2042 Multi-Level Modeling: Nested (Spring 2020)

Group Project Part 1

Group 1

April 26, 2020

Instructions

- We will use the `classroom.csv` data for this project.
  a. `math1st` will be the outcome of interest for this first part.
     – Recall that `math1st` = `mathkind` + `mathgain`
  b. Read in the data (R: store as `dat`)
  c. Fit all models using **REML** (not the default in R)
  d. It's best if you use `lmerTest::lmer` rather than `lme4::lmer` to call the MLM function. The former provides p-values for fixed effects in the summary.
  e. There are 2 common error messages one can get from lmer calls: failed to converge (problem with hessian: negative eigenvalue; max|grad| = ...); and singularity. They may both be problematic in a real problem, but the latter suggests that a variance component is on the boundary of the parameter space.
     – In your discussion/writeup, consider the latter to be a "convergence problem" and ignore the former.

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##     lmer

## The following object is masked from 'package:stats':
##
##     step
```

```r
# Load ------------------------------------------------------------------------
# setwd()
# getwd()

dat <- read.csv(
  "data/classroom.csv",
  header = TRUE
)

# Create a variable and named as math1st
dat$math1st <- dat$mathkind + dat$mathgain
```

Question 1 of 12

Estimate an Unconditional Means Model (UMM) with random intercepts for **both** schools and classrooms (nested in schools).

a. Report the ICC for schools and the ICC for classrooms.

b. **Write out this model** using your preferred notation, but use the same choice of notation for the remainder of your project.

- Be mindful and explicit about any assumptions made.

Solution

```r
unconditional_model<-lmer(math1st~(1|schoolid/classid),data=dat)
print(summary(unconditional_model))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ (1 | schoolid/classid)
##    Data: dat
##
## REML criterion at convergence: 11944.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1872 -0.6174 -0.0204  0.5821  3.8339
##
## Random effects:
##  Groups           Name        Variance Std.Dev.
##  classid:schoolid (Intercept)   85.47   9.245
##  schoolid         (Intercept)  280.69  16.754
##  Residual                      1146.79  33.864
## Number of obs: 1190, groups:  classid:schoolid, 312; schoolid, 107
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  522.540      2.037 104.403   256.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ICC_class<-280.69/(280.69+1146.79)
ICC_school<- 85.47/(85.47+1146.79)
```

a) $ICC_{school} = \frac{\sigma^2_\zeta}{\sigma^2_\zeta+\sigma^2_\epsilon} = \frac{85.47}{85.47+1146.79} = 0.069$, $ICC_{class} = \frac{\sigma^2_\zeta}{\sigma^2_\zeta+\sigma^2_\epsilon} = \frac{280.69}{280.69+1146.79} = 0.197$

b) The unconditional model fitting on math1st with random intercepts for schoolid and classid is

$$MATH1ST_{ijk} = b_0 + \eta_{jk} + \zeta_k + \epsilon_{ijk}$$

, where

$$\eta_{jk} \sim N(0, \sigma^2_\eta)$$

,

$$\zeta_k \sim N(0, \zeta^2_\zeta)$$

, independently of each other and

$$\epsilon_{ijk} \sim N(0, \sigma^2_\epsilon)$$

, j represents classrooms and k represents schools.

Question 2 of 12

Add **all** school level predictors.

a. Report if adding the predictors **as a block** is justified.

b. Report change in σ^2_ζ.

Solution

```
model_all_school<-lmer(math1st~housepov+(1|schoolid)+(1|classid),data=dat)
print(summary(model_all_school))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ housepov + (1 | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 11927.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1142 -0.6011 -0.0350  0.5600  3.8154
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  classid  (Intercept)   82.36   9.075
##  schoolid (Intercept)  250.93  15.841
##  Residual             1146.96  33.867
## Number of obs: 1190, groups:  classid, 312; schoolid, 107
##
## Fixed effects:
```

```
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  531.294      3.341 102.807 159.023   <2e-16 ***
## housepov     -45.783     14.236 111.060  -3.216   0.0017 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr)
## housepov -0.810
```

```
anova(model_all_school,unconditional_model,refit=F)
```

```
## Data: dat
## Models:
## unconditional_model: math1st ~ (1 | schoolid/classid)
## model_all_school: math1st ~ housepov + (1 | schoolid) + (1 | classid)
##                     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## unconditional_model  4 11953 11973 -5972.3    11945
## model_all_school     5 11937 11963 -5963.7    11927 17.186      1   3.39e-05
##
## unconditional_model
## model_all_school    ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a) The p-value of adding housepov variable to the model is $3.39 \times e^{-5}$, which is less than 0.05, therefore, adding the housepov is significant to the model. Also, the p-value for the coeffcient on the housepov variable is 0.0017, which also justified the significance of adding this variable.

b) After adding the all school-level variable(housepov), the σ^2_ζ dropped from 280.69 to 250.93.

Question 3 of 12

Add **all** classroom level predictors.

a. Report if adding the predictors **as a block** is justified.

b. Report change in σ^2_η and change in σ^2_ϵ.

c. Give a potential reason as to why σ^2_ϵ is reduced, but not σ^2_η?

Solution

```
library(car)
```

```
## Loading required package: carData
```

```
## Registered S3 methods overwritten by 'car':
##   method                          from
##   influence.merMod                lme4
##   cooks.distance.influence.merMod lme4
##   dfbeta.influence.merMod         lme4
##   dfbetas.influence.merMod        lme4
```

```r
model_all_class<-lmer(math1st~yearstea+mathknow+mathprep+housepov+(1|schoolid)+(1|classid),data=dat)
print(summary(model_all_class))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## math1st ~ yearstea + mathknow + mathprep + housepov + (1 | schoolid) +
##     (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10821
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5552 -0.6118 -0.0311  0.5863  3.8315
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  classid  (Intercept)   94.36   9.714
##  schoolid (Intercept)  223.31  14.943
##  Residual             1136.43  33.711
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) 532.29852    5.20495 228.85767 102.268  < 2e-16 ***
## yearstea      0.06193    0.14717 223.76570   0.421  0.67432
## mathknow      2.55143    1.44530 231.06560   1.765  0.07883 .
## mathprep     -0.75440    1.42809 203.20755  -0.528  0.59790
## housepov    -41.62117   14.08834 109.83230  -2.954  0.00383 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) yearst mthknw mthprp
## yearstea -0.264
## mathknow -0.052  0.030
## mathprep -0.666 -0.175  0.004
## housepov -0.568  0.077  0.082  0.032
```

```r
linearHypothesis(model_all_class,c('yearstea','mathknow','mathprep'))
```

```
## Linear hypothesis test
##
## Hypothesis:
## yearstea = 0
## mathknow = 0
## mathprep = 0
##
## Model 1: restricted model
## Model 2: math1st ~ yearstea + mathknow + mathprep + housepov + (1 | schoolid) +
##     (1 | classid)
```

```
##
##   Df  Chisq Pr(>Chisq)
## 1
## 2  3 3.4804     0.3233
```

a) Based the wald test, we can conclude that the p-value for adding classroom-level variable is 0.3233, which is not significant at the level of 0.05.

b) The σ_η^2 increased from 82.36 to 94.36, while the σ_ϵ^2 dropped from 1146.96 to 1136.43.

c) The residuals of the model are decreased because some of the variance are explained by adding classroom-level variable.However, to the classroom-level effect, some added variable might not be significant or there might exists some correlation between the added classroom-level variables.

