

Chapter 15

Sampling from finite populations

C. R. Henderson

1984 - Guelph

1 Finite e

The populations from which samples have been drawn have been regarded as infinite in preceding chapters. Thus if a random sample of n is drawn from such a population with variance σ^2 , the variance-covariance matrix of the sample vector is $\mathbf{I}_n \sigma^2$. Suppose in contrast, the population has only t elements and a random sample of n is drawn. Then the variance-covariance matrix of the sample is

$$\begin{pmatrix} 1 & & -1/(t-1) \\ & \ddots & \\ -1/(t-1) & & 1 \end{pmatrix} \sigma^2. \quad (1)$$

If $t = n$, that is, the sample is the entire population, the variance-covariance matrix is singular. As an example, suppose that the population of observations on a fixed animal is a single observation on each day of the week. Then the model is

$$\mathbf{y}_i = \mu + \mathbf{e}_i. \quad (2)$$

$$\text{Var}(\mathbf{e}_i) = \begin{pmatrix} 1 & -1/6 & \cdots & -1/6 \\ -1/6 & 1 & \cdots & -1/6 \\ \vdots & \vdots & & \vdots \\ -1/6 & -1/6 & & 1 \end{pmatrix} \sigma^2. \quad (3)$$

Suppose we take n random observations. Then BLUE of μ is

$$\hat{\mu} = \bar{y},$$

and

$$\text{Var}(\hat{\mu}) = \frac{7-n}{6n} \sigma^2,$$

which equals 0 if $n = 7$. In general, with a population size, t , and a sample of n ,

$$\text{Var}(\hat{\mu}) = \frac{t-n}{n(t-1)} \sigma^2,$$

which goes to σ^2/n when t goes to infinity, the latter being the usual result for a sample of n from an infinite population with $Var = \mathbf{I}\sigma^2$.

Suppose now that in this same problem we have a random sample of 3 unrelated animals with 2 observations on each and wish to estimate μ and to predict \mathbf{a} when the model is

$$y_{ij} = \mu + a_i + e_{ij},$$

$$Var(\mathbf{a}) = \mathbf{I}_3,$$

$$Var(\mathbf{e}) = 6 \begin{pmatrix} 1 & -1/6 & 0 & 0 & 0 & 0 \\ -1/6 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1/6 & 0 & 0 \\ 0 & 0 & -1/6 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1/6 \\ 0 & 0 & 0 & 0 & -1/6 & 1 \end{pmatrix}.$$

Then

$$\mathbf{R}^{-1} = \begin{pmatrix} 6 & 1 & 0 & 0 & 0 & 0 \\ 1 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{pmatrix} / 35.$$

The BLUP equations are

$$\begin{pmatrix} 1.2 & .4 & .4 & .4 \\ .4 & 1.4 & 0 & 0 \\ .4 & 0 & 1.4 & 0 \\ .4 & 0 & 0 & 1.4 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = .2 \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{pmatrix}.$$

2 Finite u

We could also have a finite number of breeding values from which a sample is drawn. If these are unrelated and are drawn at random from a population with t animals

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & & -1/t \\ & \ddots & \\ -1/t & & 1 \end{pmatrix} \sigma_a^2. \quad (4)$$

If q are chosen not at random, we can either regard the resulting elements of \mathbf{a} as fixed or we may choose to say we have a sample representing the entire population. Then

$$Var(\mathbf{a}) = \begin{pmatrix} 1 & & -1/q \\ & \ddots & \\ -1/q & & 1 \end{pmatrix} \sigma_{a^*}^2, \quad (5)$$

where σ_{a*}^2 probably is smaller than σ_a^2 . Now \mathbf{G} is singular, and we need to compute BLUP by the methods of Section 5.10. We would obtain exactly the same results if we assume \mathbf{a} fixed but with levels that are unpatterned, and we then proceed to biased estimation as in Chapter 9, regarding the average values of squares and products of elements of \mathbf{a} as

$$\mathbf{P} = \begin{pmatrix} 1 & & -1/q \\ & \ddots & \\ -1/q & & 1 \end{pmatrix} \sigma_{a*}^2. \quad (6)$$

3 Infinite By Finite Interactions

Much controversy has surrounded the problem of an appropriate model for the interactions in a 2 way mixed model. One commonly assumed model is that the interactions have $Var = \mathbf{I} \sigma_\gamma^2$. An alternative model is that the interactions in a row (rows being random and columns fixed) sum to zero. Then variance of interactions, ordered columns in rows, is

$$\begin{pmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{B} \end{pmatrix} \sigma_{\gamma*}^2 \quad (7)$$

where \mathbf{B} is $c \times c$ with 1's on the diagonal and $-1/(c-1)$ on all off-diagonals, where $c =$ number of columns. We will show in Chapter 17 how with appropriate adjustment of σ_r^2 (= variance of rows) we can make them equivalent models. See Section 1.5 for definition of equivalence of models.

