

# Multilevel logistic mixed models

Alexander McLain

## 1 Guatemalan Immunization Study

The objective of this study is to identify important family- and community-level factors that affect whether Guatemalan children are immunized. A nationally representative sample of 5160 mothers, between 15 and 44 years old were interviewed.

### *The Variables*

Level 1 (children):

- immun: dummy variable for child being immunized, the response variable.
- kid2p: child at least 2 years old at the time of the interview.

Level 2 (mothers)

- mom: identifier for mother
- Ethnicity
  - indNoSpa: mother is indigenous, not Spanish speaking
  - indSpa: mother is indigenous, Spanish speaking
- Mother's education (dummy variables with 'no education' as reference category)
  - monEdPri: mother has primary education
  - monEdSec: mother has secondary education
- Husband's education (dummy variables with 'no education' as reference category)
  - husEdPri: husband has primary education
  - husEdSec: husband has secondary education
  - husEdDK: husband's education is not known

Level 3 (communities)

- cluster: identifier for communities
- rural: dummy variable for community being rural
- pcInd81: percentage of population that was indigenous in 1981

```
library(tidyverse)
library(haven)
library(lme4)
guatemala <- read_dta("guatemala.dta")
str(guatemala)
```

```
## tibble [2,159 x 19] (S3: tbl_df/tbl/data.frame)
##  $ kid      : num [1:2159] 2 269 272 273 274 275 276 277 278 280 ...
##  ..- attr(*, "format.stata")= chr "%9.0g"
##  $ mom      : num [1:2159] 2 185 186 187 188 188 189 190 190 191 ...
##  ..- attr(*, "format.stata")= chr "%9.0g"
##  $ cluster  : num [1:2159] 1 36 36 36 36 36 36 36 36 36 ...
##  ..- attr(*, "format.stata")= chr "%9.0g"
##  $ immun    : num [1:2159] 1 0 0 0 0 1 1 1 1 1 ...
##  ..- attr(*, "format.stata")= chr "%9.0g"
```

```
## $ kid2p : num [1:2159] 1 1 1 1 1 1 1 0 1 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ mom25p : num [1:2159] 0 0 0 0 1 0 0 1 1 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ order23 : num [1:2159] 0 1 0 1 0 1 1 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ order46 : num [1:2159] 0 0 0 0 1 0 0 0 1 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ order7p : num [1:2159] 0 0 0 0 0 0 0 1 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ indNoSpa: num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ indSpa : num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ momEdPri: num [1:2159] 0 1 1 1 1 1 0 1 1 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ momEdSec: num [1:2159] 1 0 0 0 0 0 1 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ husEdPri: num [1:2159] 0 1 0 1 0 0 1 1 1 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ husEdSec: num [1:2159] 1 0 1 0 0 0 0 0 0 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ husEdDK : num [1:2159] 0 0 0 0 1 1 0 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ momWork : num [1:2159] 0 1 1 1 1 1 1 1 1 1 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ rural : num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ pcInd81 : num [1:2159] 0.1075 0.0437 0.0437 0.0437 0.0437 ...
## ..- attr(*, "format.stata")= chr "%9.0g"
```

```
guatemala <- zap_formats(guatemala)
str(guatemala)
```

```
## tibble [2,159 x 19] (S3: tbl_df/tbl/data.frame)
## $ kid : num [1:2159] 2 269 272 273 274 275 276 277 278 280 ...
## $ mom : num [1:2159] 2 185 186 187 188 188 189 190 190 191 ...
## $ cluster : num [1:2159] 1 36 36 36 36 36 36 36 36 36 ...
## $ immun : num [1:2159] 1 0 0 0 0 1 1 1 1 1 ...
## $ kid2p : num [1:2159] 1 1 1 1 1 1 1 0 1 1 ...
## $ mom25p : num [1:2159] 0 0 0 0 1 0 0 1 1 1 ...
## $ order23 : num [1:2159] 0 1 0 1 0 1 1 0 0 0 ...
## $ order46 : num [1:2159] 0 0 0 0 1 0 0 0 1 1 ...
## $ order7p : num [1:2159] 0 0 0 0 0 0 0 1 0 0 ...
## $ indNoSpa: num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## $ indSpa : num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## $ momEdPri: num [1:2159] 0 1 1 1 1 1 0 1 1 1 ...
## $ momEdSec: num [1:2159] 1 0 0 0 0 0 1 0 0 0 ...
## $ husEdPri: num [1:2159] 0 1 0 1 0 0 1 1 1 0 ...
## $ husEdSec: num [1:2159] 1 0 1 0 0 0 0 0 0 1 ...
## $ husEdDK : num [1:2159] 0 0 0 0 1 1 0 0 0 0 ...
## $ momWork : num [1:2159] 0 1 1 1 1 1 1 1 1 1 ...
## $ rural : num [1:2159] 0 0 0 0 0 0 0 0 0 0 ...
## $ pcInd81 : num [1:2159] 0.1075 0.0437 0.0437 0.0437 0.0437 ...
```

```
head(guatemala)
```