Question 4 of 12

Add (nearly) **all** student level predictors (but not `mathgain` or `mathkind`, as these are outomes in this context).

a. Report if justified statistically **as a block** of predictors.

b. Report change in variance components for all levels.

c. Give a potential reason as to why the school level variance component drops from prior model.

d. **Write out this model** using your chosen notation.

Soltuion

```
model_all_student<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(1|cla
print(summary(model_all_student))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  classid  (Intercept)   93.89   9.689
##  schoolid (Intercept)  169.45  13.017
##  Residual             1064.96  32.634
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
```

```
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63041    5.31209  275.39010 101.585  < 2e-16 ***
## sex           -1.21419    2.09483 1022.42110  -0.580    0.562
## minority     -16.18676    3.02605  704.47787  -5.349 1.20e-07 ***
## ses           10.05076    1.54485 1066.56211   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80861   0.080    0.936
## mathknow       1.35004    1.39168  234.49768   0.970    0.333
## mathprep      -0.27705    1.37583  205.27111  -0.201    0.841
## housepov     -17.64850   13.21755  113.87814  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
```

```
anova(model_all_class,model_all_student, refit=F)
```

```
## Data: dat
## Models:
## model_all_class: math1st ~ yearstea + mathknow + mathprep + housepov + (1 | schoolid) +
## model_all_class:     (1 | classid)
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_class    8 10837 10877 -5410.5    10821
## model_all_student 11 10752 10806 -5364.8    10730 91.446      3  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
linearHypothesis(model_all_student,c('sex','minority','ses'))
```

```
## Linear hypothesis test
##
## Hypothesis:
## sex = 0
## minority = 0
## ses = 0
##
## Model 1: restricted model
## Model 2: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (1 | classid)
##
##   Df  Chisq Pr(>Chisq)
## 1
## 2  3 85.055  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a) Based on the wald test, we can conclude that the p-value for adding the student-level predictors(sex,minority,ses) is $2.2 \times e^{-16}$. This implies that adding the student-level predictors to the model is significant at the 0.05 level.

b) Comparing to the previous model before adding the student-level predictors, the σ^2_η dropped from 94.36 to 93.89, the σ^2_ϵ dropped from 1136.43 to 1064.96 and the σ^2_ζ dropped from 223.31 to 169.45.

c) The school level variance dropped might due to adding the individual-level predictors that may be associated with group (school) effects in the aggregate.Thus, adding the individual-level predictors caused the decrease in the school level variance.

d) The model after adding the student-level predcitors is

$$MATH1ST_{ijk} = b_0 + b_1 SES_{ijk} + b_2 SEX_{ijk} + b_3 MINORITY_{ijk} + b_4 YEARSTEA_{jk} + b_5 MATHKNOW_{jk} + b_6 MATHPI$$

, where

$$\eta_{jk} \sim N(0, \sigma^2_\eta)$$

,

$$\zeta_k \sim N(0, \zeta^2_\zeta)$$

, independently of each other and

$$\epsilon_{ijk} \sim N(0, \sigma^2_\epsilon)$$

, j represents classrooms and k represents schools.

Question 5 of 12

a. Try to add a random slope for each **teacher level** predictor (varying at the **school level**; one by one separately - not all together).

b. Report the model fit or lack of fit.

c. Why is it a bad idea to include a random slope on the `housepov` effect?

d. Retry the above, allowing the slopes to be correlated with the random intercepts (still one by one).

e. Report anything unusual about the variance components (changes that are in a direction you didn't expect) and any potential explanation for why those changes occurred (*hint: what did you add to the model?*).

Solution

a)

```
#mathknow
model_mathknow<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(0+mathknow|schoolid
```

```
## boundary (singular) fit: see ?isSingular
```

```
print(summary(model_mathknow))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (0 + mathknow | schoolid) + (1 | schoolid) + (1 |      classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8580 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  classid    (Intercept) 9.389e+01  9.689914
##  schoolid   (Intercept) 1.694e+02 13.016328
##  schoolid.1 mathknow    1.700e-06  0.001304
##  Residual               1.065e+03 32.633705
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63047    5.31204  275.40357 101.586  < 2e-16 ***
## sex           -1.21417    2.09483 1022.42010  -0.580    0.562
## minority     -16.18681    3.02603  704.47306  -5.349 1.20e-07 ***
## ses           10.05075    1.54485 1066.56262   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.81110   0.080    0.936
## mathknow       1.34993    1.39168  234.50060   0.970    0.333
## mathprep      -0.27708    1.37583  205.27196  -0.201    0.841
## housepov     -17.64821   13.21718  113.88792  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(model_mathknow,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_mathknow: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_mathknow:     housepov + (0 + mathknow | schoolid) + (1 | schoolid) + (1 |
## model_mathknow:     classid)
```

```
##                   Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_mathknow    12 10754 10813 -5364.8    10730     0      1          1
```

```
#mathprep
model_mathprep<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(0+mathprep|schoolid
```

```
## boundary (singular) fit: see ?isSingular
```

```
print(summary(model_mathprep))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (0 + mathprep | schoolid) + (1 | schoolid) + (1 |      classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  classid    (Intercept) 9.388e+01 9.689e+00
##  schoolid   (Intercept) 1.694e+02 1.302e+01
##  schoolid.1 mathprep    2.171e-07 4.659e-04
##  Residual               1.065e+03 3.263e+01
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63039    5.31207  275.39223 101.586  < 2e-16 ***
## sex           -1.21419    2.09483 1022.42070  -0.580    0.562
## minority     -16.18676    3.02605  704.47629  -5.349 1.20e-07 ***
## ses           10.05076    1.54485 1066.56201   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80838   0.080    0.936
## mathknow       1.35003    1.39167  234.49786   0.970    0.333
## mathprep      -0.27705    1.37582  205.27063  -0.201    0.841
## housepov     -17.64851   13.21749  113.87941  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
```

```
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(model_mathprep,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_mathprep: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_mathprep:     housepov + (0 + mathprep | schoolid) + (1 | schoolid) + (1 |
## model_mathprep:     classid)
##                    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_mathprep    12 10754 10813 -5364.8    10730     0      1          1
```

```
#yearstea
model_yearstea<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(0+yearstea|schoolid
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00805454 (tol = 0.002, component 1)
```

```
print(summary(model_yearstea))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (0 + yearstea | schoolid) + (1 | schoolid) + (1 |      classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8482 -0.6147 -0.0322  0.5979  3.6603
##
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  classid    (Intercept) 9.247e+01  9.6159
##  schoolid   (Intercept) 1.684e+02 12.9758
##  schoolid.1 yearstea    1.008e-02  0.1004
##  Residual               1.065e+03 32.6361
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.59885    5.30780  266.47953 101.662  < 2e-16 ***
## sex           -1.21060    2.09480 1022.21558  -0.578    0.563
## minority     -16.16715    3.02635  702.61831  -5.342 1.24e-07 ***
## ses           10.04528    1.54492 1066.09816   6.502 1.21e-10 ***
## yearstea       0.01128    0.14192  122.87740   0.079    0.937
## mathknow       1.33106    1.39155  234.33195   0.957    0.340
```