4 Finite By Finite Interactions

Suppose that we have a finite population of r rows and c columns. Then we might assume that the variance-covariance matrix of interactions is the following matrix multiplied by σ_γ^2 .

All diagonals = 1.

Covariance between interactions in the same row = $-\sigma_\gamma^2/(c-1)$.

Covariance between interactions in the same column = $-\sigma_\gamma^2/(r-1)$.

Covariance between interactions in neither the same row nor column = $\sigma_\gamma^2/(r-1)(c-1)$.

If the sample involves r rows and c columns both regarded as fixed, and there is no assumed pattern of values of interactions, estimation biased by interactions can be accomplished by regarding these as pseudo-random variables and using the above "variances" for elements of \mathbf{P} , the average value of squares and products of interactions. This methodology was described in Chapter 9.

5 Finite, Factorial, Mixed Models

In previous chapters dealing with infinite populations from which \mathbf{u} is drawn at random as well as infinite subpopulations from which subvectors \mathbf{u}_i are drawn the assumption has been that the expectations of these vectors is null. In the case of a population with finite levels we shall assume that the sum of all elements of their population = 0. This results in a variance-covariance matrix with rank $\leq t - 1$, where t = the number of elements in the population. This is because every row (and column) of the variance-covariance matrix sums to 0. If the members of a finite population are mutually unrelated (for example, a set of unrelated sires), the variance-covariance matrix usually has d for diagonal elements and $-d/(t - 1)$ for all off-diagonal elements. If the population refers to additive genetic values of a finite set of related animals, the variance-covariance matrix would be $\mathbf{A}\sigma_a^2$, but with every row (and column) of \mathbf{A} summing to 0 and σ_a^2 having some value different from the infinite model value.

With respect to a factorial design with 2 factors with random and finite levels the following relationship exists. Let γ_{ij} represent the interaction variables. Then

$$\sum_{i=1}^{q_1} \gamma_{ij} = 0 \text{ for all } j = 1, \dots, q_2,$$

and

$$\sum_{j=1}^{q_2} \gamma_{ij} = 0 \text{ for all } i = 1, \dots, q_1, \quad (8)$$

where q_1 and q_2 are the numbers of levels of the first and second factors in the two populations.

Similarly for 3 factor interactions, γ_{ijk} ,

$$\begin{aligned} \sum_{k=1}^{q_3} \gamma_{ijk} &= 0 \text{ for all } i = 1, \dots, q_1, j = 1, \dots, q_2, \\ \sum_{j=1}^{q_2} \gamma_{ijk} &= 0 \text{ for all } i = 1, \dots, q_1, k = 1, \dots, q_3, \quad \text{and} \\ \sum_{i=1}^{q_1} \gamma_{ijk} &= 0 \text{ for all } j = 1, \dots, q_2, k = 1, \dots, q_3. \end{aligned} \quad (9)$$

This concept can be extended to any number of factors. The same principles regarding interactions can be applied to nesting factors if we visualize nesting as being a factorial design with planned disconnectedness. For example, let the first factor be sires and the second dams with 2 sires and 5 dams in the experiment. In terms of a factorial design the subclass numbers (numbers per litter, eg.) are

Sires	Dams				
	1	2	3	4	5
1	5	9	8	0	0
2	0	0	0	7	10

If this were a variance component estimation problem, we could estimate σ_s^2 and σ_e^2 but not σ_d^2 and σ_{sd}^2 . We can estimate $\sigma_d^2 + \sigma_{sd}^2$ and this would usually be called $\sigma_{d/s}^2$.

