2042 MLM Final Group Project (Spring 2020)

Part 2

Group 1

May 15, 2020

```r
# Load ---------------------------------------------------------------------------
dat <- read.csv("data/classroom.csv")
dat$math1st <- dat$mathkind + dat$mathgain
```

Team members and division of work

Group 1 Team Members:

Frank Jiang, Lisa Song, Yuyue Hua, Seeun Jang, Tong Jin

Division of Work:

Frank Jiang: Group project part 1

Lisa Song: Group project part 2

Yuyue Hua: Group project part 2

Seeun Jang: Group project part 1

Tong Jin: The mini project

All team members: Review all submissions

Question 1

Refit the model in Part 1 that has all fixed effects as well as random intercepts (in schools and classrooms). Recall that `math1st = mathkind + mathgain` is the outcome.

The model is `math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex + minority + (1|schoolid/classid), REML = T)`

```
# Fit the model with all fixed effects and random intercepts
fit1 <- lmerTest::lmer(
  math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex + minority +
    (1|schoolid/classid),
  data = dat,
  REML = T
)
# Report the model fit
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
```

a. Construct the residual that removes only the 'fixed effects' then subtract it from the outcome; call this residual `resFE`

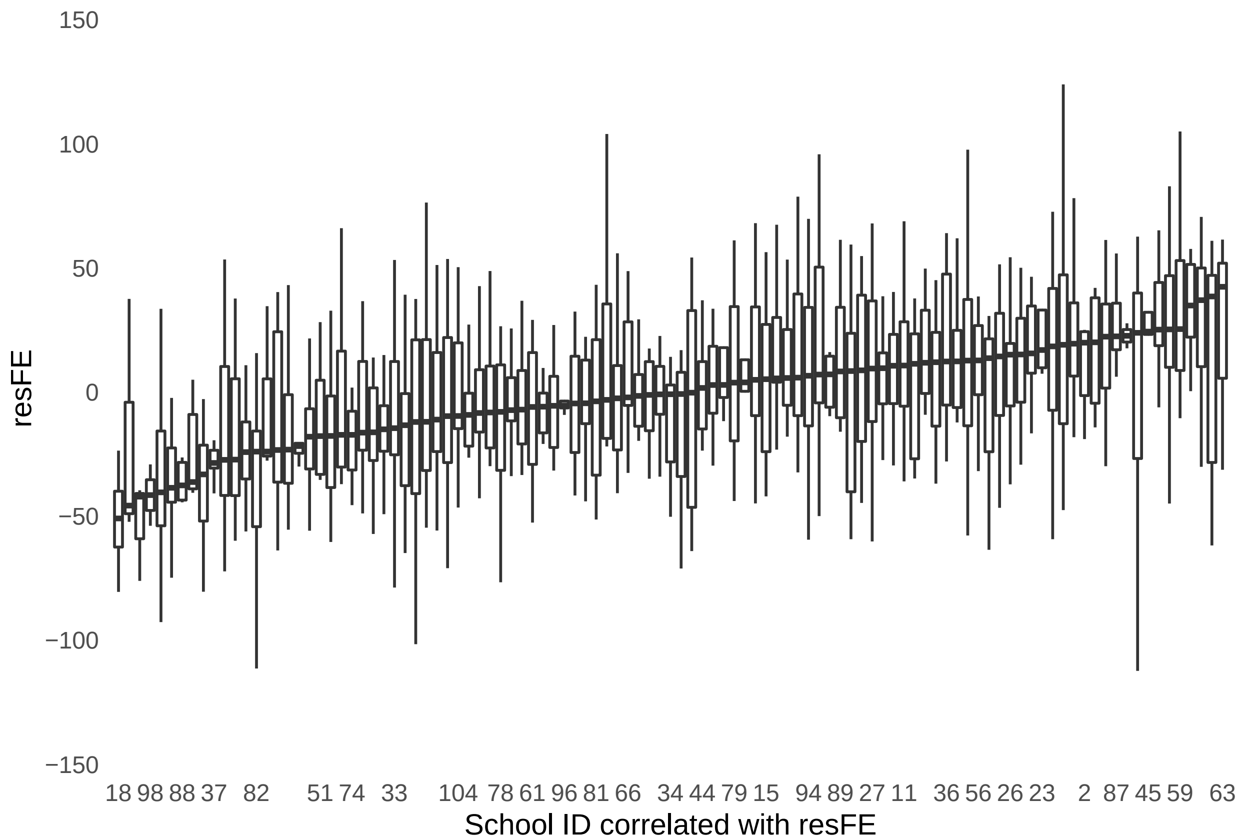
   i. R hint 1: `predict` has an option to generate the prediction based on the fixed effects only.
   ii. R hint 2: If you decide to add a column to your data frame with resFE, note that predict only generates predictions for cases uses in the model *after listwise deletion.*

```r
# Make prediction
yhat <- predict(fit1, re.form = ~0)

# Drop rows containing missing values
dat <- dat[complete.cases(dat), ]

# Create a variable to store residual
dat$resFE <- dat$math1st - yhat
```

Question 2

Show that the residual is not independent within schools in some manner.

```r
# Visualize that the residual is not independent within schools
ggplot(dat, aes(x = reorder(schoolid, resFE, FUN = median), y = resFE)) +
  geom_boxplot(outlier.alpha = 0) +
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank()
  ) +
  xlab("School ID correlated with resFE") +
  scale_x_discrete(guide = guide_axis(check.overlap = TRUE))
```



The above graph shows that as reordered school id increases, the `resFE` gradually increases. There is a positive correlation between them.

Question 3

Construct the residual that utilizes the BLUPs for the random effects using the R command `residuals`.