```
## mathprep      -0.26584    1.37588  204.90504  -0.193    0.847
## housepov     -17.72082   13.21686  113.58577  -1.341    0.183
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.191
## minority -0.320 -0.010
## ses      -0.121  0.020  0.162
## yearstea -0.258  0.015  0.023 -0.027
## mathknow -0.082  0.006  0.115 -0.007  0.028
## mathprep -0.632 -0.006  0.001  0.053 -0.172  0.003
## housepov -0.450 -0.007 -0.179  0.082  0.070  0.057  0.037
## convergence code: 0
## Model failed to converge with max|grad| = 0.00805454 (tol = 0.002, component 1)
```

```
anova(model_yearstea,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_yearstea: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_yearstea:     housepov + (0 + yearstea | schoolid) + (1 | schoolid) + (1 |
## model_yearstea:     classid)
##                    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_yearstea    12 10754 10813 -5364.8    10730 0.007      1     0.9336
```

b) According to the LRT test conducted above, we can conclude that the p-value for adding random slope for teacher level variable mathknow, mathprep and yearstea are 1, 1, and 0.934. This implies that there is no siginicance variation for adding those random slopes. Thus, the model for adding all three variables does not fit.

c) Because housepov are school-level variable, thus it only varies at the school level. Including housepov as a random slope on school level only create an redundant school-level random effects.

d)

```
#mathknow
model_mathknow_cor<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(mathknow|schooli
print(summary(model_mathknow_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (mathknow | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6131 -0.0324  0.5969  3.6603
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  classid  (Intercept) 9.393e+01  9.6915
##  schoolid (Intercept) 1.693e+02 13.0118
##           mathknow    9.182e-04  0.0303  0.97
##  Residual             1.065e+03 32.6341
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.64041    5.31203  275.38950 101.588  < 2e-16 ***
## sex           -1.21328    2.09485 1021.79809  -0.579    0.563
## minority     -16.19378    3.02608  703.80362  -5.351 1.18e-07 ***
## ses           10.04788    1.54488 1062.12259   6.504 1.20e-10 ***
## yearstea       0.01114    0.14141  226.85275   0.079    0.937
## mathknow       1.35458    1.39201  214.62535   0.973    0.332
## mathprep      -0.27754    1.37599  201.27744  -0.202    0.840
## housepov     -17.64141   13.21242  103.98189  -1.335    0.185
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.082  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.173  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.057  0.038
```

```
anova(model_mathknow_cor,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_mathknow_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_mathknow_cor:     housepov + (mathknow | schoolid) + (1 | classid)
##                    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student  11 10752 10806 -5364.8    10730
## model_mathknow_cor 13 10756 10820 -5364.8    10730 3e-04      2     0.9998
```

```
#mathprep
model_mathprep_cor<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(mathprep|school
```

```
## boundary (singular) fit: see ?isSingular
```

```r
print(summary(model_mathprep_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (mathprep | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8542 -0.6034 -0.0221  0.5915  3.6475
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   78.46   8.858
##  schoolid (Intercept)  552.68  23.509
##           mathprep      15.88   3.985   -1.00
##  Residual             1064.27  32.623
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  538.60872    5.60800  159.90159  96.043  < 2e-16 ***
## sex           -1.16756    2.08698 1023.14887  -0.559    0.576
## minority     -16.46421    2.99522  663.67376  -5.497 5.52e-08 ***
## ses           10.14167    1.53961 1060.93433   6.587 7.04e-11 ***
## yearstea      -0.02586    0.13948  223.50197  -0.185    0.853
## mathknow       1.29874    1.37192  229.68410   0.947    0.345
## mathprep       0.04067    1.34844  139.04795   0.030    0.976
## housepov     -14.01327   12.88649  116.07068  -1.087    0.279
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.183
## minority -0.275 -0.013
## ses      -0.121  0.024  0.161
## yearstea -0.260  0.023  0.025 -0.033
## mathknow -0.071  0.002  0.107 -0.001  0.048
## mathprep -0.692 -0.008 -0.035  0.061 -0.155  0.012
## housepov -0.461  0.003 -0.187  0.095  0.089  0.027  0.107
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```r
anova(model_mathprep_cor,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
```

```
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_mathprep_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_mathprep_cor:     housepov + (mathprep | schoolid) + (1 | classid)
##                    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student  11 10752 10806 -5364.8    10730
## model_mathprep_cor 13 10751 10816 -5362.3    10725 4.8144      2    0.09007 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#yearstea
model_yearstea_cor<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(yearstea|schooli
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0406922 (tol = 0.002, component 1)
```

```
print(summary(model_yearstea_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (yearstea | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10723.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7469 -0.6028 -0.0286  0.6038  3.8423
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  classid  (Intercept)   38.8149  6.2302
##  schoolid (Intercept)  363.7664 19.0727
##           yearstea       0.5471  0.7396  -0.78
##  Residual             1066.2528 32.6535
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  538.96287    5.48722  223.18317  98.221  < 2e-16 ***
## sex           -1.33373    2.08780 1024.43110  -0.639    0.523
## minority     -16.44519    2.99647  669.80484  -5.488 5.77e-08 ***
## ses           10.14914    1.53875 1062.65856   6.596 6.66e-11 ***
## yearstea       0.02194    0.15756   75.91038   0.139    0.890
## mathknow       1.04741    1.34469  210.43084   0.779    0.437
## mathprep       0.04601    1.34639  191.56330   0.034    0.973
## housepov     -17.14168   13.45317  119.73031  -1.274    0.205
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.185
```

```
## minority -0.305 -0.012
## ses      -0.119  0.022  0.168
## yearstea -0.369  0.009  0.032 -0.019
## mathknow -0.085  0.008  0.122 -0.001  0.012
## mathprep -0.606 -0.004 -0.007  0.049 -0.139  0.014
## housepov -0.455 -0.004 -0.170  0.079  0.084  0.049  0.050
## convergence code: 0
## Model failed to converge with max|grad| = 0.0406922 (tol = 0.002, component 1)
```

```
anova(model_yearstea_cor,model_all_student,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_yearstea_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_yearstea_cor:     housepov + (yearstea | schoolid) + (1 | classid)
##                     Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student   11 10752 10806 -5364.8    10730
## model_yearstea_cor  13 10750 10814 -5361.8    10724 5.8248      2    0.05434 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

e) There are a couple things I noticed interesting after allowing for correlation between random slopes and intercepts. For MATHKNOW, the variance component seems not to have a huge influence on the variance component as the variance for class-level effect and residual only slightly varies.

For Mathprep, the class-level variance dropped by 20. However, the school-level intercept variance is around 500, which is a significant large increase. This might due to the correlation added to the model between mathprep random slope and school-level intercept.

For Yearstea, the class-level effect variance dropped from 90 to around 40. This variance are explained by the school-level random intercept that are added to the model. Also, the random slope on teacher-level variance are close to 0.54, which are around 0. This might due to the added yeastea that aggregate to have an effect and influence on the variation between school-level.

