STAT 714 fa 2025 Lec 06

Variance component estimation in random and mixed effects models

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

$$y = Xb + e$$

$$\text{Cov}\, e = \sigma^2 I_n,\ \sigma^2 V$$

Mixed model setup

A *linear mixed model* has the form

fixed effects

$$\mathbf{y} = \boxed{\mathbf{X}\mathbf{b}} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

random effects

$$u \sim N(0, G)$$

$$e \sim N(0, R)$$

where

- $\mathbf{X}$ is an $n \times p$ design matrix
- $\mathbf{b}$ is a $p \times 1$ vector of parameters describing *fixed effects*
- $\mathbf{Z}$ is an $n \times q$ design matrix
- $\mathbf{u} \sim \text{Normal}(\mathbf{0}, \mathbf{G})$ is a $q \times 1$ vector of *random effects*
- $\mathbf{e} \sim \text{Normal}(\mathbf{0}, \mathbf{R})$ is an $n \times 1$ vector of error terms independent of $\mathbf{u}$.

Letting $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \mathbf{R}$, we have $\mathbf{y} \sim \text{Normal}(\mathbf{X}\mathbf{b}, \mathbf{V})$.

$$\text{Cov}\, y = \text{Cov}(Zu + e) = \underbrace{ZGZ^T + R}_{V}$$

Typically $\mathbf{V} = \mathbf{V}(\boldsymbol{\theta})$ depends on some parameters $\boldsymbol{\theta}$ which we wish to estimate.

One-way random effects model

For responses Y_{ij}, assume

indep.

$$Y_{ij} = \underline{\mu} + \underline{A_i} + \varepsilon_{ij}, \quad i = 1, \dots, a, \quad j = 1, \dots, n_i$$

where

- ε_{ij} are independent Normal$(0, \sigma^2_\varepsilon)$
- A_i are independent Normal$(0, \sigma^2_A)$
- A_i and ε_{ij} are independent

$$\underset{\sim}{Y} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ \vdots \\ Y_{a1} \\ \vdots \\ Y_{an_a} \end{bmatrix} \qquad X b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \mu$$

$$Z \underset{\sim}{u} = \underbrace{\begin{bmatrix} \underset{\sim}{1}_{n_1} & & \\ & \ddots & \\ & & \underset{\sim}{1}_{n_a} \end{bmatrix}}_{a} \begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix} \qquad \underset{\sim}{e} = \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \vdots \\ \varepsilon_{a1} \\ \vdots \\ \varepsilon_{an_a} \end{bmatrix}$$

Goals:

1. Estimate μ, σ^2_A, and σ^2_ε.
2. "Predict" the realized values of the random effects $A_1, \dots, A_a$.
3. Test H_0: $\sigma^2_A = 0$ versus H_1: $\sigma^2_A > 0$.

Exercise: Put equations in matrix form $\mathbf{Y} = \underline{\mathbf{X}}\mathbf{b} + \mathbf{Zu} + \mathbf{e}$.

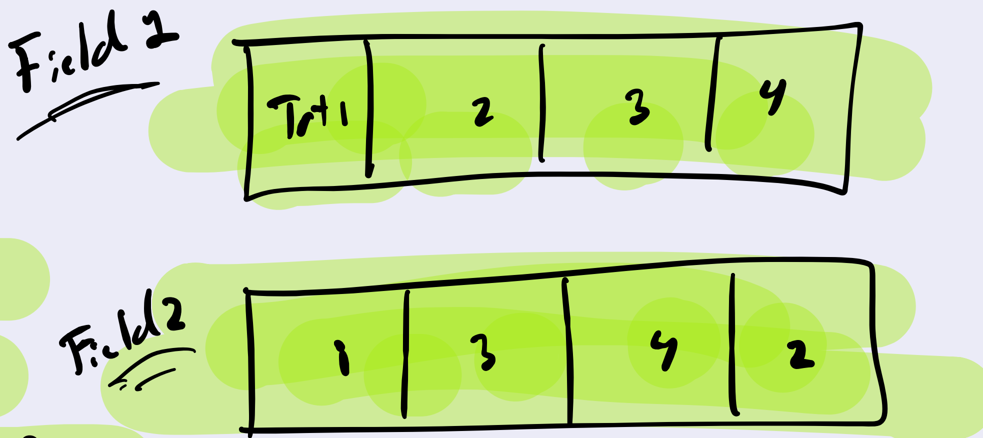
Two-way mixed effects model (Randomized complete block design)

For responses Y_{ijk}, assume

$$Y_{ijk} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + \varepsilon_{ijk}, \quad \boxed{i = 1, \dots, a,} \quad \boxed{j = 1, \dots, b,} \quad \boxed{k = 1, \dots, n_{ij}}$$

field field×trt

where



- μ is a mean
- α_i are treatment effects
- ε_{ijk} are independent Normal$(0, \sigma^2_\varepsilon)$
- B_i are independent Normal$(0, \sigma^2_B)$
- $(\alpha B)_{ij}$ are independent Normal$(0, \sigma^2_{AB})$
- B_i, $(AB)_{ij}$, and ε_{ij} are independent

Exercise: Put equations in matrix form $\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$.

$a = b = 2$ $\quad n_{ij} = n = 2$ $\quad \underset{\sim}{y} = X\underset{\sim}{b} + Z\underset{\sim}{u} + \underset{\sim}{e}$

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \\ 1 & 1 & \\ 1 & 1 & \\ 1 & 1 & \\ 1 & & 1 \\ 1 & & 1 \\ 1 & & 1 \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} 1 & & 1 & & & \\ 1 & & 1 & & & \\ & 1 & & 1 & & \\ & 1 & & 1 & & \\ 1 & & & & 1 & \\ 1 & & & & 1 & \\ & 1 & & & & 1 \\ & 1 & & & & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \end{bmatrix}$$

$\underset{\sim}{y} = X\underset{\sim}{b} + Z\underset{\sim}{u} + \underset{\sim}{e}$

$\underset{\sim}{u} \sim N(\underset{\sim}{0}, G)$, $\underset{\sim}{e} \sim N(\underset{\sim}{0}, R)$ $\quad \underset{\sim}{u} \perp\!\!\!\perp \underset{\sim}{e}$ $\xrightarrow{\text{indep.}}$ $\operatorname{Cov} \underset{\sim}{y} = \underbrace{ZGZ^T + R}_{V}$

$\underset{\sim}{y} \sim N(X\underset{\sim}{b}, V)$

Maximum likelihood estimation of variance components

Let $\mathbf{y} \sim \text{Normal}(\mathbf{Xb}, \mathbf{V})$, where $\mathbf{V} = \mathbf{V}(\boldsymbol{\theta})$ for parameters $(\mathbf{b}, \boldsymbol{\theta}) \in \mathbb{R}^p \times \Theta$. Then

$$(\hat{\mathbf{b}}_{\text{mle}}, \hat{\boldsymbol{\theta}}_{\text{mle}}) = \underset{\boxed{(\mathbf{b},\boldsymbol{\theta}) \in \mathbb{R}^p \times \Theta}}{\operatorname{argmax}} \ \ell(\mathbf{b}, \boldsymbol{\theta}; \mathbf{y}),$$

$\underset{n \times p}{X}$

where

$$\ell(\mathbf{b}, \boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{y} - \mathbf{Xb})^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{Xb})$$

is the log-likelihood function for $(\mathbf{b}, \boldsymbol{\theta})$ based on $\mathbf{y}$.

