CELL MEANS FORMULATION OF MIXED MODELS IN THE ANALYSIS OF VARIANCE

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York 14853

BU-862-M October 1984

Abstract

For a broad definition of balanced data from mixed models it is shown that the BLUE (best linear unbiased estimator) of an estimable function of the fixed effects is the same as the ordinary least squares estimator; in particular, estimates of cell means in a cell means formulation (for the fixed effects) of a mixed model therefore provide the BLUEs. Application to unbalanced data is shown for randomized complete blocks with not necessarily the same number of observations in each treatment-by-block combination; and for a special case of this, balanced incomplete blocks.

1. INTRODUCTION

a. Fixed effects models

Analysis of variance models have traditionally been formulated in terms of additive main effects and additive interaction effects that usually result in there being more parameters in the model than there are means to estimate them from. For example, suppose y_{ijk} is the k'th observation on treatment i of variety j in a two-factor experiment concerned with fertilizer treatments and plant varieties. Then a traditional analysis of variance model is of the form

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \tag{1}$$

Paper BU-862-M in the Biometrics Unit series.

where μ is a general mean, α_i is the effect on the response variable due to the i'th treatment, β_j is the effect due to the j'th variety, γ_{ij} is the interaction effect between treatment i and variety j, and e_{ijk} is the residual error term defined as $e_{ijk} = y_{ijk} - E(y_{ijk})$ for

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

where E denotes expectation over repeated sampling. For an experiment of a treatments and b varieties and n observations per cell, the number of observed cell means $\bar{y}_{ij.} = \sum_{k=1}^{n} y_{ijk}/n$ (for n observations per cell) is ab. But the model equation (1) has more parameters than this, namely 1 + a + b + ab. Thus (1) exemplifies what is known as an over-parameterized model.

In contrast to (1) there has in recent years been a growing interest in modeling y_{ijk} solely in terms of its underlying population mean, i.e., in taking

$$E(y_{ijk}) = \mu_{ij} \quad \text{and} \quad y_{ijk} = \mu_{ij} + e_{ijk} \tag{2}$$

where the y_{ijk} for k = 1, $\cdots$, n are deemed to be a random sample of n observations from a population having mean μ_{ij}. This formulation is known as the cell means model. It has been promoted extensively by Speed and Hocking and co-workers [e.g., Speed (1969), Hocking and Speed (1975), Speed and Hocking (1976), and Speed, Hocking and Hackney (1978)] and its feature of having exactly the same number of parameters to estimate as there are observed cell means has proven to be particularly useful, especially for unbalanced data, namely those having unequal numbers of observations in the subclasses. Compared to (10), we find that with (2) estimation is easier, estimable functions are simpler, and a variety of hypotheses commonly considered are more easily described and understood.

Urquhart and Weeks (1978) exemplify these advantages in an analysis of weight gains in beef cattle.

The use of (2) as an alternative to (1) tacitly implies incorporation of interactions as part of the model. When wanting to use a no-interaction form of the cell means model it is necessary to use (2) together with restrictions of the form

$$\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0 \quad , \tag{3}$$

which specify absence of interaction.

Analysis of variance models like (1), where estimation of (and testing of hypotheses about) parameters are the features of interest, are known as fixed effects models, and in such models the customary assumptions about variances and covariances are that each observation has the same variance and that every pair of observations has zero covariance. The dispersion matrix $\underset{\sim}{V}$ of the vector of observations $\underset{\sim}{y}$ then, has the form

$$\underset{\sim}{V} = \sigma^2 \underset{\sim}{I} \quad , \tag{4}$$

$\underset{\sim}{I}$ being an identity matrix and σ^2 being the variance of every observation. An assumption about $\underset{\sim}{V}$ more general than (4) is that it is simply a symmetric, positive semi-definite matrix; and in many cases that it be not just positive semi-definite but positive definite, and hence non-singular.

b. Mixed models

Variations of (1) are models where some or all of the α_i, β_j and δ_{ij} terms are assumed not to be parameters to be estimated, but are modeled as being random variables with zero means and some assumed variance-covariance structure. For example, suppose in the no-interaction form of (1), with n = 1, namely

