{(3)}_{11}$	1.03		2.42	
$\psi^{(3)}_{22}$			1.80	
$\psi^{(3)}_{21}$			−1.52	
Log likelihood	−1335.04		−1330.83	

The random coefficient models fits significantly better than random intercept using a LRT at 5% level

Results

- Overall, results do not change when we introduce a level-3 random slope
- The variance of level-3 random intercept increases from 1.03 to 2.42
- The variance of level-3 slope can be interpreted as the residual variability in the effectiveness campaign across communities and is estimated as 1.80
- The estimated correlation between the random intercept and slope is equal to -0.73 which suggests that the immunization campaign was less effective in communities where the immunization rates are high for children that were too young to be immunized during the campaign (x2=0)

Prediction

- We can obtain the empirical Bayes predictions of the random effects using the stata command (gllapred)

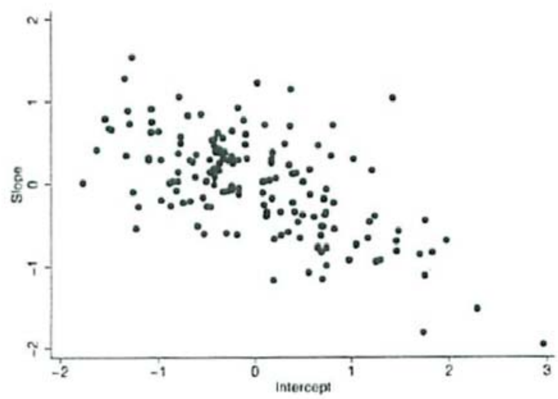

Figure 7.4: Predicted community-level random slopes versus random intercepts