Exercise: Let $Y_{ij} = \mu + A_i + \varepsilon_{ij}$, $i = 1, \ldots, a$, $j = 1, \ldots, n$, with

$$A_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad \varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

Get MLEs of μ, σ_A^2, σ_ε^2.

$$\underset{\sim}{Y} = X \underset{\sim}{b} + Z \underset{\sim}{u} + \underset{\sim}{e}$$

$$\underset{\sim}{Y} = \underbrace{\underset{\sim}{1}_{na}}_{X} \underbrace{\mu}_{\underset{\sim}{b}} + \underbrace{\begin{bmatrix} \underset{\sim}{1}_n & & \\ & \ddots & \\ & & \underset{\sim}{1}_n \end{bmatrix}}_{\underset{na \times a}{Z}} \underbrace{\begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix}}_{\underset{a \times 1}{\underset{\sim}{u}}} + \underset{\sim}{e}$$

$$\text{Cov}\, \underset{\sim}{u} = \sigma_A^2 I_a$$

$$\text{Cov}\, \underset{\sim}{e} = \sigma_\varepsilon^2 I_{na}$$

$$\text{Cov}\, \underset{\sim}{Y} = Z G Z^T + R$$

$$= \begin{pmatrix} \underset{\sim}{1}_n & & \\ & \ddots & \\ & & \underset{\sim}{1}_n \end{pmatrix} \sigma_A^2 \begin{pmatrix} \underset{\sim}{1}_n^T & & \\ & \ddots & \\ & & \underset{\sim}{1}_n^T \end{pmatrix} + \sigma_\varepsilon^2 I_{na}$$

$$= \sigma_A^2 \begin{pmatrix} \underset{\sim}{1}_n \underset{\sim}{1}_n^T & & \\ & \ddots & \\ & & \underset{\sim}{1}_n \underset{\sim}{1}_n^T \end{pmatrix} + \sigma_\varepsilon^2 \begin{pmatrix} I_n & & \\ & \ddots & \\ & & I_n \end{pmatrix}$$

$$J_n = \underset{\sim}{1}_n \underset{\sim}{1}_n^T = \begin{bmatrix} 1 & 1 & & 1 \\ 1 & 1 & & \\ \vdots & & & \\ 1 & & & 1 \end{bmatrix}$$

$$= \sigma_A^2 \begin{pmatrix} J_n & & \\ & \ddots & \\ & & J_n \end{pmatrix} + \sigma_\varepsilon^2 \begin{pmatrix} I_n & & \\ & \ddots & \\ & & I_n \end{pmatrix}$$

$$= \begin{bmatrix} \sigma_\varepsilon^2 I_n + \sigma_A^2 J_n & & \\ & \ddots & \\ & & \sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \end{bmatrix}$$

$$= I_a \otimes \left[\sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \right] = V$$

Kronecker product

$$V = I_a \otimes \left[\sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \right].$$

$$V^{-1} = I_a \otimes \left[\sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \right]^{-1}$$

4

$$\left[\sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \right]^{-1} = \frac{1}{\sigma_\varepsilon^2} \left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} J_n \right)$$

$$\underline{V^{-1}} = I_a \otimes \frac{1}{\sigma_\varepsilon^2} \left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} J_n \right)$$

Useful: $(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a}(\mathbf{I}_n - \frac{b}{a+nb}\mathbf{J}_n)$ and $|a\mathbf{I}_n + b\mathbf{J}_n| = a^{n-1}(a + nb)$.

$$|\underset{\sim}{V}| = \left| \sigma_\varepsilon^2 I_n + \sigma_A^2 J_n \right|^a \qquad \begin{vmatrix} \square & & & \\ & \square & & \\ & & \ddots & \\ & & & \square \end{vmatrix}$$

$$= \left[(\sigma_\varepsilon^2)^{n-1} (\sigma_\varepsilon^2 + n\sigma_A^2) \right]^a$$

$$\ell(\mathbf{b}, \boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{y} - \mathbf{Xb})^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{Xb})$$

For a given V, maximize in $\underset{\sim}{b}$ is solution to

$$(X^T V^{-1} X)\, \underset{\sim}{b}_{gls} = X^T V^{-1} \underset{\sim}{y} .$$

$$X \underset{\sim}{b}_{gls} = X (X^T V^{-1} X)^- X^T V^{-1} X \underset{\sim}{\hat{b}}_{gls} \qquad X^T V^{-1} \underset{\sim}{y}$$

$$= X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{y}$$

$$y - X \underset{\sim}{b}_{gls} = \underset{\sim}{y} - X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{y}$$

$\Rightarrow$

$$\boxed{(y - X \underset{\sim}{b}_{gls})^T V^{-1} (\underset{\sim}{y} - X \underset{\sim}{b}_{gls})}$$

$$= (y - X(X^T V^{-1} X)^- X^T V^{-1} y) V^{-1} (y - X(X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{y})$$

$$= \underset{\sim}{y}^T V^{-1} \underset{\sim}{y} - 2 y^T V^{-1} X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{y}$$

$$+ \underset{\sim}{y}^T V^{-1} x \boxed{(x^T V^{-1} x)^- x^T V^{-1}} x (x^T V^{-1} x)^- x^T V^{-1} \underset{\sim}{y}$$

$$\boxed{= \underset{\sim}{y}^T V^{-1} \underset{\sim}{y} - y^T V^{-1} x (x^T V^{-1} x)^- x^T V^{-1} \underset{\sim}{y}}$$

$$\ell(\sigma_\varepsilon^2, \sigma_A^2 ; \underset{\sim}{y}) = \text{const.} - \frac{1}{2} \log |V|$$

$$- \frac{1}{2} \left(\underset{\sim}{y}^T V^{-1} \underset{\sim}{y} - y^T V^{-1} x (x^T V^{-1} x)^- \boxed{x^T V^{-1} \underset{\sim}{y}} \right)$$

$$\underset{\sim}{y}^T V^{-1} \underset{\sim}{y} = \left[\underset{\sim}{y}_1^T \cdots \underset{\sim}{y}_a^T\right] \begin{bmatrix} \frac{1}{\sigma_\varepsilon^2} \left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \underbrace{J_n}_{\underset{\sim}{1}_n \underset{\sim}{1}_n^T} \right) & & \\ & \ddots & \\ & & \end{bmatrix} \begin{bmatrix} \underset{\sim}{y}_1 \\ \vdots \\ \underset{\sim}{y}_a \end{bmatrix}$$

$$\underset{na \times 1}{\underset{\sim}{y}} = \begin{bmatrix} \underset{\sim}{y}_1 \\ \vdots \\ \underset{\sim}{y}_a \end{bmatrix} \begin{matrix} \}n \\ \\ \}n \end{matrix}$$

$$= \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \left(\overbrace{\underset{\sim}{y}_i^T \underset{\sim}{y}_i}^{\sum_{j=1}^{n} y_{ij}^2} - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \overbrace{\underset{\sim}{y}_i^T \underset{\sim}{1}_n}^{y_{i\cdot} = n\bar{y}_{i\cdot}} \overbrace{\underset{\sim}{1}_n^T \underset{\sim}{y}_i}^{n\bar{y}_{i\cdot}} \right)$$

$$= \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \sum_{j=1}^{n} y_{ij}^2 - \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} n\bar{y}_{i\cdot}^2$$

$$= \frac{1}{\sigma_\varepsilon^2} \sum_{i=1}^{a} \sum_{j=1}^{n} y_{ij}^2 - \frac{1}{\sigma_\varepsilon^2} \sum_{i=1}^{a} n\bar{y}_{i\cdot}^2$$