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \quad , \tag{5}$$

that the β_j for $j = 1, \cdots, b$, are modeled as random variables with zero mean $E(\beta_j) = 0 \ \forall \ j$. The β_j are then called random effects and, along with the random error terms e_{ij}, usually have the following variance-covariance structure attributed to them:

$$\text{var}(\beta_i) = \sigma^2_\beta \ \forall \ j \ , \quad \text{cov}(\beta_j, \beta_{j'}) = 0 \ \forall \ j \neq j' \tag{6}$$

$$\text{var}(e_{ij}) = \sigma^2_e \ \forall \ i,j, \quad \text{cov}(e_{ij}, e_{i'j'}) = 0 \text{ except for } i=i' \text{ and } j=j'$$

and

$$\text{cov}(\beta_j, e_{ij'}) = 0 \ \forall \ i,j,j' \ .$$

Then with μ and the α_i in (5) being fixed effects and the β_j being random effects, (5) is known as a mixed model. And the variances σ^2_β and σ^2_e of (6) are the variance components. The structure of (6) then leads to $\underset{\sim}{V}$ having elements that are either zero, $\sigma^2_\beta + \sigma^2_e$, or σ^2_β; in general to elements that are either zero, or one of the variance components or a sum of them.

Example 1 In the case of 2 treatments and 3 blocks, where an element of a matrix that is zero is shown as a dot,

$$\underset{\sim}{V} = \text{var}\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta & \cdot & \cdot \\ \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta & \cdot \\ \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot & \sigma^2_\beta \\ \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot & \cdot \\ \cdot & \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e & \cdot \\ \cdot & \cdot & \sigma^2_\beta & \cdot & \cdot & \sigma^2_\beta+\sigma^2_e \end{bmatrix} .$$

Despite merits of the cell means formulation of fixed effects models, such as (2) as an alternative to (1), minimal formulation has been made to mixed models such as (5) and (6). Indeed, Steinhorst (1982), for the randomized complete blocks design, writes that he is "··· at a loss to see how μ_{ij} carries the right meaning if blocks are random ···." And regarding the split-plot design he continues "The cell-means model is not of much help in such cases. The classic split-plot model ··· cannot be replaced by a variation of $y_{ijk} = \mu_{ijk} + e_{ijk}$." In contrast to such remarks, we show in this paper that all of the cases (and more) that Steinhorst refers to can be formulated as cell means models. We also show that for a broad class of balanced data situations the BLUE (best linear unbiased estimator) of a cell mean in a mixed model is always the OLS (ordinary least squares) estimator. And for the randomized complete blocks model with random blocks (as is usual), we show extension to unbalanced data: an explicit (matrix-vector) expression is developed for estimating the treatment means.

c. A general mixed model

The elements of the mixed model (5) are of two kinds: μ and α_i that are fixed effects, and β_i and e_{ij} that are random variables. Generalizing this dichotomy for a vector of observations $\underset{\sim}{y}$ we write

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}\underset{\sim}{u} \tag{7}$$

where $\underset{\sim}{\beta}$ is a vector of fixed effects and $\underset{\sim}{u}$ is a vector of random effects, including error terms. The matrices and vectors of (7) are partitioned thus:

$$\begin{matrix} \underset{\sim}{X} = [\underset{\sim}{X}_1 \;\; \underset{\sim}{X}_2 \;\; \cdots \;\; \underset{\sim}{X}_d \;\; \cdots \;\; \underset{\sim}{X}_f] \\ \underset{\sim}{\beta} = [\underset{\sim}{\beta}_1' \;\; \underset{\sim}{\beta}_2' \;\; \cdots \;\; \underset{\sim}{\beta}_d' \;\; \cdots \;\; \underset{\sim}{\beta}_f']' \end{matrix} \quad \text{and} \quad \begin{matrix} \underset{\sim}{Z} = [\underset{\sim}{Z}_1 \;\; \underset{\sim}{Z}_2 \;\; \cdots \;\; \underset{\sim}{Z}_q \;\; \cdots \;\; \underset{\sim}{Z}_r] \\ \underset{\sim}{u} = [\underset{\sim}{u}_1' \;\; \underset{\sim}{u}_2' \;\; \cdots \;\; \underset{\sim}{u}_q' \;\; \cdots \;\; \underset{\sim}{u}_r']' \end{matrix} \tag{8}$$

