assumptions need: u & e ($\Rightarrow y$) follow a Normal distrib.

21. Best Linear Unbiased Prediction (BLUP) of Random Effects in the Normal Linear Mixed Effects Model

Predict random effect u; based on the structure of the model and observed data what is the most likely value of u?

C. R. Henderson

- Born April 1, 1911, in Coin, Iowa, in Page County – same county of birth as Jay Lush
- Dean H.H. Kildee visited in 1929 and convinced Henderson to come to Iowa State College.
- Returned to ISU after the war for Ph.D. with Jay Lush in animal breeding.
- Professor at Cornell until 1976
- Known for "Henderson's Mixed Model Equations" and use of BLUP in animal breeding.
- Elected member of the National Academy of Sciences

Consider our linear mixed effects model

$$* \quad \boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e},$$

where

$$* \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R} \end{bmatrix} \right).$$

Var(u)

Var(e)

Given data $\boldsymbol{y}$ *, what is our best guess for the unobserved vector $\boldsymbol{u}$?

- Because $\boldsymbol{u}$ is a random vector rather than a fixed parameter, we talk about predicting $\boldsymbol{u}$ rather than estimating $\boldsymbol{u}$.
- We seek a Best Linear Unbiased Predictor (BLUP) for $\boldsymbol{u}$, which we will denote by $\widehat{\boldsymbol{u}}$.

To be a BLUP, we require...

1. $\widehat{\boldsymbol{u}}$ to be a linear function of $\boldsymbol{y}$,
2. $\widehat{\boldsymbol{u}}$ to be unbiased for $\boldsymbol{u}$ so that $\mathrm{E}(\widehat{\boldsymbol{u}} - \boldsymbol{u}) = \boldsymbol{0}$, and
3. $\mathrm{Var}(\widehat{\boldsymbol{u}} - \boldsymbol{u})$ to be no "larger" than the $\mathrm{Var}(\boldsymbol{v} - \boldsymbol{u})$, where $\boldsymbol{v}$ is any other linear and unbiased predictor.

$$\left[\mathrm{Var}(y)\right]^{-1} \qquad y - \hat{y}$$

- In 6110 we prove that the BLUP of $\boldsymbol{u}$ is

$$\underbrace{\boldsymbol{G}}_{\mathrm{Var}(u)} \underbrace{\boldsymbol{Z}^\top}_{(\text{model matrix of } u)^\top} \boldsymbol{\Sigma}^{-1} \boxed{(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}).}$$

- This BLUP can be viewed as an approximation of

$$\mathrm{E}(\boldsymbol{u}|\boldsymbol{y}) = \boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}).$$

- To derive this expression for $\mathrm{E}(\boldsymbol{u}|\boldsymbol{y})$, we will use the following result about conditional distributions for multivariate normal vectors.

When the BLUP is based on $\hat{\Sigma}$ $\Rightarrow$ empirical BLUP

$$Var(e) = \Sigma_i = \underline{\sigma}^2 \cdot I$$

Suppose

$$\begin{bmatrix} \underline{\boldsymbol{w}_1} \\ \underline{\boldsymbol{w}_2} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \underbrace{\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}}_{\Sigma_i} \right)$$

where

$$\boldsymbol{\Sigma} \equiv \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \text{ is positive definite.}$$

Then the conditional distribution of $\boldsymbol{w}_2$ given $\boldsymbol{w}_1$ is

$$(\boldsymbol{w}_2 | \underline{\boldsymbol{w}_1}) \sim \mathcal{N}(\underbrace{\boldsymbol{\mu}_2 + \boxed{\boldsymbol{\Sigma}_{21}} \boldsymbol{\Sigma}_{11}^{-1} (\boldsymbol{w}_1 - \boldsymbol{\mu}_1)}_{\text{mean}}, \underbrace{\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}}_{\text{Variance}}).$$

Now note that

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{bmatrix}$$

Thus,

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{u} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{Z} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{G} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{Z}^\top & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \right)$$

$$\stackrel{d}{=} \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \overbrace{\boldsymbol{ZGZ}^\top + \boldsymbol{R}}^{Var(y)} & \overbrace{\boldsymbol{ZG}}^{Cov(y, u)} \\ \underbrace{\boldsymbol{GZ}^\top}_{\Sigma_{21}} & \underbrace{\boldsymbol{G}}_{Var(u)} \end{bmatrix} \right).$$

