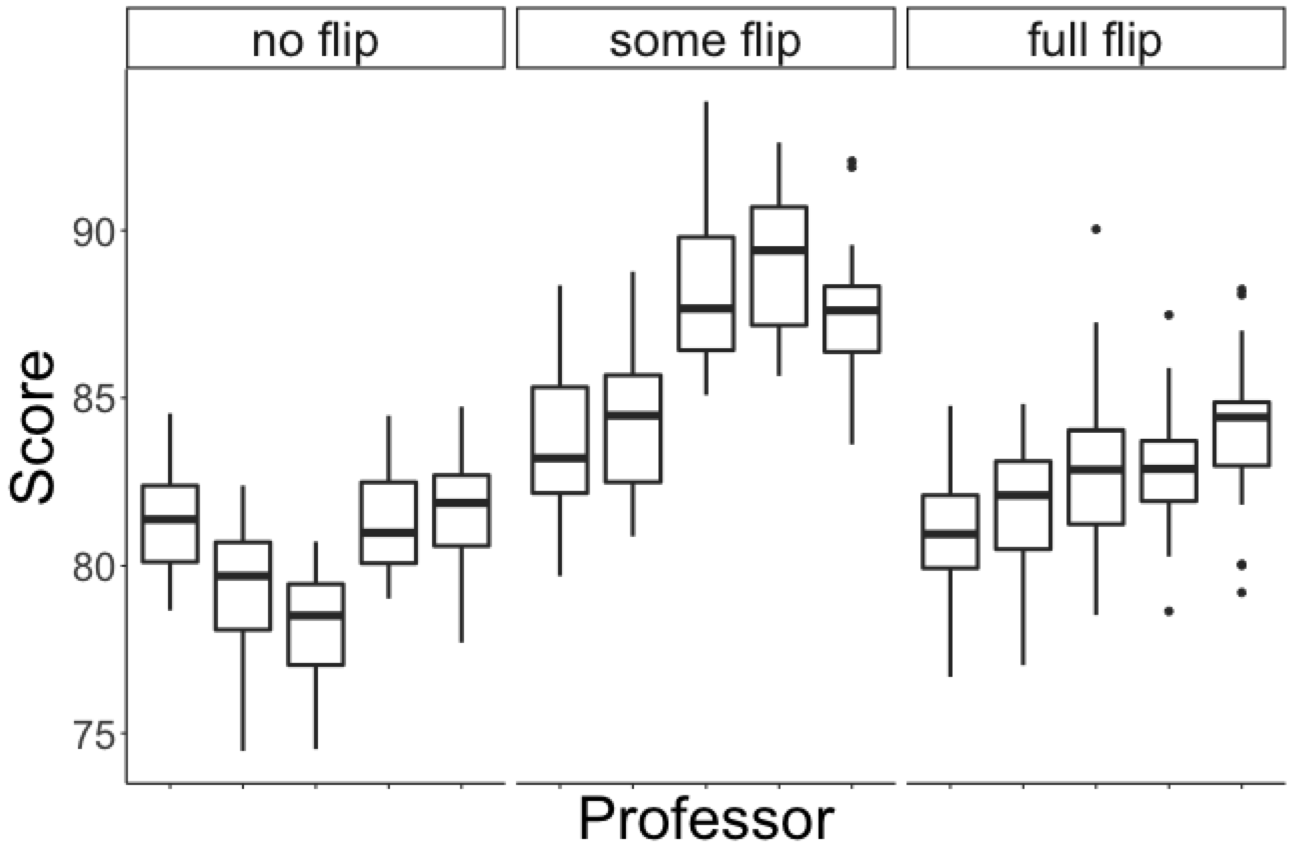
Beginning linear mixed effects models

Data: flipped classrooms?

Data set has 375 rows (one per student), and the following variables:

- `professor`: which professor taught the class (1 -- 15)
- `style`: which teaching style the professor used (no flip, some flip, fully flipped)
- `score`: the student's score on the final exam

Visualizing the data



Mixed effects model

Linear mixed effects model: Let $Score_{ij}$ be the score of student j in class i

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

> Which terms are the fixed effects?

Mixed effects model

Linear mixed effects model: Let $Score_{ij}$ be the score of student j in class i

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

> Which term is the random effect?

Mixed effects model

Linear mixed effects model: Let $Score_{ij}$ be the score of student j in class i

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

> Which term captures variability between students?