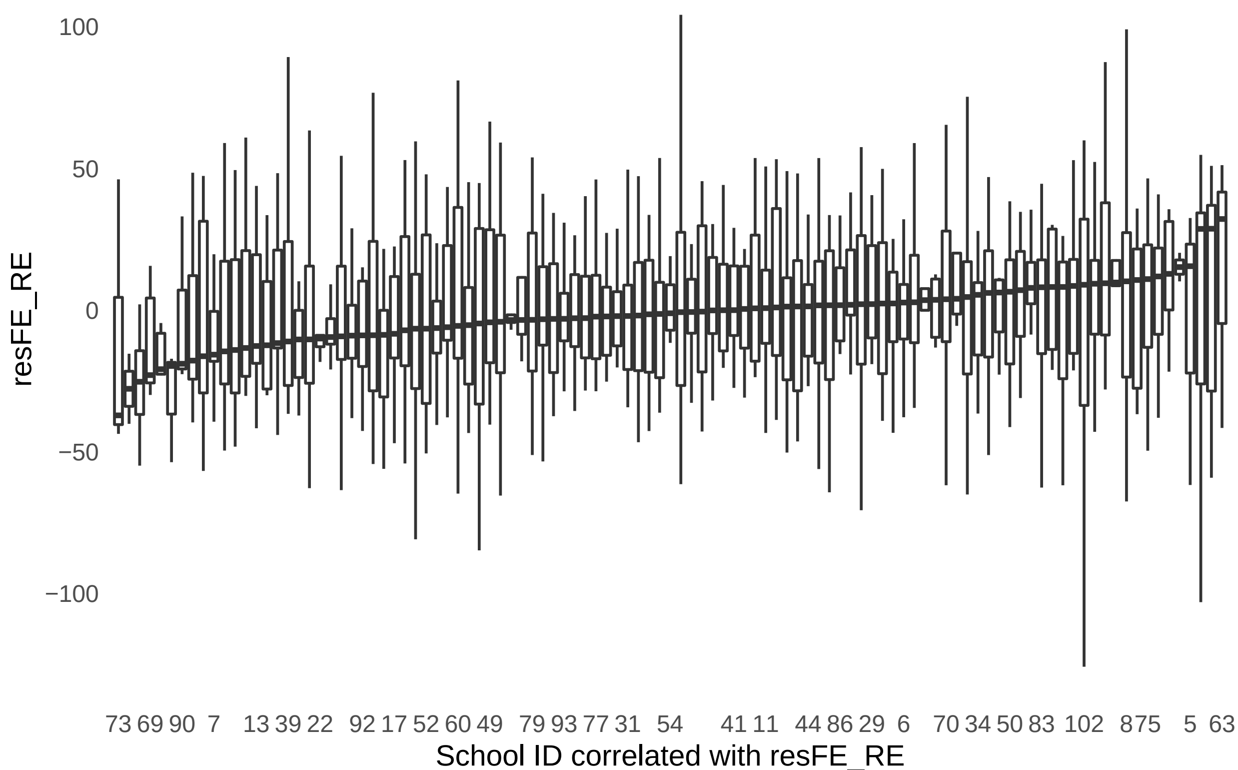
```
i. Call the new residual resFE_RE
```

```r
dat$resFE_RE <- residuals(fit1)
```

Question 4

Show that these new residuals, `resFE_RE`, are MUCH LESS (if not completely un-) correlated within schools, using the same method as before (boxplot?)

```
ggplot(dat, aes(x = reorder(schoolid, resFE_RE, FUN = median), y = resFE_RE)) +
  geom_boxplot(outlier.alpha = 0) +
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank()
  ) +
  xlab("School ID correlated with resFE_RE") +
  scale_x_discrete(guide = guide_axis(check.overlap = TRUE))
```



Response: The new residuals, resFE_RE, are much less correlated within school. This is because the boxplots are more centered around zero (median for each school would be much closer to zero if the errors are independent).

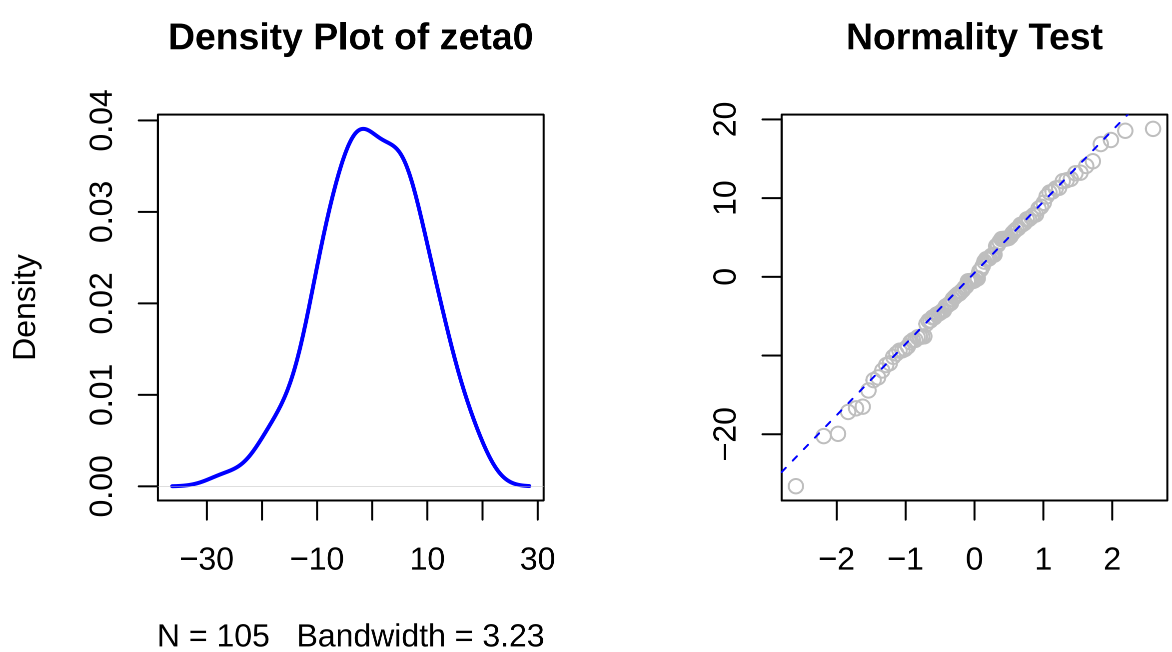
Question 5

a. Generate the two sets of BLUPs (for random effects ζ_0 and η_0)

```
ranefs <- ranef(fit1)
zeta0 <- ranefs$schoolid[, 1]
eta0 <- ranefs$classid[, 1]
```

