Chapter 8
Unbiased Methods for G and R Unknown

C. R. Henderson

1984 - Guelph

Previous chapters have dealt with known $\mathbf{G}$ and $\mathbf{R}$ or known proportionality of these matrices. In these cases BLUE, BLUP, exact sampling variances, and exact tests of hypotheses exist. In this chapter we shall be concerned with the unsolved problem of what are "best" estimators and predictors when $\mathbf{G}$ and $\mathbf{R}$ are unknown even to proportionality. We shall construct many unbiased estimators and predictors and under certain circumstances compute their variances. Tests of hypotheses pose more serious problems, for only approximate tests can be made. We shall be concerned with three different situations regarding estimation and prediction. These are described in Henderson and Henderson (1979) and in Henderson, Jr. (1982).

1. Methods of estimation and prediction not involving $\mathbf{G}$ and $\mathbf{R}$.

2. Methods involving $\mathbf{G}$ and $\mathbf{R}$ in which assumed values, say $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are used in the computations and these are regarded as constants.

3. The same situation as 2, but $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are regarded more realistically as estimators from data and consequently are random variables.

1 Unbiased Estimators

Many unbiased estimators of $\mathbf{K}'\boldsymbol{\beta}$ can be computed. Some of these are much easier than GLS or mixed models with $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ used. Also some of them are invariant to $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$. The first, and one of the easiest, is ordinary least squares (OLS) ignoring $\mathbf{u}$.

Solve for $\boldsymbol{\beta}^o$ in

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^o \ = \ \mathbf{X}'\mathbf{y}. \tag{1}$$

Then $E(\mathbf{K}'\boldsymbol{\beta}^o) = E[\mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y}] = \mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{K}'\boldsymbol{\beta}$ if $\mathbf{K}'\boldsymbol{\beta}$ is estimable. The variance of $\mathbf{K}'\boldsymbol{\beta}^o$ is

$$\mathbf{K}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{K}, \tag{2}$$

and this can be evaluated easily for chosen $\tilde{\mathbf{G}}$, $\tilde{\mathbf{R}}$, but it is valid only if $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are regarded as fixed.

A second estimator is analogous to weighted least squares. Let $\mathbf{D}$ be a diagonal matrix formed from the diagonals of $(\mathbf{Z}\tilde{\mathbf{G}}\mathbf{Z}' + \tilde{\mathbf{R}})$. Then solve

$$\mathbf{X}'\mathbf{D}^{-1}\mathbf{X}\boldsymbol{\beta}^o \; = \; \mathbf{X}'\mathbf{D}^{-1}\mathbf{y}. \tag{3}$$

$\mathbf{K}'\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$ if estimable.

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) \; = \; \mathbf{K}'(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{D}^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{D}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^-\mathbf{K}. \tag{4}$$

A third possibility if $\tilde{\mathbf{R}}^{-1}$ is easy to compute, but $\tilde{\mathbf{V}}^{-1}$ is not easy, is to solve

$$\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X}\boldsymbol{\beta}^o \; = \; \mathbf{X}\tilde{\mathbf{R}}^{-1}\mathbf{y}. \tag{5}$$

$$Var(\mathbf{K}'\boldsymbol{\beta}^o) \; = \; \mathbf{K}'(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{R}}^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\tilde{\mathbf{R}}^{-1}\mathbf{X}(\mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X})^-\mathbf{K}. \tag{6}$$

These methods all would seem to imply that the diagonals of $\mathbf{G}^{-1}$ are large relative to diagonals of $\mathbf{R}^{-1}$.