$\Sigma = \Sigma_{11}$ on slide 7

As a result,

$$\begin{aligned} \mathrm{E}(\boldsymbol{u}|\boldsymbol{y}) &= \underline{\mathbf{0}} + \boldsymbol{G}\boldsymbol{Z}^\top \underbrace{(\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R})}_{\Sigma}{}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\ &= \boldsymbol{G}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}). \end{aligned}$$

replace with $X\hat{\beta}_\Sigma$

BLUP of $\boldsymbol{u}$: replace $\boldsymbol{X}\boldsymbol{\beta}$ in the expression above with its BLUE $\boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}$ to obtain

$$\begin{aligned} \boldsymbol{G}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}) &= \boldsymbol{G}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{y}) \\ &= \boldsymbol{G}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1})\boldsymbol{y}. \end{aligned}$$

For the usual case in which

$$\boldsymbol{G} \quad \text{and} \quad \boldsymbol{\Sigma} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R}$$

are unknown, we replace the matrices by estimates and approximate the BLUP of $\boldsymbol{u}$ by

$$\widehat{\boldsymbol{G}}\boldsymbol{Z}^\top\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{\Sigma}}}).$$

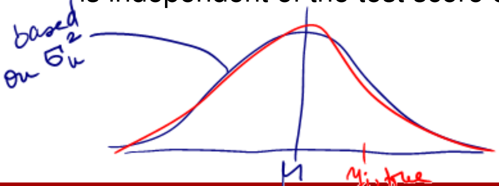
This approximation to the BLUP is sometimes called an EBLUP, where "E" stands for *empirical*.

- Often we wish to make predictions of quantities like $C\boldsymbol{\beta} + \boldsymbol{D}\boldsymbol{u}$ for some estimable $C\boldsymbol{\beta}$.

- The BLUP of such a quantity is $C\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} + \boldsymbol{D}\widehat{\boldsymbol{u}}$, the BLUE of $C\boldsymbol{\beta}$ plus $\boldsymbol{D}$ times the BLUP of $\boldsymbol{u}$.

$$\underbrace{\underset{\text{BLUE}}{C\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}} + \underset{\text{BLUP}}{\boldsymbol{D}\widehat{\boldsymbol{u}}}}_{\text{BLUP of } C\beta + D\underset{\sim}{u}}$$

- Suppose reading ability for a population of students are normally distributed with a mean μ and variance σ_u^2.
- Suppose a reading ability test was given to an i.i.d. sample of such students. sample from $N(\mu, \sigma_u^2)$
- Suppose that, given the true reading ability of a student at that time, the test score for that student is normally distributed with a mean equal to the student's reading ability and a variance σ_e^2 and is independent of the test score of any other student.



given the true μ for the student our observed test score comes from a $N(\mu, \sigma_e^2$

- Suppose it is known that $\boxed{\sigma_u^2/\sigma_e^2} = 9$.

$\bar{y} = 86$

- If the sample mean of the students' test scores was 86, what is the best prediction of the reading ability of a student who scored 96 on the test?

the student's true reading ability

$\boxed{y_{i,\text{true}} = \mu + u_i}$ we observe

however $y_i^* = 96$

u & e are indep.

- Suppose $u_1, \ldots, u_n \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_u^2)$ independent of $e_1, \ldots, e_n \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_e^2)$.

- If we let $\mu + u_i$ denote the reading ability of student i $(i = 1, \ldots, n)$, then the reading ability scores of the students are $\mathcal{N}(\mu, \sigma_u^2)$ as in the statement of the problem.

- If we let $y_i = \mu + u_i + e_i$ denote the test score of student i $(i = 1, \ldots, n)$, then $(y_i | \mu + u_i) \sim \mathcal{N}(\mu + u_i, \sigma_e^2)$ as in the problem statement.

We have $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$, where

$$\boldsymbol{X} = \mathbf{1}, \quad \boldsymbol{\beta} = \mu, \quad \boldsymbol{Z} = \boldsymbol{I}, \quad \boldsymbol{G} = \sigma_u^2 \boldsymbol{I}, \quad \boldsymbol{R} = \sigma_e^2 \boldsymbol{I}, \quad \text{and}$$

$$\boldsymbol{\Sigma} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R} = \boxed{(\sigma_u^2 + \sigma_e^2)\boldsymbol{I}.}$$

Thus,

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} = (\boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{y} = (\mathbf{1}^\top \mathbf{1})^- \mathbf{1}^\top \boldsymbol{y} = \bar{y}.$$

in our ex.
observed
mean = 86

and

end lecture
36
4-25-25

$$\boldsymbol{G}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \boldsymbol{I}.$$