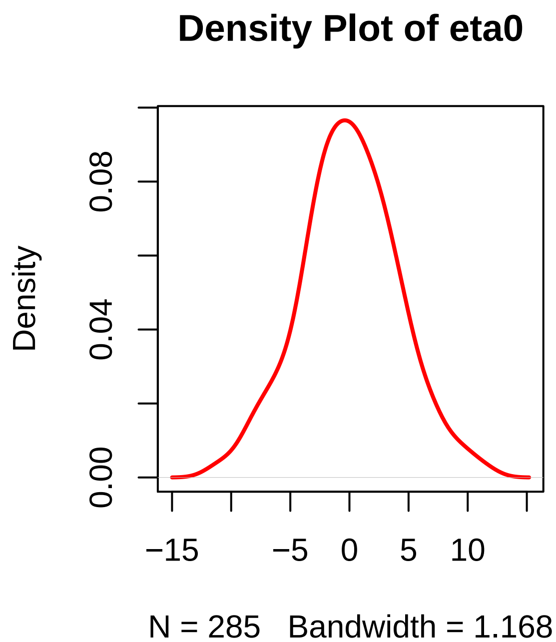
b. Examine these for normality (include evidence)

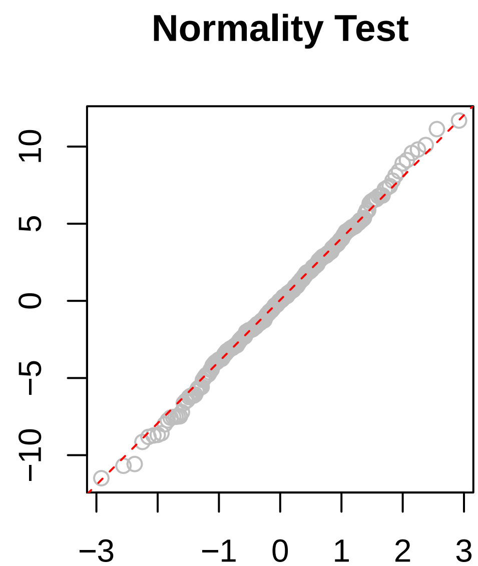
```
# Use QQ-plot to examine zeta0 and eta0
par(mfrow = c(1, 2), pty = "s")
plot(density(zeta0),
     lty = 1, lwd = 2, col = "Blue", main = "")
title("Density Plot of zeta0", cex = 0.8)
qqnorm(zeta0,
       col = "Gray", ann = FALSE)
title("Normality Test", cex = 0.8)
qqline(zeta0,
       lty = 2, lwd = 1, col = "Blue")
```




```
par(mfrow = c(1, 2), pty = "s")
plot(density(eta0),
     lty = 1, lwd = 2, col = "Red", main = "")
```

```r
title("Density Plot of eta0", cex = 0.8)
qqnorm(eta0,
       col = "Gray", ann = FALSE)
title("Normality Test", cex = 0.8)
qqline(eta0,
       lty = 2, lwd = 1, col = "Red")
```





Response: It appears that the BLUPs for classroom effects are fairly normal. Their density plot has a bell-shaped, symmetric distribution and their Q-Q plot has most of the points falling about the straight line. The BLUPs for school effects appear slightly less normal. Their density plot has a less symmetric distribution and the points on their Q-Q plot deviate a bit more from forming a straight line.

Question 6

Returning to the classroom data.

a. Fit a slightly more complicated model with the same fixed effects, but now add a random slope for minority, correlated with the random intercept, at the school level (keep the classroom level random intercept).

```
fit2 <- lmerTest::lmer(
  math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex + minority +
    (minority|schoolid) + (1|schoolid:classid), data = dat)
print(summary(fit2))
```

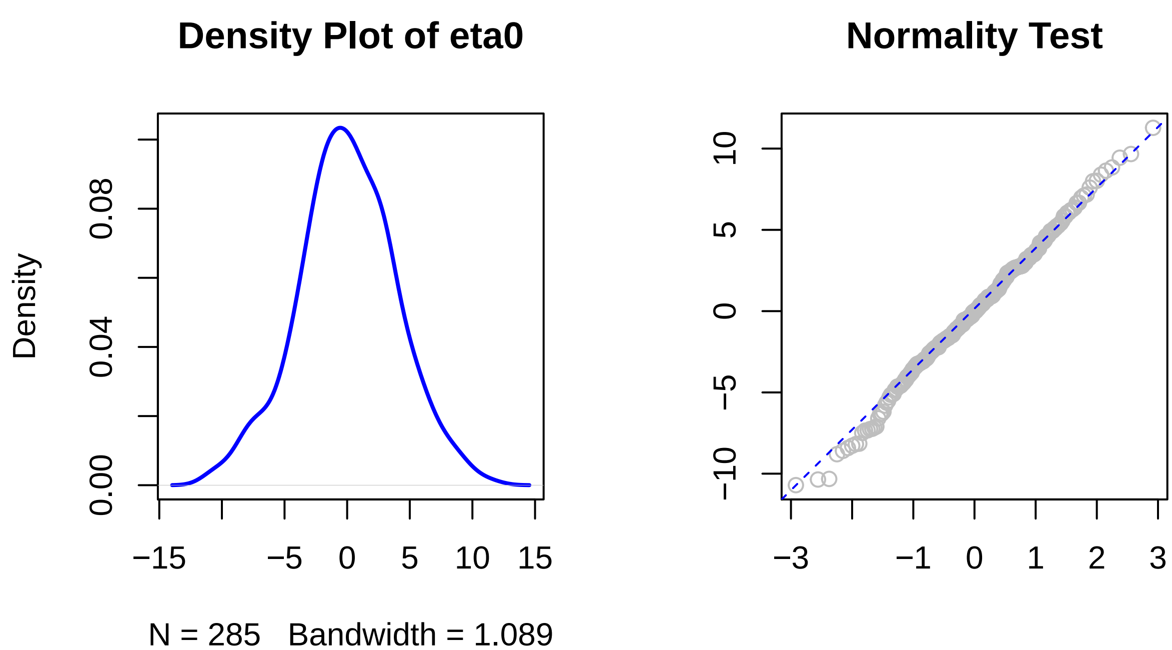
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ housepov + yearstea + mathprep + mathknow + ses + sex +
##     minority + (minority | schoolid) + (1 | schoolid:classid)
##    Data: dat
##
## REML criterion at convergence: 10717.5
##
## 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
##  schoolid: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:  schoolid: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 ***
## housepov     -16.06251   12.57477   99.99134  -1.277    0.204
## yearstea      -0.00437    0.13765  217.17884  -0.032    0.975
## mathprep      -0.29178    1.33537  198.06922  -0.218    0.827
## mathknow       1.63216    1.35929  224.78144   1.201    0.231
## ses            9.43095    1.54335 1063.13485   6.111 1.39e-09 ***
## sex           -0.86278    2.08382 1021.81437  -0.414    0.679
## minority     -16.37547    3.89604   58.24604  -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. Construct the residual (individual, level 1) and the BLUPs for the remaining random effects. Call the new residual resFE_RE as before.

