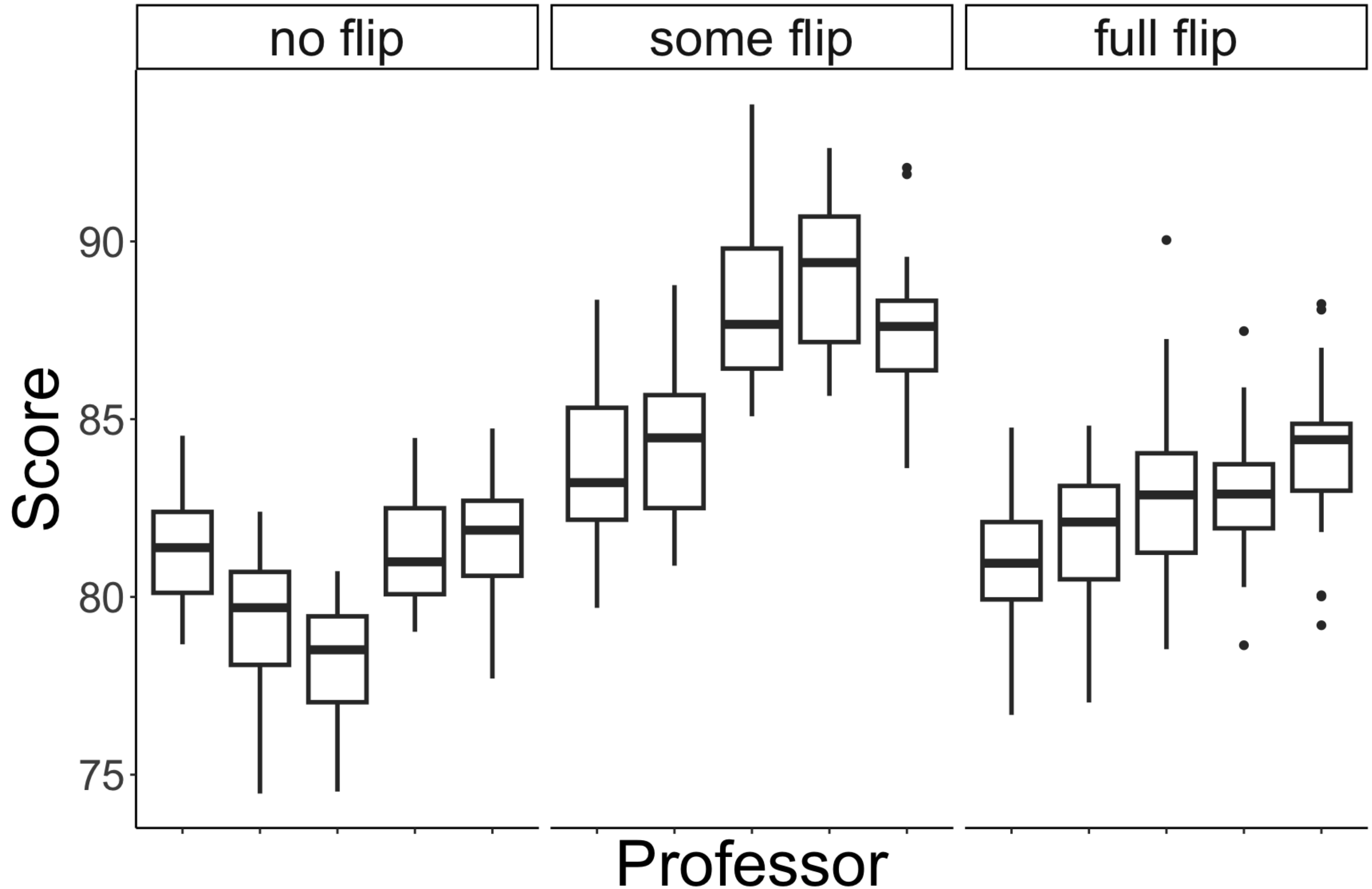
Lecture 31

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

group   obs. within group

$$\text{Score}_{ij} = \beta_0 + \underbrace{\beta_1\,\text{SomeFlipped}_i + \beta_2\,\text{FullyFlipped}_i}_{\text{fixed effects}} + u_i + \varepsilon_{ij}$$

noise (individual students)