$$+ \frac{1}{\sigma_\varepsilon^2} \sum_{i=1}^{a} n\bar{y}_{i\cdot}^2 - \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} n\bar{y}_{i\cdot}^2$$

$$= \frac{1}{\sigma_\varepsilon^2} \boxed{\sum_{i=1}^{a} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{i\cdot} \right)^2} \quad \text{SSE}$$

$$+ \underbrace{\frac{1}{\sigma_\varepsilon^2} \left(1 - \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \right)}_{\frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2}} \sum_{i=1}^{a} n\bar{y}_{i\cdot}^2$$

$$= \frac{1}{\sigma_\varepsilon^2} \text{SSE} + \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} \sum_{i=1}^{n} n\bar{y}_{i\cdot}^2$$

$$X^T V^{-1} X = \left[\underset{\sim}{1}_n^T \cdots \underset{\sim}{1}_n^T \right] \begin{bmatrix} \ddots & & \\ & \frac{1}{\sigma_\varepsilon^2}\left(I - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \overset{\underset{\sim}{1}_n \underset{\sim}{1}_n^T}{J_n} \right) & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \underset{\sim}{1}_n \\ \vdots \\ \underset{\sim}{1}_n \end{bmatrix}$$

$$X = \underset{\sim}{1}_{na}$$

$$= \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \left(\overbrace{\underset{\sim}{1}_n^T \underset{\sim}{1}_n}^{n} - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \overbrace{\underset{\sim}{1}_n^T \underset{\sim}{1}_n}^{n} \overbrace{\underset{\sim}{1}_n^T \underset{\sim}{1}_n}^{n} \right)$$

$$= \frac{na}{\sigma_\varepsilon^2} \left(1 - \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \right)$$

$$= \frac{na}{\sigma_\varepsilon^2} \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_A^2}$$

$$= \frac{na}{\sigma_\varepsilon^2 + n\sigma_A^2}$$

$$X^T V^{-1} \underset{\sim}{y} = \begin{bmatrix} \underset{\sim n}{1}^T \cdots \underset{\sim n}{1}^T \end{bmatrix} \begin{bmatrix} \ddots & & \\ & \frac{1}{\sigma_\varepsilon^2}\left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \overset{\underset{\sim n}{1}\,\underset{\sim n}{1}^T}{J_n} \right) & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \underset{\sim}{y}_1 \\ \vdots \\ \underset{\sim}{y}_a \end{bmatrix}$$

$$= \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} \left(\overset{n\bar{y}_{i\cdot}}{\overbrace{\underset{\sim n}{1}^T \underset{\sim}{y}_i}} - \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \, n\bar{y}_{i\cdot} \right)$$

$$= \sum_{i=1}^{a} \frac{1}{\sigma_\varepsilon^2} n\bar{y}_{i\cdot} \left(1 - \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \right)$$

$$= \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} \underbrace{\sum_{i=1}^{a} n\bar{y}_{i\cdot}}_{y_{\cdot\cdot} = na\bar{y}_{\cdot\cdot}}$$

$$= \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} na\bar{y}_{\cdot\cdot}$$

$$\underset{\sim}{y}^T V^{-1} X \left(X^T V^{-1} X \right)^{-1} X^T V^{-1} \underset{\sim}{y}$$

$$= \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} na\bar{y}_{\cdot\cdot} \left[\frac{na}{\sigma_\varepsilon^2 + n\sigma_A^2} \right]^{-1} \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} na\bar{y}_{\cdot\cdot}$$

$$= \frac{na}{\sigma_\varepsilon^2 + n\sigma_A^2} \bar{y}_{\cdot\cdot}^2$$

$$\ell(\sigma_\varepsilon^2, \sigma_A^2 ; \underset{\sim}{y}) = \text{const.} - \frac{1}{2} \log |V|$$

$$- \frac{1}{2} \left(\underline{\underset{\sim}{y}^T V^{-1} \underset{\sim}{y}} - \underbrace{y^T V^{-1} x (x^T V^{-1} x)^{-} \boxed{x^T V^{-1} y}} \right)$$

$$= \text{const.} - \frac{1}{2} \log \left(\left[(\sigma_\varepsilon^2)^{n-1} (\sigma_\varepsilon^2 + n \sigma_A^2) \right]^a \right)$$

$$- \frac{1}{2} \left(\frac{1}{\sigma_\varepsilon^2} SSE + \underbrace{\frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} \sum_{i=1}^{a} n \bar{y}_{i\cdot}^2 - \frac{na}{\sigma_\varepsilon^2 + n\sigma_A^2} \bar{y}_{\cdot\cdot}^2}_{\frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} \underbrace{\sum_{i=1}^{a} n (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}_{SSA}} \right)$$

$$\ell(\sigma_\varepsilon^2, \sigma_A^2 ; y)$$

$$= \text{const.} - \frac{1}{2} a(n-1) \log \sigma_\varepsilon^2 - \frac{1}{2} a \log (\sigma_\varepsilon^2 + n \sigma_A^2)$$

$$- \frac{1}{2} \left(\frac{1}{\sigma_\varepsilon^2} SSE + \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2} SSA \right)$$

Set $\lambda = \sigma_\varepsilon^2 + n \sigma_A^2$

$$\frac{\partial}{\partial \sigma_\varepsilon^2} \ell(\ ) = - \frac{a(n-1)}{2} \frac{1}{\sigma_\varepsilon^2} + \frac{1}{2 \sigma_\varepsilon^4} SSE \overset{set}{=} 0$$

$$\Rightarrow \quad \hat{\sigma}_\varepsilon^2 = \frac{SSE}{a(n-1)}$$

$$\frac{\partial}{\partial \lambda} \ell(\cdot) = -\frac{1}{2}\frac{a}{\lambda} + \frac{1}{2}\frac{1}{\lambda^2} SSA \overset{set}{=} 0$$

$$\Rightarrow \quad \hat{\lambda} = \frac{SSA}{a}$$

$$\hat{\sigma}_\varepsilon^2 + n\hat{\sigma}_A^2 = \frac{SSA}{a}$$

$$\boxed{\begin{aligned} \hat{\sigma}_\varepsilon^2 &= \frac{SSE}{a(n-1)} \\ \hat{\sigma}_A^2 &= \frac{1}{n}\left[\frac{SSA}{a} - \frac{SSE}{a(n-1)}\right] \end{aligned}}$$

possibly < 0.

Exercise: Let $Y_{ij} = \mu + A_i + \varepsilon_{ij}$, $i = 1, \dots, a$, $j = 1, \dots, n$, with

$$A_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad \varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

Show that the MLE of $(\mu, \sigma_A^2, \sigma_\varepsilon^2)$ is given by

MLE

$$(\hat{\mu}, \hat{\sigma}_A^2, \hat{\sigma}_\varepsilon^2) = \begin{cases} \left(\bar{y}_{..}, \dot{\sigma}_A^2, \dfrac{\text{SSE}}{a(n-1)}\right), & \text{if } \dot{\sigma}_A^2 > 0 \\ \left(\bar{y}_{..}, 0, \dfrac{\text{SSE} + \text{SSA}}{an}\right), & \underline{\text{if } \dot{\sigma}_A^2 \leq 0} \end{cases}$$

where $\text{SSA} = n\sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$ and $\text{SSE} = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$ and

$$\boxed{\dot{\sigma}_A^2 = \frac{1}{n}\left[\frac{\text{SSA}}{a} - \frac{\text{SSE}}{a(n-1)}\right].}$$

Useful: $(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a}(\mathbf{I}_n - \frac{b}{a+nb}\mathbf{J}_n)$ and $|a\mathbf{I}_n + b\mathbf{J}_n| = a^{n-1}(a + nb)$.

