Linear Mixed Effect Models

Practical Using R

Kamarul Imran Musa   Afiq Amsyar

2025-09-20

Table of contents

1 Linear mixed effect model for continuous outcome data: hands-on

These practical is largely based on the R practical notes that I took and adapted from the Centre for Multilevel Modelling at the Bristol University

1.1 Dataset

We will use the Scottish School Leavers Survey dataset.

This dataset was provided to Bristol University by Linda Croxford (Centre for Educational Sociology, University of Edinburgh).

These data have seven cohorts of young people which makes this dataset suitable for even a longitudinal data analysis.

Students (`caseid`) are nested within school (`schoolid`). This means:

- Level 1 is students (`caseid`)
- Level 2 is schools (`schoolid`)

1.2 Variables

The dependent variable is total attainment score (`score`). And the scores range from 0 (min) to 75 (max).

The explanatory variables are:

1. `cohort90`: 1984, 1986, 1988, 1990, 1996, 1998.
2. `female`: Sex of student (1 = female, 0 = male)
3. `sclass`: Social class, defined as the higher class of mother or father (1 = managerial and professional, 2 = intermediate, 3 = working, 4 = unclassified)
4. `schtype`: School type, distinguishing independent schools from state-funded schools (1 = independent, 0 = state-funded)
5. `schurban`: Urban-rural classification of school (1 = urban, 0 = town or rural)
6. `schdenom`: School denomination (1 = Roman Catholic, 0 = non-denominational)

The cohort90 variable is calculated by subtracting 1990 from each value. Thus values range from -6 (corresponding to 1984) to 8 (1998), with 1990 coded as zero.

1.3 Steps

1. Load relevant R packages
2. Read data
3. Do data wrangling
4. Explore data
5. Perform multilevel models
   - random intercept
   - random slope

1.4 Load packages

- General packages

```
library(haven)
library(tidyverse)
library(broom.mixed)
library(here)
library(gtsummary)
library(DT)
library(kableExtra)
```

- R packages specific to run multilevel analysis
  - `lmerTest` : to get p-value estimations that are not part of the standard `lme4` packages

```
library(lme4)
library(lmerTest)
```

1.5 Read data

We will read data

- a file named `score_lme.dta` inside the folder

```
score_d <- read_dta('score_lme.dta')
glimpse(score_d)
```

```
Rows: 33,988
Columns: 9
$ caseid   <dbl> 18, 17, 19, 20, 21, 13, 16, 14, 15, 12, 12865, 6509, 12866, 1~
$ schoolid <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
$ score    <dbl> 0, 10, 0, 40, 42, 4, 0, 0, 14, 27, 18, 23, 24, 0, 25, 4, 11, ~
$ cohort90 <dbl> -6, -6, -6, -6, -6, -6, -6, -6, -6, -6, -2, -4, -2, -2, -4, -~
$ female   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0~
$ sclass   <dbl> 2, 2, 4, 3, 2, 2, 3, 4, 3, 2, 2, 1, 2, 3, 2, 3, 2, 4, 3, 3, 3~
$ schtype  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
$ schurban <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
$ schdenom <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
```

1.6 Data wrangling

Convert numerical variables to factor variables

```r
score_d <-
  score_d %>%
  mutate(female2 = factor(female,
                          labels = c('male', 'female')),
         class2 = factor(sclass,
                         labels = c('managerial+prof', 'intermediate',
                                    'working', 'unclassified')),
         schtype2 = factor(schtype,
                           labels = c('state-funded', 'independent')),
         schurban2 = factor(schurban,
                            labels = c('town/rural', 'urban')),
         schdenom2 = factor(schdenom,
                            labels = c('state-funded', 'independent'))) %>%
  select(-c(female, sclass, schtype, schurban, schdenom))
```

