

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Papers and Publications in Animal  
Science

Animal Science Department

---

January 1982

## Estimation of Components of Variance by Method 3 and Henderson's New Method

G. F. S. Hudson  
*Cornell University*

L. Dale Van Vleck  
*University of Nebraska-Lincoln*, [dvan-vleck1@unl.edu](mailto:dvan-vleck1@unl.edu)

Follow this and additional works at: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

---

Hudson, G. F. S. and Van Vleck, L. Dale, "Estimation of Components of Variance by Method 3 and Henderson's New Method" (1982). *Faculty Papers and Publications in Animal Science*. 387.  
<https://digitalcommons.unl.edu/animalscifacpub/387>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Estimation of Components of Variance by Method 3 and Henderson's New Method

G. F. S. HUDSON and L. D. VAN VLECK

Department of Animal Science  
Cornell University  
Ithaca, NY 14853

## ABSTRACT

An algorithm for Method 3 for estimation of components of variance in a mixed model including fixed herd-year-season, genetic groups, and random sires nested within groups is presented. Computational aspects of the algorithm are discussed and compared with those of a new method of Henderson's. The new method involves equating quadratic forms based on approximate best linear unbiased prediction solutions to their expected values. Relative computing effort for the two methods depended on the ease of eliminating equations for fixed effects and on the number of traits. Numerical estimates from the two methods were similar. When the Method 3 estimate of the sire component of variance was not close to zero, the new method rapidly converged during iteration even when the a priori variance component ratio was not close to the final estimate.

## INTRODUCTION

Prediction of random effects in mixed linear models requires knowledge of components of variance associated with those random effects and with random error. True population parameters are never known; hence, estimates are required. In practical animal breeding problems, the data are often large and unbalanced. Although many variance component estimators exist which are theoretically optimal in one or more ways, most are impractical because of excessive computational requirements. Henderson's mixed model equations allow simple

algorithms to be developed for applying certain estimation methods to particular models (3, 6, 7). However, in general, computational simplicity may require invalid assumptions concerning the model. Henderson (4) recently presented a new unbiased method applicable to any model. The objectives of this paper are 1) to present an algorithm for Henderson's Method 3 (2) for a particular sire evaluation model and 2) to compare the computational feasibility of, and estimates from, Method 3 and Henderson's new method.

## DATA AND METHODS

Records from Ayrshire, Guernsey, Holstein, Jersey, and Brown Swiss cows sired by artificial insemination (AI) in the northeast United States were obtained from New York Dairy Records Processing Laboratory in Ithaca. All milk and fat records were first lactation, 305-day, 2X, mature equivalent (ME) with at least 907 kg milk and 32 kg fat and less than 15,876 kg milk and 680 kg fat. Data were divided into five opportunity groups, each including only records of cows having the opportunity to survive to 36, 48, 60, 72, or 84 mo. The record of a cow in any particular opportunity group also was included in earlier opportunity groups. Stayability records were coded 1 if the cow survived to a given age and 0 otherwise. For example, a cow sold for dairy at 68 mo of age had the opportunity to survive the 60 mo. The survival score was 1 for survival to 36, 48, and 60 mo. A cow sold for beef or otherwise disposed at 68 mo of age having the opportunity to survive to 72 and 84 mo, if the herd remained on test, was scored zero for survival to 72 and 84 mo. Thus, each data set contained milk and fat records and up to five stayability records for each cow.

The model was

$$y = Xh + Wg + Zs + e,$$

---

Received February 2, 1981.