6 Covariance Matrices

Consider the model

$$\mathbf{y} = \mathbf{X}\beta + \sum_i \mathbf{Z}_i \mathbf{u}_i + \text{possible interactions} + \mathbf{e}. \quad (10)$$

The \mathbf{u}_i represent main effects. The i^{th} factor has t_i levels in the population. Under the traditional mixed model for variance components all $t_i \rightarrow$ infinity. In that case $Var(\mathbf{u}_i) = \mathbf{I}\sigma_i^2$ for all i , and all interactions have variance-covariance that are \mathbf{I} times a scalar. Further, all subvectors of \mathbf{u}_i and those subvectors for interactions are mutually uncorrelated.

Now with possible finite t_i

$$Var(\mathbf{u}_i) = \begin{pmatrix} 1 & & -1/(t_i - 1) \\ & \ddots & \\ -1/(t_i - 1) & & 1 \end{pmatrix} \sigma_i^2. \quad (11)$$

This notation denotes one's for diagonals and all off-diagonal elements = $-1/(t_i - 1)$. Now denote by γ_{gh} the interactions between levels of \mathbf{u}_g and \mathbf{u}_h . Then there are $t_g t_h$ interactions in the population and the variance-covariance matrix has the following form, where i denotes the level of the g^{th} factor and j the level of the h^{th} factor. The diagonals are $Var(\gamma_{gh})$.

$$\begin{aligned} \text{All elements } ij \text{ with } ij' &= -Var(\gamma_{gh})/(t_h - 1). \\ \text{All elements } ij \text{ with } i'j &= -Var(\gamma_{gh})/(t_g - 1). \\ \text{All elements } ij \text{ with } i'j' &= -Var(\gamma_{gh})/(t_g - 1)(t_h - 1). \end{aligned} \quad (12)$$

i' denotes not equal to i , etc.

To illustrate suppose we have two levels of a first factor and 3 levels of a second. The variance-covariance matrix of

$$\begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{pmatrix} = \begin{pmatrix} 1 & -1/2 & -1/2 & -1 & 1/2 & 1/2 \\ & 1 & -1/2 & 1/2 & -1 & 1/2 \\ & & 1 & 1/2 & 1/2 & -1 \\ & & & 1 & -1/2 & -1/2 \\ & & & & 1 & -1/2 \\ & & & & & 1 \end{pmatrix} \sigma_{\gamma_{gh}}^2$$

Suppose that $t_g \rightarrow$ infinity. Then the four types of elements of the variance-covariance matrix would be

$$[1, -1/(t_h - 1), 0, 0] \text{Var}(\gamma_{gh}).$$

This is a model sometimes used for interactions in the two way mixed model with levels of columns fixed.

Now consider 3 factor interactions, γ_{fgh} . Denote by i, j, k the levels of $\mathbf{u}_f, \mathbf{u}_g$, and \mathbf{u}_h , respectively. The elements of the variance-covariance matrix except for the scalar, $\text{Var}(\gamma_{fgh})$ are as follows.

$$\begin{aligned} \text{all diagonals} &= 1. \\ ijk \text{ with } ijk' &= -1/(t_h - 1). \\ ijk \text{ with } ij'k &= -1/(t_g - 1). \\ ijk \text{ with } i'jk &= -1/(t_f - 1). \\ ijk \text{ with } ij'k' &= 1/(t_g - 1)(t_h - 1) \\ ijk \text{ with } i'jk' &= 1/(t_f - 1)(t_h - 1) \\ ijk \text{ with } i'j'k &= 1/(t_f - 1)(t_g - 1) \\ ijk \text{ with } i'j'k' &= 1/(t_f - 1)(t_g - 1)(t_h - 1) \end{aligned} \tag{13}$$

To illustrate, a mixed model with $\mathbf{u}_g, \mathbf{u}_h$ fixed and $t_f \rightarrow$ infinity, the above become 1, $-1/(t_h - 1), -1/(t_g - 1), 0, k \ 1/(t_g - 1)(t_h - 1), 0, 0, 0$. If levels of all factors \rightarrow infinity, the variance-covariance matrix is $\mathbf{IVar}(\gamma_{fgh})$.