Exercise: Show that the MLEs can be found in two steps:

1. Find $\hat{\boldsymbol{\theta}}_{\text{mle}}$ by maximizing (in many cases numerically) the profile likelihood

$$\ell(\boldsymbol{\theta};\mathbf{y}) = \text{const.} - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\mathbf{y}^T(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}^T\mathbf{V}^{-1})\mathbf{y}.$$

2. Find $\hat{\mathbf{b}}_{\text{mle}}$ satisfying $(\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X})\hat{\mathbf{b}}_{\text{mle}} = \mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{y}$, where $\hat{\mathbf{V}} = \mathbf{V}(\hat{\boldsymbol{\theta}}_{\text{mle}})$.

$Y_1,\dots,Y_n \overset{\text{ind}}{\sim} N(\mu,\sigma^2)$

$\hat{\sigma}^2_{\text{mle}} = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ $\qquad$ $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2$

Restricted maximum likelihood REML principle

The *REML* approach to estimating variance comps. is to first remove fixed effects.

That is, for $\mathbf{y} \sim \text{Normal}(\mathbf{X}\mathbf{b}, \mathbf{V}(\boldsymbol{\theta}))$, estimate $\boldsymbol{\theta}$ based on $(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{y}$.

More precisely:

$\underset{\sim}{y} = x\underset{\sim}{b} + Z\underset{\sim}{u} + \underset{\sim}{\varepsilon}$

1. Reparameterize $\mathbf{X}$ so that it has full-column rank (simplifies our exposition).
2. Then let $\boxed{(\mathbf{I} - \mathbf{P}_\mathbf{X}) = \mathbf{A}\mathbf{A}^T}$ with $\mathbf{A}^T\mathbf{A} = \mathbf{I}_{n-p}$, with $p = \operatorname{rank} \mathbf{X}$.
3. Then the REML approach estimates $\boldsymbol{\theta}$ based on $\boxed{\mathbf{w} = \mathbf{A}^T\mathbf{y}.}$

Exercise:

1. Obtain the "restricted" log-likelihood $\boxed{\ell_R(\boldsymbol{\theta}; \mathbf{w}) \text{ for } \boldsymbol{\theta} \text{ based on } \mathbf{w}.}$
2. Show that $\ell_R(\boldsymbol{\theta}; \mathbf{w})$ is equal to

$$\ell_R(\boldsymbol{\theta}; \mathbf{y}) = \text{const.} - \frac{1}{2}\log|\mathbf{V}| - \boxed{\frac{1}{2}\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}|} - \frac{1}{2}(\mathbf{y} - \mathbf{X}\mathbf{b}_{\text{gls}})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\mathbf{b}_{\text{gls}}),$$

where $\mathbf{b}_{\text{gls}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$.

REML log likelihood has one extra term.

See pgs. 192,193 of Monahan (2008).

Summary of REML approach

For $\mathbf{y} \sim \text{Normal}(\mathbf{X}\mathbf{b}, \mathbf{V}(\boldsymbol{\theta}))$, estimate $\boldsymbol{\theta}$ and $\mathbf{b}$ in two steps:

1. Obtain $\hat{\boldsymbol{\theta}}_R$ by maximizing the restricted log-likelihood

$$\ell_R(\boldsymbol{\theta};\mathbf{y}) = \text{const.} \boxed{-\frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}| - \frac{1}{2}(\mathbf{y}-\mathbf{X}\mathbf{b}_{\text{gls}})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\mathbf{b}_{\text{gls}})}$$

2. Set $\hat{\mathbf{V}} = \mathbf{V}(\hat{\boldsymbol{\theta}}_R)$ and compute $\hat{\mathbf{b}}_R = (\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1}\mathbf{y}$.

Can be useful to write $\ell_R(\boldsymbol{\theta};\mathbf{y})$ (scaling and ignoring the constant) as

$$\boxed{\ell_R(\boldsymbol{\theta};\mathbf{y}) = -\log|\mathbf{V}| - \log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}| - \mathbf{y}^T(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}^T\mathbf{V}^{-1})\mathbf{y}.}$$

to find REML's, use this.

Exercise: Let $Y_{ij} = \mu + A_i + \varepsilon_{ij}$, $i = 1, \dots, a$, $j = 1, \dots, n$, with

$$A_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad \varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

Show that the REML estimators of $(\sigma_A^2, \sigma_\varepsilon^2)$ are given by

$$(\hat{\sigma}_A^2, \hat{\sigma}_\varepsilon^2) = \begin{cases} \left(\dot{\sigma}_A^2, \dfrac{\text{SSE}}{a(n-1)}\right), & \text{if } \dot{\sigma}_A^2 > 0 \\ \left(0, \dfrac{\text{SSE} + \text{SSA}}{an - 1}\right), & \text{if } \dot{\sigma}_A^2 \leq 0, \end{cases}$$

where $\text{SSA} = n\sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$ and $\text{SSE} = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$ and

$$\dot{\sigma}_A^2 = \frac{1}{n}\left[\frac{\text{SSA}}{a-1} - \frac{\text{SSE}}{a(n-1)}\right].$$

$$\ell_R(\theta; \mathbf{y}) = \boxed{-\log|\mathbf{V}|} - \boxed{\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}|} - \boxed{\mathbf{y}^T(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}^T\mathbf{V}^{-1})\mathbf{y}}.$$

$$\log\left|X^T V^{-1} X\right| = \log\left(\frac{na}{\sigma_\varepsilon^2 + n\sigma_A^2}\right)$$

$$=$$

$$\ell(\sigma_\varepsilon^2, \sigma_A^2; y)$$

$$= \boxed{-\;a(n-1)\log\sigma_\varepsilon^2 - a\log(\sigma_\varepsilon^2 + n\sigma_A^2)}$$

$$\boxed{-\;\underbrace{\log(na)}_{\text{const.}} + \log(\sigma_\varepsilon^2 + n\sigma_A^2)}$$

$$-\;\boxed{\left(\frac{1}{\sigma_\varepsilon^2}SSE + \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2}SSA\right)}$$

$$= -\;a(n-1)\log\sigma_\varepsilon^2 - (a-1)\log(\sigma_\varepsilon^2 + n\sigma_A^2)$$

$$-\;\frac{1}{\sigma_\varepsilon^2}SSE - \frac{1}{\sigma_\varepsilon^2 + n\sigma_A^2}SSA + \text{const.}$$

$$\vdots$$

1. Random and mixed effects models
2. Variance component estimation, ML and REML
3. Prediction of realized values of random effects
4. Example of ANOVA and hypothesis testing in mixed models

$$Y_{ij} = \mu + A_i + \varepsilon_{ij}$$

$$A_1, \dots, A_a \overset{ind}{\sim} N(0, \sigma_A^2)$$

$\mathbb{E}[v \mid \underset{\sim}{y}]$

Let $\mathbf{c}^T\mathbf{b}$ be an estimable contrast in the model $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$.

Consider predicting the realized value of $v = \mathbf{c}^T\mathbf{b} + \mathbf{d}^T\mathbf{u}$.

rand variable fixed random

Best linear unbiased predictor (BLUP)

1. A predictor $\boxed{\tilde{v}}$ of v is an *unbiased predictor* if $\mathbb{E}\tilde{v} = \mathbb{E}v$.
2. A *linear predictor* of v has the form $\tilde{v} = a_0 + \mathbf{a}^T\mathbf{y}$.
3. A linear predictor $\boxed{\tilde{v}}$ is called the *best linear unbiased predictor* of $\boxed{v}$ if

$$\underline{\mathbb{E}(v - \tilde{v})^2} \leq \mathbb{E}(v - \tilde{v}^*)^2$$

for all linear unbiased predictors $\tilde{v}^*$ of v.

