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

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


A problem that reportedly sparked Henderson’s interest in BLUP

Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean $\mu$ and variance $\sigma_u^2$.

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 $\sigma_e^2$ 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

\[ y = X\beta + Zu + e, \]

where

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

Given data \( y \), what is our best guess for the unobserved vector \( u \)?
Because $u$ is a random vector rather than a fixed parameter, we talk about predicting $u$ rather than estimating $u$.

We seek a Best Linear Unbiased Predictor (BLUP) for $u$, which we will denote by $\hat{u}$. 
To be a BLUP, we require...

1. \( \hat{u} \) to be a linear function of \( y \),

2. \( \hat{u} \) to be unbiased for \( u \) so that \( E(\hat{u} - u) = 0 \), and

3. \( \text{Var}(\hat{u} - u) \) to be no “larger” than the \( \text{Var}(v - u) \), where \( v \) is any other linear and unbiased predictor.
In 611, we prove that the BLUP of $u$ is

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

This BLUP can be viewed as an approximation of

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

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

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

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

Then the conditional distribution of $w_2$ given $w_1$ is

$$(w_2 | w_1) \sim \mathcal{N}(\mu_2 + \Sigma_{21} \Sigma_{11}^{-1} (w_1 - \mu_1), \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}).$$
Now note that

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

Thus,

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

\[
\overset{d}{=} N \left( \begin{bmatrix}
  X\beta \\
  0
\end{bmatrix}, \begin{bmatrix}
  ZGZ' + R & ZG \\
  GZ' & G
\end{bmatrix} \right).
\]
Thus, \( E(u|y) = 0 + GZ'(ZGZ' + R)^{-1}(y - X\beta) \)

\[ = GZ'\Sigma^{-1}(y - X\beta). \]

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

\[ GZ'\Sigma^{-1}(y - X\hat{\beta}_\Sigma) = GZ'\Sigma^{-1}(y - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y) \]

\[ = GZ'\Sigma^{-1}(I - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1})y. \]
For the usual case in which

\[
G \text{ and } \Sigma = ZGZ' + R
\]

are unknown, we replace the matrices by estimates and approximate the BLUP of \( u \) by

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

This approximation to the BLUP is sometimes called an EBLUP, where “E” stands for empirical.
Often we wish to make predictions of quantities like $C\beta + Du$ for some estimable $C\beta$.

The BLUP of such a quantity is $C\hat{\beta}_\Sigma + D\hat{u}$, the BLUE of $C\beta$ plus $D$ times the BLUP of $u$. 
Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean $\mu$ and variance $\sigma_u^2$.

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 $\sigma_e^2$ 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 \( y = X\beta + Zu + e \), where

\[
X = 1, \quad \beta = \mu, \quad Z = I, \quad G = \sigma_u^2 I, \quad R = \sigma_e^2 I, \quad \text{and} \\
\Sigma = ZGZ' + R = (\sigma_u^2 + \sigma_e^2)I.
\]

Thus,

\[
\hat{\beta}_\Sigma = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y = (1'1)^{-1}1'y = \bar{y}.
\]

and

\[
GZ'\Sigma^{-1} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}I.
\]
Thus, the BLUP for $u$ is

\[ \hat{u} = GZ'\Sigma^{-1}(y - X\hat{\beta}_\Sigma) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}(y - 1\bar{y}). \]

The $i^{th}$ element of this vector is

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

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

\[ \hat{\mu} + \hat{u}_i = \bar{y} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}(y_i - \bar{y}) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}y_i + \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2}\bar{y}. \]
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 $0.9(130) + 0.1(100) = 127$. 