Finally let us look at 4 factor interactions γ_{efgh} with levels of $\mathbf{u}_e, \mathbf{u}_f, \mathbf{u}_g, \mathbf{u}_h$ denoted by i, j, k, m , respectively. Except for the scalar $\text{Var}(\gamma_{efgh})$ the variance-covariance matrix has elements like the following.

$$\begin{aligned} \text{all diagonals} &= 1. \\ ijk m \text{ with } ijk m' &= -1/(t_h - 1), \text{ and} \\ ijk m \text{ with } ijk' m &= -1/(t_g - 1), \text{ and} \\ &\text{etc.} \end{aligned}$$

$$\begin{aligned}
ijkm \text{ with } ijk'm' &= 1/(t_g - 1)(t_h - 1), \text{ and} \\
ijkm \text{ with } ij'km' &= 1/(t_f - 1)(t_h - 1) \\
&\textit{etc.} \\
ijkm \text{ with } ij'k'm' &= 1/(t_f - 1)(t_g - 1)(t_h - 1) \text{ and} \\
ijkm \text{ with } i'jk'm' &= 1/(t_e - 1)(t_g - 1)(t_h - 1) \\
&\textit{etc.} \\
ijk \text{ with } i'j'k'm' &= 1/(t_e - 1)(t_f - 1)(t_g - 1)(t_h - 1). \tag{14}
\end{aligned}$$

Note that for all interactions the numerator is 1, the denominator is the product of the $t - 1$ for subscripts differing, and the sign is plus if the number of differing subscripts is even, and negative if the number of differing subscripts is odd. This set of rules applies to any interactions among any number of factors.

7 Estimability and Predictability

Previous chapters have emphasized the importance of consideration of estimability when \mathbf{X} does not have full column rank, and this is usually the case in application. Now if we apply the same rules given in Chapter 2 for checking estimability and find that an element of β , eg. μ , is estimable, the resulting estimate can be meaningless in sampling from finite populations. To illustrate suppose we have a model,

$$y_{ij} = \mu + s_i + e_{ij}.$$

Suppose that the s_i represent a random sample of 2 from a finite population of 5 correlated sires. Now \mathbf{X} is a column vector of 1's and consequently μ is estimable by our usual rules. It seems obvious, however, that an estimate of μ has no meaning except as we define the population to which it refers. If we estimate μ by GLS does $\hat{\mu}$ refer to the mean averaged over the 2 sires in the sample or averaged over the 5 sires in the population? Looking at the problem in this manner suggests that we have a problem in prediction. Then the above question can be formulated as two alternatives, namely prediction of $\mu + \frac{1}{5} \sum_{i=1}^5 s_i$ versus prediction of $\mu + \frac{1}{2} \sum_{i=1}^2 s_i$, where the second alternative involves summing over the 2 sires in the sample. Of course we could, if we choose, predict $\mu + \mathbf{k}'\mathbf{s}$, where \mathbf{k} is any vector with 5 elements and with $\mathbf{k}'\mathbf{1} = 1$. The variance of $\hat{\mu}$, the GLS estimator or the solution to μ in mixed model equations, is identical to the variance of error of prediction of $\mu + .2 \sum_{i=1}^5 s_i$ and not equal to the variance of error of prediction of $\mu + .5 \sum_{i=1}^2 s_i$. Let us illustrate with some data. Suppose there are 20, 5 observations on sires 1, 2 respectively. Suppose $\mathbf{R} = 50 \mathbf{I}$ and

$$\mathbf{G} = \begin{pmatrix} 4 & -1 & -1 & -1 & -1 \\ & 4 & -1 & -1 & -1 \\ & & 4 & -1 & -1 \\ & & & 4 & -1 \\ & & & & 4 \end{pmatrix}.$$

Then the mixed model coefficient matrix (not including s_3, s_4, s_5) is

$$\frac{1}{30} \begin{pmatrix} 15 & 12 & 3 \\ & 20 & 2 \\ & & 11 \end{pmatrix}$$

with inverse

$$\frac{1}{9} \begin{pmatrix} 36 & -21 & -6 \\ & 26 & 1 \\ & & 26 \end{pmatrix}.$$

This gives the solution

$$\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 6 & 3 \\ 2 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}.$$

The variance of error of prediction of $\mu + .5(s_1 + s_2)$ is

$$(1 \ .5 \ .5) (\text{Inverse matrix}) (1 \ .5 \ .5)' = 2.5.$$

This is not equal to 4, the variance of $\hat{\mu}$ from the upper left diagonal of the inverse.