Exercise: In the one-way random effects model, give $\mathbf{c}$ and $\mathbf{d}$ such that the corresponding v are the realized treatment means.

$$Y_{ij} = \mu + A_i + \varepsilon_{ij}$$

$$\underset{\sim}{Y} = X \underset{\sim}{b} + Z \underset{\sim}{u} + \underset{\sim}{\varepsilon}$$

$$\begin{bmatrix} \underset{\sim}{Y}_1 \\ \vdots \\ \underset{\sim}{Y}_a \end{bmatrix} = \begin{bmatrix} \underset{\sim}{1}_n \\ \vdots \\ \underset{\sim}{1}_n \end{bmatrix} [\mu] + \begin{bmatrix} \underset{\sim}{1}_n & & \\ & \ddots & \\ & & \underset{\sim}{1}_n \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix} + \begin{bmatrix} \underset{\sim}{\varepsilon}_1 \\ \vdots \\ \underset{\sim}{\varepsilon}_a \end{bmatrix}$$

Set $\underline{\underline{V = \mu + A_i}} = \underset{\sim}{c}^T \underset{\sim}{b} + \underset{\sim}{d}^T \underset{\sim}{u}$

$\underset{1\times 1}{\underset{\sim}{c}} = [1] \qquad \underset{\sim}{d} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow$ position i

$\underset{\sim}{u}$

predictor $\hat{\underset{\sim}{u}}(\underset{\sim}{y})$, $\underset{\sim}{y}$ observed data.

Exercise: Let $\mathbf{u}$ be a random vector we wish to predict and $\hat{\mathbf{u}}$ be a predictor based on an observed random vector $\mathbf{y}$. Show that the criterion

$$\mathbb{E}(\hat{\mathbf{u}} - \mathbf{u})^T \mathbf{A} (\hat{\mathbf{u}} - \mathbf{u})$$

is minimized when $\hat{\mathbf{u}} = \mathbb{E}[\mathbf{u}|\mathbf{y}]$, where $\mathbf{A}$ is some positive definite matrix.

The best predictor of $\underset{\sim}{u}$ based on $\underset{\sim}{y}$ is $\hat{\underset{\sim}{u}} = \mathbb{E}[\underset{\sim}{u}|\underset{\sim}{y}]$.

$Y_{ij} = \mu + A_i + \varepsilon_{ij}$

predict $\mu + A_i$?

Result (BLUP for the mixed model)

In the mixed model setup $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$, *where*

$$\mathbf{u} \overset{\text{ind}}{\sim} \text{Normal}(\mathbf{0}, \mathbf{G}) \quad \textit{and} \quad \mathbf{e} \overset{\text{ind}}{\sim} \text{Normal}(\mathbf{0}, \mathbf{R}),$$

the BLUP for $\boxed{v = \mathbf{c}^T\mathbf{b} + \mathbf{d}^T\mathbf{u},}$ *where* $\mathbf{c}^T\mathbf{b}$ *is estimable, is given by*

$$\tilde{v} = \mathbf{c}^T\hat{\mathbf{b}}_{\text{gls}} + \mathbf{d}^T\mathbf{G}\mathbf{Z}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}_{\text{gls}}).$$

Note: $\tilde{v}$ is infeasible (need $\mathbf{V}$). In practice replace $\mathbf{V}$ with $\hat{\mathbf{V}}$. No longer BLUP 🤷‍♂️.

Prove the above in two steps:

1. Show that $\mathbb{E}[v|\mathbf{y}] = \mathbf{c}^T\mathbf{b} + \mathbf{d}^T\mathbf{G}\mathbf{Z}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\mathbf{b})$.
2. Minimize $\mathbb{E}(v - (a_0 + \mathbf{a}^T\mathbf{y}))^2$, ensuring that $a_0 + \mathbf{a}^T\mathbf{y}$ is unbiased.

$$\underset{\sim}{v} = \underset{\sim}{1}_{na}\,\mu + \begin{bmatrix} \underset{\sim}{1}_n & & \\ & \ddots & \\ & & \underset{\sim}{1}_n \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix} + \underset{\sim}{e}$$

Exercise: Let $Y_{ij} = \mu + A_i + \varepsilon_{ij}$, $i = 1, \dots, a$, $j = 1, \dots, n$, with

$$A_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad \varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

Show that the BLUP for $v_i = \boxed{\mu + A_i}$ is given by

$$\tilde{v}_i = \left(\frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \right) \bar{y}_{i.} + \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \right) \bar{y}_{..}, \quad i = 1, \dots, a.$$

$$v_i = \underset{\sim}{c}^T \underset{\sim}{b} + \underset{\sim}{d}^T \underset{\sim}{u}$$

$$\underset{\sim}{c} = 1 \qquad \underset{\sim}{d} = \underset{\sim}{e}_i$$

$$\underset{\sim}{b} = \mu \qquad \underset{\sim}{u} = \begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix}$$

$$\hat{\mu} = \hat{\underset{\sim}{b}}_{gls} = (X^T V^{-1} X)^{-1} X^T V^{-1} \underline{y}$$

$$= \frac{n\sigma_A^2 + \sigma_\varepsilon^2}{na} \quad \frac{na}{n\sigma_A^2 + \sigma_\varepsilon^2} \quad \bar{y}_{\cdot\cdot}$$

$$= \bar{y}_{\cdot\cdot}$$

$$\tilde{v} = \overset{\hat{\mu}}{\mathbf{c}^T\hat{\mathbf{b}}_{\mathrm{gls}}} + \mathbf{d}^T\underline{\mathbf{G}}\mathbf{Z}^T\mathbf{V}^{-1}(\mathbf{y} - \overset{\underset{\sim}{1}_{na}\bar{y}_{\cdot\cdot}}{\mathbf{X}\hat{\mathbf{b}}_{\mathrm{gls}}}).$$

$$\tilde{v}_i = \bar{y}_{\cdot\cdot} + \underset{\sim}{e}_i^T [\sigma_A^2 I_a] \begin{bmatrix} \underset{\sim}{1}_n^T & & \\ & \ddots & \\ & & \underset{\sim}{1}_n^T \end{bmatrix} \begin{bmatrix} \ddots & & \\ & \frac{1}{\sigma_\varepsilon^2}\left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} J_n\right) & \\ & & \ddots \end{bmatrix}$$

$$\times \begin{bmatrix} \underset{\sim}{y}_1 - \underset{\sim}{1}_n \bar{y}_{\cdot\cdot} \\ \vdots \\ \underset{\sim}{y}_a - \underset{\sim}{1}_n \bar{y}_{\cdot\cdot} \end{bmatrix}$$

$$= \bar{y}_{\cdot\cdot} + \sigma_A^2 \, \underset{\sim}{1}_n^T \left[\frac{1}{\sigma_\varepsilon^2}\left(I_n - \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} J_n\right)\right] (\underset{\sim}{y}_i - \underset{\sim}{1}_n \bar{y}_{\cdot\cdot})$$

$$= \bar{y}_{\cdot\cdot} + \frac{\sigma_A^2}{\sigma_\varepsilon^2}\left[(\underset{\sim}{1}^T \underset{\sim}{y}_i - n\bar{y}_{\cdot\cdot}) - \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} (\underset{\sim}{1}^T \underset{\sim}{y}_i - n\bar{y}_{\cdot\cdot}) \right]$$

$$= \bar{y}_{\cdot\cdot} + \frac{\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2} \, \overset{n\bar{y}_{i\cdot}}{\overbrace{\underset{\sim}{1}^T \underset{\sim}{y}_i}} - n\bar{y}_{\cdot\cdot}$$

$$= \bar{y}_{\cdot\cdot} + \frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2}\left(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}\right)$$

$$= \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + n\sigma_A^2}\right)\bar{y}_{\cdot\cdot} + \left(\frac{n\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_A^2}\right)\bar{y}_{i\cdot}$$