Other methods would seem to imply just the opposite, that is, the diagonals of $\mathbf{G}^{-1}$ are small relative to $\mathbf{R}^{-1}$. One of these is OLS regarding $\mathbf{u}$ as fixed for purposes of computation. That is solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}. \tag{7}$$

Then if $\mathbf{K}'\boldsymbol{\beta}$ is estimable under a fixed $\mathbf{u}$ model, $\mathbf{K}'\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$. However, if $\mathbf{K}'\boldsymbol{\beta}$ is estimable under a random $\mathbf{u}$ model, but is not estimable under a fixed $\mathbf{u}$ model, $\mathbf{K}'\boldsymbol{\beta}^o$ may be biased. To forestall this, find a function $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ that is estimable under a fixed $\mathbf{u}$ model. Then $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o$ is an unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$.

$$Var(\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o) \; = \; [\mathbf{K}' \;\; \mathbf{M}']\mathbf{C}\mathbf{W}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{W}\mathbf{C} \begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \tag{8}$$

$$= \; (\mathbf{K}' \;\; \mathbf{M}')\mathbf{C}\mathbf{W}'\mathbf{R}\mathbf{W}\mathbf{C} \begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \; + \; \mathbf{M}'\mathbf{G}\mathbf{M}, \tag{9}$$

where $\mathbf{C}$ is a g-inverse of the matrix of (7) and $\mathbf{W} \; = \; (\mathbf{X} \;\; \mathbf{Z})$.

The method of (9) is simpler than (8) if $\mathbf{R}$ has a simple form compared to $\mathbf{Z}\mathbf{G}\mathbf{Z}'$. In fact, if $\mathbf{R} \; = \; \mathbf{I}\sigma_e^2$, the first term of (9) becomes

$$(\mathbf{K}' \;\; \mathbf{M}')\mathbf{C} \begin{pmatrix} \mathbf{K} \\ \mathbf{M} \end{pmatrix} \sigma_e^2. \tag{10}$$

Analogous estimators would come from solving

$$\begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{y} \end{pmatrix}. \tag{11}$$

Another one would use $\mathbf{D}^{-1}$ in place of $\mathbf{R}^{-1}$ where $\mathbf{D}$ is a diagonal matrix formed from the diagonals of $\tilde{\mathbf{R}}$. In both of these last two methods $\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o$ would be the estimator of $\mathbf{K}'\boldsymbol{\beta}$, and we require that $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ be estimable under a fixed $\mathbf{u}$ model.

From (11)

$$\begin{aligned} Var(\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}^o) &= (\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'\tilde{\mathbf{R}}^{-1}(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}\begin{pmatrix}\mathbf{K}\\ \mathbf{M}\end{pmatrix} \\ &= (\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'\tilde{\mathbf{R}}^{-1}\mathbf{R}\tilde{\mathbf{R}}^{-1}\mathbf{W}\mathbf{C}\begin{pmatrix}\mathbf{K}\\ \mathbf{M}\end{pmatrix} \\ &\quad + \mathbf{M}'\mathbf{G}\mathbf{M}. \end{aligned} \tag{12}$$

When $\mathbf{D}^{-1}$ is substituted for $\tilde{\mathbf{R}}^{-1}$ the expression in (12) is altered by making this same substitution.

Another method which is a compromise between (1) and (11) is to ignore a subvector of $\mathbf{u}$, say $\mathbf{u}_2$, then compute by OLS regarding the remaining subvector of $\mathbf{u}$, say $\mathbf{u}_1$, as fixed. The resulting equations are

$$\begin{pmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_1\\ \mathbf{Z}_1'\mathbf{X} & \mathbf{Z}_1'\mathbf{Z}_1\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}^o\\ \mathbf{u}_1^o\end{pmatrix} = \begin{pmatrix}\mathbf{X}'\mathbf{y}\\ \mathbf{Z}_1'\mathbf{y}\end{pmatrix}. \tag{13}$$

