Intro to mixed effects models

Recap: data and motivation

We have data from a 2010 study on performance anxiety in 37 undergraduate music majors. For each musician, data was collected on anxiety levels before different performances (between 2 and 15 performances were measured for each musician), with variables including:

- `id`: a unique identifier for the musician
- `na`: negative affect score (a measure of anxiety)
- `perform_type`: whether the musican was performing in a large ensemble, small ensemble, or solo

> How can we model the relationship between performance type and anxiety?

Recap: a mixed effects model

fixed effects

$$\underbrace{Anxiety_{ij}}_{} = \beta_0 + \underbrace{u_i}_{\text{random effect}} + \beta_1\, SmallEnsemble_{ij} + \beta_2\, Solo_{ij} + \underbrace{\varepsilon_{ij}}_{\text{noise}}$$

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

$\beta_0 + u_i$ = intercept for musician i

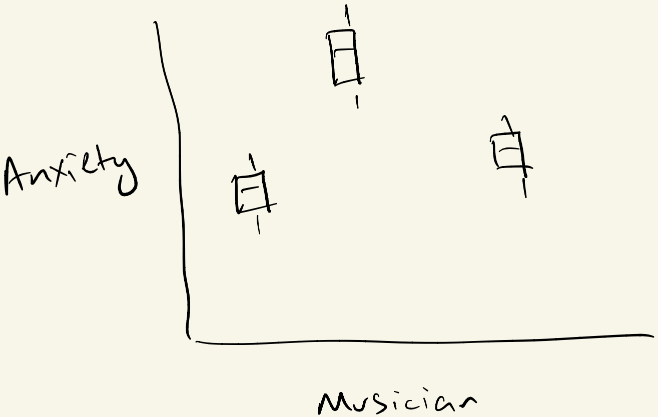
σ_u^2 = variability between musicians

σ_ε^2 = variability between performances w/in a musician

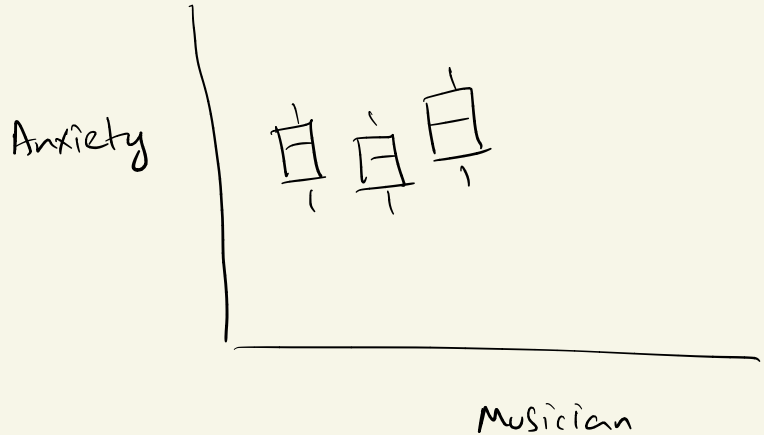
Intra-class correlation (ICC) $= \dfrac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$

$= \dfrac{\text{between-group variance}}{\text{total variance}}$

High ICC



Low ICC



$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \quad \leftarrow \text{high when } \sigma_u^2 \text{ is high \underline{relative} to } \sigma_\varepsilon^2$$

$$\in [0,1]$$

Fitting the model in R

for mixed effects

package

fixed effects

random effects

```r
library(lme4)
m1 <- lmer(na ~ perform_type + (1|id),
           data = music)
summary(m1)
```

random intercept for each grap (id)

```
...
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  5.56    2.358
##  Residual             21.75    4.664
## Number of obs: 497, groups:  id, 37
## Fixed effects:
##                            Estimate Std. Error t value
## (Intercept)                 14.9654     0.5920  25.278
## perform_typeSmall Ensemble   0.7709     0.7210   1.069
## perform_typeSolo             2.0142     0.5521   3.648
...
```

$\hat{\sigma}_u^2 = 5.56$

$\hat{\sigma}_\varepsilon^2 = 21.75$

$\hat{\beta}_0$

$\hat{\beta}_1$

$\hat{\beta}_2$

$\beta_0 + u_i + \beta_1 \text{Small}_{ij} + \ldots$

$\underbrace{\text{intercept for musician } i}_{\text{average anxiety before a large ensemble performance}} = \beta_0 + u_i$

Interpretation

$\beta_0 = E[\beta_0 + u_i]$

```
...
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  14.9654     0.5920  25.278
## perform_typeSmall Ensemble    0.7709     0.7210   1.069
## perform_typeSolo              2.0142     0.5521   3.648
...
```

> How would we interpret the estimated fixed effects?

$\hat{\beta}_0 = 14.97$ = estimated expected anxiety for a (randomly selected) musician before a large ensemble performance

$\hat{\beta}_1 = 0.77$ = estimated difference in a musician's anxiety before small ensemble performances vs large ensemble performances

Prediction

```
...
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  14.9654     0.5920  25.278
## perform_typeSmall Ensemble    0.7709     0.7210   1.069
## perform_typeSolo              2.0142     0.5521   3.648
...
```

> What is the estimated anxiety for Musician 1 before a solo performance?