kid	mom	cluster	imm	kid2	mom2	order23	order46	order7p	NoSpa	Spa	mom	End	Bri	Ed	Pu	Ed	DK	Wom	pcInd81
2	2	1	1	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0.1075042
269	185	36	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0.0437295
272	186	36	0	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0.0437295
273	187	36	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0.0437295
274	188	36	0	1	1	0	1	0	0	0	1	0	0	0	1	1	0	0	0.0437295
275	188	36	1	1	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0.0437295

```
# Frequency of dchild-level observations
table( table( guatemala$kid))
```

```
table( table( guatemala$kid))
```

$$\frac{1}{2159}$$

```
# Frequency of mother-level observations
table( table( guatemala$mom))
```

```
table( table( guatemala$mom))
```

1	2	3
1063	500	32

```
# Frequency of cluster-level observations
data.frame( table( table( guatemala$cluster)) )
```

```
data.frame( table( table( guatemala$cluster)) )
```

Var1	Freq
1	2
2	7
3	7
4	6
5	6
6	7
7	6
8	8
9	11
10	13
11	13
12	5
13	2
14	5
15	6
16	7
17	6
18	4
19	2
20	3
21	6
22	3

Var1	Freq
23	5
24	5
25	5
26	2
27	2
29	1
30	2
32	1
34	1
50	1
55	1

## 1.1 Unconditional two- and three-level Model

```
form <- immun ~ (1|mom)
(mod1 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | mom)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2925.317 2936.672 -1460.659 2921.317    2157
## Random effects:
## Groups Name      Std.Dev.
## mom      (Intercept) 1.114
## Number of obs: 2159, groups: mom, 1595
## Fixed Effects:
## (Intercept)
##      -0.2778
```

```
form <- immun ~ (1|cluster)
(mod2 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | cluster)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2880.986 2892.341 -1438.493 2876.986    2157
## Random effects:
## Groups Name      Std.Dev.
## cluster (Intercept) 0.7573
## Number of obs: 2159, groups: cluster, 161
## Fixed Effects:
## (Intercept)
##      -0.2339
```

```
form <- immun ~ (1|mom) + (1|cluster)
(mod3 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | mom) + (1 | cluster)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2860.950 2877.982 -1427.475 2854.950    2156
## Random effects:
## Groups Name          Std.Dev.
## mom      (Intercept) 0.9081
## cluster (Intercept) 0.8339
## Number of obs: 2159, groups: mom, 1595; cluster, 161
## Fixed Effects:
## (Intercept)
##      -0.2754
anova(mod1, mod2, mod3)
```

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mod1	2	2925.317	2936.672	-1460.659	2921.317	NA	NA	NA
mod2	2	2880.987	2892.341	-1438.493	2876.987	44.33067	0	NA
mod3	3	2860.950	2877.982	-1427.475	2854.950	22.03636	1	2.7e-06

## 1.2 Unconditional three-level Model with nested RE's

```
form <- immun ~ (1|cluster/mom)
(mod4 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | cluster/mom)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2860.950 2877.982 -1427.475 2854.950    2156
## Random effects:
## Groups Name          Std.Dev.
## mom:cluster (Intercept) 0.9081
## cluster      (Intercept) 0.8339
## Number of obs: 2159, groups: mom:cluster, 1595; cluster, 161
## Fixed Effects:
## (Intercept)
##      -0.2754
```