Exercise: Let $Y_{ij} = \alpha + \beta x_{ij} + A_i + B_i x_{ij} + \varepsilon_{ij}$, $i = 1, \dots, n$, $j = 1, \dots, r$, where

$$A_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad B_i \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_B^2), \quad \varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

With the `sleepstudy` data in the R package `lme4`, obtain the following:

1. The REML estimates for σ_A^2, σ_B^2, and σ_ε^2.
2. Estimates of the fixed effects parameters α and β.
3. Predictions for the realized values of the random effects.

```
# pull data from lme4 package
library(lme4)
y <- sleepstudy$Reaction/10 # rescale for numerical stability with optim()
x <- sleepstudy$Days

# construct design matrices X and Z
n <- 18; r <- 10
X <- cbind(rep(1,n*r),x)
Z <- matrix(0,n*r,2*n)
for(i in 1:n){

  ind <- ((i-1)*r + 1):(i*r)
  Z[ind,i] <- 1
  Z[ind,n + i] <- x[ind]

}
```

$$\ell_R(\boldsymbol{\theta};\mathbf{y}) = \boxed{-\log|\mathbf{V}|} - \boxed{\log|\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}|} - \boxed{\mathbf{y}^T(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}^T\mathbf{V}^{-1})\mathbf{y}}.$$

```
# define a function which is the negative restricted log-likelihood
negllR_slp <- function(th,y,X,Z){

  sgA <- th[1]; sgB <- th[2]; sge <- th[3]

  # build V
  G <- cbind(diag(rep(sgA^2,n)),matrix(0,n,n))
  G <- rbind(G,cbind(matrix(0,n,n),diag(rep(sgB^2,n))))
  V <- Z %*% G %*% t(Z) + diag(rep(sge^2,n*r))

  # compute restricted log-likelihood
  Vin <- solve(V); A <- t(X) %*% Vin %*% X
  d1 <- det(V); d2 <- det(A)
  quad <- t(y) %*% ( Vin - Vin %*% X %*% solve(A) %*% t(X) %*% Vin) %*% y
  return(log(d1) + log(d2) + quad)

}
```

$V = ZGZ^T + I\sigma_\varepsilon^2$

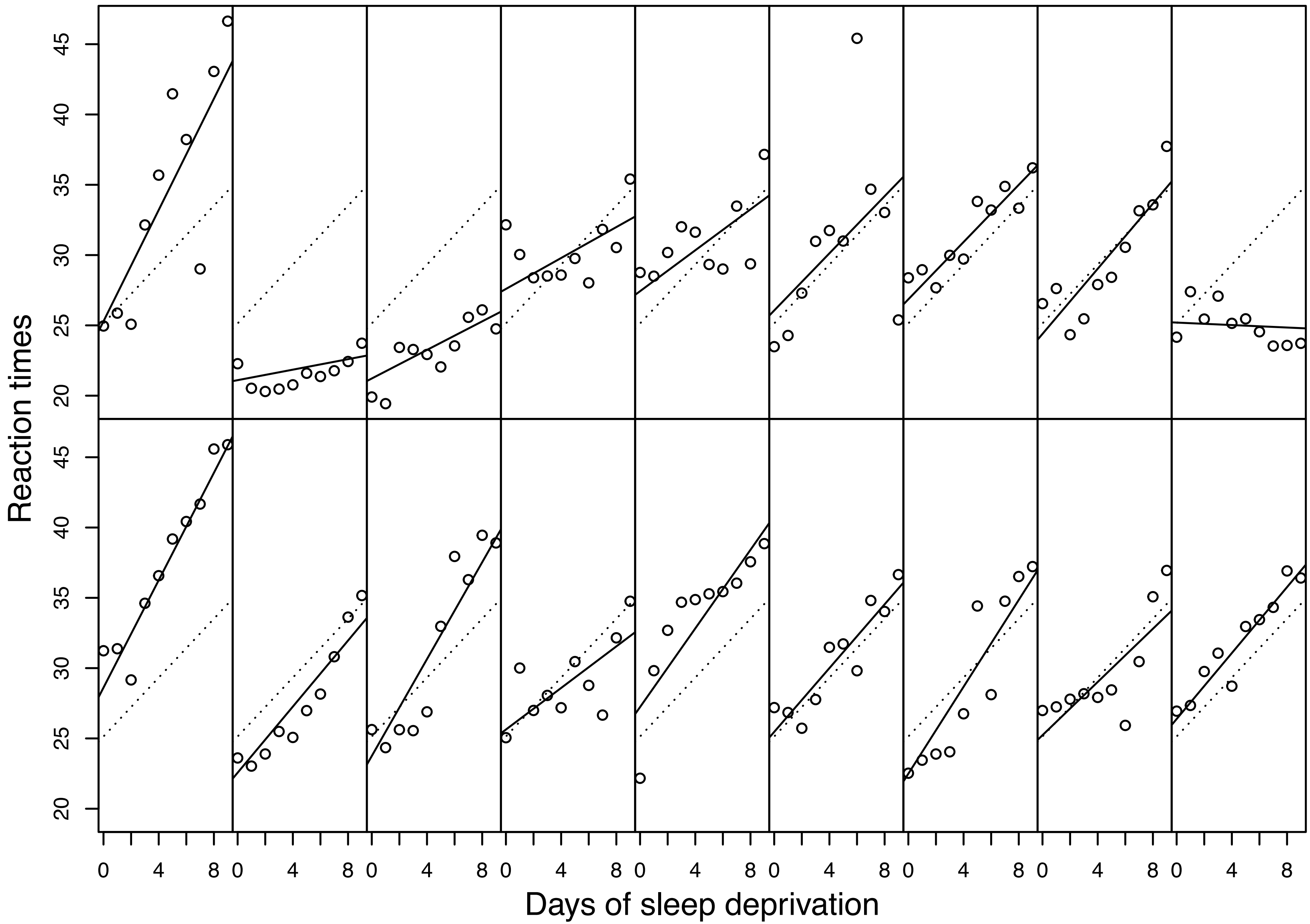
```
# use optim() to find the REML estimators (this is a clumsy way)
optim_out <- optim(par = c(1,1,1), fn = negllR_slp, y = y, X = X, Z = Z)
th_hat <- optim_out$par
sgA_hat <- th_hat[1]; sgB_hat <- th_hat[2]; sge_hat <- th_hat[3]
```