1.7 Explore data (EDA)

- Summarize data

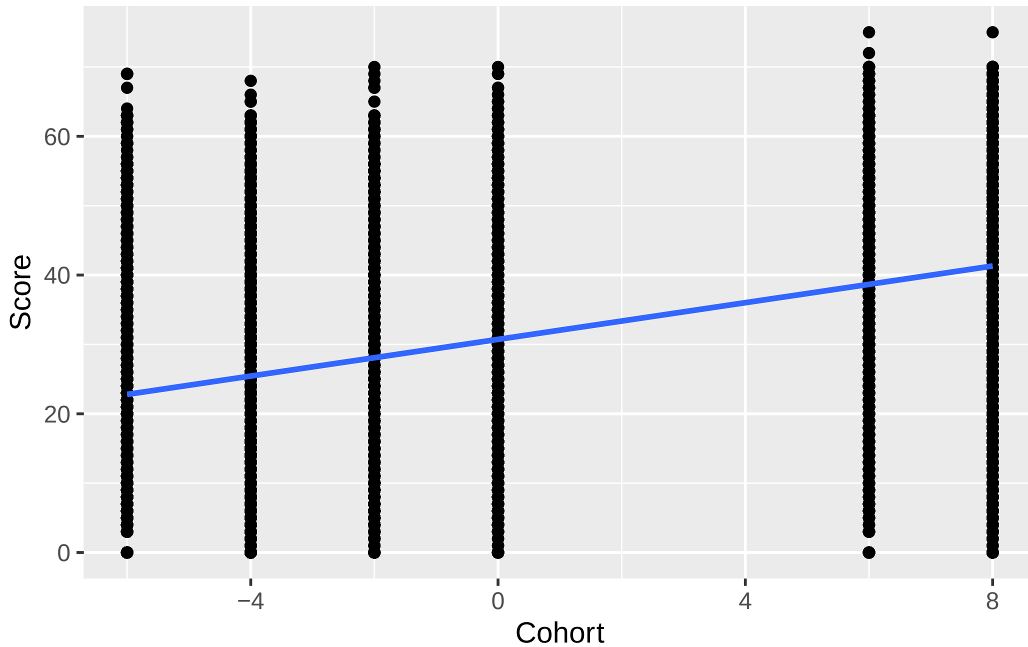
```r
score_d %>%
  select(-caseid, -schoolid) %>%
  tbl_summary(by = female2) %>%
  as_gt()
```

Characteristic	**male** N = 16,055[1]	**female** N = 17,933[1]
Score	32 (17, 43)	34 (20, 46)
Cohort		
-6	3,272 (20%)	3,206 (18%)
-4	3,016 (19%)	3,309 (18%)
-2	2,511 (16%)	2,734 (15%)
0	2,016 (13%)	2,355 (13%)
6	1,965 (12%)	2,279 (13%)
8	3,275 (20%)	4,050 (23%)
class2		
managerial+prof	5,326 (33%)	5,847 (33%)
intermediate	4,744 (30%)	5,250 (29%)
working	4,379 (27%)	5,107 (28%)
unclassified	1,606 (10%)	1,729 (9.6%)
schtype2		
state-funded	15,265 (95%)	17,183 (96%)
independent	790 (4.9%)	750 (4.2%)
schurban2		
town/rural	4,691 (29%)	5,181 (29%)
urban	11,364 (71%)	12,752 (71%)
schdenom2		
state-funded	13,573 (85%)	15,057 (84%)
independent	2,482 (15%)	2,876 (16%)

[1]Median (Q1, Q3); n (%)

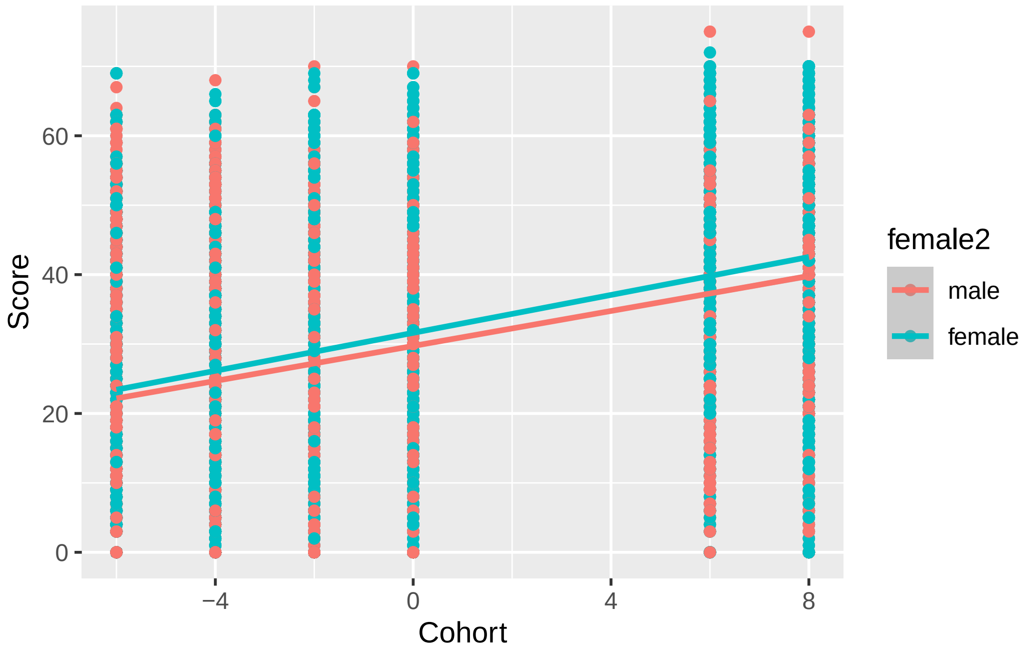
- Plot variable `score`

```
score_d %>%
  ggplot(aes(x = cohort90, y = score)) +
  geom_point() +
  geom_smooth(method = lm)
```