$(\mathbf{Z}_1 \ \mathbf{Z}_2)$ is a partitioning of $\mathbf{Z}$ corresponding to $\mathbf{u}' = (\mathbf{u}_1' \ \mathbf{u}_2')$. Now to insure unbiasedness of the estimator of $\mathbf{K}'\boldsymbol{\beta}$ we need to find a function,

$$\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}_1,$$

that is estimable under a fixed $\mathbf{u}_1$ model. Then the unbiased estimator of $\mathbf{K}'\boldsymbol{\beta}$ is

$$\mathbf{K}'\boldsymbol{\beta}^o + \mathbf{M}'\mathbf{u}_1^o,$$

The variance of this estimator is

$$(\mathbf{K}' \ \mathbf{M}')\mathbf{C}\mathbf{W}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{W}\mathbf{C} \begin{pmatrix}\mathbf{K}\\ \mathbf{M}\end{pmatrix}. \tag{14}$$

$\mathbf{W} = (\mathbf{X} \ \mathbf{Z}_1)$, and $\mathbf{Z}\mathbf{G}\mathbf{Z}'$ refers to the entire $\mathbf{Z}\mathbf{u}$ vector, and $\mathbf{C}$ is some g-inverse of the matrix of (13).

Let us illustrate some of these methods with a simple example.

$$\mathbf{X}' = [1 \ \ 1 \ \ 1 \ \ 1 \ \ 1],$$

$$\mathbf{Z}' = \begin{pmatrix}1 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}, \quad \mathbf{R} = 15\mathbf{I}, \quad \mathbf{G} = 2\mathbf{I},$$

$$\mathbf{y}' = [6\ \ 8\ \ 7\ \ 5\ \ 7].$$

$$Var(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R} = \begin{pmatrix} 17 & 2 & 0 & 0 & 0 \\ & 17 & 0 & 0 & 0 \\ & & 17 & 2 & 0 \\ & & & 17 & 0 \\ & & & & 17 \end{pmatrix}.$$

$\boldsymbol{\beta}$ is estimable. By the method of (1) we solve

$$5\boldsymbol{\beta}^o = 33.$$

$$\boldsymbol{\beta}^o = 6.6.$$

$$Var(\boldsymbol{\beta}^o) = .2\ (1\ \ 1\ \ 1\ \ 1\ \ 1)\ \ \text{Var}\ (\mathbf{y}) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \ .2 = 3.72.$$

By the method of (7) the equations to be solved are

$$\begin{pmatrix} 5 & 2 & 2 & 1 \\ & 2 & 0 & 0 \\ & & 2 & 0 \\ & & & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} 33 \\ 14 \\ 12 \\ 7 \end{pmatrix}.$$

A solution is (0, 7, 6, 7). Because $\boldsymbol{\beta}$ is not estimable when $\mathbf{u}$ is fixed, we need some function with $\mathbf{k}' = 1$ and $\mathbf{m}'$ such that $(\mathbf{k}'\ \ \mathbf{m}') \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{pmatrix}$ is estimable. A possibility is $(3\ \ 1\ \ 1\ \ 1)/3$. The resulting estimate is $20/3 \neq 6.6$, our previous estimate. To find the variance of the estimator by method (8) we can use a g-inverse.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ & .5 & 0 & 0 \\ & & .5 & 0 \\ & & & 1 \end{pmatrix}.$$

$$(\mathbf{k}'\ \mathbf{m}')\mathbf{CW}' = \frac{1}{3}\ (3\ 1\ 1\ 1) \begin{pmatrix} 0 & 0 & 0 & 0 \\ & .5 & 0 & 0 \\ & & .5 & 0 \\ & & & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{6}\ (1\ 1\ 1\ 1\ 2).$$

Then $Var(\boldsymbol{\beta}^o) = 4 \neq 3.72$ of previous result. By the method of (9) we obtain $3.333 + .667 = 4$ also.

BLUE would be obtained by using the mixed model equations with $\mathbf{R} = 15\mathbf{I}$, $\mathbf{G} = 2\mathbf{I}$ if these are the true values of $\mathbf{R}$ and $\mathbf{G}$. The resulting equations are

$$\frac{1}{15}\begin{pmatrix} 5 & 2 & 2 & 1 \\ 2 & 9.5 & 0 & 0 \\ 2 & 0 & 9.5 & 0 \\ 1 & 0 & 0 & 8.5 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} 33 \\ 14 \\ 12 \\ 7 \end{pmatrix} /15.$$

$$\boldsymbol{\beta}^o = 6.609.$$

The upper 1×1 of a g-inverse is 3.713, which is less than for any other methods, but of course depends upon true values of $\mathbf{G}$ and $\mathbf{R}$.

