MLM Nested Main Section B

Xinming Dai; Checked by Chongjun Liao, Jeremy Lu, Yu Wang

Compiled on Thu May 12 10:34 EDT

Question 0: read data and process missingness

```
dat <- read_csv("classroom.csv")
```

```
## Rows: 1190 Columns: 12
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (12): sex, minority, mathkind, mathgain, ses, yearstea, mathknow, housep...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# construct new outcome math1st
dat <- dat %>%
    mutate(math1st = mathkind + mathgain)

# remove missing data
dat <- dat %>%
    filter(complete.cases(dat))
```

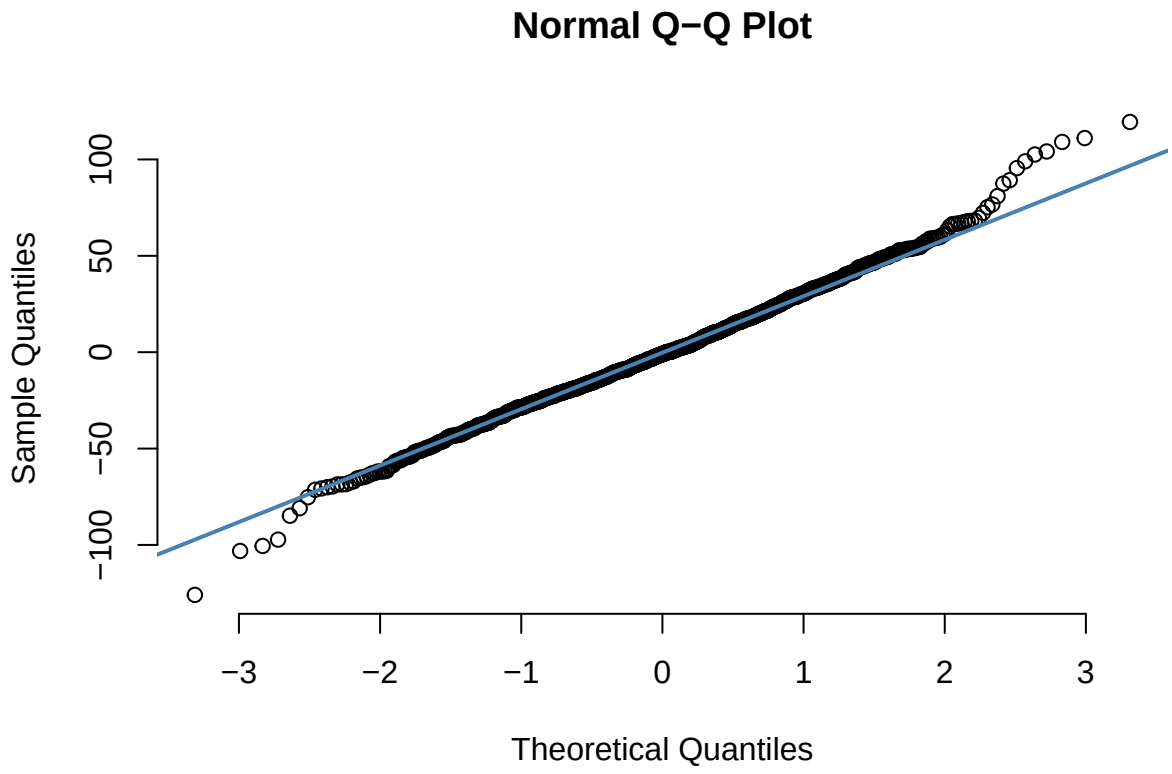
Question 1

```
# fit a model
fit1 <- lmerTest::lmer(math1st ~ housepov + yearstea + mathprep + mathknow + ses +
    sex + minority + (1 | schoolid/classid), data = dat)
summary(fit1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex +
##     minority + (1 | schoolid/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:schoolid (Intercept)   93.89   9.689
##  schoolid         (Intercept)  169.45  13.017
```

```
##  Residual                      1064.96  32.634
## Number of obs: 1081, groups:  classid:schoolid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  539.63041    5.31209  275.39010 101.585  < 2e-16 ***
## housepov     -17.64850   13.21755  113.87814  -1.335    0.184
## yearstea       0.01129    0.14141  226.80861   0.080    0.936
## mathprep      -0.27705    1.37583  205.27111  -0.201    0.841
## mathknow       1.35004    1.39168  234.49768   0.970    0.333
## ses           10.05076    1.54485 1066.56211   6.506 1.18e-10 ***
## sex           -1.21419    2.09483 1022.42110  -0.580    0.562
## minority     -16.18676    3.02605  704.47787  -5.349 1.20e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) houspv yearst mthprp mthknw ses    sex
## housepov -0.451
## yearstea -0.259  0.071
## mathprep -0.631  0.038 -0.172
## mathknow -0.083  0.058  0.029  0.004
## ses      -0.121  0.082 -0.028  0.053 -0.007
## sex      -0.190 -0.007  0.016 -0.006  0.007  0.020
## minority -0.320 -0.178  0.024  0.001  0.115  0.162 -0.011
```

