ESTIMATION OF FIXED EFFECTS IN MIXED MODELS WITH UNBALANCED DATA

Walter R. Harvey, Ohio State University

1. INTRODUCTION

The theory involved in the estimation of fixed effects under the general linear model is well known and is covered in considerable detail in many textbooks, e.g., see Graybill (1976) and Searle (1971). Efficient analytical procedures have been developed in recent years for the estimation of fixed effects which allow researchers to obtain essentially the same kind of information whether the data are balanced or unbalanced. Some of these procedures, which may be considered as extensions of Yates (1934) "method of fitting constants," are as follows:

1. The fitting of individual constants for interaction effects so that the "weighted squares of means" analysis described by Yates (1934) can be obtained directly and can also be extended to more complex models.
2. The computation of adjusted or least-squares sums of squares by a direct method (Harvey, 1970) which makes use of the estimates of the fixed effects and segments of the inverse of the variance-covariance matrix, rather than from differences in total reductions in sums of squares when different models are fitted.
3. The estimation of partial regression coefficients, both for individual classes or subclasses or on a pooled basis, for continuous independent variables (Harvey, 1964 and 1977b). Actually, the degrees of freedom for the individual class or subclass regressions may be partitioned in the same manner as the class or subclass effects.
4. Orthogonal polynomial fitting even when unequal intervals exist and/or the treatment means are correlated (Harvey and Swiger, 1978).
5. The computation of sums of squares for single degree of freedom contrasts among adjusted treatment means and the application of mean separation procedures (Kramer, 1957).

When completing an ordinary multiple regression analysis the X matrix of the underlying general linear model,

$$y = X\beta + e, \qquad (1.1)$$

is of full rank and the partial regressions (including the Y intercept, β_0) are estimated from

$$\hat{\beta} = (X'X)^{-1}X'y. \qquad (1.2)$$

Also,

$$Var(\hat{\beta}) = (X'X)^{-1}\sigma_e^2. \qquad (1.3)$$

Of course, when the model (1.1) contains design or discrete independent variables the normal equations have no unique solution and only linear functions of the underlying effects for such variables can be estimated. If one chooses a basis set of estimable functions of β to estimate, say $M\beta$, then

$$\hat{M\beta} = (L'X'XL)^{-1}L'X'y \qquad (1.4)$$

where L is the conditional inverse of M (see Harvey, 1970 and Graybill, 1976, pages 493-495). In this case,

$$Var(\hat{M\beta}) = (L'X'XL)^{-1}\sigma_e^2. \qquad (1.5)$$

Hence, it is easily seen that under the usual assumptions regarding normality, covariances of terms in the model and homogeneity of variances one has no difficulty in obtaining (i) estimates of differences among fixed effects, (ii) estimates of partial regression coefficients for continuous independent variables, (iii) adjusted treatment means or (iv) standard errors and tests of significance for any linear comparison of the estimated linear functions when the underlying model contains only fixed effects (other than random error, of course).

When the appropriate model for the analysis of a set of data contains one or more sets of random effects and disproportionate subclass frequencies exist, several alternative methods of analysis are currently available for obtaining unbiased estimates of linear functions of the fixed effects. Several of these methods utilize differences among random classes or subclasses that are partially confounded with the fixed effects to more accurately estimate the linear functions desired for fixed effects. However, to do this one must have reasonably accurate *a priori* estimates of the variance components for all sets of random effects, including error, or these must be estimated iteratively with the estimation of the linear functions of fixed effects. The estimation procedure to be considered here for mixed models might be referred to as the "least-squares" procedure. The estimates of the "fixed effects" are those obtained from a Method 3 analysis of Henderson (1953) as defined by Harvey (1977a). In these "least-squares" analyses, sums of squares and "constant" estimates for the fixed effects are adjusted for unequal subclass frequencies with respect to only the other effects included in the model sufficient to achieve expected mean squares (EMS) that are the same as for a balanced design except for the coefficients of the variance components.