Class activity, Part I

https://sta214-f22.github.io/class_activities/ca_lecture_26.html

Class activity

> Why is a mixed effect model useful for this data?

Class activity

What is the population model?

Class activity

> What is the population model?

$$Price_{ij} = \beta_0 + \beta_1 Satisfaction_{ij} + u_i + \varepsilon_{ij}$$

$u_i \overset{iid}{\sim} N(0, \sigma_u^2) \quad \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

where $Price_{ij}$ is the price of rental j in neighborhood i

- u_i is a *random intercept*
- We use subscripts i and j for $Price_{ij}$, $Satisfaction_{ij}$, and ε_{ij} because they are different for every observation in the data
- We only need a subscript i (neighborhood) for u_i, because there is one random intercept per neighborhood

Class activity

$$Price_{ij} = \beta_0 + \beta_1 Satisfaction_{ij} + u_i + \varepsilon_{ij}$$

$u_i \overset{iid}{\sim} N(0, \sigma_u^2) \quad \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

where $Price_{ij}$ is the price of rental j in neighborhood i

> What are the effect of interest, group effect, and individual effect?

Class activity

$$Price_{ij} = \beta_0 + \beta_1 Satisfaction_{ij} + u_i + \varepsilon_{ij}$$

$$u_i \overset{iid}{\sim} N(0, \sigma_u^2) \qquad \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$$

where $Price_{ij}$ is the price of rental j in neighborhood i

> What are the effect of interest, group effect, and individual effect?

- effect of interest: β_1 (slope for relationship between satisfaction and price)
- group effect: u_i (random effect for neighborhood)
- individual effect: ε_{ij} (variation between rentals in a neighborhood)

Fitting mixed effects models

```
library(lme4)
m1 <- lmer(score ~ style + (1|professor),
           data = teaching)
summary(m1)
```

- `lme4` is the R package we will use to fit mixed effects models
- `lmer` is like the `lm` function, but for mixed effects
- `style` is the teaching style (fixed effects)
- `(1|professor)` indicates we have a random intercept (the `1` indicates the intercept) for `professor` (indicated by `|professor`)

Fitting mixed effects models

```r
library(lme4)
```

```r
m1 <- lmer(score ~ style + (1|professor),
           data = teaching)
summary(m1)
```

```
...
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  professor (Intercept) 21.365   4.622
##  Residual               4.252   2.062
...
```

- $\widehat{\sigma}^2_{\varepsilon} = 4.25$
- $\widehat{\sigma}^2_{u} = 21.37$

Fitting mixed effects models

```
m1 <- lmer(score ~ style + (1|professor),
           data = teaching)
summary(m1)
```

```
...
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)       77.657      2.075  37.419
## stylesome flip    11.073      2.935   3.773
## stylefull flip     2.805      2.935   0.956
...
```

- $\widehat{\beta}_0 = 77.66$
- $\widehat{\beta}_1 = 11.07$
- $\widehat{\beta}_2 = 2.81$

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}^2_\varepsilon = 4.25, \quad \widehat{\sigma}^2_u = 21.37$

> How do I interpret $\widehat{\beta}_0 = 77.66$?

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}_\varepsilon^2 = 4.25, \quad \widehat{\sigma}_u^2 = 21.37$

> How do I interpret $\widehat{\beta}_1 = 11.07$ and $\widehat{\beta}_2 = 2.81$?

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}_\varepsilon^2 = 4.25, \quad \widehat{\sigma}_u^2 = 21.37$

> How do I interpret $\widehat{\beta}_1 = 11.07$ and $\widehat{\beta}_2 = 2.81$?

We expect that, on average, scores in some-flipped classes are 11.07 points higher than for no-flipped classes.

We expect that, on average, scores in fully-flipped classes are 2.81 points higher than for no-flipped classes.

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}_\varepsilon^2 = 4.25, \quad \widehat{\sigma}_u^2 = 21.37$

> How do I interpret $\widehat{\sigma}_u^2 = 21.37$?

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}_\varepsilon^2 = 4.25, \quad \widehat{\sigma}_u^2 = 21.37$

> How do I interpret $\quad \widehat{\sigma}_u^2 = 21.37$?