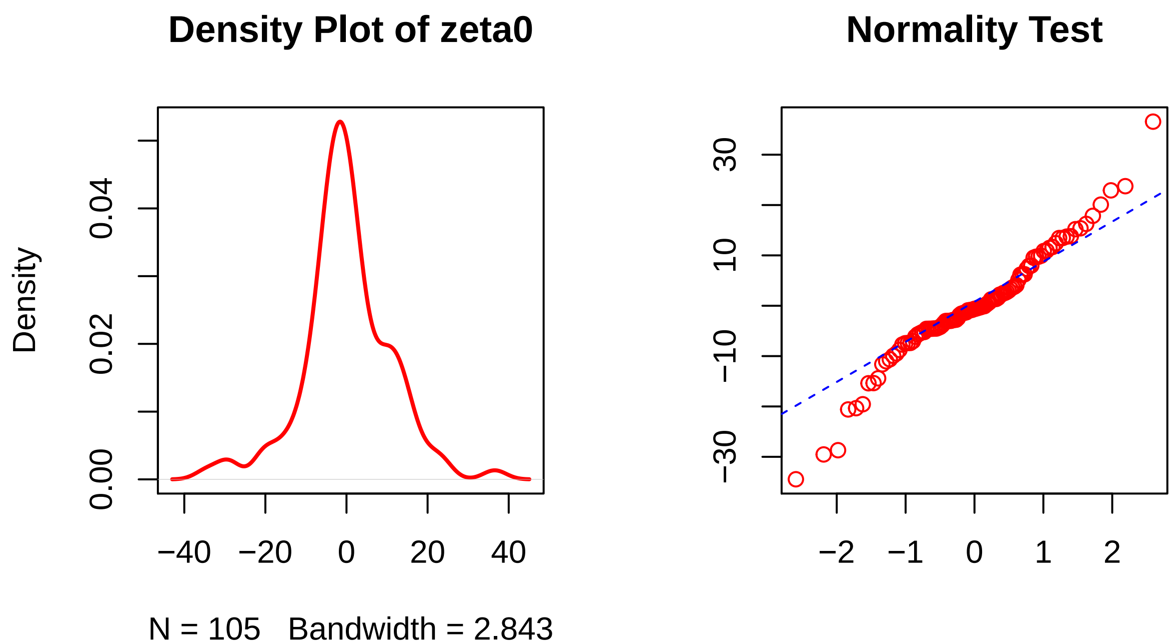
```
# Residual
resFE_RE <- residuals(fit2)

# BLUPs
ranefs.fit2 <- lme4::ranef(fit2)
eta0.fit2 <- ranefs.fit2$'schoolid:classid'[, 1]
zeta0.fit2 <- ranefs.fit2$schoolid[, 1]
zeta1.fit2 <- ranefs.fit2$schoolid[,2]
```

c. Examine all error estimates (individual level residuals, BLUPs (school and classroom level) for normality.

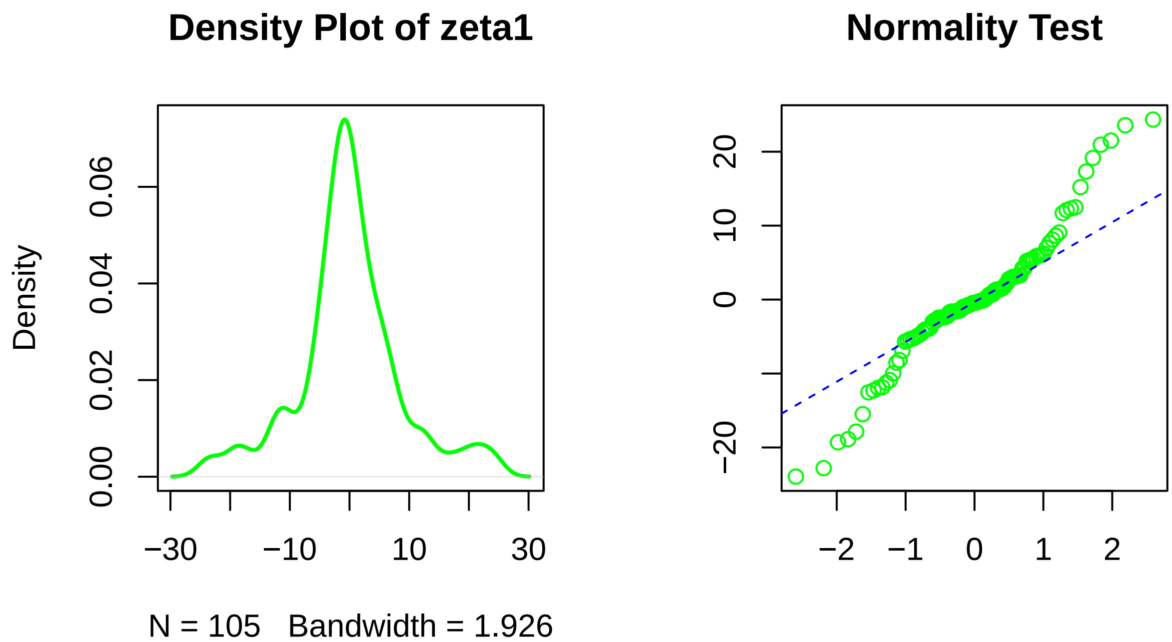
```
# Examine eta0
par(mfrow = c(1, 2), pty = "s")
plot(density(eta0.fit2),
     lty = 1, lwd = 2, col = "Blue", main = "")
title("Density Plot of eta0", cex = 0.8)
qqnorm(eta0.fit2,
       col = "Gray", ann = FALSE)
title("Normality Test", cex = 0.8)
qqline(eta0.fit2,
       lty = 2, lwd = 1, col = "Blue")
```



```r
# Examine zeta0
par(mfrow = c(1, 2), pty = "s")
plot(density(zeta0.fit2),
     lty = 1, lwd = 2, col = "Red", main = "")
title("Density Plot of zeta0", cex = 0.8)
qqnorm(zeta0.fit2,
       col = "Red", ann = FALSE)
title("Normality Test", cex = 0.8)
qqline(zeta0.fit2,
       lty = 2, lwd = 1, col = "Blue")
```



