Review: cross-level interactions in logistic regression

1) Model with random intercept and random slope and random effects distributions do not depend on the covariates (latent variable formulation)

$$y^*_{ik} = \eta_{1k} + \eta_{2k} x_{ik} + \varepsilon_{ik} \qquad \text{Stage 1}$$

$$\eta_{1k} = \gamma_1 + \varsigma_{1k}$$

$$\text{Stage 2}$$

$$\eta_{2k} = \gamma_2 + \varsigma_{2k}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

$$y^*_{ik} = \gamma_1 + \varsigma_{1k} + (\gamma_2 + \varsigma_{2k}) x_{ik} + \varepsilon_{ik}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

In lab today

Model 1: What is the effect of $kid2p_{ik}$ accounting for the between-community heterogeneity?

$$\log\left(\frac{p(y_{ik}=1)}{1-p(y_{ik}=1)}\right) = \eta_{ik}$$

$$\eta_{ik} = \beta_{0k} + \beta_{1k}kid2p_{ik}$$

$$\beta_{0k} = \beta_0 + U_{k0}$$

$$\beta_{1k} = \beta_1 + U_{k1}$$

β_{0k}: community-specific intercept, i.e., baseline log odds of being immunized (<2y)

β_{1k}: community-specific slope of $kid2p_{ik}$, i.e., log OR being immunized comparing >=2y versus <2y.

The equivalent 1-line writing of η_{ijk} is:

$$\eta_{ik} = \beta_0 + \beta_1 kid2p_{ik} + U_{k0} + U_{k1}kid2p_{ik}$$

β_0: overall intercept (fixed effects)

β_1: main effect of $kid2p_{ik}$ (fixed effects)

gllamm model

```
------------------------------------------------------------------------------
     immun |     exp(b)    Std. Err.       z     P>|z|    [95% Conf. Interval]
-----------+------------------------------------------------------------------
     kid2p |   3.073634    .4969815     6.94    0.000     2.238823    4.219728
------------------------------------------------------------------------------
```

Variances and covariances of random effects

```
------------------------------------------------------------------------------
***level 2 (cluster)

    var(1): 1.2882633 (.47966448)
    cov(2,1): -.65561142 (.39690843) cor(2,1): -.71194885

    var(2): .65824989 (.36732232)
------------------------------------------------------------------------------
```

2) Model with random intercept and random slope and random intercept distribution depends on a level-2 covariate

$$y^*_{ik} = \eta_{1k} + \eta_{2k} x_{ik} + \varepsilon_{ik} \qquad \text{Stage 1}$$

$$\eta_{1k} = \gamma_{11} + \gamma_{12} z_k + \varsigma_{1k}$$

$$\eta_{2k} = \gamma_2 + \varsigma_{2k} \qquad \text{Stage 2}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

$$y^*_{ik} = \gamma_{11} + \varsigma_{1k} + (\gamma_2 + \varsigma_{2k}) x_{ik} + \gamma_{12} z_k + \varepsilon_{ik}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

Model 2: Does community-level covariates explain the between-community heterogeneity in the baseline log odds of being immunized?

$$\log\left(\frac{p(y_{ik}=1)}{1-p(y_{ik}=1)}\right) = \eta_{ik}$$

$$\eta_{ik} = \beta_{0k} + \beta_{1k}kid2p_{ik}$$

$$\beta_{0k} = \beta_0 + \beta_2 rural_k + \beta_3 pcInd81_k + U_{k0}$$

$$\beta_{1k} = \beta_1 + U_{k1}$$

The equivalent 2-stage writing of η_{ijk} is:

$$\eta_{ik} = \beta_0 + (\beta_1 + U_{k1})kid2p_{ik} + \beta_2 rural_k + \beta_3 pcInd81_k + U_{k0}$$

$\beta_{0k}, \beta_{1k}, \beta_0, \beta_1$: Same as above.

β_2 : main effect of $rural_k$ (fixed effects)

β_3 : main effect of $pcInd81_k$ (fixed effects)

```
gllamm model

-----------------------------------------------------------------------------
      immun |     exp(b)    Std. Err.       z    P>|z|     [95% Conf. Interval]
------------+----------------------------------------------------------------
      kid2p |    2.984958    .4724544     6.91   0.000     2.188826    4.070662
      rural |    .5294077    .0867878    -3.88   0.000     .3839278    .7300136
    pcInd81 |    .3842638    .0782185    -4.70   0.000      .257848    .5726576
-----------------------------------------------------------------------------
Variances and covariances of random effects
-----------------------------------------------------------------------------
***level 2 (cluster)

   var(1): .85945899 (.36518027)
   cov(2,1): -.4942948 (.33061796) cor(2,1): -.68798101

   var(2): .60061203 (.34310316)
-----------------------------------------------------------------------------
```

The variance of the random intercept decrease, indicating that the community-level covariates $rural_k$ and $pcInd81_k$ explain the between-community variability in baseline log odd of being immunized. The statistical significance of the main effects of $rural_k$ and $pcInd81_k$ also suggests this conclusion.