2 Unbiased Predictors

The method for prediction of $\mathbf{u}$ used by most animal breeders prior to the recent general acceptance of the mixed model equations was selection index (BLP) with some estimate of $\mathbf{X}\boldsymbol{\beta}$ regarded as a parameter value. Denote the estimate of $\mathbf{X}\boldsymbol{\beta}$ by $\mathbf{X}\tilde{\boldsymbol{\beta}}$. Then the predictor of $\mathbf{u}$ is

$$\tilde{\mathbf{u}} = \tilde{\mathbf{G}}\mathbf{Z}'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}). \tag{15}$$

$\tilde{\mathbf{G}}$ and $\tilde{\mathbf{V}}$ are estimated $\mathbf{G}$ and $\mathbf{V}$.

This method utilizes the entire data vector and the entire variance-covariance structure to predict. More commonly a subset of $\mathbf{y}$ was chosen for each individual element of $\mathbf{u}$ to be predicted, and (15) involved this reduced set of matrices and vectors.

Now if $\mathbf{X}\tilde{\boldsymbol{\beta}}$ is an unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$, $E(\tilde{\mathbf{u}}) = \mathbf{0} = E(\mathbf{u})$ and is unbiased. Even if $\mathbf{G}$ and $\mathbf{R}$ were known, (15) would not represent a predictor with minimum sampling variance. We have already found that for this $\tilde{\boldsymbol{\beta}}$ should be a GLS solution. Further, in selection models (discussed in chapter 13), usual estimators for $\boldsymbol{\beta}$ such as OLS or estimators ignoring $\mathbf{u}$ are biased, so $\tilde{\mathbf{u}}$ is no longer an unbiased predictor.

Another unbiased predictor, if computed correctly, is "regressed least squares" first reported by Henderson (1948). Solve for $\mathbf{u}^o$ in equations (16).

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \mathbf{u}^o \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix} \tag{16}$$

Take a solution for which $E(\mathbf{u}^o) = \mathbf{0}$ in a fixed $\boldsymbol{\beta}$ but random $\mathbf{u}$ model. This can be done by "absorbing" $\boldsymbol{\beta}^o$ to obtain a set of equations

$$\mathbf{Z}'\mathbf{P}\mathbf{Z}\ \mathbf{u}^o = \mathbf{Z}'\mathbf{P}\mathbf{y}, \tag{17}$$

where

$$\mathbf{P} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'].$$

Then any solution to $\mathbf{u}^o$, usually not an unique solution, has expectation $\mathbf{0}$, because $E[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}']\mathbf{y} = (\mathbf{X} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X})\boldsymbol{\beta} = (\mathbf{X} - \mathbf{X})\boldsymbol{\beta} = \mathbf{0}$. Thus $\mathbf{u}^o$ is an unbiased predictor, but not a good one for selection, particularly if the amount of information differs greatly among individuals.

Let some g-inverse of $\mathbf{Z}'\mathbf{P}\mathbf{Z}$ be defined as $\mathbf{C}$. Then

$$Var(\mathbf{u}^o) = \mathbf{C}\mathbf{Z}'\mathbf{P}(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})\mathbf{P}\mathbf{Z}\mathbf{C}, \tag{18}$$

$$Cov(\mathbf{u}, \mathbf{u}^o) = \mathbf{G}\mathbf{Z}'\mathbf{P}\mathbf{Z}\mathbf{C}. \tag{19}$$

Let the i^{th} diagonal of (18) be v_i, and the i^{th} diagonal of (19) be c_i, both evaluated by some estimate of $\mathbf{G}$ and $\mathbf{R}$. Then the regressed least square prediction of u_i is

$$c_i u_i^o / v_i. \tag{20}$$

This is BLP of u_i when the only observation available for prediction is u_i^o. Of course other data are available, and we could use the entire $\mathbf{u}^o$ vector for prediction of each u_i. That would give a better predictor because (18) and (19) are not diagonal matrices.