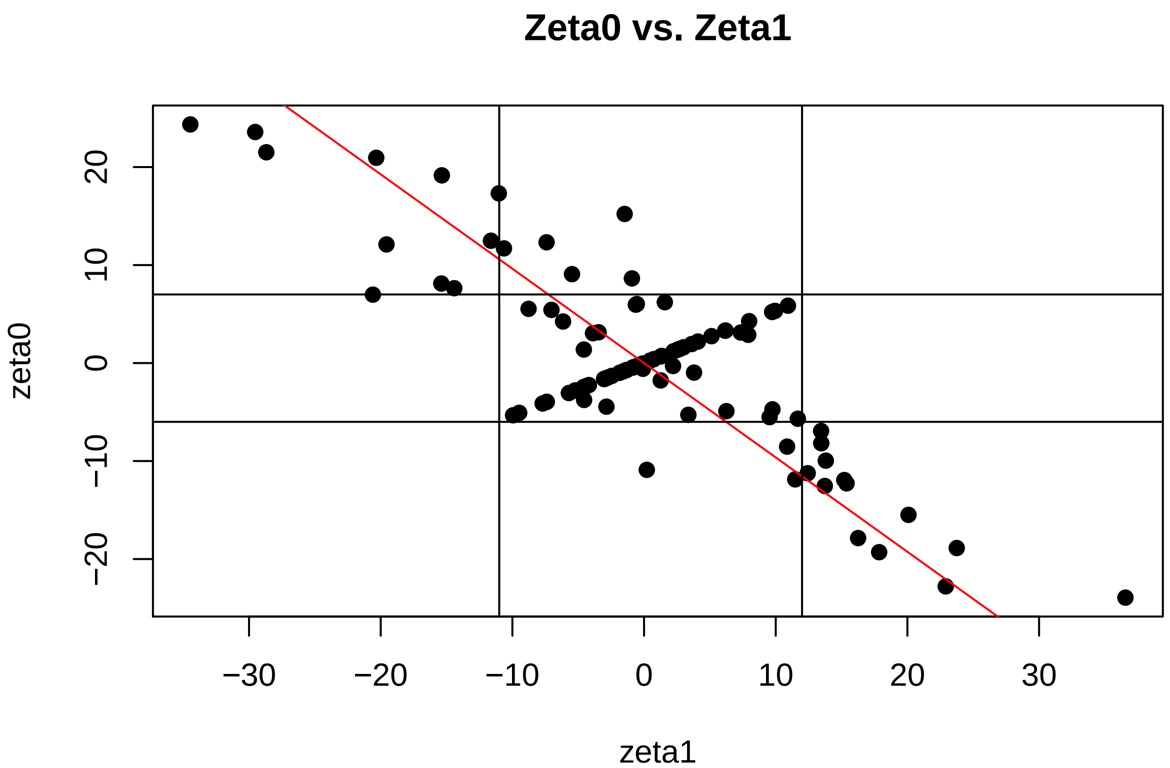

```r
# Examine zeta1
par(mfrow = c(1, 2), pty = "s")
plot(density(zeta1.fit2),
     lty = 1, lwd = 2, col = "Green", main = "")
title("Density Plot of zeta1", cex = 0.8)
qqnorm(zeta1.fit2,
       col = "Green", ann = FALSE)
title("Normality Test", cex = 0.8)
qqline(zeta1.fit2,
       lty = 2, lwd = 1, col = "Blue")
```



Response: It appears that the individual level residuals and BLUPs for classroom effects are fairly normal. Their density plots have a bell-shaped, symmetric distribution and their Q-Q plots have most of the points falling about the straight line. The BLUPs for school effects (both the random slope for minority and the random intercept) are less normal. Their density plots have a less symmetric distribution and the points on their Q-Q plots deviate more from forming a straight line.

d. Plot zeta0 vs. zeta1 to see whether the estimated correlation is consistent with the observed. **Briefly comment**.

```
plot(zeta0.fit2, zeta1.fit2, main="Zeta0 vs. Zeta1",
   xlab="zeta1", ylab="zeta0", pch=19)
abline(lm(zeta0.fit2~zeta1.fit2), col="red")
abline(7,0)
abline(-6,0)
abline(v=-11)
abline(v=12)
```



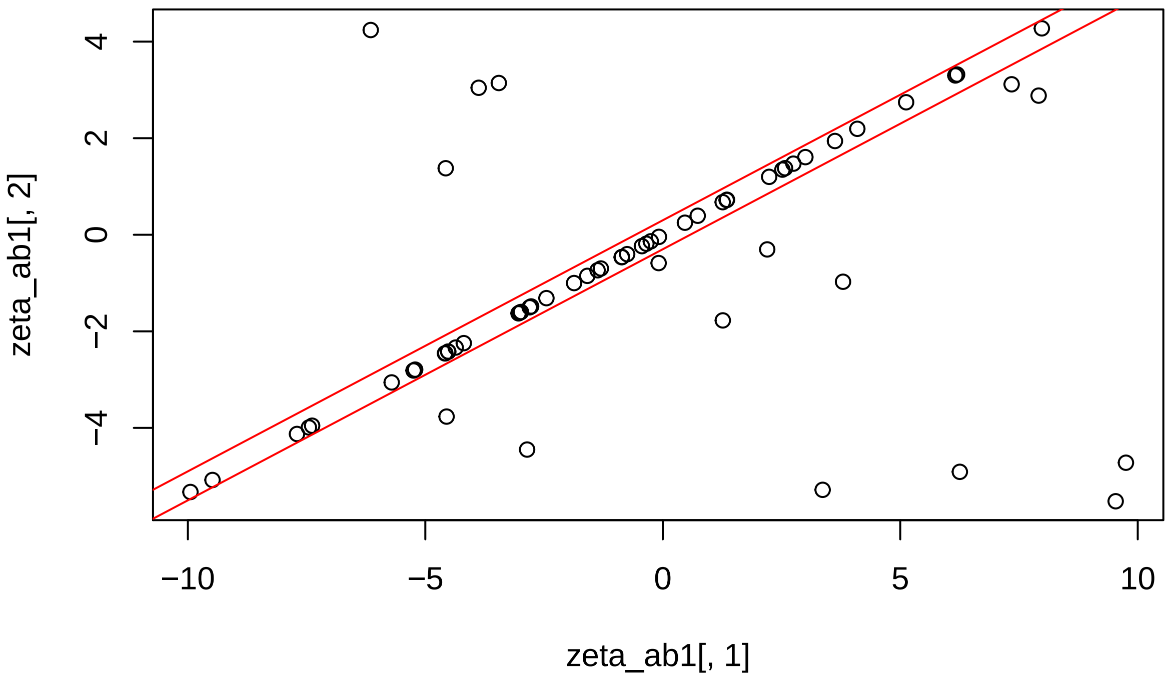

```r
# Estimated correlation
cor(zeta0.fit2,zeta1.fit2)
```

```
## [1] -0.7852153
```

