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

C. R. Henderson

- Born April 1, 1911, in Coin, Iowa, in Page County – same county of birth as Jay Lush
- Page County Farm Bureau Picnic (12 and under, 14 and under, 16 and under changed to 10-12, 13-14, 15-16)
- Dean H.H. Kildee visited in 1929 and convinced Henderson to come to Iowa State College.

C. R. Henderson

- ISC track 4 x 220 indoor world record
- 1933 ISC Field House indoor 440 record of 51.7 (stood for 30 years)
- Outdoor 440 record of 48.6 when world record was 47.4
- MS in nutrition from ISC
- 1942 U.S. Army Sanitary Corps. Nutrition research for troops.

C. R. Henderson

- Returned to ISU after the war for Ph.D. with Jay Lush in animal breeding.
- Professor at Cornell until 1976
- Know for “Henderson’s Mixed Model Equations” and use of BLUP in animal breeding.
- Elected member of the National Academy of Sciences

Henderson's Ph.D. Students Included

- Shayle Searle (who taught Henderson matrix algebra)
- David Harville (professor emeritus, Department of Statistics, ISU, linear models expert)

Henderson's Advice to Beginning Scientists

- Study methods of your predecessors.
- Work hard.
- Do not fear to try new ideas.
- Discuss your ideas with others freely.
- Be quick to admit errors. Progress comes by correcting mistakes.
- Always be optimistic. Nature is benign.
- Enjoy your scientific work. It can be a great joy.

Sources

- L. D. Van Vleck (1998). Charles Roy Henderson, 1911-1989: a brief biography. Journal of Animal Science. 76, 2959-2961.
- L. D. Van Vleck (1991). C. R. Henderson: Farm Boy, Athlete, and Scientist. Journal of Dairy Science. 74, 4082-4096.

A problem that reportedly sparked Henderson's interest in BLUP

We present here a variation of the original problem 23 on page 164 of Mood, A. M. (1950), *Introduction to the Theory of Statistics*, New York: McGraw-Hill.

- Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean μ and variance σ_u^2.
- Suppose an IQ test was given to an i.i.d. sample of such students.
- Suppose that, given the IQ of a student, the test score for that student is normally distributed with a mean equal to the student's IQ and a variance σ_e^2 and is independent of the test score of any other student.

- Suppose it is known that $\sigma_u^2/\sigma_e^2 = 9$.
- If the sample mean of the students' test scores was 100, what is the best prediction of the IQ of a student who scored 130 on the test?

Consider our linear mixed effects model

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

where

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

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 $\hat{\boldsymbol{u}}$.

To be a BLUP, we require...

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

- In 611, we prove that the BLUP of $\boldsymbol{u}$ is

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

- This BLUP can be viewed as an approximation of

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

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

Suppose

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

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

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

$$(\boldsymbol{w}_2 | \boldsymbol{w}_1) \sim N(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{w}_1 - \boldsymbol{\mu}_1),\ \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}).$$

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,

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

$$\stackrel{d}{=} N\left(\begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{ZGZ}' + \boldsymbol{R} & \boldsymbol{ZG} \\ \boldsymbol{GZ}' & \boldsymbol{G} \end{bmatrix} \right).$$

$$\text{Thus, } E(\boldsymbol{u}|\boldsymbol{y}) = \boldsymbol{0} + \boldsymbol{G}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R})^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

$$= \boldsymbol{G}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}).$$

To get the BLUP of $\boldsymbol{u}$, we replace $\boldsymbol{X}\boldsymbol{\beta}$ in the expression above with its BLUE $\boldsymbol{X}\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}$ to obtain

$$\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}) = \boldsymbol{G}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y})$$

$$= \boldsymbol{G}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1})\boldsymbol{y}.$$

For the usual case in which

$$\boldsymbol{G} \text{ and } \boldsymbol{\Sigma} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R}$$

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

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

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

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

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

- Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean μ and variance σ_u^2.

- Suppose an IQ test was given to an i.i.d. sample of such students.

- Suppose that, given the IQ of a student, the test score for that student is normally distributed with a mean equal to the student's IQ and a variance σ_e^2 and is independent of the test score of any other student.

- Suppose it is known that $\sigma_u^2/\sigma_e^2 = 9$.
- If the sample mean of the students' test scores was 100, what is the best prediction of the IQ of a student who scored 130 on the test?

- Suppose $u_1, \ldots, u_n \overset{i.i.d.}{\sim} N(0, \sigma_u^2)$ independent of $e_1, \ldots, e_n \overset{i.i.d.}{\sim} N(0, \sigma_e^2)$.
- If we let $\mu + u_i$ denote the IQ of student i $(i = 1, \ldots, n)$, then the IQs of the students are $N(\mu, \sigma_u^2)$ 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 $(y_i | \mu + u_i) \ \sim \ N(\mu + u_i, \sigma_e^2)$ as in the problem statement.

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

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

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

Thus,

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = (\mathbf{1}'\mathbf{1})^{-}\mathbf{1}'\boldsymbol{y} = \bar{y}_{\cdot}$$

and

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

Thus, the BLUP for $\boldsymbol{u}$ is

$$\hat{\boldsymbol{u}} = \boldsymbol{G}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\boldsymbol{\Sigma}}) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}(\boldsymbol{y} - \mathbf{1}\bar{y}_\cdot).$$

The i^{th} element of this vector is

$$\hat{u}_i = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}(y_i - \bar{y}_\cdot).$$

Thus, the BLUP for $\mu + u_i$ (the IQ of student i) is

$$\hat{\mu} + \hat{u}_i = \bar{y}_\cdot + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}(y_i - \bar{y}_\cdot) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}y_i + \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2}\bar{y}_\cdot$$

Note that the BLUP is a convex combination of the individual score and the overall mean score.

$$\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} y_i + \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2} \bar{y}.$$

Because $\frac{\sigma_u^2}{\sigma_e^2}$ is assumed to be 9, the weights are

$$\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} = \frac{\frac{\sigma_u^2}{\sigma_e^2}}{\frac{\sigma_u^2}{\sigma_e^2} + 1} = \frac{9}{9+1} = 0.9$$

and

$$\frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2} = 0.1.$$

We would predict the IQ of a student who scored 130 on the test to be $0.9(130) + 0.1(100) = 127$.