where  $y$  is a vector of milk, fat, or stayability records,  $h$  is a vector of fixed herd-year-season effects,  $g$  is a vector of fixed genetic group effects,  $s$  is a vector of random effects of sires nested within groups, and  $e$  is a vector of random residual effects. The  $X$ ,  $W$ , and  $Z$  are known incidence matrices of zeros and ones. For a common sire component of variance for all genetic groups, expectations are

$$E \begin{bmatrix} y \\ s \\ e \end{bmatrix} = \begin{bmatrix} Xh + Wg \\ 0 \\ 0 \end{bmatrix},$$

with variance-covariance matrix

$$\text{Var} \begin{bmatrix} y \\ s \\ e \end{bmatrix} = \begin{bmatrix} ZZ'o_s^2 + Io_e^2 & Zo_s^2 & 0 \\ & Io_s^2 & 0 \\ \text{Symmetric} & & Io_e^2 \end{bmatrix}.$$

where the  $I$ 's are identity matrices of appropriate order. Both Method 3 and Henderson's new method involve equating various sums of squares or reductions in sums of squares from fitting submodels to their expected values. Each sum of squares is computed as if the model is completely fixed (i.e., as if  $E(y) = Xh + Wg + Zs$ ) by least squares methodology. Both procedures begin with absorption of herd-year-season equations, after which the least squares equations for genetic groups and sires can be represented by

$$\begin{bmatrix} C_{11} & C_{12} \\ C'_{12} & C_{22} \end{bmatrix} \begin{bmatrix} \hat{g} \\ s \end{bmatrix} = \begin{bmatrix} r_g \\ r_s \end{bmatrix} \quad [1]$$

where

$$\begin{aligned} C_{11} &= W'MW, \\ C_{12} &= W'MZ, \\ C_{22} &= Z'MZ, \\ r_g &= W'My, \\ r_s &= Z'My, \text{ and} \\ M &= I - X(X'X)^{-1}X'. \end{aligned}$$

Computationally, only  $C_{22}$  and  $r_s$  need to be formed during absorption of herd-year-season

equations. The  $C_{11}$ ,  $C_{12}$ , and  $r_g$  can be obtained by summing appropriate elements of  $C_{22}$  and  $r_s$  because of nesting of sire effects within groups.

The same unbiased estimator of  $\sigma_e^2$  was used for both methods, which was obtained from the residual mean square after fitting the full model (8)

$$\hat{\sigma}_e^2 = [y'y - R(h,g,s)] / [N - r(XWZ)], \quad [2]$$

where  $y'y$  is the sum of squares of observations,  $R(h,g,s)$  is the reduction from fitting the full model (8),  $N$  is the number of observations, and  $r(XWZ)$  is the rank of  $(XWZ)$  which is the number of herd-year-seasons + number of sires - 1.

The Method 3 estimator of  $\sigma_s^2$  is

$$\hat{\sigma}_s^2 = [R(s|h,g) - (s - g)\hat{\sigma}_e^2] / [\text{tr}(C_{22} - C'_{12}C_{11}^{-1}C_{12})], \quad [3]$$

where  $R(s|h,g)$  is the reduction in sum of squares from fitting sites after herd-year-seasons and genetic groups,  $s$  is the number of sires,  $g$  is the number of groups, and  $C_{11}^{-1}$  is a generalized inverse of  $C_{11}$  (8). Because sires are nested within genetic groups,

$$R(g,s|h) = R(s|h). \quad [4]$$

Then

$$R(h,g,s) = R(s|h) + R(h) \quad [5]$$

and

$$R(s|h,g) = R(s|h) - R(g|h). \quad [6]$$

The three required reductions are computed as:  $R(h)$  = uncorrected herd-year-season sum of squares, computed during the absorption,

$$\begin{aligned} R(s|h) &= \hat{s}'r_s, \\ \text{with } s &\text{ a solution to } C_{22}\hat{s} = r_s \end{aligned} \quad [7]$$

and

$$\begin{aligned} R(g|h) &= \hat{g}'r_g, \\ \text{with } g &\text{ a solution to } C_{11}\hat{g} = r_g. \end{aligned} \quad [8]$$

Henderson's new method (4) requires elim-

ination of all equations corresponding to fixed effects. Thus, group equations must be absorbed into sire equations so that  $\mathbf{P}\hat{\mathbf{s}} = \mathbf{r}$ , where

$$\mathbf{P} = \mathbf{C}_{22} - \mathbf{C}'_{12} \mathbf{C}_{11}^{-1} \mathbf{C}_{12} \quad [9]$$

and

$$\mathbf{r} = \mathbf{r}_s - \mathbf{C}'_{12} \mathbf{C}_{11}^{-1} \mathbf{r}_g. \quad [10]$$