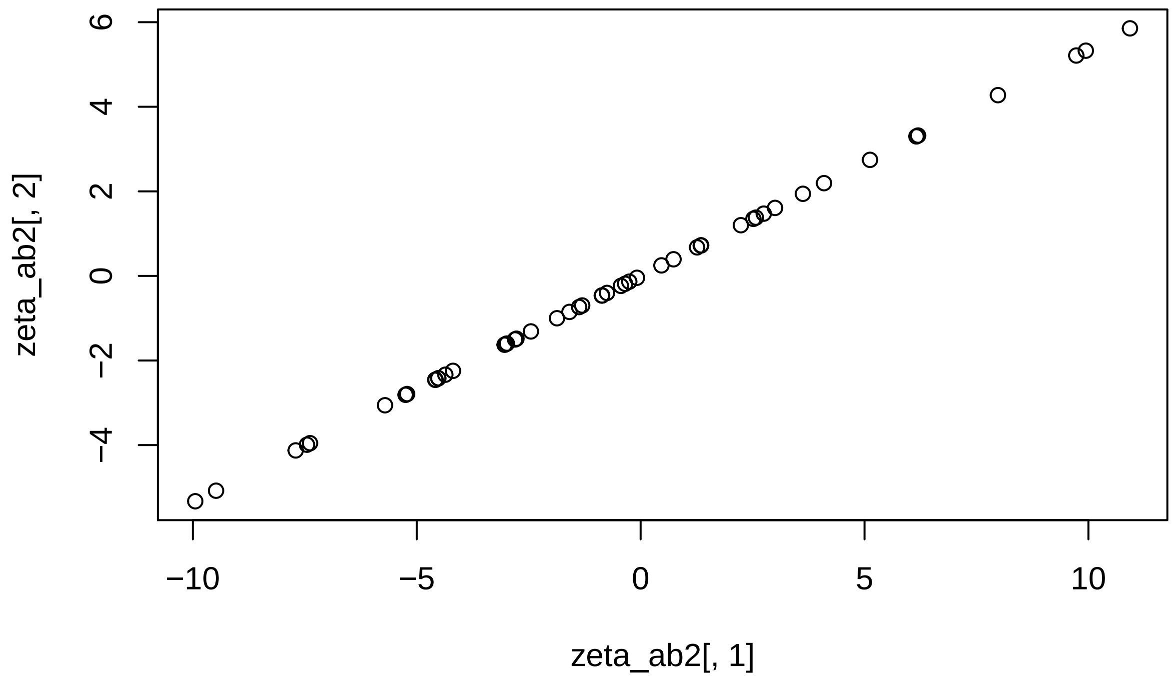
Response: The estimated correlation is -0.785 but it is not consistent with some of the points in the scatterplot that appear to be positively correlated.

e. Track down those odd points in the scatterplot. What schools are they? Do they have anything in common?

```r
# Insert code if you want to examine odd points
zeta <- ranefs.fit2$schoolid
zeta$schooid <- row.names(zeta)
zeta_ab1 <- zeta[zeta0.fit2 > -10 & zeta0.fit2 < 10 &
                   zeta1.fit2 > -10 & zeta1.fit2 < 5, ]
plot(zeta_ab1[,1], zeta_ab1[,2])
abline(a=0.3,b=0.52,col=2)
abline(a=-0.3,b=0.52,col=2)
```




```
zeta_ab2 <- zeta[0.3+0.52*zeta0.fit2>zeta1.fit2 & -0.3+0.52*zeta0.fit2<zeta1.fit2,]
plot(zeta_ab2[,1],zeta_ab2[,2])
```




```
#Part of Schoolids for the weird points
head(zeta_ab2[,3])
```

```
## [1] "1"  "4"  "5"  "9"  "10" "12"
```

```
dat2<-dat[dat$schoolid %in% zeta_ab2[,3],]

library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:car':
##
##     recode

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
head(dat2 %>% group_by(schoolid) %>% summarize(mean(minority)))
```

```
## # A tibble: 6 x 2
##   schoolid `mean(minority)`
##      <int>            <dbl>
## 1        1                1
## 2        4                1
## 3        5                1
## 4        9                1
## 5       10                1
## 6       12                1
```

Response: The odd points are from `schoolids` 1, 4, 5, 9, 10, 12, 14, 16, 17, 19, 20, 22, 23, 24, 26, 28, 33, 38, 42, 43, 45, 46, 47, 52, 57, 60, 61, 69, 73, 78, 79, 80, 84, 86, 89, 90, 96, 98, 100, 102, 103, and 106. We can see that the odd points are from schools where high proportion of students or even all of them are minorities.