Question 6 of 12

a. Try to add a random slope for each **student level** predictor (varying at the **classroom level**; one by one separately - not all together).

b. Why is it a bad idea to include a classroom-level variable with random slopes at the classroom level?

c. Retry the above, allowing the slopes to be correlated with the random intercepts. Report findings.

Solution

a)

```
#ses
model_ses<- lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(ses||classi
print(summary(model_ses))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (ses || classid)
##    Data: dat
##
## REML criterion at convergence: 10727.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7163 -0.6032 -0.0331  0.5855  3.6840
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  classid   ses           49.60   7.043
##  classid.1 (Intercept)   87.11   9.333
##  schoolid  (Intercept)  171.02  13.077
##  Residual              1043.44  32.302
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.71226    5.30641  274.46506 101.710  < 2e-16 ***
## sex           -1.37733    2.09334 1022.81814  -0.658    0.511
## minority     -16.29362    3.02464  703.33746  -5.387 9.78e-08 ***
## ses           10.14363    1.64248  176.39731   6.176 4.41e-09 ***
## yearstea       0.01103    0.14117  226.97682   0.078    0.938
## mathknow       1.36796    1.38563  229.40643   0.987    0.325
## mathprep      -0.27938    1.37171  204.89332  -0.204    0.839
## housepov     -17.50879   13.21775  113.44882  -1.325    0.188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.321 -0.011
## ses      -0.108  0.020  0.145
## yearstea -0.259  0.014  0.025 -0.026
## mathknow -0.082  0.006  0.111  0.002  0.029
## mathprep -0.631 -0.005  0.002  0.050 -0.172  0.005
## housepov -0.451 -0.007 -0.180  0.081  0.070  0.058  0.040
```

```
anova(model_all_student, model_ses,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_ses: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_ses:     housepov + (1 | schoolid) + (ses || classid)
##                    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_ses         12 10752 10812 -5364.0    10728 1.5969      1     0.2063
```

```r
#sex
model_sex<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(sex||classid)
print(summary(model_sex))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (sex || classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups    Name        Variance  Std.Dev.
##  classid   sex         3.310e-05  0.005753
##  classid.1 (Intercept) 9.387e+01  9.688824
##  schoolid  (Intercept) 1.695e+02 13.017987
##  Residual              1.065e+03 32.633681
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63033    5.31211  275.37965 101.585  < 2e-16 ***
## sex           -1.21421    2.09483 1022.41528  -0.580    0.562
## minority     -16.18672    3.02607  704.48078  -5.349 1.20e-07 ***
## ses           10.05076    1.54485 1066.56152   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80606   0.080    0.936
## mathknow       1.35013    1.39167  234.49478   0.970    0.333
## mathprep      -0.27702    1.37582  205.26984  -0.201    0.841
## housepov     -17.64878   13.21784  113.87028  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
```

```r
anova(model_all_student,model_sex,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
```

```
## model_sex: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_sex:     housepov + (1 | schoolid) + (sex || classid)
##                   Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_sex         12 10754 10813 -5364.8    10730     0      1          1
```

```
#minority
model_minority<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(minorit
```

```
## boundary (singular) fit: see ?isSingular
```

```
print(summary(model_minority))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (minority || classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8580 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  classid   minority       0.00   0.00
##  classid.1 (Intercept)   93.89   9.69
##  schoolid  (Intercept)  169.45  13.02
##  Residual              1064.95  32.63
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63042    5.31210  275.38909 101.585  < 2e-16 ***
## sex           -1.21419    2.09483 1022.42137  -0.580    0.562
## minority     -16.18678    3.02605  704.47894  -5.349 1.20e-07 ***
## ses           10.05075    1.54484 1066.56222   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80896   0.080    0.936
## mathknow       1.35004    1.39168  234.49773   0.970    0.333
## mathprep      -0.27705    1.37583  205.27155  -0.201    0.841
## housepov     -17.64848   13.21758  113.87764  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
```

```
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(model_all_student,model_sex,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_sex: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_sex:     housepov + (1 | schoolid) + (sex || classid)
##                    Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_sex         12 10754 10813 -5364.8    10730     0      1          1
```

b) Including classroom-level variable with random slopes at the classroom-level will lead to the same effect for each group, which is each classroom. Thus, adding a random slope for the same level does not show the relationship between the outcome variable on different group.

c)

```
#ses with correlation
model_ses_cor<- lmer(math1st~ses+minority+sex+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(ses|cla
print(summary(model_ses_cor))
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: math1st ~ ses + minority + sex + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (ses | classid)
##    Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  10774.5  10839.3  -5374.2  10748.5     1068
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5854 -0.6116 -0.0351  0.5966  3.6376
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   80.69   8.983
##           ses           41.73   6.460   0.79
##  schoolid (Intercept)  165.45  12.863
##  Residual             1046.69  32.353
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.51147    5.20418  275.40194 103.669  < 2e-16 ***
## ses           10.05748    1.63617  171.09752   6.147 5.39e-09 ***
## minority     -16.06530    3.01603  716.43423  -5.327 1.34e-07 ***
```

```
## sex           -1.31305    2.08233 1023.20639  -0.631    0.528
## yearstea       0.01739    0.13926  230.92401   0.125    0.901
## mathknow       1.34919    1.35728  225.63141   0.994    0.321
## mathprep      -0.38392    1.33079  184.55792  -0.288    0.773
## housepov     -16.27372   12.92786  114.17229  -1.259    0.211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) ses    minrty sex    yearst mthknw mthprp
## ses      -0.086
## minority -0.326  0.143
## sex      -0.188  0.022 -0.014
## yearstea -0.267 -0.024  0.021  0.014
## mathknow -0.078  0.014  0.109  0.007  0.029
## mathprep -0.626  0.057  0.004 -0.009 -0.164 -0.001
## housepov -0.447  0.079 -0.182 -0.009  0.074  0.059  0.037
```

```
anova(model_all_student,model_ses_cor)
```

```
## refitting model(s) with ML (instead of REML)

## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_ses_cor: math1st ~ ses + minority + sex + yearstea + mathknow + mathprep +
## model_ses_cor:     housepov + (1 | schoolid) + (ses | classid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10774 10829 -5376.1    10752
## model_ses_cor     13 10774 10839 -5374.2    10748 3.8027      2     0.1494
```

```
#sex with correlation
model_sex_cor<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(sex|clas
print(summary(model_sex_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (sex | classid)
##    Data: dat
##
## REML criterion at convergence: 10729
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7562 -0.6134 -0.0307  0.5916  3.7116
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)  130.2   11.411
##           sex           31.5    5.612   -0.67
```

```
##  schoolid (Intercept)  169.9   13.035
##  Residual             1056.3   32.502
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  5.400e+02  5.332e+00  2.723e+02 101.279  < 2e-16 ***
## sex         -1.197e+00  2.123e+00  2.158e+02  -0.564    0.573
## minority    -1.619e+01  3.028e+00  7.042e+02  -5.347 1.21e-07 ***
## ses          1.010e+01  1.544e+00  1.065e+03   6.539 9.61e-11 ***
## yearstea     3.053e-03  1.416e-01  2.270e+02   0.022    0.983
## mathknow     1.306e+00  1.391e+00  2.315e+02   0.939    0.349
## mathprep    -3.460e-01  1.374e+00  2.014e+02  -0.252    0.801
## housepov    -1.829e+01  1.323e+01  1.145e+02  -1.382    0.170
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.203
## minority -0.321 -0.009
## ses      -0.123  0.020  0.164
## yearstea -0.258  0.015  0.024 -0.027
## mathknow -0.085  0.003  0.116 -0.005  0.029
## mathprep -0.628 -0.008  0.003  0.054 -0.174  0.005
## housepov -0.452 -0.005 -0.178  0.083  0.072  0.060  0.040
```

```
anova(model_all_student,model_sex_cor,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_sex_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_sex_cor:     housepov + (1 | schoolid) + (sex | classid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_sex_cor     13 10755 10820 -5364.5    10729 0.5003      2     0.7787
```

```
#minority with correlation
model_minority_cor<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(1|schoolid)+(min
print(summary(model_minority_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (minority | classid)
##    Data: dat
##
## REML criterion at convergence: 10726.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -3.9036 -0.6221 -0.0295  0.6033  3.4574
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)  225.4   15.01
##           minority     171.3   13.09    -0.82
##  schoolid (Intercept)  157.4   12.55
##  Residual             1045.3   32.33
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.73593    5.38021  270.70839 100.319  < 2e-16 ***
## sex           -1.01014    2.08966 1015.73459  -0.483    0.629
## minority     -16.48615    3.21756  183.24221  -5.124 7.55e-07 ***
## ses            9.89350    1.54595 1062.82952   6.400 2.33e-10 ***
## yearstea      -0.01636    0.14285  234.25604  -0.115    0.909
## mathknow       1.45697    1.39354  234.05425   1.046    0.297
## mathprep      -0.13522    1.37018  203.97781  -0.099    0.921
## housepov     -17.34685   12.91273  103.34823  -1.343    0.182
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.188
## minority -0.368 -0.009
## ses      -0.117  0.021  0.149
## yearstea -0.265  0.015  0.025 -0.023
## mathknow -0.079  0.009  0.108  0.001  0.038
## mathprep -0.618 -0.005 -0.004  0.051 -0.171 -0.006
## housepov -0.435 -0.009 -0.171  0.085  0.080  0.061  0.037
```

```
anova(model_all_student,model_minority_cor,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_minority_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_minority_cor:     housepov + (1 | schoolid) + (minority | classid)
##                    Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student  11 10752 10806 -5364.8    10730
## model_minority_cor 13 10752 10817 -5363.2    10726 3.1967      2     0.2022
```