Henderson's mixed model equations (3) would be formed by adding  $\alpha = \sigma_e^2/\sigma_s^2$  to the diagonal elements of  $\mathbf{P}$ :

$$(\mathbf{P} + \mathbf{I}\alpha) \mathbf{s}^0 = \mathbf{r}. \quad [11]$$

The parameter  $\alpha$  is unknown, so an estimate,  $\hat{\alpha}$ , must be used. In this study,  $\hat{\alpha}$  was the ratio of the Method 3 estimates. The new method involves approximating sire solutions by  $\tilde{\mathbf{s}} = \mathbf{D}^{-1}\mathbf{r}$ , where  $\mathbf{D}$  is a diagonal matrix with diagonal elements equal to those of  $(\mathbf{P} + \mathbf{I}\hat{\alpha})$  in [11].

Then, since

$$\mathbf{E}(\tilde{\mathbf{s}}'\tilde{\mathbf{s}}) = \text{tr}\mathbf{D}^{-2}(\mathbf{P}^2\sigma_s^2 + \mathbf{P}\sigma_e^2), \quad [12]$$

an unbiased estimator of  $\sigma_s^2$  is

$$\hat{\sigma}_s^2 = (\tilde{\mathbf{s}}'\tilde{\mathbf{s}} - \text{tr}\mathbf{D}^{-2}\mathbf{P}\hat{\sigma}_e^2)/(\text{tr}\mathbf{D}^{-2}\mathbf{P}^2). \quad [13]$$

Another estimator also was examined, which we discovered had been proposed by C. R. Henderson (personal communication, 1980). Since

$$\mathbf{E}(\tilde{\mathbf{s}}'\mathbf{r}) = \text{tr}\mathbf{D}^{-1}(\mathbf{P}^2\sigma_s^2 + \mathbf{P}\sigma_e^2), \quad [14]$$

an unbiased estimator of  $\sigma_s^2$  is

$$\hat{\sigma}_s^2 = (\tilde{\mathbf{s}}'\mathbf{r} - \text{tr}\mathbf{D}^{-1}\mathbf{P}\hat{\sigma}_e^2)/(\text{tr}\mathbf{D}^{-1}\mathbf{P}^2). \quad [15]$$

The expectation of  $\tilde{\mathbf{s}}'\tilde{\mathbf{s}}$  in [12] involves  $\mathbf{D}^{-2}$ , whereas that of  $\tilde{\mathbf{s}}'\mathbf{r}$  in [14] includes  $\mathbf{D}^{-1}$ . Hereinafter, the two estimators are referred to as D2 and D1, respectively.

Both D1 and D2 estimators appear to lend themselves to iteration. Therefore, iterative estimates based on D1 and D2 also were obtained. Successive estimates of  $\alpha$  were calculated from  $\hat{\sigma}_s^2$  or  $\hat{\sigma}_s^2$  from [15] or [13] with  $\hat{\sigma}_e^2$  from [2].

Components of covariance between two traits, say  $a$  and  $b$ , were estimated by

$$\hat{\sigma}_{a,b} = .5(\hat{\sigma}_{(a+b)}^2 - \hat{\sigma}_a^2 - \hat{\sigma}_b^2),$$

where  $\hat{\sigma}_{(a+b)}^2$  refers to the variance component resulting from summing the two traits  $a$  and  $b$ . To reduce errors from rounding when the two traits were stayability (a zero or one trait) and a production trait (kg), milk and fat records were standardized. Thus, records in the summation were of similar magnitude.