3) Model with random intercept and random slope and random effects distributions depend on a level-2 covariate

$$y_{ik}^* = \eta_{1k} + \eta_{2k} x_{ik} + \varepsilon_{ik} \qquad \text{Stage 1}$$

$$\eta_{1k} = \gamma_{11} + \gamma_{12} z_k + \varsigma_{1k}$$

$$\eta_{2k} = \gamma_{21} + \gamma_{22} z_k + \varsigma_{2k} \qquad \text{Stage 2}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

$$y_{ik}^* = \gamma_{11} + \varsigma_{1k} + (\gamma_{21} + \varsigma_{2k}) x_{ik} + \gamma_{12} z_k + \gamma_{22} z_k x_{ik} + \varepsilon_{ik}$$

$$\varsigma_{1k} \sim N(0, \tau_1^2)$$

$$\varsigma_{2k} \sim N(0, \tau_2^2)$$

Model 3:Does community-level covariates explain the between-community heterogeneity in both the baseline log odds of being immunized and the log OR being immunized comparing >=2y versus <2y?

$$\log\left(\frac{p(y_{ik}=1)}{1-p(y_{ik}=1)}\right) = \eta_{ik}$$

$$\eta_{ik} = \beta_{0k} + \beta_{1k}kid2p_{ik}$$

$$\beta_{0k} = \beta_0 + \beta_2 rural_k + \beta_3 pcInd81_k + U_{k0}$$

$$\beta_{1k} = \beta_1 + \beta_4 rural_k + \beta_5 pcInd81_k + U_{k1}$$

The equivalent 2-stage writing of η_{ijk} is:

$$\eta_{ik} = \beta_0 + \beta_2 rural_k + \beta_3 pcInd81_k + U_{k0} + (\beta_1 + \beta_4 rural_k + \beta_5 pcInd81_k + U_{k1})kid2p_{ik}$$

$$\eta_{ik} = \beta_0 + \beta_1 kid2p_{ik} + \beta_2 rural_k + \beta_3 pcInd81_k + \beta_4 rural_k * kid2p_{ik} + \beta_5 pcInd81_k * kid2p_{ik} + U_{k0} + U_{k1} * kid2p_{ik}$$

β_{0k}, β_{1k}, β_0, β_1, β_2, β_3 : Same as above.

β_4 :cross-level interaction between $rural_k$ and $kid2p_{ik}$ (fixed effects)

β_5 :cross-level interaction between $pcInd81_k$ and $kid2p_{ik}$ (fixed effects)

```
. gllamm immun kid2p rural pcInd81 int_2p_ru int_2p_pc, family(binomial)
link(logit) i(cluster) nrf(2) eqs(inter slope) nip(4 4) adapt eform
gllamm model


------------------------------------------------------------------------
      immun |    exp(b)    Std. Err.     z      P>|z|    [95% Conf. Interval]
------------+-----------------------------------------------------------
      kid2p |   2.311586   .7539445    2.57    0.010    1.219784   4.380635
      rural |   .5115291   .1639525   -2.09    0.036    .2729278   .9587223
    pcInd81 |   .2402431   .0980197   -3.50    0.000    .1079839    .534494
  int_2p_ru |   1.045638   .3464314    0.13    0.893    .5462218   2.001676
  int_2p_pc |   1.755981    .727406    1.36    0.174    .7796755    3.95481
------------------------------------------------------------------------
Variances and covariances of random effects
------------------------------------------------------------------------
***level 2 (cluster)

   var(1): .95682725 (.39271689)
   cov(2,1): -.56582712 (.34798787) cor(2,1): -.72621719

   var(2): .63445517 (.34985769)
------------------------------------------------------------------------
```

The variance of the random slope remains approximately the same, indicating that the community-level covariates $rural_k$ and $pcInd81_k$ do not explain the between-community variability in the log OR being immunized comparing >=2y versus <2y. This can be also inferred from the non-statistically significant (cross-level) interaction between $kid2p_{ik}$ and the community-level variables $rural_k$ and $pcInd81_k$.