For allowing correlation on SES variable, the classroom-level variance dropped by 10 and the interaction effect variance between student-level variable SES and classroom-level are around 41. The p-value are 0.14 for comparing the new model with all predictors model. For allowing correlation on SEX variable, the classroom-level variance increased by 40 and intereaction effect variance between student-level variable Sex and classroom-level are around 31. The p-value are 0.7875 implies that it might not be a good fit to allow correlation on Sex variable with random slope.

For allowing correlation on Minority variable, the classroom-level variance significantly increased from 94 to 225, while the interaction effect variance between student-level variable minority and classroom-level are

around 171 with a correlatio of -0.82. The p-value are 0.208. This significant increase might due to the aggregate effect of student-level predictors minority on the classroom-level effect.

Question 7 of 12

a. Try to add a random slope for each **student level** predictor (varying at the **school level**; one by one separately - not all together).

b. Retry the above, allowing the slopes to be correlated with the random intercepts.

c. Report anything unusual about the variance components (changes that are unexpected).

Solution

a)

```
#ses varying at school-level
model_ses_school<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(ses||schoolid)+(1|
print(summary(model_ses_school))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (ses || schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6138 -0.6185 -0.0289  0.5798  3.7130
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  classid    (Intercept)   88.56   9.411
##  schoolid   ses           72.50   8.515
##  schoolid.1 (Intercept)  168.00  12.961
##  Residual               1035.11  32.173
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.13752    5.27926  270.52802 102.124  < 2e-16 ***
## sex           -1.40187    2.08169 1011.29089  -0.673    0.501
## minority     -16.52526    3.02191  700.07600  -5.468 6.32e-08 ***
## ses            9.78982    1.82216   79.01650   5.373 7.61e-07 ***
## yearstea       0.03079    0.14052  223.94252   0.219    0.827
## mathknow       1.35586    1.38461  232.19737   0.979    0.328
## mathprep      -0.19799    1.35995  198.59551  -0.146    0.884
## housepov     -16.94575   13.21161  112.81447  -1.283    0.202
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.323 -0.010
## ses      -0.091  0.017  0.124
## yearstea -0.260  0.018  0.024 -0.019
## mathknow -0.079  0.006  0.110  0.006  0.028
## mathprep -0.628 -0.007  0.001  0.042 -0.172  0.002
## housepov -0.451 -0.007 -0.180  0.076  0.070  0.056  0.041
```

```
anova(model_all_student,model_ses_school,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_ses_school: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_ses_school:     housepov + (ses || schoolid) + (1 | classid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_ses_school  12 10749 10809 -5362.4    10725 4.6972      1    0.03021 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#sex varying at school-level
model_sex_school<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(sex||schoolid)+(1|
print(summary(model_ses_school))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (ses || schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6138 -0.6185 -0.0289  0.5798  3.7130
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  classid    (Intercept)   88.56   9.411
##  schoolid   ses           72.50   8.515
##  schoolid.1 (Intercept)  168.00  12.961
##  Residual               1035.11  32.173
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.13752    5.27926  270.52802 102.124  < 2e-16 ***
## sex           -1.40187    2.08169 1011.29089  -0.673    0.501
## minority     -16.52526    3.02191  700.07600  -5.468 6.32e-08 ***
```

```
## ses            9.78982    1.82216   79.01650   5.373 7.61e-07 ***
## yearstea       0.03079    0.14052  223.94252   0.219    0.827
## mathknow       1.35586    1.38461  232.19737   0.979    0.328
## mathprep      -0.19799    1.35995  198.59551  -0.146    0.884
## housepov     -16.94575   13.21161  112.81447  -1.283    0.202
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.323 -0.010
## ses      -0.091  0.017  0.124
## yearstea -0.260  0.018  0.024 -0.019
## mathknow -0.079  0.006  0.110  0.006  0.028
## mathprep -0.628 -0.007  0.001  0.042 -0.172  0.002
## housepov -0.451 -0.007 -0.180  0.076  0.070  0.056  0.041
```

```
anova(model_all_student,model_sex_school,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_sex_school: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_sex_school:     housepov + (sex || schoolid) + (1 | classid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student 11 10752 10806 -5364.8    10730
## model_sex_school  12 10753 10813 -5364.4    10729 0.6137      1     0.4334
```

```
#minority varying at school-level
model_minority_school<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(minority||sch
```

```
## boundary (singular) fit: see ?isSingular
```

```
## Warning: Model failed to converge with 1 negative eigenvalue: -1.3e+01
```

```
print(summary(model_minority_school))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (minority || schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6134 -0.0321  0.5971  3.6598
##
```

```
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  classid    (Intercept) 9.388e+01  9.689369
##  schoolid   minority    1.777e-06  0.001333
##  schoolid.1 (Intercept) 1.694e+02 13.017176
##  Residual               1.065e+03 32.633690
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63040    5.31208  275.39129 101.586  < 2e-16 ***
## sex           -1.21419    2.09483 1022.42090  -0.580    0.562
## minority     -16.18676    3.02605  704.47702  -5.349 1.20e-07 ***
## ses           10.05076    1.54485 1066.56207   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80855   0.080    0.936
## mathknow       1.35003    1.39168  234.49782   0.970    0.333
## mathprep      -0.27705    1.37582  205.27092  -0.201    0.841
## housepov     -17.64850   13.21752  113.87888  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(model_all_student,model_minority_school,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_minority_school: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_minority_school:     housepov + (minority || schoolid) + (1 | classid)
##                        Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student      11 10752 10806 -5364.8    10730
## model_minority_school 12 10754 10813 -5364.8    10730     0      1          1
```