- Plot score based on `female2`

```
score_d %>%
  ggplot(aes(x = cohort90, y = score,
             col = female2, group = female2)) +
  geom_point() +
  geom_smooth(method = lm)
```



1.8 Random intercept model

We will use **lme4** package.

We will start with the simplest model. It is a constant only model or also known as the null model. There will be no explanatory variables. In the argument, we will set the estimation using maximum likelihood estimates (MLE).

The null model which we will name as `m0`. For this one, we will set the random effect comes from the schools.

This is basically a random intercept with constant-only model

```
m0 <-
  lmer(score ~ 1 + (1 | schoolid),
       data = score_d, REML = FALSE)
summary(m0)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ 1 + (1 | schoolid)
   Data: score_d
```

```
      AIC       BIC    logLik -2*log(L)  df.resid 
 286545.1  286570.4 -143269.5  286539.1     33985 

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-2.9763 -0.7010  0.1017  0.7391  3.0817 

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolid (Intercept)  61.02    7.812  
 Residual             258.36   16.073  
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
            Estimate Std. Error       df t value Pr(>|t|)    
(Intercept)  30.6006     0.3694 451.5326   82.83   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The overall score mean attainment (across schools) is estimated as 30.6. The mean score for each schools j is estimated as $30.60 + \hat{U}_{0j}$ where $\hat{U}_{0j}$ is the school residuals (Level-2 residuals)

The intra-class correlation (ICC) is $\frac{61.02}{61.02+258.4}$ or 19 percent:

```
61.02/(61.02+258.4)
```

```
[1] 0.1910337
```

It seems the cluster effect is rather big (19 percent).

You may use `tidy()` to make a more pleasant table

```
tidy(m0)
```

```
# A tibble: 3 x 8
  effect   group    term            estimate std.error statistic    df    p.value
  <chr>    <chr>    <chr>              <dbl>     <dbl>     <dbl> <dbl>      <dbl>
1 fixed    <NA>     (Intercept)         30.6     0.369      82.8  452.  3.55e-275
2 ran_pars schoolid sd__(Intercep~       7.81   NA          NA      NA  NA       
3 ran_pars Residual sd__Observati~      16.1    NA          NA      NA  NA       
```

The log-likelihood is evaluated using adaptive Gauss-Hermite approximation. If the value for it is 1, then it reduces to Laplacian approximation.

This default approximation can be changed by using the `nAGQ = n` option, where `n` is an integer greater than `one`, representing the number of points used for evaluating the adaptive Gauss-Hermite approximation. The greater the value of `n`, the more accurate the evaluation of the log-likelihood, but the longer it takes to fit the model.

There are 2 variances, level-1 variance (students) = 258.4 and level-2 variance (schools) = 61.0

Hence, $\frac{61.0}{61.0+258.4} = 0.19$ or 19% of the variance in score attainment can be attributed to the differences between schools.

Let's check the random effect:

```
head(ranef(m0)$schoolid)
```

```
  (Intercept)
1  -11.841271
2    3.206333
3    3.396003
4   -7.415008
5    3.426227
6   12.433731
```

So, for the 1st school or `schooid` number l, the baseline score is `30.6 - 11.84` equals 18.76

1.9 More on random intercept model

We will model the effect of a student-level variable **cohort** in the model. Which means we expect there are variations in `score` at baseline between schools. But we do not expect variations in the change in `score` between schools over time.

$$score_{ij} = \beta_{00} + \beta_1 cohort90_{ij} + u_{01j} + e_{ij}$$

```
ri <- lmer(score ~ cohort90 + (1 | schoolid),
           data = score_d,
           REML = FALSE)
summary(ri)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + (1 | schoolid)
   Data: score_d

      AIC       BIC    logLik -2*log(L)  df.resid
 280921.6  280955.3 -140456.8  280913.6     33984

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1487 -0.7242  0.0363  0.7339  3.7097

Random effects:
 Groups   Name        Variance Std.Dev.
 schoolid (Intercept)  45.99    6.781
 Residual             219.29   14.808
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept) 3.056e+01  3.225e-01 4.326e+02   94.74   <2e-16 ***
cohort90    1.215e+00  1.553e-02 3.392e+04   78.24   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
         (Intr)
cohort90 -0.002
```

Or

