Review of Linear Mixed Models

$$\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{I}\sigma^2)$$

Objectives

- A sample model
- Parts of a model
- Estimation
  - Fixed effects

  - Variance
  - Standard Errors
  - Confidence Intervals

- Maximum Likelihood

- Hypothesis Testing

A simple model

- 2×3 FAT
- RCBD with 2 Replications

$$y_{ijk} = \overbrace{\mu + \alpha_i + \beta_j + \gamma_{ij} + B_k}^{\text{Systematic}} + \overbrace{e_{ijk}}^{\text{Random}}$$

- Mean (Fixed Blocks)

$$\mathrm{E}(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij} + B_k$$

- Variance

$$\mathrm{var}(y_{ijk}) = \sigma^2$$

all other covariances are zero

- Distribution

Normal

Matrix Form

$$
\begin{pmatrix} y_{111} \\ y_{112} \\ y_{121} \\ \vdots \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{232} \end{pmatrix}
\begin{array}{ll}
\mathbf{1}_2 \otimes \mathbf{1}_3 \otimes \mathbf{1}_2 \mu & + \boldsymbol{I}_2 \otimes \mathbf{1}_3 \otimes \mathbf{1}_2 \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \\
= \mathbf{1}_2 \otimes \boldsymbol{I}_3 \otimes \mathbf{1}_2 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + & \boldsymbol{I}_2 \otimes \boldsymbol{I}_3 \otimes \mathbf{1}_2 \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{pmatrix} \\
+ \mathbf{1}_2 \otimes \mathbf{1}_3 \otimes \boldsymbol{I}_2 \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} & + \boldsymbol{e}
\end{array}
$$

$$
\boldsymbol{y} = \boldsymbol{X}_0 \mu + \boldsymbol{X}_1 \boldsymbol{\alpha} + \boldsymbol{X}_2 \boldsymbol{\beta} + \boldsymbol{X}_3 \boldsymbol{\gamma} + \boldsymbol{X}_4 \boldsymbol{B} + \boldsymbol{e}
$$

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e}$$

$$\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X\beta}$$

$$\mathrm{var}(\boldsymbol{y}) = \boldsymbol{I}\sigma^2$$

$$\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{X\beta}, \boldsymbol{I}\sigma^2)$$

Parts of a Model

1. Location (Mean or Expected Value)
   - Linear $\boldsymbol{X\beta}$
2. Variability (Variances and Covariances)
   - Constant Variance
   - Covariances zero
3. Distribution
   - Normal

Estimation

- Model – Framework to describe the process that generated the data
- Parameters – Identify the specific process
- Estimation – Best guess of which specific process we have
  - Certain level of uncertainty

Fixed Effects

- Point estimate

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$$

  I will restrict my attention to symmetric G-inverses

- Distribution

  1. Linear function of a multivariate normal random vector $\Rightarrow \widehat{\boldsymbol{\beta}}$ is also MVN

$$\widehat{\boldsymbol{\beta}} \sim \mathrm{N}((\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}, (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\sigma^2)$$

2. If we assume that $(\boldsymbol{X}'\boldsymbol{X})^-$ is reflexive

$$\widehat{\boldsymbol{\beta}} \sim \mathrm{N}((\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}, (\boldsymbol{X}'\boldsymbol{X})^-\sigma^2)$$

3. If $\boldsymbol{K}'\boldsymbol{\beta}$ is estimable, then

$$\boldsymbol{K}'\widehat{\boldsymbol{\beta}} \sim \mathrm{N}(\boldsymbol{K}'\boldsymbol{\beta}, \boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{K}\sigma^2)$$

Variance

$$\widehat{\sigma}^2 = \frac{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}')\boldsymbol{y}}{\operatorname{tr}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}')}$$

- Quadratic Form ($\boldsymbol{Q}$ symmetric)

$$\mathrm{E}(\widehat{\sigma}^2) = \operatorname{tr}(\boldsymbol{Q}\boldsymbol{V}) + \boldsymbol{\mu}'\boldsymbol{Q}\boldsymbol{\mu} = \sigma^2$$

$$\begin{aligned}\operatorname{var}(\widehat{\sigma}^2) &= 2\operatorname{tr}(\boldsymbol{Q}\boldsymbol{V}\boldsymbol{Q}\boldsymbol{V}) + 4\boldsymbol{\mu}'\boldsymbol{Q}\boldsymbol{V}\boldsymbol{Q}\boldsymbol{\mu} \\ &= 2\sigma^4 \frac{1}{\operatorname{tr}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}')}\end{aligned}$$