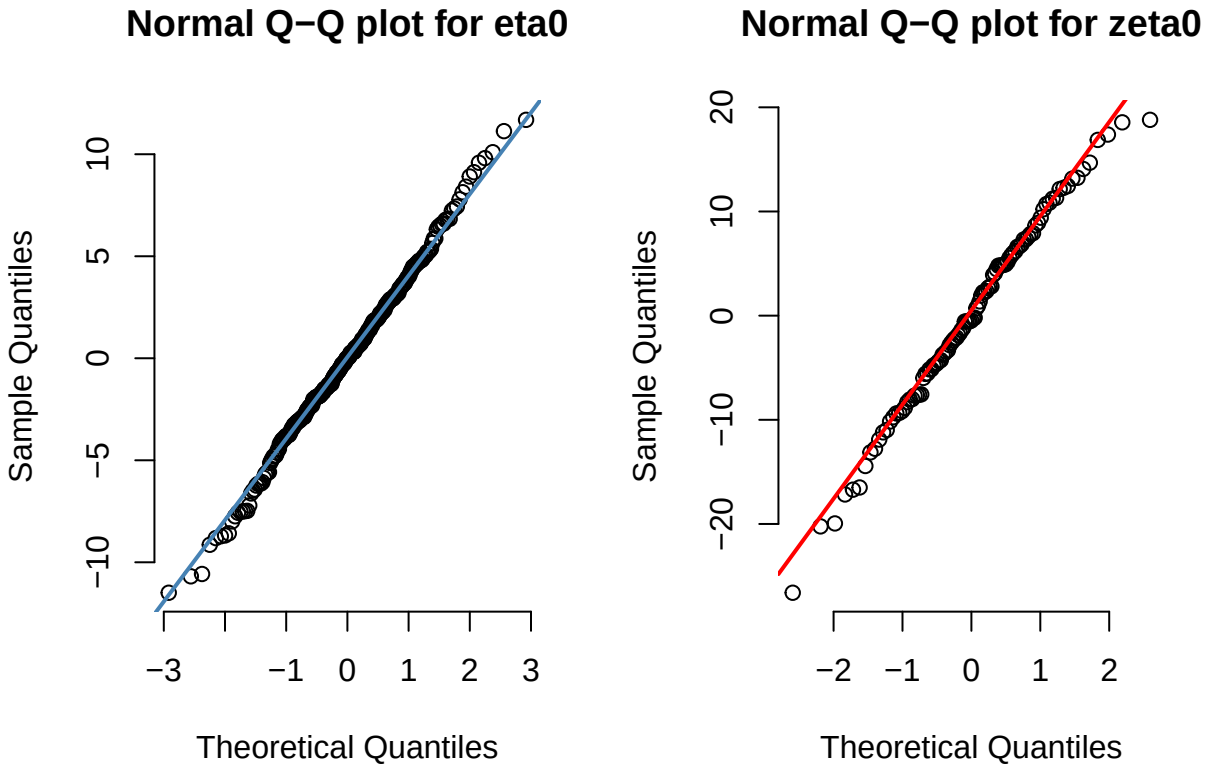
```r
# plot residuals to test normality assumption
res1 <- residuals(fit1)
qqnorm(res1, pch = 1, frame = FALSE)
qqline(res1, col = "steelblue", lwd = 2)
```



QQ plot shows that points are around the line, and thus we believe the normality assumption holds.