Question 7

Make a *person-period* file with math score (Kindergarten and First grade). That is, `math0 <- mathkind; math1 <- mathkind + mathgain` (you have to make this work in the dataframe). Using `reshape` in R, you have to be careful to specify the name of the math variable (`math0` and `math1`) as *varying*.

```
dat$math0 <- dat$mathkind
dat$math1 <- dat$mathkind+dat$mathgain
class_pp <- reshape(
  dat,
  varying = c("math0", "math1"),
  v.names = "math",
  timevar = "year",
  times = c(0, 1),
  direction = "long"
)
```

Question 8

We ignore classrooms in this analysis, but keep it in the notation.

a. Fit a model with `math` as outcome, and fixed effect for time trend (`year`), and random intercepts for schools.

```
fit.M00 <- lmer(math ~ year + (1 | schoolid),
                data = class_pp)
# Report the model fit
print(summary(fit.M00))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math ~ year + (1 | schoolid)
##    Data: class_pp
##
## REML criterion at convergence: 21794.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.0693 -0.6031  0.0030  0.6321  3.7529
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  schoolid (Intercept)  337     18.36
##  Residual             1288     35.89
## Number of obs: 2162, groups:  schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)  465.053      2.136  133.580  217.76   <2e-16 ***
## year          57.844      1.544 2055.557   37.47   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## year -0.361
```

b. Write down the model

Equation:

$$MATH_{tijk} = b_0 + \zeta_{0k} + b_1 TIME_{tijk} + \varepsilon_{tijk},$$

and we assume $\zeta_{0k} \sim \mathcal{N}(0, \sigma^2_{\zeta_0})$ and $\varepsilon_{tijk} \sim \mathcal{N}(0, \sigma^2_{\varepsilon})$, independently.

c. Add random intercepts for child

```
fit.M0 <- lmer(math ~ year + (1 | schoolid/childid),
               data = class_pp)
# Report the model fit
summary(fit.M0)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math ~ year + (1 | schoolid/childid)
##    Data: class_pp
##
## REML criterion at convergence: 21425.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7233 -0.4827  0.0125  0.4922  3.4892
##
## Random effects:
##  Groups           Name        Variance Std.Dev.
##  childid:schoolid (Intercept) 722.0    26.87
##  schoolid         (Intercept) 293.2    17.12
##  Residual                     602.2    24.54
## Number of obs: 2162, groups:  childid:schoolid, 1081; schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)  465.288      2.057  116.652   226.2   <2e-16 ***
## year          57.844      1.056 1080.000    54.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## year -0.257
```

d. Write down the model

Equation:

$$MATH_{tijk} = b_0 + \delta_{0ijk} + \zeta_{0k} + b_1 TIME_{tijk} + \varepsilon_{tijk}$$

and assume $\delta_{0ijk} \sim \mathcal{N}(0, \sigma^2_{\delta_0})$, $\zeta_{0k} \sim \mathcal{N}(0, \sigma^2_{\zeta_0})$, and $\varepsilon_{tijk} \sim \mathcal{N}(0, \sigma^2_{\varepsilon})$, independently.

Question 9

Report original and new variance estimates of $\sigma^2_{\zeta_0}$ (between schools) and σ^2_{ε} (within schools):

$\sigma^2_{\zeta_0}$: The original variance estimate is 337. The new variance estimate is 293.2. σ^2_{ε} : The original variance estimate is 1288. The new variance estimate is 602.2.

a. Compute a pseudo R^2 relating the between school variation and ignoring between students in the same school. In other words, what fraction of the between-school variance in the first model is 'explained' by the addition of a student random effect?

```
# Insert code to compute psuedo R^2 or do this inline
(337 - 293.2) / 337
```

```
## [1] 0.1299703
```

Response: The proportion of the between-school variance in the first model that is 'explained' by the addition of a student random effect is 0.1299703.

b. Does the total variation stay about the same (adding between children within schools variance as well, to the second model results) **(you should comment)**?

```
# Total variation of first model
337 + 1288
```

```
## [1] 1625
```

```
# Total variaion of second model
722 + 293.2 + 602.2
```

```
## [1] 1617.4
```

Response: Yes, the total variation stays about the same. The total variation of the first model is 1625. The total variation of the second model is 1617.4.

Question 10

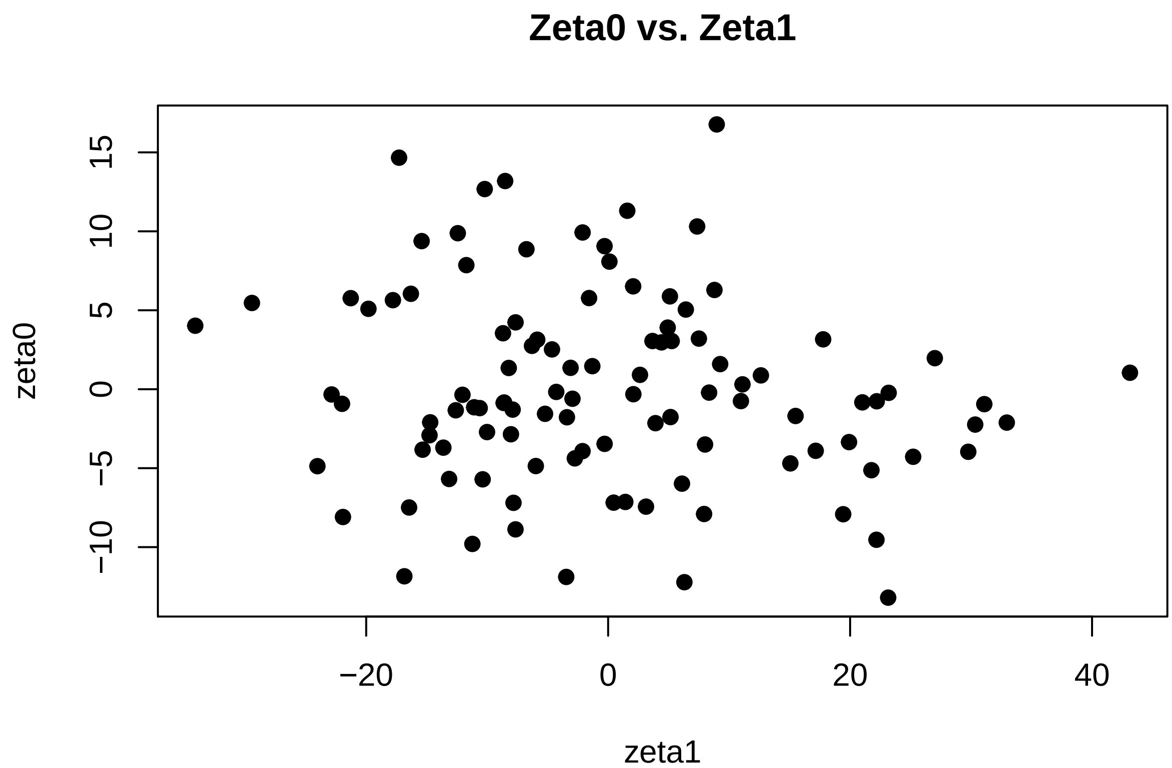
Add a random slope (ζ_1) for the trend (year) within schools (uncorrelated with random intercept (ζ_0))

```
fit.M1 <- lmer(math ~ year + (0 + year | schoolid) + (1 | schoolid/childid),
               data = class_pp)
summary(fit.M1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math ~ year + (0 + year | schoolid) + (1 | schoolid/childid)
##    Data: class_pp
##
## REML criterion at convergence: 21403.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7461 -0.4788  0.0119  0.4719  3.5976
##
## Random effects:
##  Groups           Name        Variance Std.Dev.
##  childid.schoolid (Intercept) 745.36   27.301
##  schoolid         (Intercept) 315.70   17.768
##  schoolid.1       year         85.97    9.272
##  Residual                     554.92   23.557
## Number of obs: 2162, groups:  childid:schoolid, 1081; schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  465.257      2.108 108.630  220.73   <2e-16 ***
## year          57.751      1.402 101.341   41.18   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## year -0.188
```