The average score varies from professor to professor by about 4.62 ($= \sqrt{21.37}$) points

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}^2_\varepsilon = 4.25, \quad \widehat{\sigma}^2_u = 21.37$

> How do I interpret $\widehat{\sigma}^2_\varepsilon = 4.25$?

Interpretation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

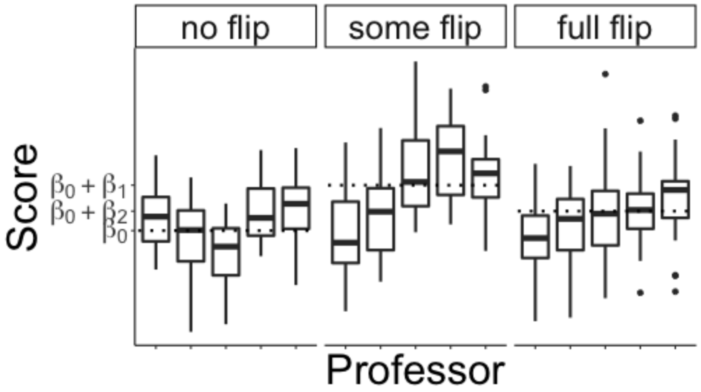
$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

- $\widehat{\beta}_0 = 77.66, \quad \widehat{\beta}_1 = 11.07, \quad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}_\varepsilon^2 = 4.25, \quad \widehat{\sigma}_u^2 = 21.37$

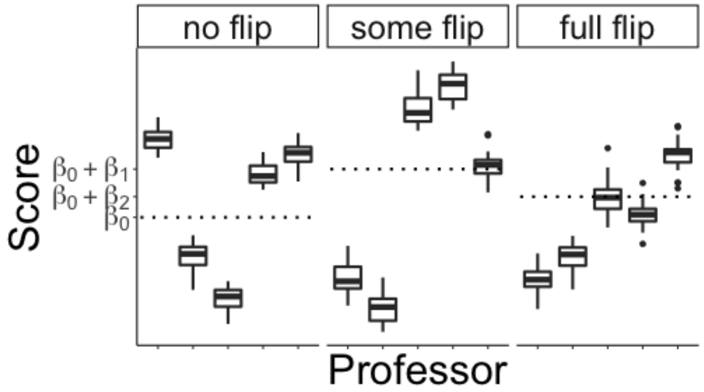
> How do I interpret $\widehat{\sigma}_\varepsilon^2 = 4.25$?

Within a class, students' scores vary by about 2.06 ($= \sqrt{4.25}$) points

Intra-class correlation



σ_ε^2 is large relative to σ_u^2



σ_ε^2 is small relative to σ_u^2

- Observations within a group are *more correlated* when σ_ε^2 is small relative to σ_u^2
- **Intra-class correlation:**

$$\rho_{group} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} = \frac{\text{between group variance}}{\text{total variance}}$$

Intra-class correlation

$$Score_{ij} = \beta_0 + \beta_1 \text{SomeFlipped}_i + \beta_2 \text{FullyFlipped}_i + u_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \qquad u_i \overset{iid}{\sim} N(0, \sigma^2_u)$$

- $\widehat{\beta}_0 = 77.66, \qquad \widehat{\beta}_1 = 11.07, \qquad \widehat{\beta}_2 = 2.81$
- $\widehat{\sigma}^2_\varepsilon = 4.25, \qquad \widehat{\sigma}^2_u = 21.37$

$$\hat{\rho}_{group} = \frac{21.37}{21.37 + 4.25} = 0.83$$

So 83% of the variation in student's scores can be explained by differences in average scores from class to class (after accounting for teaching style). That's huge!

Class activity, Part II

https://sta214-f22.github.io/class_activities/ca_lecture_26.html

Class activity

Interpret the estimate fixed effect coefficients $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

Class activity

> Interpret the estimate fixed effect coefficients $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

On average (across neighborhoods), we expect that the price of rental with 0 overall satisfaction is $27.28.

For a fixed neighborhood, an increase of 1 point in overall satisfaction is associated with an increase of $14.81 in rental price.

Class activity

Calculate and interpret the estimated intra-class correlation.

Class activity

> Calculate and interpret the estimated intra-class correlation.

$$\hat{\rho}_{group} = \frac{1048}{1048 + 6762} = 0.134$$

About 13% of the variability in price can be explained by differences in the average price between neighborhoods (after accounting for overall satisfaction).