- Distribution

$$\theta \boldsymbol{QV} = \boldsymbol{QVQV}$$

$$\frac{\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}}{\theta} \sim \chi^2(\mathrm{tr}(\boldsymbol{QV}/\theta), nc)$$

$$\frac{\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{M}_X)\boldsymbol{y}}{\sigma^2} \sim \chi^2_{\mathrm{tr}(\boldsymbol{I}-\boldsymbol{M}_X)}$$

Standard Errors

- $\boldsymbol{K}'\widehat{\boldsymbol{\beta}}$

$$\sqrt{\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{K}\sigma^2}$$

- $\widehat{\sigma}^2$

$$\sigma^2\sqrt{\frac{2}{\mathrm{tr}(\boldsymbol{I}-\boldsymbol{M}_X)}}$$

Confidence Intervals

- $\boldsymbol{K}'\boldsymbol{\beta}$

$$\boldsymbol{K}'\widehat{\boldsymbol{\beta}} \pm t_{df_e,\alpha/2}\widehat{\sigma}_{\boldsymbol{K}'\widehat{\boldsymbol{\beta}}}$$

- σ^2
  - Exact

$$\frac{df_e\widehat{\sigma}^2}{\chi^2_{df_e,1-\alpha/2}} < \sigma^2 < \frac{df_e\widehat{\sigma}^2}{\chi^2_{df_e,\alpha/2}}$$

  - Approximately

$$\widehat{\sigma}^2 \pm Z_{\alpha/2}\widehat{\sigma}_{\widehat{\sigma}^2}$$

Maximum Likelihood

$$\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{X\beta}, \boldsymbol{I\sigma^2})$$

1. Density $\longrightarrow$ Likelihood

2. Natural Log

3. Score (First Partials)

4. Information (Second Partials)

Likelihood

$$L(\boldsymbol{\beta}, \sigma^2 | \boldsymbol{y}) = \left[2\pi\sigma^2\right]^{-N/2} \exp\left[-\frac{1}{2}\frac{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})}{\sigma^2}\right]$$

Log Likelihood

$$\ell(\boldsymbol{\beta}, \sigma^2|\boldsymbol{y}) = -\frac{1}{2}[N\ln(2\pi) + N\ln(\sigma^2) + \frac{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})}{\sigma^2}]$$

Score Function

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})/\sigma^2$$

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

- Set equal to zero and solve
  1. Directly
  2. Iteratively
     - Newton Raphson (-H)
     - Fisher's Scoring (I=-E(H))

Hessian

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = -\boldsymbol{X}'\boldsymbol{X}/\sigma^2$$

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \sigma^2} = -\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})/\sigma^4$$

$$\frac{\partial^2 \ell}{\partial \sigma^2 \partial \sigma^2} = \frac{N}{2\sigma^4} - \frac{1}{\sigma^6}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

Information

$$-\operatorname{E}\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right] = \boldsymbol{X}'\boldsymbol{X}/\sigma^2$$

$$-\operatorname{E}\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \sigma^2}\right] = \mathbf{0}$$

$$-\operatorname{E}\left[\frac{\partial^2 \ell}{\partial \sigma^2 \partial \sigma^2}\right] = \frac{N}{2\sigma^4}$$

Estimation

- Newton Raphson

$$-\begin{pmatrix} \frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} & \frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \sigma^2} \\ \frac{\partial^2 \ell}{\partial \sigma^2 \partial \boldsymbol{\beta}'} & \frac{\partial^2 \ell}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}}_{[i]} - \widehat{\boldsymbol{\beta}}_{[i-1]} \\ \widehat{\sigma}^2_{[i]} - \widehat{\sigma}^2_{[i-1]} \end{pmatrix} = \begin{pmatrix} \frac{\partial \ell}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell}{\partial \sigma^2} \end{pmatrix}$$