Now let us set up equations for BLUP including all 5 sires. Since \mathbf{G} is now singular we need to use one of the methods of Section 5.10. The non-symmetric set of equations is

$$\begin{pmatrix} .5 & .4 & .1 & 0 & 0 & 0 \\ 1.5 & 2.6 & -.1 & 0 & 0 & 0 \\ 0 & -.4 & 1.4 & 0 & 0 & 0 \\ -.5 & -.4 & -.1 & 1. & 0 & 0 \\ -.5 & -.4 & -.1 & 0 & 1. & 0 \\ -.5 & -.4 & -.1 & 0 & 0 & 1. \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{s} \end{pmatrix} = \begin{pmatrix} .4 & .1 \\ 1.6 & -.1 \\ -.4 & .4 \\ -.4 & -.1 \\ -.4 & -.1 \\ -.4 & -.1 \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}.$$

Post-multiplying the inverse of this coefficient matrix by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{G} \end{pmatrix}$$

we get as the prediction error variance matrix the following

$$9^{-1} \begin{pmatrix} 36 & -21 & -6 & 9 & 9 & 9 \\ & 26 & 1 & -9 & -9 & -9 \\ & & 26 & -9 & -9 & -9 \\ & & & 36 & -9 & -9 \\ & & & & 36 & -9 \\ & & & & & 36 \end{pmatrix}.$$

The upper 3×3 submatrix is the same as the inverse when only sires 1 and 2 are included. The solution is

$$\begin{pmatrix} \hat{\mu} \\ \hat{s} \end{pmatrix} = 9^{-1} \begin{pmatrix} 6 & 3 \\ 2 & -2 \\ -2 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}.$$

$$\hat{s}_3, \hat{s}_4, \hat{s}_5 = 0$$

as would be expected because these sires are unrelated to the 2 with progeny relative to the population of 5 sires. The solution to $\hat{\mu}$, \hat{s}_1 , \hat{s}_2 are the same as before. The prediction error variance of $\mu + .2 \sum s_i$ is

$$(1 \ .2 \ .2 \ .2 \ .2 \ .2) (\text{Inverse matrix}) (1 \ .2 \ .2 \ .2 \ .2 \ .2)' = 4,$$

the value of the upper diagonal element of the inverse. By the same reasoning we find that \hat{s}_j is BLUP of $s_j - .2 \sum_{i=1}^5 s_i$ and not of $s_i - .5 (s_1 + s_2)$ for $i=1,2$. Using the former function with the inverse of the matrix of the second set of equations we obtain for s_1 the value, 2.889. This is also the value of the corresponding diagonal. In contrast the variance of the error of prediction of $s_1 - .5 (s_1 + s_2)$ is 1.389. Thus \hat{s}_j is the BLUP of $s_j - .2 \sum_{i=1}^5 s_i$.

The following rules insure that one does not attempt to predict $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ that is not predictable.

1. $\mathbf{K}'\boldsymbol{\beta}$ must be estimable in a model in which $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$.
2. Pretend that there are no missing classes or subclasses involving all levels of \mathbf{u}_i in the population.
3. Then if $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ is estimable in such a design with \mathbf{u} regarded as fixed, $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ is predictable.

Use the rules of Chapter 2 in checking estimability.

For an example suppose we have sire \times treatment design with 3 treatments and 2 sires regarded as a random sample from an infinite population of possibly related sires. Let the model be

$$y_{ijk} = \mu + s_i + t_j + \gamma_{ij} + e_{ijk}.$$

μ, t_j are fixed

$$Var(\mathbf{s}) = \mathbf{I}\sigma_s^2.$$

Var (γ) when γ are ordered treatments in sires is

$$\begin{pmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B} \end{pmatrix},$$

where

$$\mathbf{B} = \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix} \sigma_{\gamma}^2.$$

Suppose we have progeny on all 6 sire \times treatment combinations except (2,3). This creates no problem in prediction due to rule 1 above. Now we can predict for example

$$t_1 + \sum_{i=1} c_i (s_i + \gamma_{i1}) - t_2 - \sum_{i=1} d_i (s_i + \gamma_{i2})$$

where

$$\sum_i c_i = \sum_i d_i = 1.$$

That is, we can predict the difference between treatments 1 and 2 averaged over any sires in the population, including some not in the sample of 2 sires if we choose to do so. In fact, as we shall see, BLUE of $(t_1 - t_2)$ is BLUP of treatment 1 averaged equally over all sires in the population minus treatment 2 averaged equally over all sires in the population.