```r
tidy(ri, conf.int = TRUE)
```

```
# A tibble: 4 x 10
  effect   group    term  estimate std.error statistic     df    p.value conf.low
  <chr>    <chr>    <chr>    <dbl>     <dbl>     <dbl>  <dbl>      <dbl>    <dbl>
1 fixed    <NA>     (Int~    30.6     0.323       94.7   433.  2.09e-291    29.9
2 fixed    <NA>     coho~     1.21    0.0155      78.2 33925.  0             1.18
3 ran_pars school~ sd__~     6.78   NA            NA      NA  NA            NA
4 ran_pars Residu~ sd__~    14.8    NA            NA      NA  NA            NA
# i 1 more variable: conf.high <dbl>
```

1.9.1 Random effect

These are the random effect for the first 10 schools.

```
rand_ef <- ranef(ri)
head(rand_ef$schoolid, 10)
```

```
   (Intercept)
1    -6.728310
2     2.789899
3     2.034756
4    -7.631468
5     3.074107
6    11.833121
7    -1.604856
8    17.792799
9    -7.995792
10    2.778860
```

Or you may Use **broom.mixed::augment()** function to get the fitted values. You will see that the when you take the fixed value and sum with the random effect, you will get the fitted values.

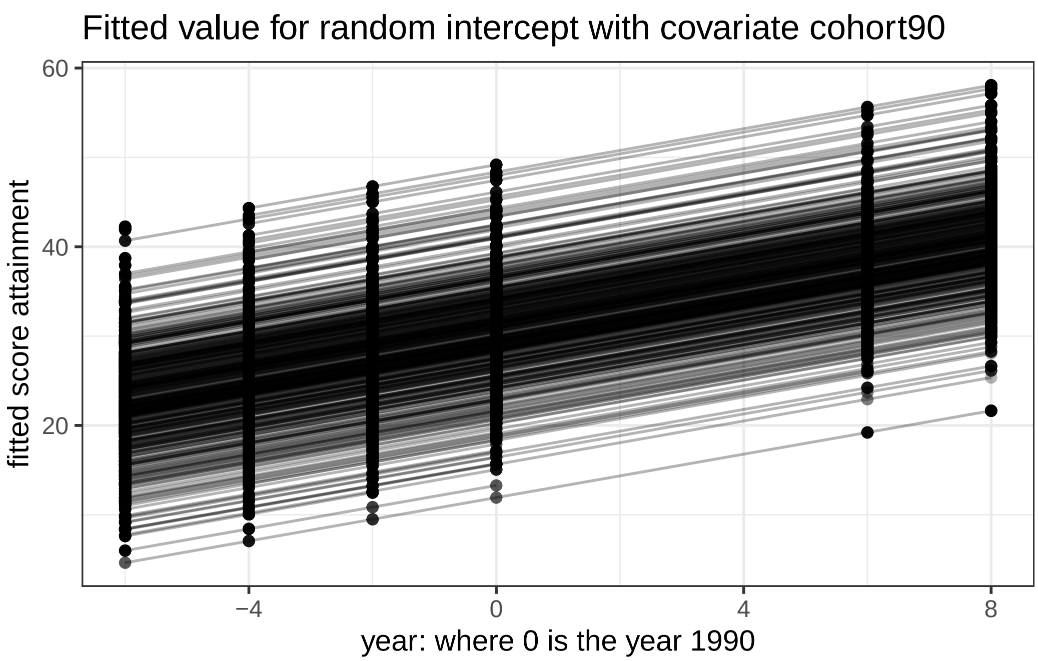
```
ri_fitted <- augment(ri)
ri_fitted %>%
  select(score, schoolid, cohort90, .fitted, .fixed) %>%
  slice(1:2, 60:62)
```

```
# A tibble: 5 x 5
  score schoolid cohort90 .fitted .fixed
  <dbl>    <dbl>    <dbl>   <dbl>  <dbl>
1     0        1       -6    16.5   23.3
2    10        1       -6    16.5   23.3
3    59        2        8    43.1   40.3
4    48        2        6    40.6   37.8
5    31        2        8    43.1   40.3
```

1.9.2 Plot random intercept

Careful examination of the top left hand corner of the graph shows that a small number of schools are observed for only one cohort.