2. CROSSCLASSIFIED SET OF RANDOM EFFECTS

Consider the following linear model:

$$y_{ijk} = \mu + a_i + F_j + e_{ijk} \qquad (2.1)$$

where the a_i are random and F_j refers to any number of sets of fixed effects including possibly partial regression coefficients for continuous independent variables. In matrix notation this model is

$$y = 1\mu + X_a a + X_F F + e \qquad (2.2)$$

and the reparameterized model may be written as follows:

$$y = X_\mu L_\mu M_\mu \beta + X_a L_a M_a a + X_F L_F M_F F + e \qquad (2.3)$$

where $\beta' = (\mu \; a \; F)$. It can be shown that $X_\mu L_\mu = \mathbf{1}$, a column vector of 1's and to simplify notation let $M_\mu\beta = \mu^*$, $X_aL_a = X_1$, $M_a a = a^*$, $X_FL_F = X_2$ and $M_FF = F^*$. The least squares equations may then be written as follows:

$$\begin{pmatrix} 1'1 & 1'X_1 & 1'X_2 \\ X_1'1 & X_1'X_1 & X_1'X_2 \\ X_2'1 & X_2'X_1 & X_2'X_2 \end{pmatrix} \begin{pmatrix} \hat{\mu}^* \\ \hat{a}^* \\ \hat{F}^* \end{pmatrix} = \begin{pmatrix} 1y \\ X_1'y \\ X_2'y \end{pmatrix} \quad (2.4)$$

Now if the inverse of the coefficient matrix is

$$C^{-1} = \begin{pmatrix} Z_{\mu\mu} & Z_{\mu a} & Z_{\mu F} \\ Z_{\mu a}' & Z_{aa} & Z_{aF} \\ Z_{\mu F}' & Z_{aF}' & Z_{FF} \end{pmatrix} \quad (2.5)$$

then the solution of equations (2.4) yields estimates of the overall mean (μ^*), the basis set of linear functions for the random effects (a^*) and the basis set of linear functions for the fixed sets of effects (F^*) as follows:

$$\begin{pmatrix} \hat{\mu}^* \\ \hat{a}^* \\ \hat{F}^* \end{pmatrix} = \begin{pmatrix} Z_{\mu\mu} & Z_{\mu a} & Z_{\mu F} \\ Z_{\mu a}' & Z_{aa} & Z_{aF} \\ Z_{\mu F}' & Z_{aF}' & Z_{FF} \end{pmatrix} \begin{pmatrix} 1y \\ X_1'y \\ X_2'y \end{pmatrix} \quad (2.6)$$

Calculation of appropriate tests of significance for estimates of differences among fixed effects and standard errors for such contrasts is straightforward since

$$\text{Var}(\hat{F}^*) = Z_{FF}\sigma_e^2. \quad (2.7)$$

Oftentimes when models contain one or more sets of crossclassified random effects along with the fixed sets of effects, the investigator is also interested in obtaining estimates of adjusted "treatment" means. Therefore, the question arises concerning the procedures that should be used to calculate (or estimate) such means as well as the corresponding standard errors.

In the one-way classification random model with unequal frequencies, Weiler and Culpin (1970) point out that the simple overall mean of the $y_{ij}(\bar{y})$ is a more accurate estimate of μ than the unweighted mean of the class means (the least-squares estimate of μ^* under the usual restrictions) when σ_a^2 is small relative to σ_e^2. The variances of these two estimators of the overall mean in this case are

$$\text{Var}(\bar{y}) = \frac{1}{n_.} (\sigma_e^2 + k\sigma_a^2) \quad (2.8)$$

where

$$k = \frac{\sum_i n_i^2}{n_.} \quad (2.9)$$

and

$$\text{Var}(\hat{\mu}^*) = \frac{\sigma_e^2}{p} \sum_i \frac{1}{n_i} + \frac{\sigma_a^2}{p} \quad (2.10)$$

where $i = 1, 2, \ldots, p$ and $\sum_i n_i = n.$.

The estimate of the overall mean obtained from the mixed model least squares analysis described above ($\hat{\mu}^*$) will be adjusted for unequal class or subclass frequencies for the random effects as well as for the fixed classes and/or subclasses. Also, the random set (or sets) of effects will usually consist of many classes. Hence, it is difficult, if not impossible, to directly complete the LS analysis as described.

To illustrate the computing procedures used in LSML76, let the partitioning of the reparameterized model (2.3) be as follows:

$$y = WB_1 + X_rB_2^* + e \quad (2.11)$$

where B_1 contains the $\mu + a_i$ effects, B_2^* contains the linear functions of the fixed effects to be estimated (F^*) and X_r is the reduced X matrix (X_FL_F) for the fixed effects (X_2). When only one set of crossclassified random effects are included in B_1, as considered here, $W'W$ is a diagonal matrix of the a_i class numbers. If the a_i effects are deleted from the model, it may be written as follows:

$$y = 1\mu^* + X_rB_2^* + e \quad (2.12)$$

The least squares equations under this reduced model are

$$\begin{pmatrix} n_{..} & 1'X_r \\ X_r'1 & X_r'X_r \end{pmatrix} \begin{pmatrix} \hat{\hat{\mu}}^* \\ \hat{\hat{B}}_2^* \end{pmatrix} = \begin{pmatrix} Y_{..} \\ X_r'y \end{pmatrix} \quad (2.13)$$