Suppose we want to predict the merit of sire 1 versus sire 2. By the rules above, $(s_1 - s_2)$ is not predictable, but

$$s_1 + \sum_{j=1}^3 c_j (t_j + \gamma_{1j}) - s_2 - \sum_{j=1}^3 d_j (t_j + \gamma_{2j})$$

is predictable if $\sum_j c_j = \sum_j d_j = 1$. That is, we can predict sire differences only if we specify treatments, and obviously only treatments 1, 2, 3. We cannot predict unbiasedly, from the data, sire differences associated with some other treatment or treatments. But note that even though subclass (2,3) is missing we can still predict $s_1 + t_3 + \gamma_{13} - s_2 - t_3 - \gamma_{23}$. In contrast, if sires as well as treatments were fixed, this function could not be estimated unbiasedly.

8 BLUE When Some u_i Are Finite

Calculation of BLUE and BLUP when there are finite levels of random factors must take into account the fact that there may be singular \mathbf{G} . Consider the simple one way

case with a population of 4 related sires. Suppose

$$\mathbf{A} = \begin{pmatrix} 1. & -.2 & -.3 & -.5 \\ & 1. & -.2 & -.6 \\ & & 1 & .5 \\ & & & 1.6 \end{pmatrix}.$$

Suppose we have progeny numbers on these sires that are 9, 5, 3, 0. Suppose the model is

$$\begin{aligned} y_{ijk} &= \mu + s_i + e_{ij}. \\ \text{Var}(\mathbf{s}) &= \mathbf{A}\sigma_s^2. \\ \text{Var}(\mathbf{e}) &= \mathbf{I}\sigma_e^2. \end{aligned}$$

Then if we wish to include all 4 sires in the mixed model equations we must resort to the methods of Sect. 5.10 since \mathbf{G} is singular. One of those methods is to solve

$$\begin{pmatrix} \left(\begin{array}{cc} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{array} \right) \begin{pmatrix} 17 & 9 & 5 & 3 & 0 \\ 9 & 9 & 0 & 0 & 0 \\ 5 & 0 & 5 & 0 & 0 \\ 3 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sigma_e^{-2} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \cdot \\ \cdot \\ \hat{s}_4 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{pmatrix} \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \\ 0 \end{pmatrix} / \sigma_e^2. \quad (15)$$

$$\hat{\mu} \text{ is BLUP of } \mu + \frac{1}{4} \sum_i s_i.$$

$$\hat{s}_j \text{ is BLUP of } s_j - \frac{1}{4} \sum_i s_i.$$

The inverse of the coefficient matrix post-multiplied by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{A}\sigma_s^2 \end{pmatrix}$$

is the variance-covariance matrix of errors of predictions of these functions.

If we had chosen to include only the 3 sires with progeny, the mixed model equations would be

$$\begin{pmatrix} 17 & 9 & 5 & 3 \\ 9 & 9 & 0 & 0 \\ 5 & 0 & 5 & 0 \\ 3 & 0 & 0 & 3 \end{pmatrix} \sigma_e^{-2} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \left(\begin{array}{ccc} 1 & -.2 & -.3 \\ -.2 & 1 & -.2 \\ -.3 & -.2 & 1 \end{array} \right)^{-1} & & \\ 0 & & & \sigma_s^{-2} \end{pmatrix}$$

$$\begin{pmatrix} \hat{\mu} \\ \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{pmatrix} \sigma_e^{-2}. \quad (16)$$

This gives the same solution to $\hat{\mu}$, \hat{s}_1 , \hat{s}_2 , \hat{s}_3 as the solution to (15), and the inverse of the coefficient matrix gives the same prediction variances. Even though s_4 is not included, $\hat{\mu}$ predicts $\mu + \frac{1}{4}\sum_{i=1}^4 s_i$, and \hat{s}_j predicts $s_j - \frac{1}{4}\sum_{i=1}^4 s_i$. \hat{s}_4 can be computed by

$$- [.5 \ .6 \ .5] \begin{pmatrix} 1 & -.2 & -.3 \\ -.2 & 1 & -.2 \\ -.3 & -.2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \end{pmatrix}.$$

As another example suppose we have a sire by treatment model with an infinite population of sires. The n_{ij} are