In fact, BLUP of $\mathbf{u}$ can be derived from $\mathbf{u}^o$. Denote (18) by $\mathbf{S}$ and (19) by $\mathbf{T}$. Then BLUP of $\mathbf{u}$ is

$$\mathbf{T}\mathbf{S}^-\mathbf{u}^o, \tag{21}$$

provided $\mathbf{G}$ and $\mathbf{R}$ are known. Otherwise it would be approximate BLUP.

This is a cumbersome method as compared to using the mixed model equations, but it illustrates the reason why regressed least squares is not optimum. See Henderson (1978b) for further discussion of this method.

3 Substitution Of Fixed Values For G And R

In the methods presented above it appears that some assumption is made concerning the relative values of $\mathbf{G}$ and $\mathbf{R}$. Consequently it seems logical to use a method that approaches optimality as $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ approach $\mathbf{G}$ and $\mathbf{R}$. This would be to substitute $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ for the corresponding parameters in the mixed model equations. This is a procedure which requires no choice among a variety of unbiased methods. Further, it has the desirable property that if $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ are fixed, the estimated sampling variance and prediction error variances are simple to express. Specifically the variances and covariances estimated for $\mathbf{G} = \tilde{\mathbf{G}}$ and $\mathbf{R} = \tilde{\mathbf{R}}$ are precisely the results in (34) to (41) in Chapter 5.

It also is true that the estimators and predictors are unbiased. This is easy to prove for fixed $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ but for estimated (random) $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ we need to invoke a result by Kackar and Harville (1981) presented in Section 4. For fixed $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{R}}$ note that after "absorbing" $\mathbf{u}$ from the mixed model equations we have

$$\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\boldsymbol{\beta}^o = \mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{y}.$$

Then

$$\begin{aligned} E(\mathbf{K}'\boldsymbol{\beta}^o) &= E(\mathbf{K}'(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{y}) \\ &= \mathbf{K}'(\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X})^-\mathbf{X}'\tilde{\mathbf{V}}^{-1}\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{K}'\boldsymbol{\beta}. \end{aligned}$$

Also

$$\hat{\mathbf{u}} = (\mathbf{Z}'\tilde{\mathbf{R}}^{-1}\mathbf{Z} + \tilde{\mathbf{G}}^{-1})^{-1}\mathbf{Z}'\tilde{\mathbf{R}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o).$$

But $\mathbf{X}\boldsymbol{\beta}^o$ is an unbiased estimator of $\mathbf{X}\boldsymbol{\beta}$, $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^o$ with expectation $\mathbf{0}$ and consequently $E(\hat{\mathbf{u}}) = \mathbf{0}$ and is unbiased.

4 Mixed Model Equations With Estimated G and R

It is not a trivial problem to find the expectations of $\mathbf{K}'\boldsymbol{\beta}^o$ and $\hat{\mathbf{u}}$ from mixed model equations with estimated $\mathbf{G}$ and $\mathbf{R}$. Kackar and Harville (1981) derived a very important result for this case. They prove that if $\mathbf{G}$ and $\mathbf{R}$ are estimated by a method having the following properties and substituted in mixed model equations, the resulting estimators and predictors are unbiased. This result requires that

1. $\mathbf{y}$ is symmetrically distributed, that is, $f(\mathbf{y}) = f(-\mathbf{y})$.

2. The estimators of $\mathbf{G}$ and $\mathbf{R}$ are translation invariant.

3. The estimators of $\mathbf{G}$ and $\mathbf{R}$ are even functions of $\mathbf{y}$.

These are not very restrictive requirements because they include a variety of distributions of $\mathbf{y}$ and most of the presently used methods for estimation of variances and covariances.