b)

```
#ses varying at school-level
model_ses_school_cor<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(ses|schoolid)+
print(summary(model_ses_school_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
```

```
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (ses | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5647 -0.6166 -0.0264  0.5888  3.7073
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   86.62   9.307
##  schoolid (Intercept)  171.12  13.081
##           ses           73.35   8.565   0.19
##  Residual             1035.89  32.185
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  538.72268    5.27645  271.16175 102.099  < 2e-16 ***
## sex           -1.40428    2.08074 1011.40155  -0.675    0.500
## minority     -16.26714    3.03575  668.89455  -5.359 1.16e-07 ***
## ses            9.72644    1.82981   78.36254   5.316 9.74e-07 ***
## yearstea       0.03616    0.14002  220.43198   0.258    0.796
## mathknow       1.26005    1.38204  230.90960   0.912    0.363
## mathprep      -0.21707    1.35647  197.11400  -0.160    0.873
## housepov     -15.89853   13.15319  111.73535  -1.209    0.229
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.188
## minority -0.325 -0.011
## ses      -0.062  0.018  0.117
## yearstea -0.259  0.017  0.021 -0.021
## mathknow -0.077  0.005  0.108  0.007  0.028
## mathprep -0.627 -0.008  0.002  0.045 -0.172  0.001
## housepov -0.449 -0.009 -0.183  0.070  0.073  0.057  0.039
```

```
anova(model_all_student,model_ses_school_cor,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_ses_school_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_ses_school_cor:     housepov + (ses | schoolid) + (1 | classid)
##                      Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_all_student    11 10752 10806 -5364.8    10730
## model_ses_school_cor 13 10750 10815 -5362.2    10724 5.1385      2    0.07659 .
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#sex varying at school-level
model_sex_school_cor<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(sex|schoolid)+
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00526733 (tol = 0.002, component 1)
```

```
print(summary(model_ses_school_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (ses | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5647 -0.6166 -0.0264  0.5888  3.7073
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   86.62   9.307
##  schoolid (Intercept)  171.12  13.081
##           ses           73.35   8.565   0.19
##  Residual             1035.89  32.185
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  538.72268    5.27645  271.16175 102.099  < 2e-16 ***
## sex           -1.40428    2.08074 1011.40155  -0.675    0.500
## minority     -16.26714    3.03575  668.89455  -5.359 1.16e-07 ***
## ses            9.72644    1.82981   78.36254   5.316 9.74e-07 ***
## yearstea       0.03616    0.14002  220.43198   0.258    0.796
## mathknow       1.26005    1.38204  230.90960   0.912    0.363
## mathprep      -0.21707    1.35647  197.11400  -0.160    0.873
## housepov     -15.89853   13.15319  111.73535  -1.209    0.229
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.188
## minority -0.325 -0.011
## ses      -0.062  0.018  0.117
## yearstea -0.259  0.017  0.021 -0.021
## mathknow -0.077  0.005  0.108  0.007  0.028
## mathprep -0.627 -0.008  0.002  0.045 -0.172  0.001
## housepov -0.449 -0.009 -0.183  0.070  0.073  0.057  0.039
```

```r
anova(model_all_student,model_sex_school_cor,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_sex_school_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_sex_school_cor:     housepov + (sex | schoolid) + (1 | classid)
##                       Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model_all_student     11 10752 10806 -5364.8    10730
## model_sex_school_cor 13 10754 10818 -5363.8    10728 1.863      2      0.394
```

```r
#minority varying at school-level
model_minority_school_cor<-lmer(math1st~sex+minority+ses+yearstea+mathknow+mathprep+housepov+(minority|s
print(summary(model_minority_school_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (minority | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10717.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8952 -0.6358 -0.0345  0.6129  3.6444
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   86.7    9.311
##  schoolid (Intercept)  381.2   19.524
##           minority     343.2   18.525   -0.83
##  Residual             1039.4   32.240
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.49369    5.65513  173.09178  95.399  < 2e-16 ***
## sex           -0.86278    2.08382 1021.81437  -0.414    0.679
## minority     -16.37547    3.89604   58.24604  -4.203 9.17e-05 ***
## ses            9.43095    1.54335 1063.13485   6.111 1.39e-09 ***
## yearstea      -0.00437    0.13765  217.17884  -0.032    0.975
## mathknow       1.63216    1.35929  224.78144   1.201    0.231
## mathprep      -0.29178    1.33537  198.06922  -0.218    0.827
## housepov     -16.06251   12.57477   99.99134  -1.277    0.204
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.172
## minority -0.494 -0.014
```

```
## ses      -0.105  0.024  0.113
## yearstea -0.253  0.014  0.027 -0.021
## mathknow -0.078  0.010  0.099 -0.005  0.024
## mathprep -0.576 -0.005 -0.002  0.052 -0.167 -0.002
## housepov -0.394 -0.013 -0.157  0.089  0.091  0.061  0.037
```

```
anova(model_all_student,model_minority_school_cor,refit=F)
```

```
## Data: dat
## Models:
## model_all_student: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_all_student:     housepov + (1 | schoolid) + (1 | classid)
## model_minority_school_cor: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_minority_school_cor:     housepov + (minority | schoolid) + (1 | classid)
##                           Df   AIC   BIC  logLik deviance  Chisq Chi Df
## model_all_student         11 10752 10806 -5364.8    10730
## model_minority_school_cor 13 10744 10808 -5358.8    10718 11.967      2
##                           Pr(>Chisq)
## model_all_student
## model_minority_school_cor    0.00252 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c) One thing that is unusual is the fact that when allowing correlation for between minority(varing at school level) random effect, the variance of school-level random intercept and minority random slope has a significant increase of 381 and 343. This might due to the high negative correlation between the random intercept and the random slope which is -0.83.

Question 8 of 12

a. Take the two predictors that had significant (at 0.05 level) random slopes, in the forms in which they worked (independent or correlated) and add both to the model, and test for need of one conditional on needing the other.

b. Is the more complex model (with both random slopes in it) justified?

c. **Write out this model** in your preferred notation.

Solution

a)