	1	2
1	0	8
2	9	2
3	6	0

$$\text{Var}(\mathbf{s}) = 2\mathbf{I}, \text{Var}(\mathbf{e}) = 10\mathbf{I},$$

Var (γ) including missing subclasses is

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & & 1 & -1 & 0 & 0 \\ & & & 1 & 0 & 0 \\ & & & & 1 & -1 \\ & & & & & 1 \end{pmatrix} / 2.$$

If we do not include γ_{11} and γ_{32} in the solution the only submatrix of \mathbf{G} that is singular is the 2x2 block pertaining to γ_{21} , γ_{22} . The GLS equations regarding \mathbf{u} as fixed are

$$\frac{1}{10} \begin{pmatrix} 8 & 0 & 0 & 0 & 8 & 8 & 0 & 0 & 0 \\ & 11 & 0 & 9 & 2 & 0 & 9 & 2 & 0 \\ & & 6 & 6 & 0 & 0 & 0 & 0 & 6 \\ & & & 15 & 0 & 0 & 9 & 0 & 6 \\ & & & & 10 & 8 & 0 & 2 & 0 \\ & & & & & 8 & 0 & 0 & 0 \\ & & & & & & 9 & 0 & 0 \\ & & & & & & & 2 & 0 \\ & & & & & & & & 6 \end{pmatrix} \begin{pmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \hat{s}_3 \\ \hat{t}_1 \\ \hat{t}_2 \\ \hat{\gamma}_{12} \\ \hat{\gamma}_{21} \\ \hat{\gamma}_{22} \\ \hat{\gamma}_{31} \end{pmatrix} = \begin{pmatrix} y_{1..} \\ y_{2..} \\ y_{3..} \\ y_{1.} \\ y_{2.} \\ y_{12.} \\ y_{21.} \\ y_{22.} \\ y_{31.} \end{pmatrix} \frac{1}{10} \quad (17)$$

Then we premultiply the 7th and 8th equations of (17) by

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} / 2$$

and add to the diagonal coefficients, (.5, .5, .5, 0, 0, 2, 1, 1, 2). The solution to the resulting equations is BLUP. If we had included γ_{12} and γ_{32} , we would premultiply the last 6 GLS equations (equations for γ) by $\text{Var}(\gamma)$ and then add to the diagonals, (.5, .5, .5, 0, 0, 1, 1, 1, 1, 1, 1). When all elements of a population are included in a BLUP solution, an interesting property becomes apparent. The same summing to 0's occurs in the BLUP solution as is true in the corresponding elements of the finite populations described in Section 4.

9 An Easier Computational Method

Finite populations complicate computation of BLUE and BLUP because non-diagonal and singular \mathbf{G} matrices exist. But if the model is that of Section 2, that is, finite populations of unrelated elements with common variance, computations can be carried out with diagonal submatrices for \mathbf{G} . The resulting $\hat{\mathbf{u}}$ do not always predict the same functions predicted by using the actual \mathbf{G} matrices, but appropriate linear functions of them do. We illustrate with a simple one way case.

$$y_{ij} = \mu + a_i + e_{ij}. \quad i = 1, 2, 3.$$

$$\text{Var}(\mathbf{a}) = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix},$$

$$\text{Var}(\mathbf{e}) = 10\mathbf{I}.$$

$$n_i = (5, 3, 2), y_i = (10, 8, 6).$$

Using singular \mathbf{G} the nonsymmetric mixed model equations are

$$\begin{pmatrix} 1. & .5 & .3 & .2 \\ .5 & 2. & -.3 & -.2 \\ -.1 & -.5 & 1.6 & -.2 \\ -.4 & -.5 & -.3 & 1.4 \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} 2.4 \\ .6 \\ 0 \\ -.6 \end{pmatrix}. \quad (18)$$

The solution is [2.4768, -.2861, .0899, .1962]. Note that

$$\sum \hat{u}_i = 0.$$

We can obtain the same solution by pretending that $Var(\mathbf{a}) = 3\mathbf{I}$. Then the mixed model equations are

$$\begin{pmatrix} 1 & .5 & .3 & .2 \\ .5 & .5 + 3^{-1} & 0 & 0 \\ .3 & 0 & .3 + 3^{-1} & 0 \\ .2 & 0 & 0 & .2 + 3^{-1} \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 1.0 \\ .8 \\ .6 \end{pmatrix} \quad (19)$$

