Intro to mixed effects models

Warm-up: class activity

https://sta712-f22.github.io/class_activities/ca_lecture_39.html

- As σ_u^2 increases
  - variance of the mixed effects does ~~better~~
    relative to the other models
  - may see bias in $\hat{\beta}$ if the model doesn't
    include groups
- Using fixed effects for each can result in increased
  variance of $\hat{\beta}$, particularly if X is correlated
  with group label

Fitting mixed effects models

Mixed effects model: $Y = X\beta + Zu + \varepsilon$

$(u \perp\!\!\!\perp \varepsilon)$

$u \sim N(0, G)$ (e.g., $G = \sigma_u^2 I$)

$\varepsilon \sim N(0, R)$ (e.g. $R = \sigma_\varepsilon^2 I$)

$$\Rightarrow \quad Zu + \varepsilon \sim N(0, ZGZ^T + R)$$

Marginal model: $Y = X\beta + \varepsilon^*$

$$\varepsilon^* \sim N(0, \underbrace{ZGZ^T + R}_{V})$$

Assume V is known: use weighted least squares

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y$$ (inverse variance weighting)

If V is unknown: estimation is more complicated (jointly estimate β & V)

Case-control example:

$$Y_{ij} = \beta_0 + u_i + \beta_1 X_i + \varepsilon_{ij} \qquad u_i \overset{iid}{\sim} N(0, \sigma_u^2)$$

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

$$X_i \in \{0, 1\}$$

control ↗ ↖ case / treatment

Controls

Groups

$1, \ldots, m_0$

Cases

Groups

$m_0+1, \ldots, m$

$\Rightarrow \quad X_1, \ldots, X_{m_0} = 0 \qquad X_{m_0+1}, \ldots, X_m = 1$

Let n_i = # obs in group i

Let $\bar{Y}_i \equiv \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$

Then $\hat{\beta}_1$ (mixed effects model) $=$

$$\frac{\displaystyle\sum_{i=m_0+1}^{m} \frac{n_i \bar{Y}_i}{n_i \sigma_u^2 + \sigma_\varepsilon^2}}{\displaystyle\sum_{i=m_0+1}^{m} \frac{n_i}{n_i \sigma_u^2 + \sigma_\varepsilon^2}} - \frac{\displaystyle\sum_{i=1}^{m_0} \frac{n_i \bar{Y}_i}{n_i \sigma_u^2 + \sigma_\varepsilon^2}}{\displaystyle\sum_{i=1}^{m_0} \frac{n_i}{n_i \sigma_u^2 + \sigma_\varepsilon^2}}$$

$$\frac{n_i \bar{Y}_i}{n_i \sigma_u^2 + \sigma_\varepsilon^2} = \frac{\bar{Y}_i}{\sigma_u^2 + \frac{\sigma_\varepsilon^2}{n_i}}$$

If instead we fit a linear model w/out group effect:

$$\hat{\beta}_1 = \frac{\displaystyle\sum_{i=m_0+1}^{m} n_i \bar{Y}_i}{\displaystyle\sum_{i=m_0+1}^{m} n_i} - \frac{\displaystyle\sum_{i=1}^{m_0} n_i \bar{Y}_i}{\displaystyle\sum_{i=1}^{m_0} n_i}$$

This gives potentially much more weight to groups with more observations

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?

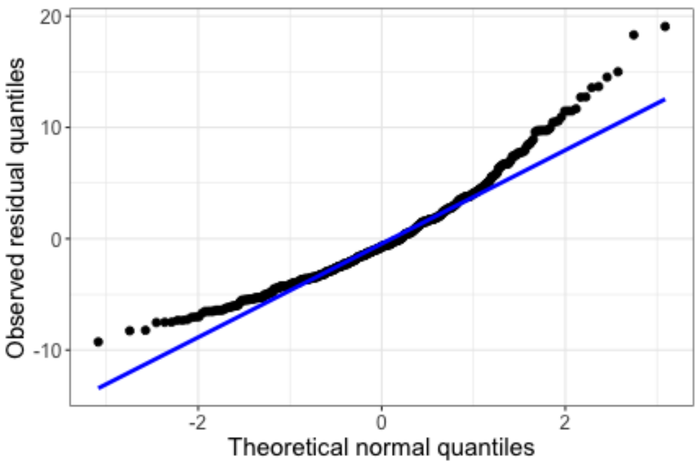
Assessing normality

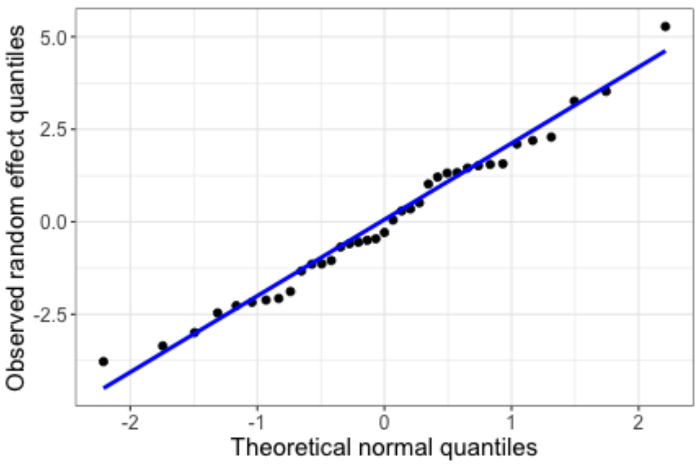
$$Anxiety_{ij} = \beta_0 + u_i + \beta_1 \ SmallEnsemble_{ij} + \beta_2 \ Solo_{ij} + \varepsilon_{ij}$$

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

> How should we check the normality assumption?

QQ plots


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


Observed random effect quantiles
5.0
2.5
0.0
-2.5
-2
-1
0
1
2
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?

$$Anxiety_{ij} = (\beta_0 + u_i) + (\beta_1 + v_i) Small_{ij} + (\beta_2 + w_i) Solo_{ij} + \varepsilon_{ij}$$

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

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho_{uv} \sigma_u \sigma_v & \rho_{uw} \sigma_u \sigma_w \\ & \sigma_v^2 & \rho_{vw} \sigma_v \sigma_w \\ & & \sigma_w^2 \end{bmatrix} \right)$$

Fitting the model

fixed effect

```
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
...
```

$\hat{\sigma}_u^2$ $\hat{\sigma}_w^2$ $\hat{\sigma}_v^2$ $\hat{\sigma}_\varepsilon^2$ $\hat{\rho}_{uv}$ $\hat{\rho}_{uw}$ $\hat{\rho}_{vw}$

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

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
...
```

$$\underbrace{12.38}_{\hat{\beta}_0 + \hat{u}_1} + \underbrace{0.66}_{\hat{\beta}_2 + \hat{w}_1}$$