The remainder or residual sum of squares under reduced model (2.12) is

$$\begin{aligned} E_1 &= y'y - \hat{\hat{\mu}}^*Y_{..} - \left(\hat{\hat{B}}_2^*\right)'X_r'y \\ &= y'y - R(\mu, F). \end{aligned} \quad (2.14)$$

Now if the equations for B_1 in the complete model (2.11) are absorbed into the equations for B_2^* the following equations are obtained:

$$(S - N'D^{-1}N)\hat{B}_2^* = (Y_2 - N'D^{-1}Y_1) \quad (2.15)$$

where $S = X_r'X_r$, $N = W'X_r$, $D^{-1} = (W'W)^{-1}$, $Y_2 = X_r'y$ and $Y_1 = W'y$. In the process of computing these reduced equations one should also compute

$$y'y - Y_1'D^{-1}Y_1, \quad (2.16)$$

the "within" random classes sum of squares for the dependent variable. The remainder sum of squares from the analysis under the complete model (2.11) is

$$\begin{aligned} E_2 &= y'y - Y_1'D^{-1}Y_1 - (\hat{B}_2^*)'(Y_2 - N'D^{-1}Y_1) \\ &= y'y - R(\mu, a, F) \end{aligned} \quad (2.17)$$

and the adjusted sum of squares for the random set of effects is

$$\begin{aligned} \text{S.Sqs.}(A) &= E_1 - E_2 \\ &= R(\mu, a, F) - R(\mu, F). \end{aligned} \quad (2.18)$$

For illustration, let us suppose B_2 contains the following fixed effects:

$$b_j + c_k + d_\ell + (bd)_{j\ell} \quad (2.19)$$

Unbiased estimates of least squares or adjusted means for the fixed classes and subclasses and approximate variances are computed as follows with LSML76 (Harvey 1977b):

$$\text{Est. of LSM}(\mu) = \frac{1}{n_{....}}(Y_{....} - 1'X_r\hat{B}_2^*) = \tilde{\mu}^*$$

$$\text{Var}(\tilde{\mu}^*) = R^{\mu\mu}(\hat{\sigma}_e^2 + k\hat{\sigma}_a^2)$$

Where: $R^{\mu\mu}$ is the inverse diagonal element for $\tilde{\mu}^*$ from the analysis under the reduced model, $k = \frac{\sum_i n_{i...}^2}{n_{....}}$ and the estimates of the variance components are computed as explained in Harvey (1977b).

$$\text{Est. of LSM}(b_j) = \tilde{\mu}^* + \hat{b}_j^*$$

$$\text{Var}(\tilde{\mu}^* + \hat{b}_j^*) = V(\tilde{\mu}^*) + V(\hat{b}_j^*) + 2\text{ Cov }\tilde{\mu}^*\hat{b}_j^*$$

$$= (R^{\mu\mu} + C^{b_jb_j} + 2R^{\mu b_j})\hat{\sigma}_e^2 + R^{\mu\mu}k\hat{\sigma}_a^2$$

Where: R^{-1} is the inverse of the coefficient matrix when the random effects are ignored and C^{-1} is the inverse of the reduced coefficient matrix after the μ + a_i equations have been absorbed.

$$\text{Est. of LSM}(c_k) = \tilde{\mu}^* + \hat{c}_k^*$$

$$\text{Var}(\tilde{\mu}^* + \hat{c}_k^*) = (R^{\mu\mu} + C^{c_kc_k} + 2R^{\mu c_k})\hat{\sigma}_e^2 + R^{\mu\mu}k\hat{\sigma}_a^2$$

$$\text{Est. of LSM}(d_\ell) = \tilde{\mu}^* + \hat{d}_\ell^*$$

$$\text{Var}(\tilde{\mu}^* + \hat{d}_\ell^*) = (R^{\mu\mu} + C^{d_\ell d_\ell} + 2R^{\mu d_\ell})\hat{\sigma}_e^2 + R^{\mu\mu}k\hat{\sigma}_a^2$$

$$\text{Est. of LSM}[(bd)_{j\ell}] = \tilde{\mu}^* + \hat{b}_j^* + \hat{d}_\ell^* + (\hat{bd})_{j\ell}^*$$

$$\text{Var}[\tilde{\mu}^* + \hat{b}_j^* + \hat{d}_\ell^* + (\hat{bd})_{j\ell}^*] = (R^{\mu\mu} + C^{b_jb_j} + C^{d_\ell d_\ell} + C^{bd_{j\ell}bd_{j\ell}} + 2R^{\mu b_j} + 2R^{\mu d_\ell} + 2R^{\mu bd_{j\ell}} + 2C^{b_jd_\ell} + 2C^{b_jbd_{j\ell}} + 2C^{d_\ell bd_{j\ell}})\hat{\sigma}_e^2 + R^{\mu\mu}k\hat{\sigma}_a^2$$