## 1.3 Three-level random intercept model

```
form <- immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 + (1|mom)
(mod5 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
```

```
##      (1 | mom)
##      Data: guatemala
##           AIC          BIC      logLik  deviance  df.resid
##  2770.280  2815.699 -1377.140  2754.280      2151
## Random effects:
## Groups Name          Std.Dev.
## mom      (Intercept) 1.286
## Number of obs: 2159, groups: mom, 1595
## Fixed Effects:
## (Intercept)      kid2p      momEdPri      husEdPri      momWork      rural
##      -0.7544      1.2758      0.2845      0.2776      0.3272      -0.7623
##      pcInd81
##      -0.8296
```

```
mod1
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | mom)
##      Data: guatemala
##           AIC          BIC      logLik  deviance  df.resid
##  2925.317  2936.672 -1460.659  2921.317      2157
## Random effects:
## Groups Name          Std.Dev.
## mom      (Intercept) 1.114
## Number of obs: 2159, groups: mom, 1595
## Fixed Effects:
## (Intercept)
##      -0.2778
```

In a linear model having the random effect variance increase in this scenario is not possible.

```
form <- immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 + (1|cluster)
(mod6 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
##      (1 | cluster)
##      Data: guatemala
##           AIC          BIC      logLik  deviance  df.resid
##  2768.005  2813.425 -1376.003  2752.005      2151
## Random effects:
## Groups Name          Std.Dev.
## cluster (Intercept) 0.6256
## Number of obs: 2159, groups: cluster, 161
## Fixed Effects:
## (Intercept)      kid2p      momEdPri      husEdPri      momWork      rural
##      -0.5105      1.0027      0.1720      0.2196      0.2041      -0.6074
##      pcInd81
##      -0.8043
```

```
mod2
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
```

```

## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ (1 | cluster)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2880.986 2892.341 -1438.493 2876.986    2157
## Random effects:
## Groups Name          Std.Dev.
## cluster (Intercept) 0.7573
## Number of obs: 2159, groups: cluster, 161
## Fixed Effects:
## (Intercept)
##      -0.2339

form <- immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
  (1|mom) + (1|cluster)
mod6 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0026005 (tol = 0.002, component 1)

summary(mod6)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
##      (1 | mom) + (1 | cluster)
## Data: guatemala
##
##      AIC      BIC    logLik deviance df.resid
## 2737.7    2788.8  -1359.9   2719.7    2150
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.6628 -0.6424 -0.3527  0.6927  2.5342
##
## Random effects:
## Groups Name          Variance Std.Dev.
## mom      (Intercept) 1.2401    1.1136
## cluster (Intercept) 0.4908    0.7006
## Number of obs: 2159, groups: mom, 1595; cluster, 161
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.6776     0.2605  -2.601 0.009299 **
## kid2p         1.2662     0.1577   8.030 9.72e-16 ***
## momEdPri      0.2450     0.1378   1.778 0.075472 .
## husEdPri      0.2822     0.1325   2.130 0.033205 *
## momWork       0.2789     0.1357   2.055 0.039832 *
## rural        -0.7573     0.1952  -3.879 0.000105 ***
## pcInd81      -1.0001     0.2461  -4.064 4.82e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