```
model_complex<- lmer(math1st~ses+minority+sex+yearstea+mathknow+mathprep+housepov+(minority|schoolid)+(0
print(summary(model_complex))
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: math1st ~ ses + minority + sex + yearstea + mathknow + mathprep +
##     housepov + (minority | schoolid) + (0 + ses | schoolid) +
##     (1 | classid)
##    Data: dat
```

```
##
##      AIC      BIC   logLik deviance df.resid
##  10763.6  10833.4  -5367.8  10735.6     1067
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6750 -0.6274 -0.0308  0.6075  3.7143
##
## Random effects:
##  Groups     Name        Variance Std.Dev. Corr
##  classid    (Intercept)   75.7    8.700
##  schoolid   ses           72.0    8.485
##  schoolid.1 (Intercept)  388.2   19.703
##             minority     322.4   17.954   -0.85
##  Residual               1008.6   31.759
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.04847    5.59366  169.61465  96.368  < 2e-16 ***
## ses            9.19972    1.80877   83.00657   5.086 2.23e-06 ***
## minority     -16.70586    3.86031   56.49063  -4.328 6.22e-05 ***
## sex           -1.03004    2.06437 1016.13264  -0.499    0.618
## yearstea       0.02201    0.13511  216.50022   0.163    0.871
## mathknow       1.64745    1.33510  224.48655   1.234    0.219
## mathprep      -0.23892    1.30286  193.31862  -0.183    0.855
## housepov     -15.32842   12.29149  101.57414  -1.247    0.215
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) ses    minrty sex    yearst mthknw mthprp
## ses      -0.081
## minority -0.508  0.088
## sex      -0.172  0.020 -0.013
## yearstea -0.255 -0.011  0.027  0.017
## mathknow -0.072  0.005  0.094  0.010  0.023
## mathprep -0.569  0.042 -0.003 -0.005 -0.165 -0.004
## housepov -0.393  0.084 -0.150 -0.014  0.094  0.060  0.040
```

```
anova(model_complex,model_ses_school)
```

```
## refitting model(s) with ML (instead of REML)

## Data: dat
## Models:
## model_ses_school: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
## model_ses_school:     housepov + (ses || schoolid) + (1 | classid)
## model_complex: math1st ~ ses + minority + sex + yearstea + mathknow + mathprep +
## model_complex:     housepov + (minority | schoolid) + (0 + ses | schoolid) +
## model_complex:     (1 | classid)
##                  Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model_ses_school 12 10772 10832 -5373.9    10748
```

```
## model_complex    14 10764 10833 -5367.8    10736 12.239      2     0.0022 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

b) According to the LRT test, we can conclude that the p-value ofr adding the minority variable random slope(varying at school-level) while allowing correlation to the model conditioning on adding ses variable random slope(varying at school-level) is 0.002. This implies that there is a need for adding both variable to the model. This justified the more complex model.

c) The model for adding both SES variable and Minority variable(varying at school-level while allowing correlation for minority) is

$MATH1ST_{ijk} = b_0 + b_1 SES_{ijk} + b_2 Minority_{ijk} + b_3 Sex_{ijk} + b_4 Yearstea_{jk} + + b_5 Mathknow_{jk} + b_6 Mathprep_{jk} + b_7 Housep$

, where

$$\eta_{0jk} \sim N(0, \sigma^2_{\eta_0})$$

,

$$\zeta_0 k \sim N(0, \zeta^2_{\zeta_0})$$

,

$$\zeta_1 k \sim N(0, \zeta^2_{\zeta_1})$$

,

$$\zeta_2 k \sim N(0, \zeta^2_{\zeta_2})$$

,

$$\epsilon_{ijk} \sim N(0, \sigma^2_{\epsilon})$$

, i represents students, j represents classrooms and k represents schools.

$$Corr(\zeta_{0k}, \zeta_{2k}) = \rho_{\zeta_0, \zeta_2}$$

, all other random terms independent of each other.

Question 9 of 12

a. For UMM, write down: `V_S`, `V_C`, `V_E` for the three variance components (simply the estimates).

b. For the most complicated (all fixed effects) random **intercepts only** model, what are: `V_S`, `V_C`, `V_E`?

c. By what fraction did these each decrease with the new predictors in the model?

Solution

```
print(summary(unconditional_model))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ (1 | schoolid/classid)
##    Data: dat
##
## REML criterion at convergence: 11944.6
##
## Scaled residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -5.1872 -0.6174 -0.0204  0.5821  3.8339
##
## Random effects:
##  Groups           Name        Variance Std.Dev.
##  classid:schoolid (Intercept)   85.47   9.245
##  schoolid         (Intercept)  280.69  16.754
##  Residual                      1146.79  33.864
## Number of obs: 1190, groups:  classid:schoolid, 312; schoolid, 107
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  522.540      2.037 104.403   256.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print(summary(model_all_student))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (1 | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10729.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8581 -0.6134 -0.0321  0.5971  3.6598
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  classid  (Intercept)   93.89   9.689
##  schoolid (Intercept)  169.45  13.017
##  Residual              1064.96  32.634
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63041    5.31209  275.39010 101.585  < 2e-16 ***
## sex           -1.21419    2.09483 1022.42110  -0.580    0.562
## minority     -16.18676    3.02605  704.47787  -5.349 1.20e-07 ***
## ses           10.05076    1.54485 1066.56211   6.506 1.18e-10 ***
## yearstea       0.01129    0.14141  226.80861   0.080    0.936
## mathknow       1.35004    1.39168  234.49768   0.970    0.333
## mathprep      -0.27705    1.37583  205.27111  -0.201    0.841
## housepov     -17.64850   13.21755  113.87814  -1.335    0.184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.320 -0.011
```

```
## ses      -0.121  0.020  0.162
## yearstea -0.259  0.016  0.024 -0.028
## mathknow -0.083  0.007  0.115 -0.007  0.029
## mathprep -0.631 -0.006  0.001  0.053 -0.172  0.004
## housepov -0.451 -0.007 -0.178  0.082  0.071  0.058  0.038
```

a) For Unconditional mean model, `V_C` is 85.47 , `V_S` is 280.69, and `V_E` is 1146.79.

b) For most complicated model(Random Intercept only), `V_C` is 93.89 , `V_S` is 169.45, and `V_E` is 1064.96.

c) For `V_C`, it increased by 9.85% For `V_S`, it decreased by 39.63% For `V_E`, it decreased by 7.14%

Question 10 of 12

Now consider the model with a random slope in `ses`.

a. What are: `V_C`, `V_S` (`ses` = 0), `V_E`?

- We need to list `ses = 0` here, or we don't know how to use the slope variance.

b. What are: `V_S` (`ses` = -0.50), `V_S` (`ses` = +0.50)?