It is easily seen that if one has a balanced set of data the variances of the least squares means given above will be computed correctly and they will not be approximations. The extent of the approximation with unbalanced data depends primarily on the difference between the approximate covariances of $\tilde{\mu}^*$ and the estimates of the linear functions of the fixed effects ($\hat{B}_2^*$) used to compute the estimates of the least squares means and the "true" covariances. In the calculations used by LSML76 to compute the variances of the least squares means the covariances between $\tilde{\mu}^*$ and $\hat{B}_2^*$ are assumed to be the same as the covariances between $\hat{\mu}^*$ and $\tilde{B}_2^*$. Also, $V(\tilde{\mu}^*)$ is assumed to be the same as $V(\hat{\mu}^*)$.

3. NESTED SET OF RANDOM EFFECTS

Consider the following linear model:

$$y_{ijk\ell} = \mu + a_i + b_{ij} + F_k + e_{ijk\ell} \qquad (3.1)$$

where a_i is a fixed set of effects, the b_{ij} are the random nested effects and F_k represents all other fixed sets of effects included in the model.

In order to describe the computing procedures used in LSML76, let the partitioning of the reparameterized model be written as follows:

$$y = WB_1 + X_rB_2^* + e \qquad (3.2)$$

where B_1 contains the $\mu + a_i + b_{ij}$ effects, B_2^* contains the linear functions of all fixed effects to be estimated other than μ and the a_i and X_r is the reduced X matrix for the fixed effects other than μ and the a_i. In this case, $W'W$ is a diagonal matrix of the AB subclass numbers. If the b_{ij} effects are deleted from model (3.1), the reduced reparameterized model is as follows:

$$y = I\mu^* + X_1a^* + X_rB_2^* + e \qquad (3.3)$$

where X_1 is the reduced X matrix for the a_i fixed set of effects and a^* is the basis set of linear functions of the a_i to be estimated. The least squares equations under the reduced model are

$$\begin{pmatrix} n_{...} & 1'X_1 & 1'X_r \\ X_1'1 & X_1'X_1 & X_1'X_r \\ X_r'1 & X_r'X_1 & X_r'X_r \end{pmatrix} \begin{pmatrix} \tilde{\mu}^* \\ \tilde{a}^* \\ \tilde{B}_2^* \end{pmatrix} = \begin{pmatrix} Y_{...} \\ X_1'y \\ X_r'y \end{pmatrix}. \qquad (3.4)$$

The remainder or residual sum of squares under the reduced model is

$$E_1 = y'y - \tilde{\mu}^*Y_{...} - (\tilde{a}^*)'X_1'y - (\tilde{B}_2^*)'X_r'y$$

$$= y'y - R(\mu, a, F). \qquad (3.5)$$

Now if the equations for B_1 in the complete model (3.2) are absorbed into the equations for B_2^* the following equations are obtained:

$$(S - N'D^{-1}N)\,\hat{B}_2^* = (Y_2 - N'D^{-1}Y_1) \qquad (3.6)$$

where $S = X_r'X_r$, $N = W'X_r$, $D^{-1} = (W'W)^{-1}$, $Y_2 = X_r'y$ and $Y_1 = W'y$. Again,

$$y'y - Y_1'D^{-1}Y_1, \qquad (3.7)$$

the "within" AB subclasses sum of squares, is computed for the dependent variable during the absorption process. The remainder sum of squares from the analysis under the complete model is

$$E_2 = y'y - Y_1'D^{-1}Y_1 - (\hat{B}_2^*)'(Y_2 - N'D^{-1}Y_1)$$

$$= y'y - R(\mu, a, b, F) \qquad (3.8)$$

and the adjusted sum of squares for the random set of nested effects is

$$\text{S.Sqs.}(B{:}A) = E_1 - E_2$$

$$= R(\mu, a, b, F) - R(\mu, a, F) \qquad (3.9)$$

To describe the procedure used in LSML76 to compute least squares means and standard errors under this model type let B_2 contain the following fixed effects:

$$c_k + (ac)_{ik} + d_\ell + (cd)_{k\ell} \qquad (3.10)$$

Estimates of the adjusted or least squares means for the fixed classes and subclasses and the corresponding variances are computed as follows:

Est. of $LSM(\mu) = \frac{1}{n_{....}} (Y_{....} - 1'X_1\tilde{a}^* - 1'X_r\hat{B}_2^*)$

$= \tilde{\mu}^*$

$Var(\tilde{\mu}^*) = R^{\mu\mu}(\hat{\sigma}_e^2 + k\hat{\sigma}_{b:A}^2)$

Where: $R^{\mu\mu}$ is the inverse diagonal element for $\tilde{\mu}^*$ from the analysis under the reduced model, $k = \frac{\sum_i\sum_j n_{ij..}^2}{n_{....}}$ and the variance components are computed from expected mean squares as explained by Harvey (1977b).