Each $\underset{\sim}{\beta}_d$ for $d = 1, 2, \cdots, f$ has as its element the h_d effects corresponding to the h_d levels of the d'th fixed effect (main effect or interaction) factor, and $\underset{\sim}{X}_d$ is the incidence matrix corresponding to $\underset{\sim}{\beta}_d$. Similarly, $\underset{\sim}{u}_q$ (of p_q elements) and $\underset{\sim}{Z}_q$ for $q = 1, 2, \cdots, r-1$ are defined for the random effect (main effect or interaction) factors analogously to β_d and X_d for fixed effect factors. For q = r, we define $u_r = e$, the vector of error terms, and accordingly $Z_r = I_N$ where N is the total number of observations, and $p_r = N$.

Example 2 Using (5) and (6) as the model for a randomized complete blocks experiment for a treatments in b blocks, μ and $[\alpha_1 \cdots \alpha_a]'$ would be β_1 and β_2 of (8), respectively, and $\beta_1 \cdots \beta_2$ and the e_{ij}-terms of (5) would be u_1 and u_2 of (8), respectively.

The variance and covariance properties of (6) generalized to u are

$$\text{var}(u_q) = \sigma^2_q I_{p_q} \quad \text{for} \quad q = 1, 2, \cdots, r$$

and

$$\text{cov}(u_q, u_{q'}') = 0_{p_q \times p_{q'}} \quad \text{for} \quad q \neq q' = 1, 2, \cdots, r \quad . \tag{9}$$

Hence from (7) the variance-covariance matrix of y is

$$V = \text{var}(y) = \text{var}(Zu) = \sum_{q=1}^{r} \sigma^2_q Z_q Z_q' \quad . \tag{10}$$

Thus (7) through (10) constitute a description of a general mixed model.

d. Estimation in the general mixed model

The ordinary least squares (OLS) estimator of an estimable function $\lambda' X\beta$ of the parameters in β in the model (7) will be denoted by (OLS)$\lambda' X\beta$ and is, as is well-known,

$$\text{OLS}(\lambda' X\beta) = \lambda' X(X'X)^{-} X'y \tag{11}$$

where $(X'X)^{-}$ is a generalized inverse of $X'X^{-}$, i.e., $(X'X)^{-}$ is any matrix satisfying

$$X'X(X'X)^{-}X'X = X'X \quad .$$

Similarly the best linear unbiased estimator (BLUE) of that same estimable $\lambda' X\beta$, to be denoted BLUE($\lambda' X\beta$), is

$$\text{BLUE}(\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\beta}) = \underset{\sim}{\lambda}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \quad , \tag{12}$$

where $\underset{\sim}{V}$ is assumed to be positive definite.

In fixed effects models, $\underset{\sim}{V} = \sigma^2\underset{\sim}{I}$, as in (4), whereupon (12) very simply reduces to (11), as is well known. An extension to $\underset{\sim}{V} = [(1-\rho)\underset{\sim}{I} + \rho\underset{\sim}{J}]\sigma^2$ is given by McElroy (1967) and, in complete generality, Zyskind (1967) has shown that these two estimators are equal, if and only if

$$\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q} \qquad \text{for some } \underset{\sim}{Q} \quad . \tag{13}$$

Graybill (1976), p. 209) also has this result, restricted to $\underset{\sim}{X}$ of full column rank. We use (13) to show for a broad definition of balanced data that for mixed models of the form (7) through (10) the BLUE of an estimable function of the fixed effects parameters is the same as the OLS estimator; and for randomized complete blocks with unbalanced data we obtain an explicit expression for the BLUE of estimable functions of treatment effects.