$14.97 + 2.01 + \hat{u}_1$

Intuition:

$\hat{u}_1 =$ (Average anxiety for musician 1 before Large performances) $- \hat{\beta}_0$

Prediction

> What is the estimated anxiety for Musician 1 before a solo performance?

```
coef(m1)
```

```
...
## $id
##    (Intercept) perform_typeSmall Ensemble perform_typeSolo
## 1     11.61227                  0.7708706         2.014226
## 2     12.78968                  0.7708706         2.014226
## 3     12.85152                  0.7708706         2.014226
...
```

$\underbrace{11.61}_{\hat{\beta}_0 + \hat{u}_1} + \underbrace{2.01}_{\hat{\beta}_2}$

Prediction

```
...
## Fixed effects:
##                               Estimate Std. Error t value
## (Intercept)                    14.9654     0.5920  25.278
## perform_typeSmall Ensemble      0.7709     0.7210   1.069
## perform_typeSolo                2.0142     0.5521   3.648
...
```

> What is the estimated anxiety for a *new* musician (not in the data) before a solo performance?

$\underset{\hat{\beta}_0}{14.97} + \underset{\hat{\beta}_2}{2.01}$

we don't know u_i should be

Since $u_i \sim N(0, \sigma_u^2)$,

guess that $u_i = 0$

Assumptions

$$Anxiety_{ij} = \beta_0 + u_i + \beta_1\ SmallEnsemble_{ij} + \beta_2\ Solo_{ij} + \varepsilon_{ij}$$

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

> What assumptions does this mixed effects model make?

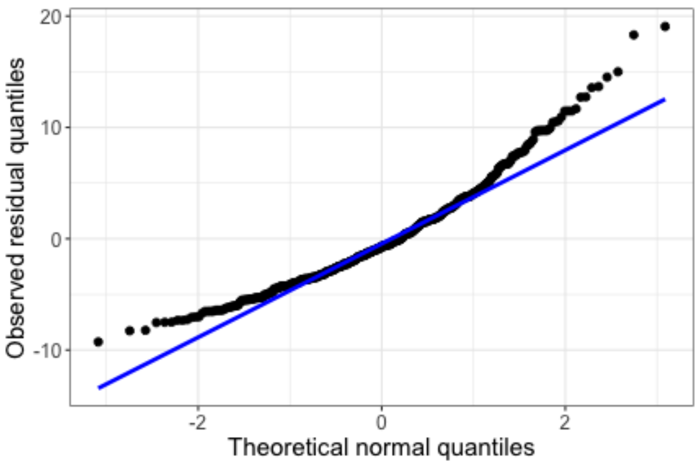
Assessing normality

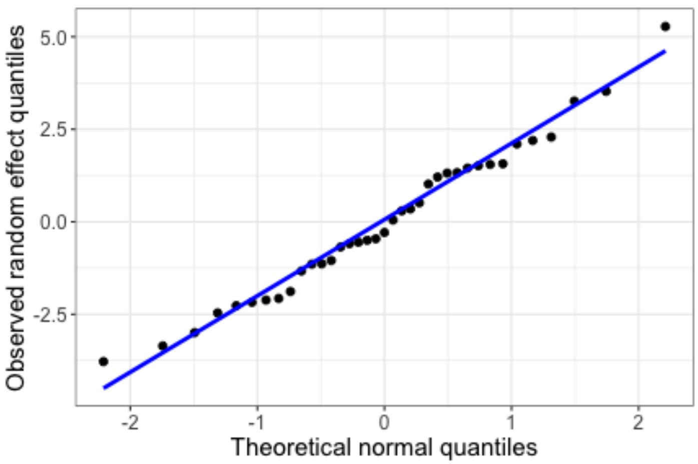
$$Anxiety_{ij} = \beta_0 + u_i + \beta_1\ SmallEnsemble_{ij} + \beta_2\ Solo_{ij} + \varepsilon_{ij}$$

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

> How should we check the normality assumption?

QQ plots


20
10
0
-10
-2
0
2
Observed residual quantiles
Theoretical normal quantiles


5.0
2.5
0.0
-2.5
-2
-1
0
1
2
Observed random effect quantiles
Theoretical normal quantiles

Changing the model

$$Anxiety_{ij} = \beta_0 + u_i + \beta_1 \; SmallEnsemble_{ij} + \beta_2 \; Solo_{ij} + \varepsilon_{ij}$$

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

> How could we change the model to allow the effect of performance type to differ between musicians?

Fitting the model

```
m2 <- lmer(na ~ perform_type + (perform_type|id),
           data = music)
summary(m2)
```

```
...
## Random effects:
##  Groups   Name                      Variance Std.Dev. Corr
##  id       (Intercept)                3.986   1.997
##           perform_typeSmall Ensemble  2.019   1.421    -0.43
##           perform_typeSolo            1.017   1.008     0.74  0.29
##  Residual                           21.288   4.614
## Number of obs: 497, groups:  id, 37
##
## Fixed effects:
##                            Estimate Std. Error t value
## (Intercept)                 15.0503     0.5436  27.685
## perform_typeSmall Ensemble   0.6996     0.7410   0.944
## perform_typeSolo             2.0134     0.5671   3.550
...
```

Prediction

> What is the estimated anxiety for Musician 1 before a solo performance?

```
coef(m2)
```

```
...
## $id
##    (Intercept) perform_typeSmall Ensemble perform_typeSolo
## 1     12.37560                 0.84623321        0.6590148
## 2     13.61693                 0.30915635        1.0413577
## 3     12.86707                 1.31366273        1.1674007
...
```