Est. of $LSM(a_i) = \tilde{\mu}^* + \tilde{a}_i^*$

$Var(\tilde{\mu}^* + \tilde{a}_i^*) = (R^{\mu\mu} + R^{a_ia_i} + 2R^{\mu a_i})$

$\times (\hat{\sigma}_e^2 + k_i\hat{\sigma}_{b:A}^2)$

Where: R^{-1} is the inverse of the coefficient matrix for the analysis under the reduced model and $k_i = \frac{\sum_j n_{ij..}}{n_{i...}}$

Est. of $LSM(c_k) = \tilde{\mu}^* + \hat{c}_k^*$

$Var(\tilde{\mu}^* + \hat{c}_k^*) = (R^{\mu\mu} + C^{c_kc_k} + 2R^{\mu c_k})\hat{\sigma}_e^2$

$+ R^{\mu\mu}k\hat{\sigma}_{b:A}^2$

Where: C^{-1} is the inverse of the reduced coefficient matrix after the $\mu + a_i + b_{ij}$ equations have been absorbed.

Est. of $LSM[(ac)_{ik}] = \tilde{\mu}^* + \tilde{a}_i + \hat{c}_k^* + (\hat{ac})_{ik}^*$

$Var[\tilde{\mu}^* + \tilde{a}_i^* + \hat{c}_k^* + (\hat{ac})_{ik}^*] = \lambda_{ik}(\hat{\sigma}_e^2 + k_{ik}\hat{\sigma}_{b:A}^2)$

Where: λ_{ik} is the appropriate linear function of the inverse elements from R^{-1} and C^{-1} and

$$k_{ik} = \frac{\sum_j n_{ijk.}^2}{n_{i.k.}} - \frac{\sum_j n_{ij..}^2}{n_{i...}} \cdot \frac{1}{c}$$

and c is the number of C classes.

Est. of $LSM(d_\ell) = \tilde{\mu}^* + \hat{d}_\ell^*$

$Var(\tilde{\mu}^* + \hat{d}_\ell^*) = (R^{\mu\mu} + C^{d_\ell d_\ell} + 2R^{\mu d_\ell})\hat{\sigma}_e^2$

$+ R^{\mu\mu}k\hat{\sigma}_{b:A}^2$

Est. of $LSM[(cd)_{k\ell}] = \tilde{\mu}^* + \hat{c}_k^* + \hat{d}_\ell^* + (\hat{cd})_{k\ell}$

$Var[\tilde{\mu}^* + \hat{c}_k^* + \hat{d}_\ell^* + (\hat{cd})_{k\ell}^*] = \lambda_{k\ell}\hat{\sigma}_e^2 + R^{\mu\mu}k\hat{\sigma}_{b:A}^2$

Where: $\lambda_{k\ell}$ is the appropriate linear function of the inverse elements from R^{-1} and C^{-1}.

The assumptions are made in these calculations that the covariances between $\tilde{\mu}^*$ and $\hat{B}_2^*$ are the same as the covariances between $\tilde{\mu}^*$ and $\tilde{B}_2^*$ and that the covariances between $\tilde{a}^*$ and $\hat{B}_2^*$ are the same as the covariances between $\tilde{a}^*$ and $\tilde{B}_2^*$. Also, $V(\tilde{\mu}^*)$ is assumed to be the same as $V(\hat{\mu}^*)$ and an approximate value is used for k_{ik}.

Under this model type the denominator mean square for testing the significance of differences among the $\hat{a}_i^*$ is B:A when equal subclass frequencies exist. With unbalanced data, the ratio MS(A)/MS(B:A) is at best only an approximate "F" test. However, since

$$V(\tilde{a}_i^*) = R^{a_ia_i}(\hat{\sigma}_e^2 + k_i\hat{\sigma}_{b:A}^2)$$

and

$$Cov(\tilde{a}_i^*, \tilde{a}_{i'}^*) = R^{a_ia_{i'}}[\hat{\sigma}_e^2 + ½(k_i + k_{i'})\hat{\sigma}_{b:A}^2]$$

one may easily calculate standard errors and "t" ratios for linear contrasts among the $\tilde{a}_i^*$. The number of degrees of freedom to assign to these standard errors or "t" ratios probably lies somewhere between the degrees of freedom for estimating $\hat{\sigma}_e^2$ and the degrees of freedom for B:A. An approximate formula for obtaining these degrees of freedom is needed.