random effect for each class/professor

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

$\varepsilon_{ij} \perp\!\!\!\perp u_i$

Parameters: $\beta_0, \beta_1, \beta_2, \sigma_\varepsilon^2, \sigma_u^2$

Fitting mixed effects models

response

fixed effects

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

mixed effects model

(random) intercept which depends on professor

i.e. u_i $i =$ professor

$Y \sim 1$

Fitting mixed effects models

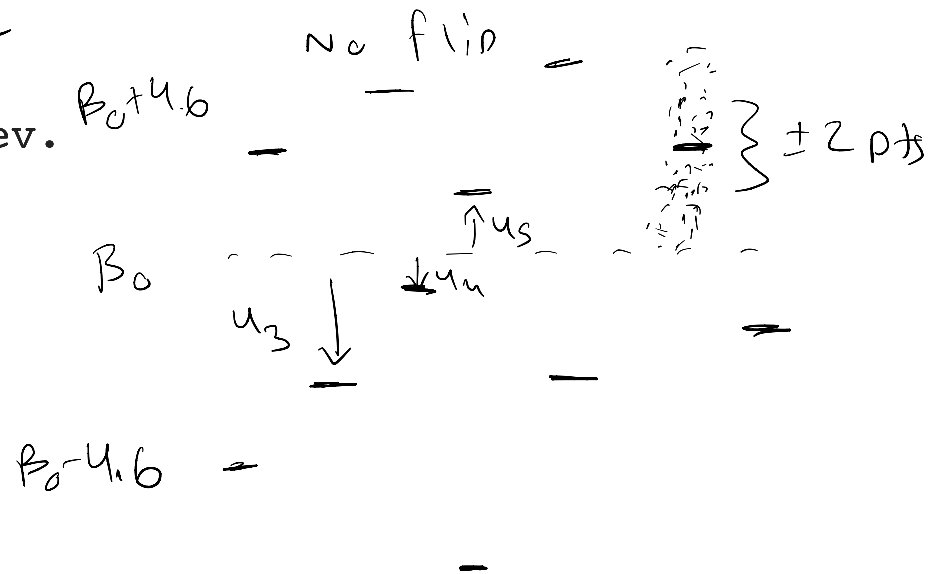
```r
library(lme4)
```

```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
...
```

$\hat{\sigma}_u^2$

$\hat{\sigma}_\varepsilon^2$



$\hat{\sigma}_u^2 = 21.365$ $\quad \hat{\sigma}_u = 4.622$

On average, scores within a teaching style, scores between professors vary by ± 4 points

$\hat{\sigma}_\varepsilon^2 = 4.252$ $\quad \hat{\sigma}_\varepsilon = 2.06$ $\Rightarrow$ scores for individual students (within a class) tend to be around ± 2 from class average

Fitting mixed effects models

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

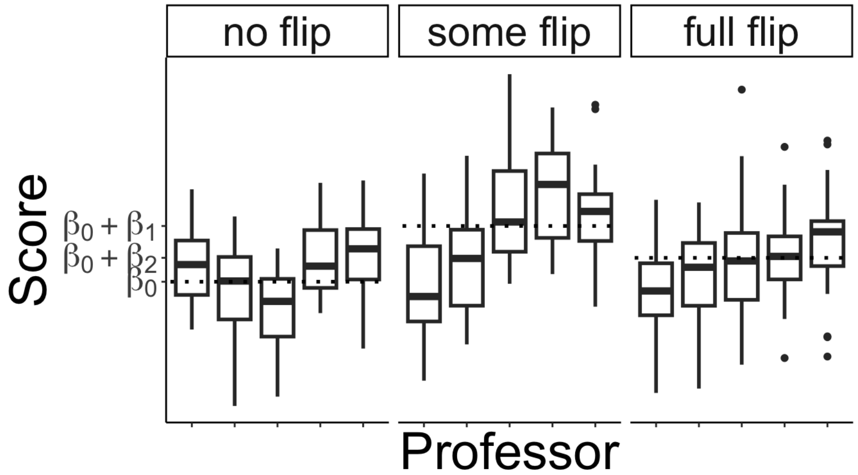
```
...
Fixed effects:
                Estimate Std. Error t value
(Intercept)       77.657      2.075  37.419
stylesome flip    11.073      2.935   3.773
stylefull flip     2.805      2.935   0.956
...
```