- Fisher's Scoring

$$\begin{pmatrix} \boldsymbol{X}'\boldsymbol{X}/\sigma^2 & 0 \\ 0 & \frac{N}{2\sigma^4} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}}_{[i]} - \widehat{\boldsymbol{\beta}}_{[i-1]} \\ \widehat{\sigma}^2_{[i]} - \widehat{\sigma}^2_{[i-1]} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})/\sigma^2 \\ \frac{1}{2\sigma^4}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) - \frac{N}{2\sigma^2} \end{pmatrix}$$

- Likelihood evaluated at MLE

$$\ell(\widehat{\boldsymbol{\beta}}, \widehat{\sigma^2}|\boldsymbol{y}) = -\frac{1}{2}[N\ln(2\pi) + N\ln(\widehat{\sigma}^2) + N]$$

Hypothesis Testing

- t-tests

$$\frac{\boldsymbol{k}'\widehat{\boldsymbol{\beta}}}{\widehat{\sigma}_{\boldsymbol{k}'\widehat{\boldsymbol{\beta}}}}$$

- F-tests $dfh = \mathrm{rank}(\boldsymbol{K})$

$$\frac{\widehat{\boldsymbol{\beta}}'\boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{K})^{-}\boldsymbol{K}'\widehat{\boldsymbol{\beta}/dfh}}{\widehat{\sigma^2}}$$

- χ^2 tests

$$\frac{\widehat{\boldsymbol{\beta}}'\boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{K})^{-}\boldsymbol{K}'\widehat{\boldsymbol{\beta}}}{\sigma^2}$$

- Wald Statistics (Approximately χ^2 with $df = dfh$)

$$W = \boldsymbol{\theta}' \begin{pmatrix} \boldsymbol{K}' & 0 \end{pmatrix} \left[\begin{pmatrix} \boldsymbol{K} \\ 0 \end{pmatrix} \boldsymbol{I}^{-}(\boldsymbol{\theta}) \begin{pmatrix} \boldsymbol{K}' & 0 \end{pmatrix} \right]^{-1} \begin{pmatrix} \boldsymbol{K} \\ 0 \end{pmatrix} \boldsymbol{\theta}$$

$$= \frac{\widehat{\boldsymbol{\beta}}' \boldsymbol{K} (\boldsymbol{K}' (\boldsymbol{X}' \boldsymbol{X})^{-} \boldsymbol{K})^{-} \boldsymbol{K}' \widehat{\boldsymbol{\beta}}}{\widehat{\sigma}^2}$$

- Likelihood Ratio Test (Approximately χ^2 with $df = rank(K)$)

$$
\begin{aligned}
G &= -2 \ln \left[\frac{L(\boldsymbol{K}'\boldsymbol{\beta} = 0)}{L} \right] \\
&= -2 \left[\ell(\boldsymbol{K}'\boldsymbol{\beta} = 0) - \ell \right] \\
&= N \ln \left[\frac{\widehat{\sigma}^2_R}{\widehat{\sigma}^2} \right]
\end{aligned}
$$

$$
\widehat{\boldsymbol{\beta}}_R = [\boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{K})^- \boldsymbol{K}']\widehat{\boldsymbol{\beta}}
$$

$$
\widehat{\sigma}^2_R = \widehat{\sigma}^2 + \frac{\widehat{\boldsymbol{\beta}}' \boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{K})^- \boldsymbol{K}'\widehat{\boldsymbol{\beta}}}{N}
$$

$$
G = N \ln \left[1 + \frac{\widehat{\boldsymbol{\beta}}' \boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{K})^- \boldsymbol{K}\widehat{\boldsymbol{\beta}}}{N\widehat{\sigma}^2} \right]
$$

- Likelihood Ratio Test when σ^2 is known

$$
\begin{aligned}
G &= -2\left[\ell(\boldsymbol{K}'\boldsymbol{\beta}=0)-\ell\right] \\
&= -2[-.5\{(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}_R)'(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}_R)-(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})\}/\sigma^2] \\
&= \frac{\boldsymbol{y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{K})^-\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}}{\sigma^2} \\
&= \frac{\widehat{\boldsymbol{\beta}}'\boldsymbol{K}(\boldsymbol{K}'(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{K})^-\boldsymbol{K}'\widehat{\boldsymbol{\beta}}}{\sigma^2}
\end{aligned}
$$