Contextual Effects

Kalyango Jovan

25/02/2022

Library:

```
#MLM packages
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
```

CONTEXTUAL ANALYSIS

For contextual analysis, we are presented with the same dataset exam, the analysis is done from the perspective of contextual effect, the file exam.csv holds data from pupils in English Schools. The dependent variable is exam-scores, and the independent variables are reading ability (LRT, London Reading Test) and average reading ability in school (AvsLRT).

```
#load the data
exam <- read.csv(file = "exam.csv", header = TRUE)
head(exam)
```

```
##   ï..School Student Examscore LRTscore  AvsLRT
## 1         1       1   0.26132  0.61906 0.16617
## 2         1       2   0.13407  0.20580 0.16617
## 3         1       3  -1.72390 -1.36460 0.16617
## 4         1       4   0.96759  0.20580 0.16617
## 5         1       5   0.54434  0.37110 0.16617
## 6         1       6   1.73490  2.18940 0.16617
```

1). Model 3:

The model that includes a random intercept, the predictors at level 1 (in this case,LRT - reading ability per student), and the predictors at level 2 (in this case, AvsLRT - the average reading ability in the school). Include the predictor LRT using grand mean centering.

```
exam$lrt_gmc <- exam$LRTscore - mean(exam$LRTscore) #grand mean centering
#model with fixed predictors for both levels
model_3_gmc <- lmer(Examscore~1 + lrt_gmc + AvsLRT + (1|ï..School), REML = FALSE, data = exam)
summary(model_3_gmc)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: Examscore ~ 1 + lrt_gmc + AvsLRT + (1 | ï..School)
##    Data: exam
##
##      AIC      BIC   logLik deviance df.resid
##   9357.6   9389.1  -4673.8   9347.6     4054
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7156 -0.6316  0.0273  0.6851  3.2603
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  ï..School (Intercept) 0.07606  0.2758
##  Residual              0.56591  0.7523
## Number of obs: 4059, groups:  ï..School, 65
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) 1.307e-02  3.687e-02 6.015e+01   0.355  0.72417
## lrt_gmc     5.595e-01  1.254e-02 3.991e+03  44.628  < 2e-16 ***
## AvsLRT      3.583e-01  1.103e-01 6.498e+01   3.250  0.00183 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) lrt_gm
## lrt_gmc  0.001
## AvsLRT   0.075 -0.114
```

a) What does the regression coefficient for LRT represent in terms of contextual effects?

The regression coefficient for LRT represents the **within school effect** of individual reading ability on the exam score.So with every point higher a student has on individual reading scores compared to her/his peers, the student scores 0.56 points higher on the exam score.

b) What does the regression coefficient for AvsLRT represent in terms of contextual effects?

As the **individual level predictor** is added grand mean centered, the regression coefficient for AvsLRT represents the **difference between the within school effect** of individual reading ability on the exam score and the between school effect of average reading ability on the exam score. So on **top of the within effect** of the individual/student reading score on exam, if a stuent is within a school that on average scores 1 point higher on reading score, the student on average gains an added **0.36** points on the exam score.

2) Calculate the predicted exam score for 2 children, child X and Y

- Child X has score -2, so he/she has a score below average
- Child Y has score 2, so he/she has a score above average
- School A has an average of -1, so the school is below average
- School B has an average of 1, so the school is above average

What are the predicted exam scores for the two children under the two scenarios?

estimated exam scores: intercept + y10 * Childscore + y01 * Schoolaverage = estimate

- child X on school A: 0.01 + 0.56 * -2.00 + 0.36 * -1.00 = -1.46
- child X on school B: 0.01 + 0.56 * -2.00 + 0.36 * 1.00 = -0.75
- child Y on school A: 0.01 + 0.56 * 2.00 + 0.36 * -1.00 = 0.77
- child Y on school B: 0.01 + 0.56 * 2.00 + 0.36 * 1.00 = 1.49

The two students can belong to any of the schools i.e., Two way to Four results

X can go to A or B

```
X_Aestimate <- 0.01 + 0.56*-2 + 0.36*-1
X_Bestimate <- 0.01 + 0.56*-2 + 0.36*1
X_Aestimate
```

```
## [1] -1.47
```

```
X_Bestimate
```

```
## [1] -0.75
```

Y can go to A or B

```
Y_Aestimate <- 0.01 + 0.56*2 + 0.36*-1
Y_Bestimate <- 0.01 + 0.56*2 + 0.36*1
Y_Aestimate
```

```
## [1] 0.77
```

```r
Y_Bestimate
```

```
## [1] 1.49
```

3. Again fit model 3 (the model that includes a random intercept, and the predictors at level 1 and 2), but this time include the predictor LRT using **cluster** mean centering (i.e., within cluster centering).

```r
exam$lrt_groupmean <- ave(exam$LRTscore, exam$ï..School)
exam$lrt_wcc <- exam$LRTscore - exam$lrt_groupmean

model_3_wcc <- lmer(Examscore~1 + lrt_wcc + AvsLRT + (1|ï..School), REML = FALSE, data = exam)
summary(model_3_wcc)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: Examscore ~ 1 + lrt_wcc + AvsLRT + (1 | ï..School)
##    Data: exam
##
##      AIC      BIC   logLik deviance df.resid
##   9357.6   9389.1  -4673.8   9347.6     4054
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7156 -0.6316  0.0273  0.6851  3.2603
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  ï..School (Intercept) 0.07606  0.2758
##  Residual              0.56591  0.7523
## Number of obs: 4059, groups:  ï..School, 65
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept) 1.206e-02  3.687e-02 6.015e+01   0.327    0.745
## lrt_wcc     5.595e-01  1.254e-02 3.991e+03  44.628  < 2e-16 ***
## AvsLRT      9.178e-01  1.095e-01 6.331e+01   8.379 7.51e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) lrt_wc
## lrt_wcc 0.000
## AvsLRT  0.075  0.000
```

a) What does the regression coefficient for LRT represent in terms of contextual effects? Interpret this coefficient.

The LRT regression coefficient represents the **within effects** of the model, i.e. the effect that the LRT reading score of an individual student has on their exam score. One point above the class average results in a higher predicted exam score. In this case, for every point a student scores higher on LRT, on average the predicted exam score of this student increases by 0.56 points.

b) What does the regression coefficient for AvsLRT represent in terms of contextual effects? Interpret this coefficient.

The AvsLRT regression coefficient represents the **between effects** of the model, i.e. the effect that belonging to **different schools** has one the exam score of students. One point higher in average reading score results in a higher predicted (mean school) exam score. In this case, for every point a school scores higher on the AvsLRT, the predicted exam score of the average student in this class increases by 0.92 points.

4. Does the within effect differ significantly from the between effect? Interpret

Yes, the between effect is 0.358 points larger compared to the within effect, which is statistically significant, $t(63) = 3.250, p = 0.0018$.