4. CROSSCLASSIFIED SET OF RANDOM EFFECTS WHICH INTERACT WITH A SET OF FIXED EFFECTS

Let us write this mixed model as follows:

$$y_{ijk\ell} = \mu + a_i + b_j + (ab)_{ij} + F_k + e_{ijk\ell} \quad (4.1)$$

where a_i is a set of random effects, b_j is a set of fixed effects and F_k represents all other sets of fixed effects that may be included in the model. Interactions of the a_i with all fixed effects included in F_k are assumed to be nonexistent. The standard randomized block split-plot design falls into this class of mixed models.

In order to describe the computing procedures used in LSML76, let the reparameterized model be written as before, i.e.,

$$y = WB_1 + X_rB_2^* + e \quad (4.2)$$

where, in this case, B_1 contains the $\mu + a_i + b_j + (ab)_{ij}$ effects, B_2^* contains the linear functions of all fixed effects to be estimated other than μ and the b_j, and X_r is the reduced X matrix (i.e., XL) for the other fixed effects. In this case W'W is a diagonal matrix of the AB subclass numbers. If the a_i and $(ab)_{ij}$ sets of effects are deleted from model (4.1) the reduced reparameterized model may be written as follows:

$$y = 1\mu^* + X_1b^* + X_rB_2^* + e \quad (4.3)$$

where X_1 is the reduced X matrix for the b_j fixed set of effects and b* is the basis set of linear functions of the b_j to be estimated. The least squares equations under model (4.3) are the same as under model (3.3), i.e., equations (3.4). The remainder or residual sum of squares is

$$E_1 = y'y - \tilde{\mu}^*Y_{...} - (\hat{b}^*)'X_1'y - (\hat{B}_2^*)'X_r'y$$
$$= y'y - R(\mu, b, F). \quad (4.4)$$

Now if the $(ab)_{ij}$ set of effects in model (4.1) are deleted the reduced model may be partitioned as in model (4.2), i.e.,

$$y = WB_1 + X_rB_2^* + e \quad (4.5)$$

where, in this case, B_1 contains the $\mu + a_i$ effects and B_2^* now also contains the linear functions being estimated for the b_j effects. Note that W'W is now a diagonal matrix of the numbers in the A random classes. Absorbing the B_1 equations (the $\mu + a_i$) into the equations for B_2^* gives

$$(S - N'D^{-1}N)\tilde{B}_2^* = (Y_2 - N'D^{-1}Y_1) \quad (4.6)$$

for the reduced set of equations for the fixed effects. The remainder sum of squares from this analysis is

$$E_2 = y'y - Y_1'D^{-1}Y_1 - (\tilde{B}_2^*)'(Y_2 - N'D^{-1}Y_1)$$
$$= y'y - R(\mu, a, b, F). \quad (4.7)$$

The "within A classes" sum of squares ($y'y - Y_1D^{-1}Y_1$) is computed during the absorption process.

Now if the equations for B_1 in the complete model (4.2), i.e., the $\mu + a_i + b_j + (ab)_{ij}$ equations are absorbed into the equations for B_2^*, the following equations are obtained:

$$(S - N'D^{-1}N)\hat{B}_2^* = (Y_2 - N'D^{-1}Y_1) \quad (4.8)$$

The remainder sum of squares from this analysis is

$$E_3 = y'y - Y_1'D^{-1}Y_1 - (\hat{B}_2^*)'(Y_2 - N'D^{-1}Y_1)$$
$$= y'y - R(\mu, a, b, ab, F). \quad (4.9)$$

The "within AB subclasses" sum of squares ($y'y - Y_1D^{-1}Y_1$) is computed during the absorption process.