```r
ggplot(ri_fitted, aes(cohort90, .fitted, group = schoolid )) +
  geom_point(alpha = 0.3) +
  geom_line(alpha = 0.3) +
  ylab('fitted score attainment') +
  xlab('year: where 0 is the year 1990') +
  ggtitle('Fitted value for random intercept with covariate cohort90') +
  theme_bw()
```



1.10 Random slope model

This will allow different slopes. Which means the change of scores over time is different between schools.

In our example, for random intercept model, we allowed for school effect on the mean score by allowing the intercept of the regression of attainment on cohort to vary randomly across schools. We assumed, however, that cohort changes in attainment are the same for all schools, i.e. the slope of the regression line was assumed fixed across schools.

We will now extend the random intercept model fitted before to allow both the intercept and the slope to vary randomly across schools.

1.10.1 Random effect

$$score_{ij} = \beta_{00} + \beta_{10}cohort90_{ij} + u_{01j} + u_{11j}cohort90_{ij} + e_{ij}$$

```r
rs <- lmer(score ~ cohort90 + (1 + cohort90 | schoolid),
           data = score_d, REML = FALSE)
summary(rs)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + (1 + cohort90 | schoolid)
   Data: score_d

      AIC       BIC    logLik -2*log(L)  df.resid
 280698.2  280748.8 -140343.1  280686.2     33982

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1008 -0.7202  0.0387  0.7264  3.5220

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolid (Intercept)  42.8573  6.5465
          cohort90      0.1606  0.4008  -0.39
 Residual             215.7393 14.6881
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept) 30.60967    0.31344 426.77096   97.66   <2e-16 ***
cohort90     1.23391    0.02532 316.40136   48.74   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
         (Intr)
cohort90 -0.266
optimizer (nloptwrap) convergence code: 0 (OK)
Model failed to converge with max|grad| = 0.00542676 (tol = 0.002, component 1)
```

However, our model failed to converge, so we may switch to `bobyqa` optimizer

```r
rs <- lmer(score ~ cohort90 + (1 + cohort90 | schoolid), data = score_d,
           control = lmerControl(optimizer = 'bobyqa'),
           REML = FALSE)
summary(rs)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + (1 + cohort90 | schoolid)
   Data: score_d
Control: lmerControl(optimizer = "bobyqa")

      AIC       BIC    logLik -2*log(L)  df.resid
 280698.2  280748.8 -140343.1  280686.2     33982

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1008 -0.7202  0.0387  0.7264  3.5220

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolid (Intercept)  42.8582  6.5466
          cohort90      0.1606  0.4007  -0.39
 Residual             215.7393 14.6881
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept)  30.60963    0.31345 426.76273   97.66   <2e-16 ***
cohort90      1.23390    0.02531 316.45087   48.74   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
         (Intr)
cohort90 -0.266
```

We can get a nicer output with `tidy`:

```r
tidy(rs)
```

```
# A tibble: 6 x 8
```

```
  effect   group    term            estimate std.error statistic    df    p.value
  <chr>    <chr>    <chr>              <dbl>     <dbl>     <dbl> <dbl>      <dbl>
1 fixed    <NA>     (Intercept)       30.6      0.313       97.7  427.  4.41e-294
2 fixed    <NA>     cohort90           1.23     0.0253      48.7  316.  3.58e-149
3 ran_pars schoolid sd__(Intercep~     6.55    NA           NA      NA  NA
4 ran_pars schoolid cor__(Interce~    -0.390   NA           NA      NA  NA
5 ran_pars schoolid sd__cohort90       0.401   NA           NA      NA  NA
6 ran_pars Residual sd__Observati~    14.7     NA           NA      NA  NA
```

These are the random effects:

```r
rand_ef_s <- ranef(rs)
head(rand_ef_s$schoolid, 10)
```

```
   (Intercept)     cohort90
1    -5.904644  0.198019882
2     2.683568  0.048683788
3     1.727237  0.149958811
4    -7.802840  0.304729063
5     3.295992 -0.453988736
6    12.116171 -0.580155542
7    -1.640374  0.008941222
8    18.047796 -0.219314405
9    -7.708394  0.184207525
10    2.764885  0.202445981
```

The fitted (average) score attainment based on the random slope model is:

```r
rs_res <- augment(rs)
rs_res %>%
  select(score, schoolid, cohort90, .fitted, .fixed) %>%
  slice(1:2, 60:62)
```

```
# A tibble: 5 x 5
  score schoolid cohort90 .fitted .fixed
  <dbl>    <dbl>    <dbl>   <dbl>  <dbl>
1     0        1       -6    16.1   23.2
2    10        1       -6    16.1   23.2
3    59        2        8    43.6   40.5
4    48        2        6    41.0   38.0
5    31        2        8    43.6   40.5
```

The intercept-slope correlation is estimated as -0.39 which means means that schools with a high intercept (above-average attainment in 1990) tend to have a flatter-than-average slope.

1.11 Comparing models between random intercept and random slope