Question 2

```
# Generate the two sets of BLUPs (for random effects zeta0 and eta0)
blups_fit1 <- ranef(fit1)

par(mfrow = c(1, 2))
# examine normality for eta0 (class-level)
eta0_fit1 <- blups_fit1$`classid:schoolid`$`(Intercept)`
qqnorm(eta0_fit1, pch = 1, frame = FALSE, main = "Normal Q-Q plot for eta0")
qqline(eta0_fit1, col = "steelblue", lwd = 2)

# examine normality for zeta0 (school-level)
zeta0_fit1 <- blups_fit1$schoolid$`(Intercept)`
qqnorm(zeta0_fit1, pch = 1, frame = FALSE, main = "Normal Q-Q plot for zeta0")
qqline(zeta0_fit1, col = "red", lwd = 2)
```




```
par(mfrow = c(1, 1))
```

QQ plot shows that both sets of BLUPs of zeta0 and eta0 are around the line, and thus we believe the normality assumption holds.

Question 3

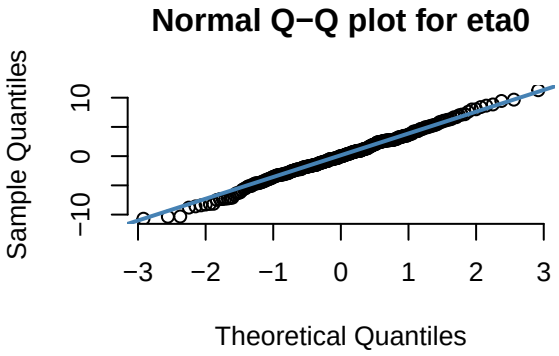
```
# a add a random slope for minority, correlated with the random intercept, at
# the school level
fit2 <- lmerTest::lmer(math1st ~ housepov + yearstea + mathprep + mathknow + ses +
    sex + minority + (minority | schoolid) + (1 | classid), data = dat)
```

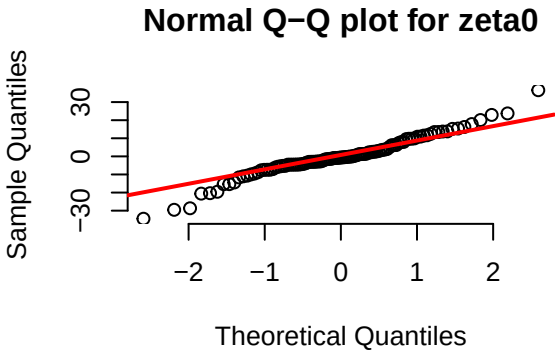
```
print(summary(fit2))
```

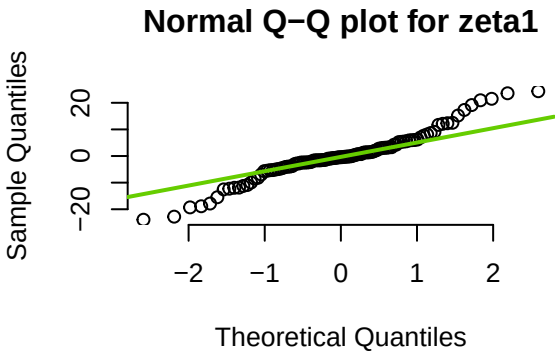
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex +
##     minority + (minority | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10717.5
##
## Scaled 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)   86.69   9.311
##  schoolid (Intercept)  381.20  19.524
##           minority     343.13  18.524   -0.83
##  Residual             1039.39  32.240
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  5.395e+02  5.655e+00  1.731e+02  95.399  < 2e-16 ***
## housepov    -1.606e+01  1.257e+01  9.999e+01  -1.277    0.204
## yearstea    -4.368e-03  1.376e-01  2.172e+02  -0.032    0.975
## mathprep    -2.918e-01  1.335e+00  1.981e+02  -0.218    0.827
## mathknow     1.632e+00  1.359e+00  2.248e+02   1.201    0.231
## ses          9.431e+00  1.543e+00  1.063e+03   6.111 1.39e-09 ***
## sex         -8.628e-01  2.084e+00  1.022e+03  -0.414    0.679
## minority    -1.638e+01  3.896e+00  5.824e+01  -4.203 9.17e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) houspv yearst mthprp mthknw ses    sex
## housepov -0.394
## yearstea -0.253  0.091
## mathprep -0.576  0.037 -0.167
## mathknow -0.078  0.061  0.024 -0.002
## ses      -0.105  0.089 -0.021  0.052 -0.005
## sex      -0.172 -0.013  0.014 -0.005  0.010  0.024
## minority -0.494 -0.157  0.027 -0.002  0.099  0.113 -0.014
```