The adjusted sum of squares for the random sets of effects are computed by the "indirect" procedure as follows:

$$S.Sqs.(A) = E_1 - E_2$$
$$= R(\mu, a, b, F) - R(\mu, b, F) \quad (4.10)$$

$$S.\ Sqs.(AB) = E_2 - E_3$$
$$= R(\mu, a, b, ab, F) - R(\mu, a, b, F) \quad (4.11)$$

To describe the procedures used in LSML76 to compute least squares means and standard errors under this model type let B_2 contain the following fixed effects:

$$c_k + (bc)_{jk} \quad (4.12)$$

Estimates of the least squares means for the fixed classes and subclasses and the corresponding variances are computed as follows:

$$\text{Est. of LSM}(\mu) = \frac{1}{n_{...}}(Y_{...} - 1'X_1\tilde{b}^* - 1'X_r\hat{B}_2)$$
$$= \tilde{\mu}^*$$

$$\text{Var}(\tilde{\mu}^*) = R^{\mu\mu}(\hat{\sigma}_e^2 + k_1\hat{\sigma}_a^2 + k_2\hat{\sigma}_{aB}^2)$$

Where: $R^{\mu\mu}$ is the inverse diagonal element for $\tilde{\mu}^*$ from the analysis under the reduced model (4.3),

$$k_1 = \frac{\sum_i n_{i..}^2}{n_{...}}$$

$$k_2 = \frac{\sum\sum_{ij} n_{ij.}^2}{n_{...}}$$

and the estimates of the variance components are computed as explained by Harvey (1977b).

$$\text{Est. of LSM}(b_j) = \tilde{\mu}^* + \tilde{b}_j^*$$

$$\text{Var}(\tilde{\mu}^* + \tilde{b}_j^*) = (R^{\mu\mu} + C^{b_jb_j} + 2R^{\mu b_j})$$
$$\times (\hat{\sigma}_e^2 + k_j\hat{\sigma}_{aB}^2) + R^{\mu\mu}k_1\hat{\sigma}_a^2$$

Where: C^{-1} is the inverse of the coefficient matrix after absorbing the $\mu + a_i$ equations [$S - N'D^{-1}N$ of equations (4.6)] and

$$k_j = \frac{\sum_i n_{ij.}^2}{n_{.j.}}.$$

$$\text{Est. of LSM}(c_k) = \tilde{\mu}^* + \hat{c}_k^*$$

$$\text{Var}(\tilde{\mu}^* + \hat{c}_k^*) = (R^{\mu\mu} + H^{c_kc_k} + 2R^{\mu c_k})\hat{\sigma}_e^2$$
$$+ R^{\mu\mu}(k_1\hat{\sigma}_a^2 + k_2\hat{\sigma}_{aB}^2)$$

Where: H^{-1} is the inverse of the coefficient matrix after absorbing the $\mu + a_i + b_j + (ab)_{ij}$ equations [$S - N'D^{-1}N$ of equations (4.8)].

$$\text{Est. of LSM}[(bc)_{jk}] = \tilde{\mu}^* + \tilde{b}_j^* + \hat{c}_k^* + (\hat{bc})_{jk}^*$$

$$\text{Var}[\tilde{\mu}^* + \tilde{b}_j^* + \hat{c}_k^* + (\hat{bc})_{jk}^*]$$
$$= \lambda_{jk}(\hat{\sigma}_e^2 + k_{jk}\hat{\sigma}_a^2 + k_{jk}\hat{\sigma}_{aB}^2)$$

Where: λ_{jk} is the appropriate linear function of the inverse elements and k_{jk} is estimated as follows:

$$k_{jk} = \frac{\sum_i n_{ijk}^2}{n_{.jk}} - \frac{\sum_i n_{ij.}^2}{n_{.j.}} \cdot \frac{1}{c}$$

and c is the number of C classes.

As can easily be seen, the assumptions are made that (i) the covariances between $\tilde{\tilde{\mu}}^*$ and $\tilde{\tilde{b}}_j^*$ are the same as between $\tilde{\mu}^*$ and $\tilde{b}_j^*$, (ii) the covariances between $\tilde{\tilde{\mu}}^*$ and $\hat{B}_2^*$ are the same as between $\tilde{\mu}^*$ and $\tilde{B}_2^*$, (iii) the covariances between $\tilde{\tilde{b}}_j^*$ and $\hat{B}_2^*$ are the same as between $\tilde{\tilde{b}}_j^*$ and $\tilde{\tilde{B}}_2^*$ and (iv) the $V(\tilde{\tilde{\mu}}^*)$ is the same as $V(\tilde{\mu}^*)$.

As with the model type considered in section 4, the denominator mean square for testing the significance of differences among the $\tilde{\tilde{b}}_j^*$ (AB) gives only an approximate test when unequal subclass frequencies exist. Standard errors of linear contrasts among the $\tilde{\tilde{b}}_j^*$ are easily computed, however, since

$$V(\tilde{\tilde{b}}_j^*) = c^{b_jb_j}(\hat{\sigma}_e^2 + k_j\hat{\sigma}_{aB}^2)$$

and

$$Cov(\tilde{\tilde{b}}_j^*, \tilde{\tilde{b}}_{j'}^*) = c^{b_jb_{j'}}[\hat{\sigma}_e^2 + \tfrac{1}{2}(k_j + k_{j'})\hat{\sigma}_{aB}^2].$$

In order to use these standard errors to obtain a "t" test one again needs to know how many degrees of freedom to assign to the standard error.