```r
anova(ri, rs)
```

```
Data: score_d
Models:
ri: score ~ cohort90 + (1 | schoolid)
rs: score ~ cohort90 + (1 + cohort90 | schoolid)
   npar    AIC    BIC  logLik -2*log(L) Chisq Df Pr(>Chisq)
ri    4 280922 280955 -140457    280914
rs    6 280698 280749 -140343    280686 227.4  2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

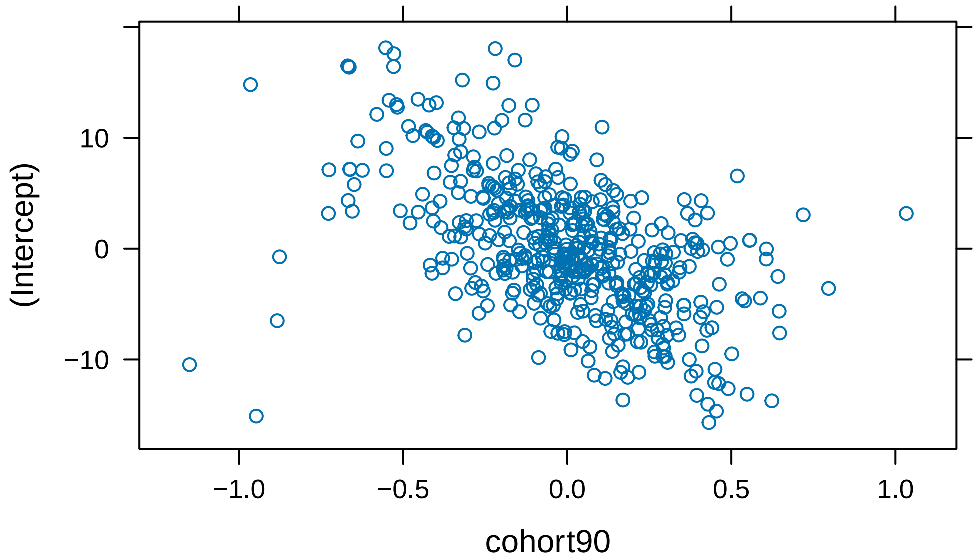
There is very strong evidence that the cohort effect differs across schools. So we prefer random slope model over random intercept.

1.12 Plot of random effects

School slope vs school intercept u_{0j} and u_{1j}

```r
plot(rand_ef_s)
```

```
$schoolid
```

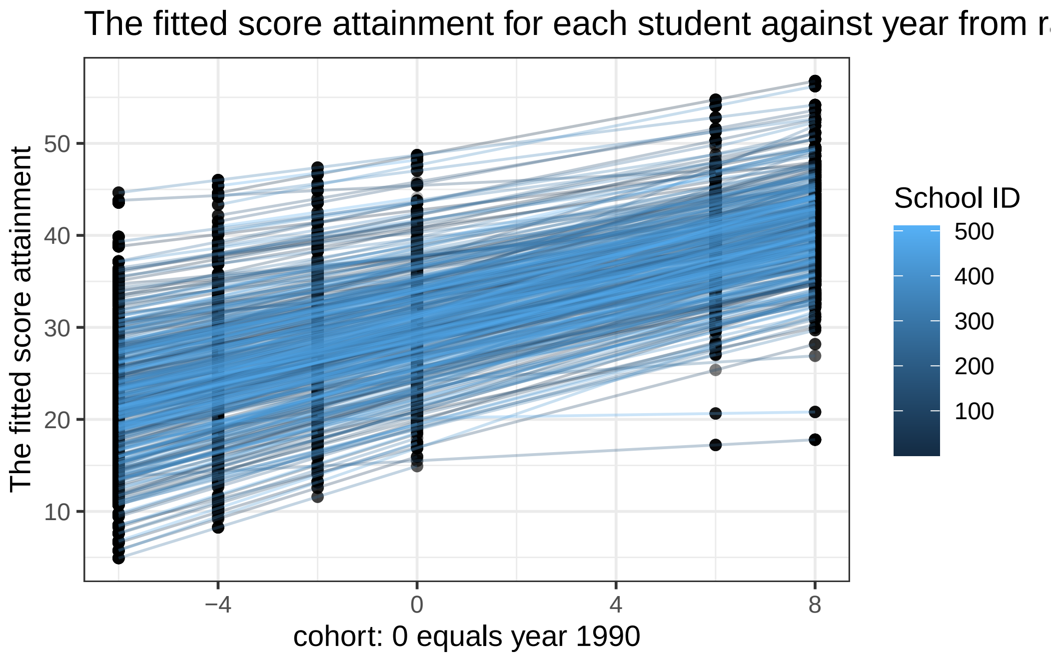


1.12.1 Model

$$\widehat{score}_{ij} = (30.610 + \hat{u_{0j}}) + (1.234 + \hat{u_{1j}})cohort90_{ij}$$