Solution

```
print(summary(model_ses_school))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (ses || schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6138 -0.6185 -0.0289  0.5798  3.7130
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  classid    (Intercept)   88.56   9.411
##  schoolid   ses           72.50   8.515
##  schoolid.1 (Intercept)  168.00  12.961
##  Residual               1035.11  32.173
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.13752    5.27926  270.52802 102.124  < 2e-16 ***
## sex           -1.40187    2.08169 1011.29089  -0.673    0.501
## minority     -16.52526    3.02191  700.07600  -5.468 6.32e-08 ***
```

```
## ses            9.78982    1.82216   79.01650   5.373 7.61e-07 ***
## yearstea       0.03079    0.14052  223.94252   0.219    0.827
## mathknow       1.35586    1.38461  232.19737   0.979    0.328
## mathprep      -0.19799    1.35995  198.59551  -0.146    0.884
## housepov     -16.94575   13.21161  112.81447  -1.283    0.202
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.190
## minority -0.323 -0.010
## ses      -0.091  0.017  0.124
## yearstea -0.260  0.018  0.024 -0.019
## mathknow -0.079  0.006  0.110  0.006  0.028
## mathprep -0.628 -0.007  0.001  0.042 -0.172  0.002
## housepov -0.451 -0.007 -0.180  0.076  0.070  0.056  0.041
```

a) For model with a random slope in `ses`, the `V_C` is 88.56, `V_S`(`ses`=0) is 168, `V_E` is 1035.11

b) `V_S` (`ses` = -0.50) is

$$168 + 0.5 \times 72.5 = 204.5$$

`V_S` (`ses` = +0.50) is

$$168 - 0.5 \times 72.5 = 131.75$$

Question 11 of 12

Now consider the model with a random slope in `minority`.

a. What are: `V_C`, `V_S` (`minority` = 0), `V_E`?

- We need to list `minority` = 0 here, or we don't know how to use the slope variance.

b. What are: `V_S` (`minority` = +0.25),`V_S` (`minority` = +0.50),`V_S` (`minority` = +0.75)?

Solution

```
print(summary(model_minority_school_cor))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ sex + minority + ses + yearstea + mathknow + mathprep +
##     housepov + (minority | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10717.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8952 -0.6358 -0.0345  0.6129  3.6444
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   86.7    9.311
##  schoolid (Intercept)  381.2   19.524
##           minority     343.2   18.525   -0.83
##  Residual             1039.4   32.240
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.49369    5.65513  173.09178  95.399  < 2e-16 ***
## sex           -0.86278    2.08382 1021.81437  -0.414    0.679
## minority     -16.37547    3.89604   58.24604  -4.203 9.17e-05 ***
## ses            9.43095    1.54335 1063.13485   6.111 1.39e-09 ***
## yearstea      -0.00437    0.13765  217.17884  -0.032    0.975
## mathknow       1.63216    1.35929  224.78144   1.201    0.231
## mathprep      -0.29178    1.33537  198.06922  -0.218    0.827
## housepov     -16.06251   12.57477   99.99134  -1.277    0.204
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) sex    minrty ses    yearst mthknw mthprp
## sex      -0.172
## minority -0.494 -0.014
## ses      -0.105  0.024  0.113
## yearstea -0.253  0.014  0.027 -0.021
## mathknow -0.078  0.010  0.099 -0.005  0.024
## mathprep -0.576 -0.005 -0.002  0.052 -0.167 -0.002
## housepov -0.394 -0.013 -0.157  0.089  0.091  0.061  0.037
```

a) For model with a random slope in `minority`(allowing correlation,varying at school-level random effect), the `V_C` is 86.7, `V_S` (`minority` $= 0$) is 381.2, `V_E` is 1039.4.

b) `V_S` (`minority` $= +0.25$) is

$$381.2 + 0.25 \times 343.2 = 467$$

`V_S` (`minority` $= +0.50$) is

$$381.2 + 0.5 \times 343.2 = 552.8$$

`V_S` (`minority` $= +0.75$) is

$$381.2 + 0.75 \times 343.2 = 638.6$$

Question 12 of 12

Now consider the model with a random slope in `ses` and `minority`.

a. What are: `V_C`, `V_S` (`minority` $= 0$, `ses` $= 0$), `V_E`?

- We need to list `minority` $= 0$ and `ses` $= 0$ here, or we don't know how to use the slope variance.

b. In the last model, what is a "likely" (+/- 1 s.d.) range for η_{0jk}.

c. Can we make a similar statement about ζ_{0k}?

d. If you had a large value for η_{0jk}, would you expect a large or small or "any" value for the two random slope terms, ζ_{1k} and ζ_{2k} for `ses` and `minority`?

e. If you had a large value for ζ_{0jk}, would you expect a large or small or "any" value for the two random slope terms, ζ_{1k} and ζ_{2k} for `ses` and `minority` (discuss each separately)?

solution

```
print(summary(model_complex))
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: math1st ~ ses + minority + sex + yearstea + mathknow + mathprep +
##     housepov + (minority | schoolid) + (0 + ses | schoolid) +
##     (1 | classid)
##    Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  10763.6  10833.4  -5367.8  10735.6     1067
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6750 -0.6274 -0.0308  0.6075  3.7143
##
## Random effects:
##  Groups     Name        Variance Std.Dev. Corr
##  classid    (Intercept)   75.7    8.700
##  schoolid   ses           72.0    8.485
##  schoolid.1 (Intercept)  388.2   19.703
##             minority     322.4   17.954   -0.85
##  Residual               1008.6   31.759
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.04847    5.59366  169.61465  96.368  < 2e-16 ***
## ses            9.19972    1.80877   83.00657   5.086 2.23e-06 ***
## minority     -16.70586    3.86031   56.49063  -4.328 6.22e-05 ***
## sex           -1.03004    2.06437 1016.13264  -0.499    0.618
## yearstea       0.02201    0.13511  216.50022   0.163    0.871
## mathknow       1.64745    1.33510  224.48655   1.234    0.219
## mathprep      -0.23892    1.30286  193.31862  -0.183    0.855
## housepov     -15.32842   12.29149  101.57414  -1.247    0.215
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) ses    minrty sex    yearst mthknw mthprp
## ses      -0.081
## minority -0.508  0.088
## sex      -0.172  0.020 -0.013
## yearstea -0.255 -0.011  0.027  0.017
```

```
## mathknow -0.072  0.005  0.094  0.010  0.023
## mathprep -0.569  0.042 -0.003 -0.005 -0.165 -0.004
## housepov -0.393  0.084 -0.150 -0.014  0.094  0.060  0.040
```

a) Model with random slope in SES(varying at school-level effect, no correlation with random intercept), Minority(varying at school-level effect, with correlation), `V_C` = 75.7, `V_S` (`minority` = 0, `ses` = 0)= 388.2, `V_E`= 1008.6

b) In the last model, η_{0jk} (+/- 1 s.d.) are likely to be around -2 to 2 since it fits a normal distribution from 0 to 76.

c) ζ_{0k} are likely to be from -5 to 5. Since it fits a normal distribution from 0 to 388.

d) If η_{0jk} is relatively large, then the two random slope ζ_{1k} and ζ_{2k} for `ses` and `minority` are not likely to be affected by this change, since ζ_{2k}, ζ_{1k} and η_{0jk} should be independent of each other.

e) If ζ_{0jk} are likely to be large, then for ζ_{1k} (`ses`) it would likely to not have an impact since those two coefficient are indepedent of each other. For ζ_{2k} (`minority`), it would likely to be relatively small since there is negative correlation of -0.85 between those two random effect coefficient.