## RESULTS AND DISCUSSION

### Computational Aspects

Obtaining sire solutions,  $\hat{\mathbf{s}}$ , in [7] for Method 3 required solving a large system of equations. The coefficient matrix was too large to store in core and, thus, to solve by obtaining the inverse. The dependency between herd-year-season equations and sire equations was removed by setting the first sire solution to zero and deleting the corresponding equation. An iterative method often used for solving equations of this type is successive overrelaxation (SOR), also referred to as modified Gauss-Seidel. The SOR can be described in general by considering the system of equations  $\mathbf{Ax} = \mathbf{b}$ , with  $\mathbf{A} = \{a_{ij}\}$ ,  $\mathbf{x} = \{x_i\}$ , and  $\mathbf{b} = \{b_i\}$  for  $i, j = 1, 2, \dots, n$ . The solution for  $\mathbf{x}$  at the  $m$ th round of iteration is  $\mathbf{x}^{(m)} = \{x_i^{(m)}\}$  with  $x_1^{(m)} = 0$ , which is the constraint to remove the dependency, and

$$\begin{aligned} x_i^{(m)} &= x_i^{(m-1)} + \omega [b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m)} \\ &\quad - \sum_{j=i}^n a_{ij}x_j^{(m-1)}] / a_{ii} \\ &\text{for } i = 2, 3, \dots, n \end{aligned} \quad [16]$$

where  $1 < \omega < 2$  is the overrelaxation parameter. When  $\omega = 1$ , SOR becomes Gauss-Seidel. Three decisions have to be made with SOR: initial solutions,  $\mathbf{x}^{(0)}$ , to begin iteration; the magnitude of  $\omega$ ; and the stopping or convergence criterion, i.e., when is  $\mathbf{x}^{(m)}$  "sufficiently close" to the actual solution to  $\mathbf{Ax} = \mathbf{b}$ . In this study  $x_i^{(0)} = b_i/a_{ii}$ . Overrelaxation parameters of 1.4, 1.6, 1.8, and 1.9 were tested for a small data set containing 87 sire equations. Most

rapid convergence occurred with  $\omega = 1.8$ . A 1.9 caused fluctuations in sire solutions between rounds whereas less than 1.8 had little effect relative to 1.

In [16] elements of  $\mathbf{x}^{(m)}$  are calculated individually whereas if an inverse to the coefficient matrix is obtained, all elements are calculated simultaneously. A compromise is to partition  $\mathbf{A}$  into  $r$  submatrices or blocks,  $\mathbf{A}_{ij}$ , and obtain solutions to each block of equations simultaneously. This method is known as block SOR (or, with  $\omega = 1$ , block Gauss-Seidel), and solution to the  $i$ th subvector of  $\mathbf{x}$ ,  $\mathbf{x}_i$ , at the  $m$ th round is

$$\begin{aligned} \mathbf{x}_i^{(m)} = & \mathbf{x}_i^{(m-1)} + \omega \mathbf{A}_{ii}^{-1} [\mathbf{b}_i \\ & - \sum_{j=1}^{i-1} \mathbf{A}_{ij} \mathbf{x}_j^{(m)} - \sum_{j=i}^r \mathbf{A}_{ij} \mathbf{x}_j^{(m-1)}] \\ & \text{for } i = 1, 2, \dots, r. \end{aligned} \quad [17]$$

The block inverses do not need to be calculated each round, but elements of inverses can be saved in place of those of the original block, i.e., replace  $\mathbf{A}_{ii}$  with  $\mathbf{A}_{ii}^{-1}$ . Block size is arbitrary and is a compromise between rapid convergence with large blocks and time required to invert the  $\mathbf{A}_{ii}$ 's. In this study, block size was 20. Initial solutions  $\mathbf{x}_i^{(0)} = \mathbf{A}_{ii}^{-1} \mathbf{b}_i$  may lead to faster convergence than  $\mathbf{x}_i^{(0)} = \mathbf{b}_i / a_{ii}$ .

Convergence of sire solutions required fewer rounds with block iteration than with standard iteration (Table 1). Computer time on an IBM 370/138 was reduced by only 9% because block iteration requires inverting the  $\mathbf{A}_{ii}$ 's in [17]. Convergence of the reduction from sires,  $\mathbf{R}(\mathbf{s}|\mathbf{h})$ , was more rapid with block iteration

than standard iteration. The final criterion for convergence was a change of less than .5% in  $\mathbf{R}(\mathbf{s}|\mathbf{h})$ , which would have little effect on the estimate of  $\sigma_s^2$ . A minimum of 20 rounds of iteration was imposed because during early rounds the reduction fluctuated from round to round.