a. Generate the BLUPs for the random effects and examine whether the independence between zeta_0 and zeta_1 is reflected in a scatterplot of these two sets of effects.

```
ranefs.fit.M1 <- ranef(fit.M1)
delta.fit.M1 <- ranefs.fit.M1$`childid:schoolid`[,1]
zeta0.fit.M1 <- ranefs.fit.M1$schoolid[,1]
zeta1.fit.M1 <- ranefs.fit.M1$schoolid[,2]

# Scatterplot of zeta_0 vs. zeta_1
plot(zeta0.fit.M1, zeta1.fit.M1,
     main="Zeta0 vs. Zeta1", xlab="zeta1", ylab="zeta0", pch=19)
```




```r
# Correlation between zeta_0 and zeta_1
cor(zeta0.fit.M1, zeta1.fit.M1)
```

```
## [1] -0.1526661
```

Response: Yes, the independence of zeta_0 and zeta_1 is reflected in the scatterplot because there does not seem to be a strong correlation between the two (-0.11).

b. Compute V_S(year = 0) and V_S (year = 1). Since there are only two years, this is a form of heteroscedasticity in the random effects.

```r
V_S_year0 = 324.79
V_S_year0
```

```
## [1] 324.79
```

```r
V_S_year1 = 324.79 + 88.67
V_S_year1
```

```
## [1] 413.46
```

i. In which year is there more between school variation, net of all else?

Response: There is more between school variation in year 1.

Question 11

If you ran the model BY YEAR, and removed the year trend from the model, would you get the same estimates for the variances between schools?

```
fit.M_K <- lmer(mathkind ~ 1 + (1 | schoolid), data = dat)
print(summary(fit.M_K))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: mathkind ~ 1 + (1 | schoolid)
##    Data: dat
##
## REML criterion at convergence: 11017.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7416 -0.5655  0.0086  0.6286  3.5648
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  schoolid (Intercept)  364.1   19.08
##  Residual             1391.4   37.30
## Number of obs: 1081, groups:  schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  465.382      2.244 101.588   207.4   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fit.M_1 <- lmer(math1st ~ 1 + (1 | schoolid), data = dat)
print(summary(fit.M_1))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math1st ~ 1 + (1 | schoolid)
##    Data: dat
##
## REML criterion at convergence: 10851.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5824 -0.6158 -0.0063  0.6191  3.8063
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  schoolid (Intercept)  279.1   16.71
##  Residual             1202.6   34.68
## Number of obs: 1081, groups:  schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)
```

```
## (Intercept)    523.2        2.0 101.9   261.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Response: No, you get different estimates for the variances between schools.

Question 12

Rerun the last nested longitudinal model, allowing correlation between intercept and slope.

a. Is the correlation significant?

```
fit.M2 <- lmer(math ~ year + (year | schoolid) + (1 | schoolid:childid), data = class_pp)
summary(fit.M2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: math ~ year + (year | schoolid) + (1 | schoolid:childid)
##    Data: class_pp
##
## REML criterion at convergence: 21391.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6737 -0.4699  0.0038  0.4683  3.4882
##
## Random effects:
##  Groups           Name        Variance Std.Dev. Corr
##  schoolid:childid (Intercept) 749.0    27.37
##  schoolid         (Intercept) 373.5    19.33
##                   year        112.4    10.60    -0.53
##  Residual                     547.8    23.41
## Number of obs: 2162, groups:  schoolid:childid, 1081; schoolid, 105
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  465.257      2.241 101.262   207.6   <2e-16 ***
## year          58.006      1.491  95.408    38.9   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## year -0.486
```

```
# LRT
anova(fit.M2, fit.M1, refit = F)
```

```
## Data: class_pp
## Models:
## fit.M1: math ~ year + (0 + year | schoolid) + (1 | schoolid/childid)
## fit.M2: math ~ year + (year | schoolid) + (1 | schoolid:childid)
##        Df   AIC   BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## fit.M1  6 21415 21449 -10702    21403
## fit.M2  7 21405 21445 -10696    21391 11.879      1  0.0005678 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Response: Yes, the LRT suggests that the correlation is significant.

b. Compute V_S (year = 0) and V_S(year = 1) for this new model (your formula should include covariance terms).

```
V_S_year0_new = 370.6
V_S_year0_new
```

```
## [1] 370.6
```

```
V_S_year1_new = 370.6 + 2*(-0.45)*10.44*19.25 + 109.1
V_S_year1_new
```

```
## [1] 298.827
```

i. Is this result (and thus model) more consistent with the separate grade analysis? You are implicitly testing model fit here.

Response: Yes, this result (and model) is more consistent with the separate grade analysis.