2. BALANCED DATA

a. A general mixed model

We deal with data categorized by a number of factors, each of which is either a main effects factor (including the possibility of nested main effects factors), or an interaction factor representing the interaction of two or more main effects factors. Suppose there are m main effects factors, with the t'th one having N_t levels, for t = 1, 2, $\cdots$, m. Then the k'th observation in the "cell" defined by the i_t'th level (for $i_t = 1, \cdots, N_t$) of the t'th main effect for t = 1, $\cdots$, m, where there are $n_{i_1 i_2 \cdots i_t \cdots i_m}$ such observations, is $y_{i_1 i_2 \cdots i_t \cdots i_m k}$ for k = 1, 2,

$\cdots, n_{i_1 i_2 \cdots i_t \cdots i_m}$. On defining $\underset{\sim}{i} = [i_1 \; i_2 \; \cdots \; i_m]$, a typical observation can then be denoted as $y_{\underset{\sim}{i}k}$ for $k = 1, 2, \cdots, n_{\underset{\sim}{i}}$. Furthermore, the total number of observations is

$$N = p_r = \sum_{\underset{\sim}{i}=\underset{\sim}{1}'_m}^{\underset{\sim}{i}=\underset{\sim}{N}'} n_{\underset{\sim}{i}} \qquad \text{for } \underset{\sim}{N}' = [N_1 \quad N_2 \; \cdots \; N_t \; \cdots \; N_m] \quad .$$

($\underset{\sim}{1}'_m$ is a row vector of m unities.)

A tight, rigorous, formal and complete definition of balanced data is elusive. Development of such a definition would, as Cornfield and Tukey (1956) write, involve "··· systematic algebra [which] can take us deep into the forest of notation. But the detailed manipulation will, sooner or later, blot out any understanding we may have started with." Nevertheless, one formulation of a model that yields a wide class of balanced data situations is as follows. It is similar to that used by Smith and Hocking (1978), Searle and Henderson (1979), Seifert (1979), Khuri (1981) and Anderson *et al.* (1984).

The balanced data models we consider are those that have $n_{\underset{\sim}{i}} = n \;\forall\; \underset{\sim}{i}$. They also have each $\underset{\sim}{X}_d$ and each $\underset{\sim}{Z}_q$ of (8) being a Kronecker product (KP, for brevity) of m + 1 matrices, each of which is either an $\underset{\sim}{I}$-matrix or a $\underset{\sim}{1}$-vector; i.e.,

each $\underset{\sim}{X}_d$ and each $\underset{\sim}{Z}_q$ is a KP of m+1 matrices that are each $\underset{\sim}{I}$ or $\underset{\sim}{1}$. (14)

The occurrence of the $\underset{\sim}{I}$-matrices and $\underset{\sim}{1}$-vectors in these KPs is as follows. First, corresponding to the scalar parameter μ in the model is $\underset{\sim}{X}_1$ which is $\underset{\sim}{1}_N$, and so every matrix in its KP is a $\underset{\sim}{1}$:

$$\underset{\sim}{X}_1 = \underset{\sim}{1}_N = \underset{\sim}{1}_{N_1} * \underset{\sim}{1}_{N_2} * \cdots * \underset{\sim}{1}_{N_t} * \cdots * \underset{\sim}{1}_{N_m} * \underset{\sim}{1}_n \quad ,$$

where * represents the operation of Kronecker multiplication. Second, corresponding to $\underset{\sim}{u}_r = \underset{\sim}{e}$ is $\underset{\sim}{Z}_r$ which is $\underset{\sim}{I}_N$ and so each of the m + 1 matrices in its KP is an $\underset{\sim}{I}$-matrix:

$$\underset{\sim}{Z}_r = \underset{\sim}{I}_N = \underset{\sim}{I}_{N_1} * \underset{\sim}{I}_{N_2} * \cdots * \underset{\sim}{I}_{N_t} * \cdots * \underset{\sim}{I}_{N_m} * \underset{\sim}{I}_n \ .$$

Finally, in the KP for each $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$ (other than $\underset{\sim}{X}_1$ and $\underset{\sim}{Z}_r$), the t'th matrix corresponds to the t'th main effects factor and is $\underset{\sim}{I}_{N_t}$ when that factor is part of the definition of the factor corresponding to $\underset{\sim}{X}_d$ or $\underset{\sim}{Z}_q$; otherwise it is $\underset{\sim}{1}_{N_t}$. This is for $t = 1, \cdots, m$. And for all $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$, other than $\underset{\sim}{Z}_r$, the (m+1)'th matrix in the KP is $\underset{\sim}{1}_n$.