An interesting consequence of substituting ML estimates of $\mathbf{G}$ and $\mathbf{R}$ for the corresponding parameters of mixed model equations is that the resulting $\mathbf{K}'\boldsymbol{\beta}^o$ are ML and the $\hat{\mathbf{u}}$ are ML of $(\mathbf{u} \mid \mathbf{y})$.

5 Tests Of Hypotheses Concerning β

We have seen that unbiased estimators and predictors can be obtained even though $\mathbf{G}$ and $\mathbf{R}$ are unknown. When it comes to testing hypotheses regarding $\boldsymbol{\beta}$ little is known except that exact tests do not exist apart from a special case that is described below. The problem is that quadratics in $\mathbf{H}'\boldsymbol{\beta}^o - \mathbf{c}$ appropriate for exact tests when $\mathbf{G}$ and $\mathbf{R}$ are known, do not have a χ^2 or any other tractable distribution when $\tilde{\mathbf{G}}$, $\tilde{\mathbf{R}}$ replace $\mathbf{G}$, $\mathbf{R}$ in the computation. What should be done? One possibility is to estimate, if possible $\mathbf{G}$, $\mathbf{R}, \boldsymbol{\beta}$ by ML and then invoke a likelihood ratio test, in which under normality assumptions and large samples, -2 log likelihood ratio is approximated by χ^2. This raises the question of what is a large sample of unbalanced data. Certainly $n \rightarrow \infty$ is not a sufficient condition. Consideration needs to be given to the number of levels of each subvector of $\mathbf{u}$ and to the proportion of missing subclasses. Consequently the value of a χ^2 approximation to the likelihood ratio test is uncertain.

A second and easier approximation is to pretend that $\tilde{\mathbf{G}} = \mathbf{G}$ and $\tilde{\mathbf{R}} = \mathbf{R}$ and proceed to an approximate test using χ^2 as described in Chapter 4 for hypothesis testing with known $\mathbf{G}$, $\mathbf{R}$ and normality assumptions. The validity of this test must surely depend, as it does in the likelihood ratio approximation, upon the number of levels of $\mathbf{u}$ and the balance and lack of missing subclasses.

One interesting case exists in which exact tests of $\boldsymbol{\beta}$ can be made even when we do not know $\mathbf{G}$ and $\mathbf{R}$ to proportionality. The requirements are as follows

1. $Var(\mathbf{e}) = \mathbf{I}\sigma_e^2$, and

2. $\mathbf{H}_0'\boldsymbol{\beta}$ is estimable under a fixed $\mathbf{u}$ model.

Solve for $\boldsymbol{\beta}^o$ in equations (7). Then

$$Var(\mathbf{H}_0'\boldsymbol{\beta}^o) = \mathbf{H}_0'\mathbf{C}_{11}\mathbf{H}_0\sigma_e^2 \tag{22}$$

where $\mathbf{C}_{11}$ is the upper $p \times p$ submatrix of a g-inverse of the coefficient matrix. Then under the null hypothesis versus the unrestricted hypothesis

$$(\mathbf{H}_0'\boldsymbol{\beta}^o)' \; [\mathbf{H}_0'\mathbf{C}_{11} \; \mathbf{H}_0]^{-1} \; \mathbf{H}_0'\boldsymbol{\beta}^o / s\hat{\boldsymbol{\sigma}}_e^2 \tag{23}$$

is distributed as F with degrees of freedom s, $n-$rank $(\mathbf{X} \;\; \mathbf{Z})$. $\hat{\sigma}_e^2$ is an estimate of σ_e^2 computed by

$$(\mathbf{y}'\mathbf{y} - (\boldsymbol{\beta}^o)'\mathbf{X}'\mathbf{y} - (\mathbf{u}^o)'\mathbf{Z}'\mathbf{y})/[n - rank(\mathbf{X} \;\; \mathbf{Z})], \tag{24}$$

and s is the number of rows, linearly independent, in $\mathbf{H}_0'$.