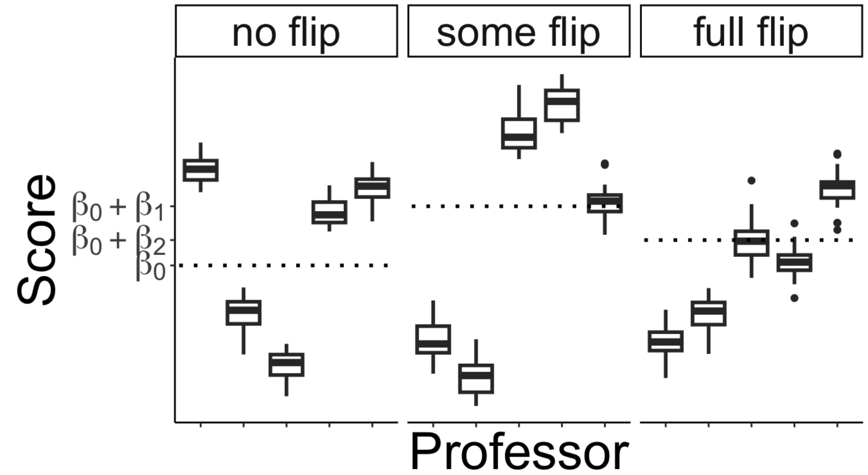
$\hat{\beta}_0$ $\hat{\beta}_1$ $\hat{\beta}_2$

$\hat{\beta}_0 = 77.657$

On average, professors teaching "no flip" classes have a score of 77.657

Intra-class correlation





σ_ε^2 is large relative to σ_u^2

σ_ε^2 is small relative to σ_u^2

Intra-class correlation:

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

$$Y = X\beta + Zu + \varepsilon \qquad N = \sum_{i=1}^{m} n_i$$

For $\quad Y_{ij} = \beta_0 + u_i + \beta_1 X_{ij} + \varepsilon_{ij} \quad ;$

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ \vdots \\ Y_{mn_m} \end{bmatrix} = \begin{bmatrix} 1 & X_{11} \\ 1 & X_{12} \\ \vdots & \vdots \\ 1 & X_{mn_m} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & 0 & & 0 \\ 1 & \vdots & & \vdots \\ 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & & 1 \\ & 0 & & \vdots \\ & \vdots & & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{mn_m} \end{bmatrix}$$

$$u \sim N(0, \sigma_u^2 I_m) \qquad \varepsilon \sim N(0, \sigma_\varepsilon^2 I_N) \qquad u \perp\!\!\!\perp \varepsilon$$

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{Var}(Zu) + \mathrm{Var}(\varepsilon) \\ &= Z\,\mathrm{Var}(u)\,Z^T + \mathrm{Var}(\varepsilon) \\ &= \sigma_u^2 ZZ^T + \sigma_\varepsilon^2 I_N \end{aligned}$$

$$\begin{aligned}
\operatorname{Var}(Y) &= \operatorname{Var}(Zu) + \operatorname{Var}(\varepsilon) \\
&= Z \operatorname{Var}(u) Z^T + \operatorname{Var}(\varepsilon) \\
&= \sigma_u^2 \, Z Z^T + \sigma_\varepsilon^2 \, I_N
\end{aligned}$$

$$= \begin{bmatrix}
\begin{matrix} \sigma_u^2+\sigma_\varepsilon^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2+\sigma_\varepsilon^2 & \cdots & \sigma_u^2 \\ \vdots & & \ddots & \\ \sigma_u^2 & & & \sigma_u^2+\sigma_\varepsilon^2 \end{matrix} & 0 \cdots\cdots 0 & & 0 \\
& \ddots & & \\
0 & & & \begin{matrix} \sigma_u^2+\sigma_\varepsilon^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2+\sigma_\varepsilon^2 & \cdots & \sigma_u^2 \\ \vdots & & \ddots & \\ \sigma_u^2 & & & \sigma_u^2+\sigma_\varepsilon^2 \end{matrix}
\end{bmatrix}$$

$$\operatorname{Var}(Y_{ij}) = \sigma_u^2 + \sigma_\varepsilon^2$$

$$\operatorname{Cov}(Y_{ij}, Y_{ik}) = \sigma_u^2$$

$$\Rightarrow \operatorname{Cor}(Y_{ij}, Y_{ik}) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$$

= intra-class correlation

$$\operatorname{Cov}(Y_{ij}, Y_{kl}) = 0$$

Intra-class correlation

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

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

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

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

https://sta712-f23.github.io/class_activities/ca_lecture_32.html