1.12.2 Plot the fitted values from random slope

```
rs_res %>%
  ggplot(aes(cohort90, .fitted, group = schoolid)) +
  geom_point(alpha = 0.3) +
  geom_line(aes(colour = schoolid), alpha = 0.3) +
  ylab('The fitted score attainment') +
  xlab('cohort: 0 equals year 1990') +
  theme_bw() +
  ggtitle('The fitted score attainment for each student against year from random slope model
```



1.13 Adding a level-1 variable to the random slope model

- Random slope for gender

We start off with a model that assumes gender has a fixed effect

$$score_{ij} = \beta_0 + \beta_1 cohort90_{ij} + \beta_2 female_{ij} + u_{0j} + u_{1j} cohort90_{ij} + e_{ij}$$

```
rs_gend <- lmer(score ~ cohort90 + female2 +
                    (1 + cohort90 | schoolid),
                  data = score_d, REML = FALSE)
summary(rs_gend)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + female2 + (1 + cohort90 | schoolid)
   Data: score_d

     AIC       BIC    logLik -2*log(L)  df.resid
 280558.1  280617.2 -140272.1  280544.1     33981
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.1617 -0.7185  0.0395  0.7282  3.5326

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 schoolid (Intercept)  42.5750  6.5250
          cohort90      0.1613  0.4016  -0.39
 Residual             214.8370 14.6573
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
               Estimate Std. Error        df t value Pr(>|t|)
(Intercept)   2.958e+01  3.241e-01 4.944e+02   91.30   <2e-16 ***
cohort90      1.227e+00  2.533e-02 3.168e+02   48.46   <2e-16 ***
female2female 1.945e+00  1.630e-01 3.364e+04   11.93   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) chrt90
cohort90    -0.253
female2feml -0.265 -0.022
optimizer (nloptwrap) convergence code: 0 (OK)
Model failed to converge with max|grad| = 0.00245207 (tol = 0.002, component 1)
```

Will use bobyqa optimizer because our model did not converge

```r
rs_gend <- lmer(score ~ cohort90 + female2 +
                  (1 + cohort90 | schoolid),
                data = score_d, REML = FALSE,
                lmerControl(optimizer = 'bobyqa'))
summary(rs_gend)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + female2 + (1 + cohort90 | schoolid)
   Data: score_d
Control: lmerControl(optimizer = "bobyqa")

      AIC       BIC    logLik -2*log(L)  df.resid
 280558.1  280617.2 -140272.1  280544.1     33981
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max 
-3.1617 -0.7185  0.0395  0.7282  3.5326 

Random effects:
 Groups   Name        Variance Std.Dev. Corr 
 schoolid (Intercept)  42.5750  6.5249       
          cohort90      0.1613  0.4016  -0.39
 Residual             214.8374 14.6573       
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
               Estimate Std. Error        df t value Pr(>|t|)    
(Intercept)   2.958e+01  3.241e-01 4.944e+02   91.30   <2e-16 ***
cohort90      1.227e+00  2.533e-02 3.169e+02   48.46   <2e-16 ***
female2female 1.945e+00  1.630e-01 3.364e+04   11.93   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) chrt90
cohort90    -0.253       
female2feml -0.265 -0.022
```

Now, we will consider a model with gender is assumed to have random slope

$$score_{ij} = \beta_0 + \beta_1 cohort90_{ij} + \beta_2 female_{ij} + u_{0j} + u_{1j} cohort90_{ij} + u_{2j} female + e_{ij}$$

```r
rs_gend_sl <- lmer(score ~ cohort90 + female2 +
                     (1 + cohort90 + female2 | schoolid),
                data = score_d, REML = FALSE)
summary(rs_gend_sl)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + female2 + (1 + cohort90 + female2 | schoolid)
   Data: score_d

      AIC       BIC    logLik -2*log(L)  df.resid 
 280558.9  280643.2 -140269.4  280538.9     33978 
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max 
-3.1571 -0.7182  0.0388  0.7268  3.5317 

Random effects:
 Groups   Name          Variance Std.Dev. Corr       
 schoolid (Intercept)    40.5649  6.3691             
          cohort90        0.1617  0.4022  -0.39      
          female2female   1.3767  1.1733   0.21 -0.11
 Residual               214.5140 14.6463             
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
               Estimate Std. Error        df t value Pr(>|t|)    
(Intercept)    29.58915    0.31768 406.81170   93.14   <2e-16 ***
cohort90        1.22777    0.02535 315.98006   48.44   <2e-16 ***
female2female   1.93141    0.17395 345.48965   11.10   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) chrt90
cohort90    -0.253       
female2feml -0.201 -0.046
optimizer (nloptwrap) convergence code: 0 (OK)
Model failed to converge with max|grad| = 0.00925414 (tol = 0.002, component 1)
```