5. OTHER TYPES OF MIXED MODELS

Of course, the number of types of mixed models that can exist or that might be most appropriate for the analysis of a set of data is infinite. By utilizing the procedures described in sections 2, 3 and 4 above the mixed model least squares computer program (LSML76) completes the analysis for the following additional three types of mixed models:

1. One set of crossclassified random effects and one set of nested random effects with the nesting being within the crossclassified random effects.

   General Model

   $y_{ijk\ell} = \mu + a_i + b_{ij} + F_k + e_{ijk\ell}$

   where the a_i and b_{ij} are random and F_k represents all fixed sets of effects.

2. Two sets of nested random effects. The second set of nested effects is nested within the first set and the first set is nested within a set of fixed effects.

   General Model

   $y_{ijk\ell m} = \mu + a_i + b_{ij} + c_{ijk} + F_\ell + e_{ijk\ell m}$

   where the a_i are fixed, the b_{ij} and c_{ijk} are random and F_ℓ refers to all other sets of fixed effects.

3. One set of nested random effects which interact with one set of fixed crossclassified effects.

   General Model

   $y_{ijk\ell m} = \mu + a_i + b_{ij} + c_k + (ac)_{ik} + (bc)_{ijk} + F_\ell + e_{ijk\ell m}$

   where a_i, c_k and $(ac)_{ik}$ are fixed; b_{ij} and $(bc)_{ijk}$ are random and F_ℓ again refers to all other fixed effects.

Other options included in LSML76 with respect to the estimation of fixed effects, whether they are included in a fixed linear model or in a mixed model, are as follows:

1. Sums of squares, tests of significance and prediction equations may be computed for orthogonal polynomials for factorial or nested sets of fixed effects.

2. Sums of squares, tests of significance and individual class or subclass regression coefficients with standard errors may be computed for continuous independent variables.

3. Non-orthogonal polynomials may be fitted automatically through the third degree for continuous independent variables.

6. SUMMARY

Several methods are available for obtaining unbiased estimates of fixed effects in mixed models. The procedure discussed here is the "least squares" method where all sums of squares and corresponding sets of linear functions of the fixed effects being estimated are adjusted for unequal subclass frequencies with respect to other effects in the model as necessary to achieve the same expected mean squares as would be obtained in a balanced design analysis except, of course, for the coefficients of the variance components. Procedures used in the mixed model least squares program, LSML76, to compute estimates of the linear functions of the fixed effects, the least squares means and the approximate standard errors are given in detail for three types of mixed models. The extent which the assumptions being made in the computation of standard errors and test of significance for "whole plot" effects and in the computation of standard errors of least squares means cause such statistics to be in error with unbalanced data is unknown. Clearly, more work is needed in this area.

REFERENCES

Graybill, Franklin A. Theory and Application of the Linear Model. Duxbury Press, North Scituate, Mass., 1976.

Harvey, Walter R. Computing procedures for a generalized least-squares analysis program. 50 pp. Mimeograph. Presented at the Analysis of Variance Conference, Ft. Collins, Colo., 1964.

Harvey, Walter R. Estimation of variance and covariance components in the mixed model. Biometrics 26:485-504. 1970.

Harvey, Walter R. Missing subclass problems using Method 3 of Henderson in mixed models. Proc. Stat. Comp. Sec. Am. Stat. Assoc. pp. 22-26, 1977.

Harvey, Walter R. User's Guide for LSML76. Mimeograph. Ohio State Univ., 76 pp., 1977.

Harvey, Walter R. and L. A. Swiger. Orthogonal polynomial fitting with arbitrary spacing and correlated means. Proc. 3rd Ann. Conf. SUGI Meetings. pp. 108-112. 1978.

Henderson, C. R. Estimation of variance and covariance components. Biometrics 9:226-252. 1953.

Kramer, Clyde Young. Extension of multiple range tests to group correlated means. Biometrics 13:13-18. 1957.

Searle, S. R. Linear Models. John Wiley and Sons, New York. 1971.

Weiler, H. and D. Culpin. On the variance of weighted means. Technometrics 12:757-774. 1970.

Yates, F. The analysis of multiple classifications with unequal numbers in the different classes. J. Amer. Stat. Assoc. 29:51-66. 1934.