Data sets in (5) involved from 6 to 28 traits, including the summed traits for covariance component estimation. Each trait results in a different right-hand-side (RHS) in [1], but the coefficient matrix in [1] is the same for all RHS in a data set. The block inverses were saved on tape (along with the off-diagonal blocks) after solution for the first RHS. Computer time required for solving subsequent RHS's was reduced by 40%.

Group solutions,  $\hat{\mathbf{g}}$ , in [8] were obtained directly as  $\mathbf{C}_{11}$  was small enough to store in core. The  $\text{tr}(\mathbf{C}_{22} - \mathbf{C}_{12} \mathbf{C}_{11}^{-1} \mathbf{C}_{12})$  in [3] was computed directly because this is the trace of  $\mathbf{P}$  in [9]. However, when  $\mathbf{P}$  is not needed, calculating the trace as  $\text{tr}(\mathbf{C}_{22}) - (\mathbf{C}_{11}^{-1} \mathbf{C}_{12} \mathbf{C}_{12}')$  may require less computer (CPU) time (1). The matrix operation  $\mathbf{A}:\mathbf{B}$  is the sum over rows and columns of element by element products, i.e.,  $\mathbf{A}:\mathbf{B} = \{a_{ij}\}:\{b_{ij}\} = \sum_{ij} a_{ij} b_{ij}$ .  $\mathbf{A}:\mathbf{B} = \text{tr}(\mathbf{AB})$  only for  $\mathbf{B}$  symmetric.

The quadratic forms  $\tilde{\mathbf{s}}'\tilde{\mathbf{s}}$  and  $\tilde{\mathbf{s}}'\tilde{\mathbf{r}}$  in [13] and [15] were easily computed as

$$\sum_i \{r_i / (p_{ii} + \hat{\alpha})\}^2$$

and

$$\sum_i r_i^2 / (p_{ii} + \hat{\alpha}),$$

TABLE 1. Comparison of standard and block Gauss-Seidel iteration (87 equations, 1,440 nonzero off-diagonals, overrelaxation parameter = 1.8, block size = 20).

Convergence criterion	Standard iteration		Block iteration	
	No. of rounds	Computer time (s)	No. of rounds	Computer time (s)
Avg absolute change in sire solutions less than .00001	250	119.9	135	108.6
Change in $\mathbf{R}(\mathbf{s} \mathbf{h})$ less than .00001	288	140.4	69	62.7

respectively, where  $r_i$  is the  $i$ th element of  $\mathbf{r}$ , and  $p_{ii}$  is the  $i$ th diagonal element of  $\mathbf{P}$ . The  $\text{tr}(\mathbf{D}^{-1}\mathbf{P})$  and  $\text{tr}(\mathbf{D}^{-1}\mathbf{P}^2)$  in [15] were calculated by

$$\sum_i p_{ii}/(p_{ii} + \hat{\alpha})$$

and

$$\sum_i [\sum_j p_{ij}^2/(p_{ii} + \hat{\alpha})],$$

respectively. Computing the traces in [13] involved squaring the denominator in each of these expressions. The only storage required in core was three arrays containing diagonals of  $\mathbf{P}$ , sum of squares of elements in each row

$$(\sum_j p_{ij}^2),$$

and RHS. The first two of these are the same for all RHS and for all  $\hat{\alpha}$ , so iterating [14] and [15] was rapid once  $\mathbf{P}$  had been calculated.

Relative computing time required for Method 3 and D1 or D2 depended on the number of equations and RHS (Table 2). The data sets with fewer equations had more RHS than larger data sets (5). The total time required for Method 3 increased with the number of RHS

because the reduction from sires,  $\mathbf{R}(\mathbf{s|h})$ , had to be calculated separately for each RHS. The major effort in computing D1 and D2 estimates was in absorption of group effects. The time required would be reduced if  $\mathbf{C}_{12}$  could be stored in core. However, once the group equations were absorbed, the new methods were rapid.

If the model had not included fixed effects other than herd-year-seasons, Methods D1 and D2 always would have been computationally faster than Method 3 since the costly absorption of group equations would not have been needed.