Again, we will be using bobyqa optimizer because our default Gaus-Hermite optimizer could not converge

```r
rs_gend_sl <- lmer(score ~ cohort90 + female2 +
                     (1 + cohort90 + female2 | schoolid),
                data = score_d, REML = FALSE,
                lmerControl(optimizer = 'bobyqa'))
summary(rs_gend_sl)
```

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
  method [lmerModLmerTest]
Formula: score ~ cohort90 + female2 + (1 + cohort90 + female2 | schoolid)
   Data: score_d
Control: lmerControl(optimizer = "bobyqa")
```

```
      AIC       BIC    logLik -2*log(L)  df.resid 
 280558.9  280643.2 -140269.4  280538.9     33978 

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-3.1572 -0.7182  0.0388  0.7267  3.5316 

Random effects:
 Groups   Name          Variance Std.Dev. Corr       
 schoolid (Intercept)    40.5580  6.3685             
          cohort90        0.1617  0.4021  -0.39      
          female2female   1.3711  1.1710   0.21 -0.11
 Residual               214.5158 14.6464             
Number of obs: 33988, groups:  schoolid, 508

Fixed effects:
               Estimate Std. Error        df t value Pr(>|t|)    
(Intercept)    29.58909    0.31766 406.88434   93.15   <2e-16 ***
cohort90        1.22777    0.02535 316.04102   48.44   <2e-16 ***
female2female   1.93145    0.17390 345.08083   11.11   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) chrt90
cohort90    -0.253       
female2feml -0.201 -0.046
```

Comparison between random slope mode for cohort and random slope model for cohort and female. Does a random slope model with both female and cohort90 differ from a random slope model with only cohort90?

```r
anova(rs_gend, rs_gend_sl)
```

```
Data: score_d
Models:
rs_gend: score ~ cohort90 + female2 + (1 + cohort90 | schoolid)
rs_gend_sl: score ~ cohort90 + female2 + (1 + cohort90 + female2 | schoolid)
           npar    AIC    BIC  logLik -2*log(L)  Chisq Df Pr(>Chisq)
rs_gend       7 280558 280617 -140272    280544                     
rs_gend_sl   10 280559 280643 -140269    280539 5.2362  3     0.1553
```

we conclude that the gender effect is the same for each school. We therefore revert to a model with a fixed coefficient for female. ## Adding a level-2 explanatory variable to the random slope model

Let us assume that although we found evidence that the effect of social class on attainment differs across schools, we will work with a simpler model by removing the random coefficients on the class dummy variables. So social class comes in as a fixed effcet in the model.

```
rs_gend_class <-
  lmer(score ~ cohort90 + female2 + class2 +
         (1 + cohort90 | schoolid), data = score_d,
       REML = FALSE, lmerControl(optimizer = 'bobyqa'))
tidy(rs_gend_class)
```

```
# A tibble: 10 x 8
   effect   group    term          estimate std.error statistic     df    p.value
   <chr>    <chr>    <chr>            <dbl>     <dbl>     <dbl>  <dbl>      <dbl>
 1 fixed    <NA>     (Intercept)     35.7       0.274     130.     709.  0
 2 fixed    <NA>     cohort90         1.18      0.0243     48.6    322.  2.21e-150
 3 fixed    <NA>     female2fema~     1.96      0.154      12.7  33710.  6.11e- 37
 4 fixed    <NA>     class2inter~    -5.21      0.197     -26.5  33720.  1.25e-152
 5 fixed    <NA>     class2worki~   -11.1       0.206     -53.7  33816.  0
 6 fixed    <NA>     class2uncla~   -14.8       0.286     -51.9  33838.  0
 7 ran_pars schoolid sd__(Interc~     4.74     NA          NA       NA  NA
 8 ran_pars schoolid cor__(Inter~    -0.317    NA          NA       NA  NA
 9 ran_pars schoolid sd__cohort90     0.388    NA          NA       NA  NA
10 ran_pars Residual sd__Observa~    13.9      NA          NA       NA  NA
```

1.14 Summary

Models that can be used in multilevel analysis:

- Observations are not independent
- Clustering effects reduces the standard error
- Random intercept
- Random slope