The phrase "part of the definition" demands explanation. It is exemplified in the 2-factor model (1), wherein the two main effects factors are each part of the definition of the interaction factor. Similarly, if nested within an α-factor there is a β-factor then the α-factor is part of the definition of that β-factor. (See also, comments B and C which follow the examples.)

Each h_d and p_q (number of levels in the d'th fixed factor and the q'th random factor, respectively) in the balanced data we have defined is the product of the numbers of columns in the $\underset{\sim}{I}$ and $\underset{\sim}{1}$ terms in the KP (14) that is $\underset{\sim}{X}_d$ and $\underset{\sim}{Z}_q$. Hence h_d is the product of the N_t values for the main effects factors that are part of the definition of the d'th fixed effect factor; p_q is a similar product for the q'th random effects factor.

Examples We give four examples that are each in terms of those of the following vectors that are appropriate: $\underset{\sim}{\alpha} = [\alpha_1, \cdots, \alpha_a]'$, $\underset{\sim}{\beta} = [\beta_1, \cdots, \beta_b]'$ or $\underset{\sim}{\beta}_+ = [\beta_{11} \cdots \beta_{1b} \ \beta_{21} \cdots \beta_{2b} \cdots \beta_{a1} \cdots \beta_{ab}]'$, $\underset{\sim}{\gamma} = [\gamma_{11} \cdots \gamma_{1b} \ \gamma_{21} \cdots \gamma_{2b} \cdots \gamma_{a1} \cdots \gamma_{ab}]'$, and $\underset{\sim}{e}$, the vector of error terms, the same order as $\underset{\sim}{y}$. Determination of which KPs are $\underset{\sim}{X}$-matrices and which are $\underset{\sim}{Z}$-matrices is governed by which factors are defined as fixed effects and which are random. This is illustrated for only example (iii).

(i) One-way classification: $y_{ij} = \mu + \alpha_i + e_{ij}$ with $i=1,\cdots,a$ and $j=1,\cdots,n$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_n)\alpha + (\underset{\sim}{I}_a * \underset{\sim}{I}_n)e \quad . \tag{15}$$

(ii) Two-way crossed classification, no interaction, and one observation per cell: $y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$ for $i=1,\cdots,a$ and $j=1,\cdots,b$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b)\underset{\sim}{\alpha} + (\underset{\sim}{1}_a * \underset{\sim}{I}_b)\underset{\sim}{\beta} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b)\underset{\sim}{e} \quad . \tag{16}$$

(iii) Two-way crossed classification, with interaction and n observations per cell: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$ with $i=1,\cdots,a$, $j=1,\cdots,b$ and $k=1,\cdots,n$.

$$\begin{aligned}\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\underset{\sim}{\alpha} + (\underset{\sim}{1}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\beta} \\ + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\gamma} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n)\underset{\sim}{e} \quad .\end{aligned} \tag{17}$$

Suppose in (17) that elements of $\underset{\sim}{\beta}$ and $\underset{\sim}{\gamma}$ were taken to be random effects. Then the terms of (8) for the general mixed model would have the following values:

$m=3$, $f=2$ with $h_1 = N_1 = 1$ and $\underset{\sim}{X}_1 = \underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n$ for $\underset{\sim}{\beta}_1 = \mu$,
and $h_2 = N_2 = a$ and $\underset{\sim}{X}_2 = \underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n$ for $\underset{\sim}{\beta}_2 = \underset{\sim}{\alpha}$;

$r=3$ with $p_1 = N_3 = b$ and $\underset{\sim}{Z}_1 = \underset{\sim}{1}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n$ for $\underset{\sim}{u}_1 = \underset{\sim}{\beta}$,
$p_2 = N_2N_3 = ab$ and $\underset{\sim}{Z}_2 = \underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n$ for $\underset{\sim}{u}_2 = \underset{\sim}{\gamma}$,
and $p_3 = N_2N_3n = abn$ and $\underset{\sim}{Z}_3 = \underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n$ for $\underset{\sim}{u}_3 = \underset{\sim}{e}$.