```

```

## Correlation of Fixed Effects:
##      (Intr) kid2p  mmEdPr hsEdPr momWrk rural
## kid2p      -0.455
## momEdPri -0.313  0.089
## husEdPri -0.230  0.082 -0.151
## momWork  -0.377  0.036  0.030  0.026
## rural    -0.471 -0.101 -0.042 -0.064  0.092
## pcInd81  -0.433 -0.121  0.174 -0.014  0.107  0.020
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.0026005 (tol = 0.002, component 1)

form <- immun ~ kid2p + momEdPri + husEdPri + kid2p*momWork +
  (1|mom) + (1|cluster)
mod6b <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial)
summary(mod6b)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + momEdPri + husEdPri + kid2p * momWork + (1 |
##      mom) + (1 | cluster)
## Data: guatemala
##
##      AIC      BIC   logLik deviance df.resid
## 2765.5    2810.9  -1374.7   2749.5     2151
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.7861 -0.6388 -0.3697  0.6910  2.1843
##
## Random effects:
## Groups Name Variance Std.Dev.
## mom (Intercept) 1.1839  1.0881
## cluster (Intercept) 0.6557  0.8097
## Number of obs: 2159, groups: mom, 1595; cluster, 161
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.57651    0.22213  -7.097 1.27e-12 ***
## kid2p         1.01289    0.19516   5.190 2.10e-07 ***
## momEdPri      0.35417    0.13658   2.593 0.00951 **
## husEdPri      0.26821    0.13209   2.031 0.04230 *
## momWork      -0.02474    0.26256  -0.094 0.92492
## kid2p:momWork 0.54661    0.28624   1.910 0.05619 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) kid2p  mmEdPr hsEdPr momWrk
## kid2p      -0.756
## momEdPri   -0.295  0.052
## husEdPri   -0.315  0.058 -0.146
## momWork    -0.509  0.536 -0.032  0.014
## kid2p:mmWrk 0.384 -0.602  0.058  0.006 -0.855

```



## 1.4 Two-level random intercept slope model

```
form <- immun ~ kid2p + (1 + kid2p|cluster)
(mod7 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + (1 + kid2p | cluster)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2807.202 2835.589 -1398.601  2797.202     2154
## Random effects:
## Groups Name          Std.Dev. Corr
## cluster (Intercept) 1.0629
##      kid2p          0.7258  -0.68
## Number of obs: 2159, groups: cluster, 161
## Fixed Effects:
## (Intercept)      kid2p
##      -1.131      1.110
```

```
anova(mod2, mod7)
```

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mod2	2	2880.987	2892.341	-1438.493	2876.987	NA	NA	NA
mod7	5	2807.202	2835.589	-1398.601	2797.202	79.78422	3	0

```
form <- immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
(1 + kid2p|cluster)
(mod8 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.416888 (tol = 0.002, component 1)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
## (1 + kid2p | cluster)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2766.472 2823.246 -1373.236  2746.472     2149
## Random effects:
## Groups Name          Std.Dev. Corr
## cluster (Intercept) 0.8640
##      kid2p          0.7291  -0.69
## Number of obs: 2159, groups: cluster, 161
## Fixed Effects:
## (Intercept)      kid2p      momEdPri      husEdPri      momWork      rural
##      -0.5822      1.0855      0.1715      0.2210      0.2081     -0.6199
##      pcInd81
##      -0.8189
## optimizer (Nelder_Mead) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

## 1.5 Three-level random intercept slope model

```
form <- immun ~ kid2p + (1 + kid2p|cluster) + (1|mom)
(mod10 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial))
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: immun ~ kid2p + (1 + kid2p | cluster) + (1 | mom)
## Data: guatemala
##      AIC      BIC    logLik deviance df.resid
## 2783.255 2817.320 -1385.628  2771.255     2153
## Random effects:
## Groups Name          Std.Dev. Corr
## mom      (Intercept) 1.0287
## cluster (Intercept) 1.1646
##      kid2p          0.7329   -0.66
## Number of obs: 2159, groups: mom, 1595; cluster, 161
## Fixed Effects:
## (Intercept)      kid2p
##      -1.359      1.333
```

```
anova(mod7,mod10)
```

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mod7	5	2807.202	2835.589	-1398.601	2797.202	NA	NA	NA
mod10	6	2783.255	2817.320	-1385.628	2771.255	25.94698	1	4e-07

```
form <- immun ~ kid2p + momEdPri + husEdPri + momWork + rural + pcInd81 +
(1 + kid2p|cluster) + (1|mom)
mod11 <- glmer(formula = form, data = guatemala, nAGQ = 1, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.481706 (tol = 0.002, component 1)
```

```
anova(mod6,mod11)
```

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mod6	9	2737.709	2788.806	-1359.855	2719.709	NA	NA	NA
mod11	11	2738.381	2800.832	-1358.190	2716.381	3.328456	2	0.1893367