$\hat{V}$

$$\boxed{(X^T\hat{V}^{-1}X)\,\hat{b}_{gls} = X^T\hat{V}^{-1}\underset{\sim}{y}}$$

```
# construct V_hat
G_hat <- cbind(diag(rep(sgA_hat^2,n)),matrix(0,n,n))
G_hat <- rbind(G_hat,cbind(matrix(0,n,n),diag(rep(sgB_hat^2,n))))
V_hat <- Z %*% G_hat %*% t(Z) + diag(rep(sge_hat^2,n*r))

# compute estimated fixed effects
b_hat <- solve(t(X) %*% solve(V_hat) %*% X) %*% t(X) %*% solve(V_hat)%*% y
A_hat <- b_hat[1]
B_hat <- b_hat[2]

# obtain predicted values of A and B (which is u)
u_pr <- G_hat %*% t(Z) %*% solve(V_hat) %*% ( y - X %*% b_hat)
A_pr <- u_pr[1:n]
B_pr <- u_pr[(n+1):(2*n)]
```

```
# compare output to that of the lmer function from the lme4 package
summary(lmer(Reaction/10 ~ Days + (Days | Subject), sleepstudy))

# make plots
par(mfrow = c(2,9),mar = c(0,0,0,0), oma = c(4,4,1,1))
xlims <- range(x); ylims <- range(y)
for( i in 1:n){

  ind <- ((i-1)*r + 1):(i*r)
  plot(y[ind]~sleepstudy$Days[ind],xlim = xlims, ylim = ylims,
       xaxt = "n",yaxt = "n")

  if(i %in% c(1,10)) axis(side = 2)
  if(i %in% c(10:18)) axis(side = 1)

  abline(A_hat + A_pr[i], B_hat + B_pr[i])
  abline(A_hat,B_hat,lty = 3)

}

mtext(side = 1, outer = TRUE, text = "Days of sleep deprivation",line=2.5)
mtext(side = 2, outer = TRUE, text = "Reaction times", line = 2.5)
```


Reaction times
20
25
30
35
40
45
0
4
8
Days of sleep deprivation

Exercise: Let $\boxed{Y_{ij}} = \boxed{\mu + A_i + \varepsilon_{ij},}$ $i = 1, \ldots, a$, $j = 1, \ldots, n$, with

$$A_i \stackrel{\text{ind}}{\sim} \text{Normal}(0, \sigma_A^2), \quad \varepsilon_{ij} \stackrel{\text{ind}}{\sim} \text{Normal}(0, \sigma_\varepsilon^2).$$

Define the sums of squares

$$\text{SST} = \textstyle\sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2, \ \text{SSA} = n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2, \ \text{SSE} = n \sum_{i=1}^a (y_{ij} - \bar{y}_{i.})^2$$

as well as the mean squares MSA = SSA $/(a-1)$ and MSE = SSE $/(a(n-1))$.

1. Show that SSA $/(\sigma_\varepsilon^2 + n\sigma_A^2) \sim \chi^2_{a-1}$ ✓
2. Show that SSE $/\sigma_\varepsilon^2 \sim \chi^2_{a(n-1)}$ ✓
3. Give $\mathbb{E}$ MSA and $\mathbb{E}$ MSE. ✓
4. Give ANOVA (method of moments) estimators of σ_A^2 and σ_ε^2.
5. Give the distribution of MSA / MSE.
6. Give a size-α test for H_0: $\sigma_A^2 = 0$ based on MSA / MSE. Give power function.

$\mathbb{E}\,\text{MSE} = \sigma_\varepsilon^2$

$\mathbb{E}\,\text{MSA} = n\sigma_A^2 + \sigma_\varepsilon^2$

$$\boldsymbol{y} = X\boldsymbol{b} + Z\boldsymbol{u} + \boldsymbol{e}$$

$$= \underbrace{\boldsymbol{1}_{na}}_{X} \underbrace{\mu}_{\boldsymbol{b}} + \underbrace{\begin{bmatrix} \boldsymbol{1}_n & & \\ & \ddots & \\ & & \boldsymbol{1}_n \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} A_1 \\ \vdots \\ A_a \end{bmatrix}}_{\boldsymbol{u}} + \boldsymbol{e}$$

$$P_{\boldsymbol{1}} \boldsymbol{y} = \boldsymbol{1}(\boldsymbol{1}^T \boldsymbol{1})^{-1} \boldsymbol{1}^T \boldsymbol{y} = \boldsymbol{1}_{na} \bar{y}_{..}$$

$$P_Z y = \begin{bmatrix} \boldsymbol{1}_n & & \\ & \ddots & \\ & & \boldsymbol{1}_n \end{bmatrix} \begin{pmatrix} n^{-1} & & \\ & \ddots & \\ & & n^{-1} \end{pmatrix} \begin{bmatrix} \boldsymbol{1}_n^T & & \\ & \ddots & \\ & & \boldsymbol{1}_n^T \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_a \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}_n \bar{y}_{1.} \\ \vdots \\ \boldsymbol{1}_n \bar{y}_{a.} \end{bmatrix}$$

$$(I - P_{\boldsymbol{1}})\boldsymbol{y} = \boldsymbol{y} - \boldsymbol{1}_{na} \bar{y}_{..} \quad \Rightarrow \quad y^T (I - P_1)\boldsymbol{y} = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$$

$$(P_Z - P_{\boldsymbol{1}})\boldsymbol{y} = \begin{bmatrix} \boldsymbol{1} \bar{y}_{1.} \\ \vdots \\ \boldsymbol{1} \bar{y}_{a.} \end{bmatrix} - \boldsymbol{1}_{na} \bar{y}_{..} = \begin{bmatrix} \boldsymbol{1}_n (\bar{y}_{1.} - \bar{y}_{..}) \\ \vdots \\ \boldsymbol{1}_n (\bar{y}_{a.} - \bar{y}_{..}) \end{bmatrix}$$

$$\Rightarrow \quad \boldsymbol{y}^T (P_Z - P_{\boldsymbol{1}}) \boldsymbol{y} = \sum_{i=1}^{n} n (\bar{y}_i - \bar{y}_{..})^2$$

$$(I - P_Z)\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_a \end{bmatrix} - \begin{bmatrix} \boldsymbol{1}_n \bar{y}_{1.} \\ \vdots \\ \boldsymbol{1}_n \bar{y}_{a.} \end{bmatrix}$$

$$\Rightarrow \quad \boldsymbol{y}^T (I - P_Z) \boldsymbol{y} = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2$$

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \underset{\sim}{y}^T (I - P_1)\underset{\sim}{y}$$

$$SSA = n\sum_{i=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2 = \underset{\sim}{y}^T (P_Z - P_{\underset{\sim}{1}})\underset{\sim}{y}$$

$$SSE = \sum_{i=1}^{n}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2 = \underset{\sim}{y}^T (I - P_Z)\underset{\sim}{y}$$

Look at SSE:

$$\underset{a\times 1}{\underset{\sim}{u}} \sim N(0, G) \qquad \underset{na\times 1}{\underset{\sim}{e}} \sim N(0, R)$$

$$G = \sigma_A^2 I_a \qquad R = \sigma_\varepsilon^2 I_{na}$$

$$\underset{\sim}{y} \sim N(\underset{\sim}{1}_{na}\mu, V), \quad V = ZGZ^T + R$$

$$= \sigma_A^2 ZZ^T + \sigma_\varepsilon^2 I_{na}$$

$$\underset{\sim}{y}^T A \underset{\sim}{y} \sim \chi_s^2 \quad (\phi = \ldots) \quad \text{if } AV \text{ idemp with rank } s.$$

$$A = \frac{1}{\sigma_\varepsilon^2}(I - P_Z)$$

$$AV = \frac{1}{\sigma_\varepsilon^2}(I - P_Z)\left[\sigma_A^2 ZZ^T + \sigma_\varepsilon^2 I_{na}\right]$$