```
# b residual
blups_fit2 <- ranef(fit2)
# BULPs
zeta0 <- blups_fit2$schoolid$`(Intercept)`
zeta1 <- blups_fit2$schoolid$minority
eta0 <- blups_fit2$classid$`(Intercept)`

# c check normality
par(mfrow = c(2, 2))
```

```r
# examine normality for eta0 (class-level)
qqnorm(eta0, pch = 1, frame = FALSE, main = "Normal Q-Q plot for eta0")
qqline(eta0, col = "steelblue", lwd = 2)

# examine normality for zeta0 (school-level)
zeta0_fit1 <- blups_fit1$schoolid$`(Intercept)`
qqnorm(zeta0, pch = 1, frame = FALSE, main = "Normal Q-Q plot for zeta0")
qqline(zeta0, col = "red", lwd = 2)

# examine normality for zeta1 (random slop)
qqnorm(zeta1, pch = 1, frame = FALSE, main = "Normal Q-Q plot for zeta1")
qqline(zeta1, col = "chartreuse3", lwd = 2)
par(mfrow = c(1, 1))
```

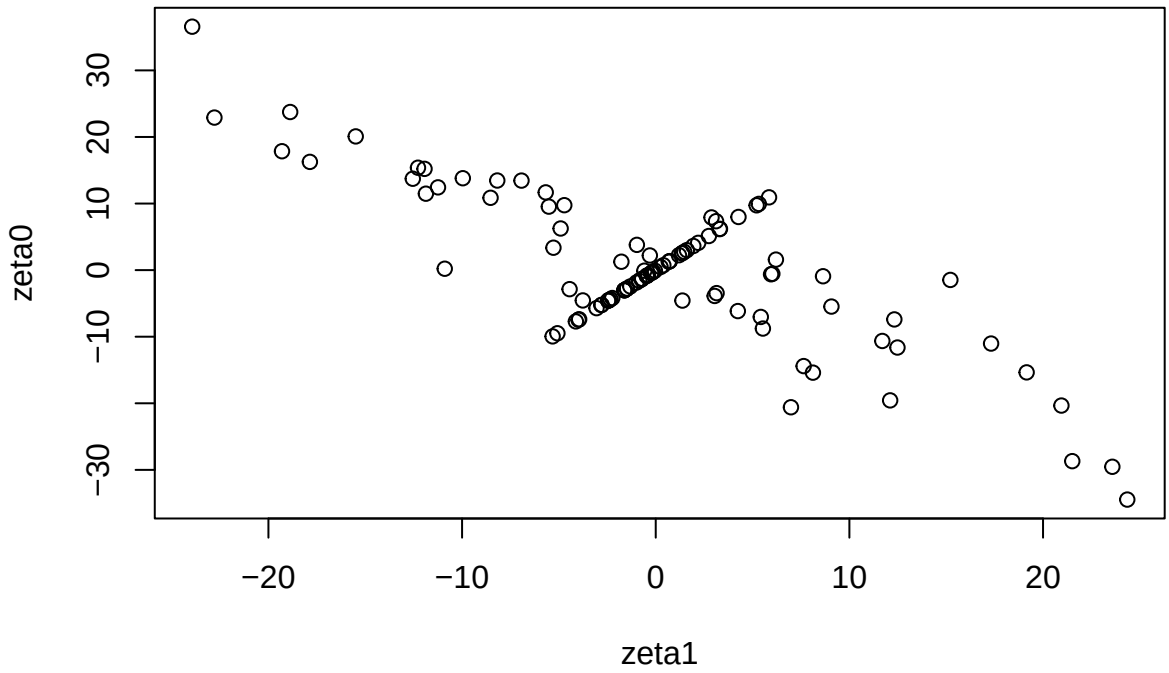






QQ plot shows that BLUPs of zeta0, eta0, and zeta1 are around the line, and thus we believe the normality assumption holds.