#### Numerical Comparison of Component Estimates

Sampling variances were not computed. Henderson (4), however, compared sampling variances of D2 with Method 3 for a particular data design. The new method had sampling variances lower than or equal to those of Method 3, depending on the true  $\alpha$  and the a priori  $\alpha$  used in D2. An estimate of approximately one-half the true  $\alpha$  yielded lowest sampling variances. The one-half may be dependent on the data set and model. He also found (C. R. Henderson, personal communication, 1980) that D2 had smaller sampling variances than D1 for a particular design.

TABLE 2. Computer times (min) for programs for each method of variance component estimation.

Estimator	Program	Opportunity group		
		36 mo (6, 2532) <sup>a</sup>	60 mo (10, 1931) <sup>a</sup>	84 mo (28, 1487) <sup>a</sup>
Method 3	Solving $\mathbf{R}(\mathbf{s h})$			
	a) Including computing block inverses	46.5	34.6	24.5
	b) Using block inverses saved on tape	26.6	20.9	15.6
	Total <sup>b</sup>	179.2	223.3	448.1
D1 and D2	Absorbing group equations	256.3	138.8	77.2
	Estimating $\sigma_s^2$ and $b^2$ <sup>c</sup>	24.8	16.5	17.9
	Total	281.1	155.3	95.1

<sup>a</sup>No. of RHS and no. of Holstein sire equations.

<sup>b</sup>Program (a) was run only for the first RHS in each data set; program (b) was run once for all but the first RHS. A data set with only one RHS would require only one run of program (a).

<sup>c</sup>All RHS, five iterations of D1 and D2.

TABLE 3. Sire components of variance for milk, fat, and stayability to 84 mo of age in Holsteins.

	Estimator				
	Method 3	D1 <sup>a</sup>	D2 <sup>a</sup>	Iterative D1 <sup>a,b</sup>	Iterative D2 <sup>a,b</sup>
Milk (kg <sup>2</sup> )	100711.52	95962.19	93319.66	95975.45	93245.16
Fat (kg <sup>2</sup> )	149.22	120.76	110.86	147.17	135.11
Stayability	.0023	.0024	.0022	.0024	.0022

<sup>a</sup>Initial  $\hat{\alpha}$  from Method 3.<sup>b</sup>Five rounds of iteration.

The numerical variance components by Method 3, D1, or D2 were similar for most traits and data sets. Most unbiased estimators applied to large data sets probably will yield similar results. For example, estimated sire components of variance for milk and fat yields and stayability to 84 mo in Holsteins are in Table 3. Differences among estimates were small, but the estimates of sire variance for fat from iterative D1 and D2 were closer to the Method 3 estimate than were those from non-iterative D1 and D2. Although in the example of Table 3 D2 gave lower estimates than Meth-

od 3 or D1, there was no apparent trend over breeds, traits, or data sets. Error components were estimated by only one method (equation [2]). Thus, estimates of heritabilities and phenotypic correlations (Table 4) were even less different among methods because those estimates involve the common estimate of  $\sigma_e^2$ .

For traits where no reasonable prior estimate of  $\alpha$  is available, a guess must be made, which may be far from optimal. Results after two poor a priori values of  $\alpha$  were used for milk yield are in Table 5. Regardless of the initial  $\hat{\alpha}$  used, convergence occurred rapidly in both iterative D1 and iterative D2. In many cases, the estimate did not change after even one round. An initial zero for D2 in the Ayrshire data yielded a high estimate of heritability after one round, but convergence was obtained after five iterations.

TABLE 4. Heritability, genetic and phenotypic correlations for milk, fat, and stayability to 84 mo of age in Holsteins.<sup>a</sup>

	Milk	Fat	Stayability
Milk	.3155 <sup>b</sup>	.8244	.4697
	.2979	.8331	.4858
	.2901	.7832	.4290
Fat	.8561	.3000	.4571
	.8568	.2961	.4755
	.8534	.2735	.4093
Stayability	.1717	.1601	.0507
	.1722	.1607	.0513
	.1702	.1583	.0487

<sup>a</sup>Heritabilities are on the diagonal, genetic correlations above the diagonal, and phenotypic correlations below the diagonal.