The inverse of (15.19) is different from the inverse of (15.18) post-multiplied by

$$\begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{G} \end{pmatrix}.$$

The inverse of (19) does not yield prediction error variances. To obtain prediction error variances of $\mu + \bar{a}$, and of $a_i - \bar{a}$, pre-multiply it by

$$\frac{1}{3} \begin{pmatrix} 3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

and post-multiply that product by the transpose of this matrix. This is a consequence of the fact that the solution to (19) is BLUP of

$$\frac{1}{3} \begin{pmatrix} 3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} \mu \\ s_1 \\ s_2 \\ s_3 \end{pmatrix}.$$

In most cases use of diagonal \mathbf{G} does not result in the same solution as using the true \mathbf{G} , and the inverse never yields directly the prediction error variance-covariance matrix.

Rules for deriving diagonal submatrices of \mathbf{G} to use in place of singular submatrices follow. For main effects say of \mathbf{u}_i with t_i levels substitute for the \mathbf{G} submatrix described in Section 6, $\mathbf{I}\sigma_{*i}^2$, where

$$\begin{aligned} \sigma_{*i}^2 &= \frac{t_i}{t_{i-1}} \sigma_i^2 - \sum_j \frac{t_i}{(t_{i-1})(t_{j-1})} \sigma_{ij}^2 + \sum_{j,k} \frac{t_i}{(t_{i-1})(t_{j-1})(t_{k-1})} \sigma_{ijk}^2 \\ &\quad - \sum_{j,k,m} \frac{t_i}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})} \sigma_{ijkm}^2 \text{ etc.} \end{aligned} \quad (20)$$

for 5 factor, 6 factor interactions.

σ_i^2 refers to the scalar part of the variance of the i^{th} factor, σ_{ij}^2 refers to 2 factor interactions involving \mathbf{u}_i , σ_{ijk}^2 refers to 3 factor interactions involving \mathbf{u}_i , etc. Note that the signs alternate

$$\sigma_{*ij}^2 = \frac{t_i t_j}{(t_{i-1})(t_{j-1})} \sigma_{ij}^2 - \sum_k \frac{t_i t_j}{(t_{i-1})(t_{j-1})(t_{k-1})} \sigma_{ijk}^2$$

$$+ \sum_{k,m} \frac{t_i t_j}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})} \sigma_{ijkm}^2 \text{ etc.} \quad (21)$$

$$\sigma_{*ijk}^2 = \frac{t_i t_j t_k}{(t_{i-1})(t_{j-1})(t_{k-1})} \sigma_{ijk}^2 - \sum_m \frac{t_i t_j t_k}{(t_{i-1})(t_{j-1})(t_{k-1})(t_{m-1})} \sigma_{ijk}^2 + \text{etc.} \quad (22)$$

Higher order interactions for σ_*^2 follow this same pattern with alternating signs. The sign is positive when the number of factors in the denominator minus the number in the numerator is even.

It appears superficially that one needs to estimate the different σ_i^2 , σ_{ij}^2 , σ_{ijk}^2 , etc., and this is difficult because non-diagonal, singular submatrices of \mathbf{G} are involved. But if one plans to use their diagonal representations, one might as well estimate the σ_*^2 directly by any of the standard procedures for the conventional mixed model for variance components estimation. Then if for pedagogical or other reasons one wishes estimates of σ^2 rather than σ_*^2 , one can use equations (20), (21), (22) that relate the two to affect the required linear transformation.

The solution using diagonal \mathbf{G} should not be assumed to be the same as would have been obtained from use of the true \mathbf{G} matrix. If we consider predictable functions as defined in Section 7 and take these same functions of the solution using diagonal \mathbf{G} we do obtain BLUP. Similarly using these functions we can derive prediction error variances using a g-inverse of the coefficient matrix with diagonal \mathbf{G} .

10 Biased Estimation

If we can legitimately assume that there is no expected pattern of values of the levels of a fixed factor and no expected pattern of values of interactions between levels of fixed factors, we can pretend that these fixed factors and interactions are populations with finite levels and proceed to compute biased estimators as though we are computing BLUP of random variables. Instead of prediction error variance as derived from the g-inverse of the coefficient matrix we obtain estimated mean squared errors.