```r
# d
plot(zeta1, zeta0)
```



Overall, zeta0 and zeta1 are negative correlated. However, some odds points are positive correlated.

```
# e
blups_fit2$schoolid$minority[3 - abs(blups_fit2$schoolid$minority) > 0]
```

```
##  [1] -1.6264791  1.1988784 -0.4602063 -0.1853924 -0.2355959 -1.5989159
##  [7] -1.0001424 -0.1355968 -2.4186766  1.6075153 -2.2437240 -0.0437638
## [13]  0.7238725 -2.4516203 -2.3348817 -1.3118949 -0.5863832  1.9406871
## [19]  0.7191746  2.7431116 -0.7352057 -1.4842833 -0.8509368  2.8806161
## [25]  1.3783179 -1.6078368 -0.9737567 -2.8103456 -1.6126150 -2.4197034
## [31]  0.2491616  2.1930359 -0.6985121 -2.7917343  1.4722214 -0.4017513
## [37] -0.3037134 -0.4639631 -2.4558712  0.6751290  1.3778922 -1.5011151
## [43]  0.3935569 -1.7735168  1.3485642
```

Question 4

a

$V_S = 93.89$, $V_C = 169.45$, and $V_E = 1064.96$.

```
# fit a model
fit3 <- lmerTest::lmer(math1st ~ housepov + yearstea + mathprep + mathknow + ses +
    sex + minority + (ses | schoolid) + (1 | classid), data = dat)
summary(fit3)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex +
##     minority + (ses | schoolid) + (1 | classid)
##    Data: dat
##
## REML criterion at convergence: 10724.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -3.5646 -0.6166 -0.0264  0.5888  3.7073
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  classid  (Intercept)   86.57   9.305
##  schoolid (Intercept)  171.18  13.083
##           ses           73.36   8.565   0.19
##  Residual             1035.90  32.185
## Number of obs: 1081, groups:  classid, 285; schoolid, 105
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  538.72222    5.27648  271.13305 102.099  < 2e-16 ***
## housepov     -15.89873   13.15396  111.71336  -1.209    0.229
## yearstea       0.03617    0.14002  220.42240   0.258    0.796
## mathprep      -0.21697    1.35642  197.10758  -0.160    0.873
## mathknow       1.26025    1.38201  230.89913   0.912    0.363
## ses            9.72646    1.82985   78.36212   5.315 9.75e-07 ***
## sex           -1.40436    2.08074 1011.40322  -0.675    0.500
## minority     -16.26698    3.03580  668.91588  -5.358 1.16e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) houspv yearst mthprp mthknw ses    sex
## housepov -0.449
## yearstea -0.259  0.073
## mathprep -0.627  0.039 -0.172
## mathknow -0.077  0.057  0.028  0.001
## ses      -0.062  0.070 -0.021  0.045  0.007
## sex      -0.188 -0.009  0.017 -0.008  0.005  0.018
## minority -0.325 -0.182  0.021  0.002  0.108  0.117 -0.011
```

c

$V_C = 86.57$, $V_{S(ses=0)} = 171.18$, and $V_E = 73.36$.

d

$V_{S(ses=-0.50)} = 171.18 + 2 * (-0.5) * 13.083 * 8.565 * 0.19 + (-0.5)^2 * 73.36 = 168.23$

$V_{S(ses=0.50)} = 171.18 + 2 * (0.5) * 13.083 * 8.565 * 0.19 + (0.5)^2 * 73.36 = 210.81$ ### e There is heteroscedasticity at school level because $V_{S(ses=0.50)}$ and $V_{S(ses=-0.50)}$ are not approximate