$$= (I_{na} - P_Z), \text{ idemp with rank}$$

$$na - a = a(n-1)$$

$$\frac{\underset{\sim}{y}^T (I - P_Z)\underset{\sim}{y}}{\sigma_\varepsilon^2} \sim \chi^2_{a(n-1)}\Big(\phi = \underbrace{\mu \underset{\sim}{1}_{na}^T (I - P_Z)\underset{\sim}{1}_{na}\mu \frac{1}{\sigma_\varepsilon^2}}_{=0}\Big)$$

$$\underset{\sim}{1}_{na} \in \text{Col } Z$$

$$= \chi^2_{a(n-1)}$$

SSA

$$SSA = n \sum_{i=1}^{n} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \underset{\sim}{y}^T (P_Z - P_{\underset{\sim}{1}}) \underset{\sim}{y}$$

$$\downarrow$$

$$P_{\underset{\sim}{1}_a} \underbrace{\begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix}}_{a \times 1} = \underset{\sim}{1}_a (\underset{\sim}{1}_a^T \underset{\sim}{1}_a) \underset{\sim}{1}_a^T \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix}$$

$$= \underset{\sim}{1}_a \frac{1}{a} \sum_{i=1}^{a} \bar{y}_{i\cdot}$$

$$= \underset{\sim}{1}_a \frac{1}{na} \sum_{i=1}^{a} n\bar{y}_{i\cdot}$$

$$= \underset{\sim}{1}_n \bar{y}_{\cdot\cdot}$$

$$(I_a - P_{\underset{\sim}{1}_a}) \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix} = \begin{bmatrix} \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{a\cdot} - \bar{y}_{\cdot\cdot} \end{bmatrix}$$

$$n \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix}^T (I_a - P_{\underset{\sim}{1}_a}) \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix} = n \sum_{i=1}^{n} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = SSA.$$

$$= \underset{\sim}{y}^T (P_Z - P_{\underset{\sim}{1}}) \underset{\sim}{y}$$

$$\sqrt{n}\begin{bmatrix}\bar{y}_{1\cdot}\\ \vdots \\ \bar{y}_{a\cdot}\end{bmatrix} \sim N\left(\sqrt{n}\,\underset{\sim}{1}_a \mu,\ I_a\left(n\sigma_A^2 + \sigma_\varepsilon^2\right)\right)$$

$$Y_{ij} = \mu + A_i + \varepsilon_{ij}$$

$$\begin{aligned}
\mathrm{Var}\left(\bar{y}_{i\cdot}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} y_{ij}\right)\\
&= \mathrm{Var}\left(\frac{1}{n}\,\underset{\sim}{1}_n^T \underset{\sim}{y}_i\right)\\
&= \frac{1}{n^2}\,\underset{\sim}{1}_n^T \,\mathrm{Cov}\left(\underset{\sim}{y}_i\right)\underset{\sim}{1}_n\\
&= \frac{1}{n^2}\,\underset{\sim}{1}_n^T\left(\sigma_A^2 J_n + \sigma_\varepsilon^2 I_n\right)\underset{\sim}{1}_n\\
&= \frac{1}{n^2}\left(n^2\sigma_A^2 + n\sigma_\varepsilon^2\right)\\
&= \frac{1}{n}\left(n\sigma_A^2 + \sigma_\varepsilon^2\right)
\end{aligned}$$

$$\mathrm{Cov}\left(\underset{\sim}{y}_i\right) = \sigma_A^2 J_n + \sigma_\varepsilon^2 I_n$$

$$\frac{SSA}{n\sigma_A^2+\sigma_\varepsilon^2} = \frac{\sqrt{n}\begin{bmatrix}\bar{y}_{1\cdot}\\ \vdots \\ \bar{y}_{a\cdot}\end{bmatrix}^T\left(I_a - P_{\underset{\sim}{1}_a}\right)\sqrt{n}\begin{bmatrix}\bar{y}_{1\cdot}\\ \vdots \\ \bar{y}_{a\cdot}\end{bmatrix}}{\left(n\sigma_A^2+\sigma_\varepsilon^2\right)} \sim \chi^2_{a-1}\left(\phi = \overbrace{\frac{\mu\,\underset{\sim}{1}_a^T\left(I_a - P_{\underset{\sim}{1}_a}\right)\underset{\sim}{1}_a\,\mu}{\left(n\sigma_A^2+\sigma_\varepsilon^2\right)}}^{=0}\right)$$

$$\mathrm{rank}\left(I_a - P_{\underset{\sim}{1}_a}\right) = a-1$$

3. Give $\mathbb{E}\,$MSA and $\mathbb{E}\,$MSE.

$$MSA = \frac{SSA}{a-1}$$

$$\mathbb{E}\, \frac{SSA}{n\sigma_A^2 + \sigma_\varepsilon^2} = a-1 \leftarrow \text{df of } \chi^2_{a-1} \text{ dist.}$$

$$\Rightarrow \mathbb{E}\, \underbrace{\frac{SSA}{a-1}}_{MSA} = n\sigma_A^2 + \sigma_\varepsilon^2$$

$$MSE = \frac{SSE}{a(n-1)}$$

$$\mathbb{E}\, \underbrace{\frac{SSE}{\sigma_\varepsilon^2}}_{\sim \chi^2_{a(n-1)}} = a(n-1) \quad \Rightarrow \quad \boxed{\mathbb{E}\, \frac{SSE}{a(n-1)} = \sigma_\varepsilon^2}$$

$$\frac{MSA}{MSE} = \frac{SSA/(a-1)}{SSE/[a(n-1)]}$$

$$= \left(\frac{n\sigma_A^2 + \sigma_\varepsilon^2}{\sigma_\varepsilon^2}\right) \frac{\overbrace{\frac{1}{n\sigma_A^2+\sigma_\varepsilon^2} SSA}^{\sim \chi^2_{a-1}} \Big/ (a-1)}{\underbrace{\frac{1}{\sigma_\varepsilon^2} SSE}_{\sim \chi^2_{a(n-1)}} \Big/ [a(n-1)]}$$

$$\frac{\sigma_\varepsilon^2}{n\sigma_A^2 + \sigma_\varepsilon^2} \frac{MSA}{MSE} \sim F_{a-1,\, a(n-1)}$$

6. Give a size-α test for H_0: $\sigma_A^2 = 0$ based on MSA / MSE. Give power function.

H_0: $\sigma_A^2 = 0$.

Reject H_0 if $\dfrac{MSA}{MSE} > F_{a-1,\, a(n-1),\, \alpha}$.

Power:

$$\gamma(\sigma_A^2) = P_{\sigma_A^2}(\text{Reject } H_0 \text{ under } \sigma_A^2)$$

$$= P_{\sigma_A^2}\left(\frac{MSA}{MSE} > F_{a-1,a(n-1),\alpha} \right)$$

$$= P_{\sigma_A^2}\left(\underbrace{\frac{\sigma_\varepsilon^2}{n\sigma_A^2+\sigma_\varepsilon^2} \frac{MSA}{MSE}}_{\sim F_{a-1,a(n-1)}} > \frac{\sigma_\varepsilon^2}{n\sigma_A^2+\sigma_\varepsilon^2} F_{a-1,a(n-1),\alpha} \right)$$

Monahan, J. F. (2008). *A primer on linear models*. CRC Press.

$$= P_{\sigma_A^2}\left(F > \frac{\sigma_\varepsilon^2}{n\sigma_A^2+\sigma_\varepsilon^2} F_{a-1,a(n-1),\alpha} \right)$$

$$= 1 - F\left(\frac{\sigma_\varepsilon^2}{n\sigma_A^2+\sigma_\varepsilon^2} F_{a-1,a(n-1),\alpha} \right)$$

cdf of $F_{a-1,a(n-1)}$ dist.