<sup>b</sup>The three estimates in each cell are by Method 3, iterative D1, iterative D2, respectively. D1 and D2 estimates after five rounds of iteration, initial  $\hat{\alpha}$  from Method 3.

## CONCLUSIONS

The choice of an estimator of components of variance depends on the model but is influenced by data available and by computing facilities. Numerical estimates did not differ much except that D1 and D2 occasionally yielded negative estimates when Method 3 estimates were positive and vice versa. The problem of what to do when negative estimates are obtained from unbiased estimators remains unresolved.

Another unanswered question addresses properties of iterative D1 and iterative D2. Both D1 and D2 are unbiased by derivation and also are translation invariant (estimates were unaffected by changes in fixed effects). However, iterative D1 and iterative D2 do not necessarily have

TABLE 5. Heritability for milk yield estimated by iteration of Henderson's new method using different a priori  $\hat{\alpha}$ 's.

Breed	Estimator	Initial $\hat{\alpha}$	$h^2$ after 1 round	$h^2$ after convergence	No. of rounds to convergence <sup>b</sup>
Ayrshire	D1	8.545 <sup>a</sup>	.42226	.42226	1
		100.0	.42316	.42226	3
		0	.42094	.42226	4
	D2	8.545 <sup>a</sup>	.41982	.41979	2
		100.0	.41889	.41979	3
		0	.67157	.41979	5
Holstein	D1	11.841 <sup>a</sup>	.29791	.29794	2
		100.0	.30285	.29794	3
		0	.29666	.29794	2
	D2	11.841 <sup>a</sup>	.29029	.29008	2
		100.0	.28801	.29008	3
		0	.23827	.29008	4

<sup>a</sup>Method 3 estimate of  $\hat{\alpha}$ .<sup>b</sup>Convergence criterion: a change in  $h^2$  of less than .00001.

these properties because  $\hat{\alpha}$  used in rounds of iteration other than the first is dependent on the data; hence  $E(\hat{\mathbf{S}}'\mathbf{r})$  and  $E(\hat{\mathbf{S}}'\hat{\mathbf{S}})$  may not be those shown in [14] and [12]. If iterative D1 and iterative D2 are biased, mean square error (MSE), which is sampling variance plus square of bias, may be higher than MSE of D1 and D2 unless iteration reduces sampling variance.

Relative computing time was dependent on number of equations and on number of right-hand-sides in each data set. As computer technology and programming skills advance, these differences in time and cost will decrease, and estimators can be chosen on a theoretical basis rather than on a practical one.

#### ACKNOWLEDGMENTS

Partial financial support of this project was provided by Eastern Artificial Insemination Cooperative, Inc., Ithaca, NY.

#### REFERENCES

- 1 Cunningham, E. P., and C. R. Henderson. 1968. An iterative procedure for estimating fixed effects and variance components in mixed model situations. *Biometrics* 24:13.
- 2 Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics* 9:226.
- 3 Henderson, C. R. 1973. Sire evaluation and genetic trends. Page 10 in *Proc. Anim. Breeding and Genetics Symp. in honor of Dr. Jay L. Lush*. Am. Soc. Anim. Sci., Champaign, IL.
- 4 Henderson, C. R. 1980. A simple method for unbiased estimation of variance components in the mixed model. 72nd Annu. Mtg. Am. Soc. Anim. Sci., Cornell University, Ithaca, NY.
- 5 Hudson, G.F.S., and L. D. Van Vleck. 1981. Relationship between production and stayability in Holstein cattle. *J. Dairy Sci.* 64:2246.
- 6 Schaeffer, L. R. 1976. Maximum likelihood estimation of variance components in dairy cattle breeding research. *J. Dairy Sci.* 59:2146.
- 7 Schaeffer, L. R., and E. B. Burnside. 1974. Survival rates of tested daughters of sires in artificial insemination. *J. Dairy Sci.* 57:1394.
- 8 Searle, S. R. 1971. *Linear models*. Wiley and Sons, New York.