(iv) Two-way nested classification: $y_{ij} = \mu + \alpha_i + \beta_{ij} + e_{ijk}$ for $i=1,\cdots,a$, $j=1,\cdots,b$ and $k=1,\cdots,n$.

$$\underset{\sim}{y} = (\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\mu + (\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)\underset{\sim}{\alpha} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{1}_n)\underset{\sim}{\beta}_+ + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_n)\underset{\sim}{e} \quad . \tag{18}$$

Comments Several comments are in order. (A) In every case $\underset{\sim}{X}_1$ for μ is $\underset{\sim}{1}$, a KP of 1-vectors; and $\underset{\sim}{Z}_r$ for $\underset{\sim}{e}$ is $\underset{\sim}{I}$, a KP of $\underset{\sim}{I}$-matrices. (B) In every case the KP that is the coefficient of $\underset{\sim}{\alpha}$ has only one $\underset{\sim}{I}$-matrix in it, namely $\underset{\sim}{I}_a$. This is so because, obviously, the definition of $\underset{\sim}{\alpha}$ involves only $\underset{\sim}{\alpha}$. The same is true of the coefficient of $\underset{\sim}{\beta}$ in (16) and (17). (C) In contrast, the KP that is the coefficient of $\underset{\sim}{\beta}_+$ in (18) has two $\underset{\sim}{I}$-matrices, $\underset{\sim}{I}_a$ and $\underset{\sim}{I}_b$. This is because $\underset{\sim}{\beta}_+$ has elements that represent the nesting of the β-factor within the α-factor. Thus the α-factor is involved in the definition of $\underset{\sim}{\beta}_+$ and so the coefficient of $\underset{\sim}{\beta}_+$ contains $\underset{\sim}{I}_a$ and $\underset{\sim}{I}_b$. Thus the coefficient of $\underset{\sim}{\beta}_+$ in (18) is the same as that of $\underset{\sim}{\gamma}$, the interaction term, in (17). Judged solely by their coefficients, $\underset{\sim}{\beta}_+$ and $\underset{\sim}{\gamma}$ would therefore appear to be the same. What makes $\underset{\sim}{\gamma}$ an interaction term is that both main effect factors that go into defining it are also present on their own in (17), but with $\underset{\sim}{\beta}_+$, only one factor that goes into defining it is present on its own in (18), and so $\underset{\sim}{\beta}_+$ represents nesting. In other words, a factor that looks like an interaction factor is such when all of its associated main effects factors are present in the model; otherwise it is a nested factor. (D) Equation (16) is a special case of (17) with $\underset{\sim}{\gamma}$ omitted and n=1 and hence, for example, $\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n = \underset{\sim}{1}_a * \underset{\sim}{1}_b * 1 = \underset{\sim}{1}_a * \underset{\sim}{1}_b$.

A final observation concerns $\underset{\sim}{V} = \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{Z}_q \underset{\sim}{Z}_q'$ of (10). It is based on the general result that $(\underset{\sim}{A} * \underset{\sim}{B})(\underset{\sim}{P} * \underset{\sim}{Q}) = \underset{\sim}{AP} * \underset{\sim}{BQ}$, given the necessary conformability requirements. Thus, for $\underset{\sim}{1}_n \underset{\sim}{1}_n' = \underset{\sim}{J}_n$ being a square matrix of order n with every element unity, we have from (14) that every $\underset{\sim}{Z}_q \underset{\sim}{Z}_q'$ is a KP of $\underset{\sim}{I}$ and $\underset{\sim}{J}$ matrices. Thus we rewrite (10) as

$$\underset{\sim}{V} = \sum_{q=1}^{r} \sigma_q^2 \text{ (the KP of } \underset{\sim}{I} \text{ and } \underset{\sim}{J} \text{ matrices that is } \underset{\sim}{Z}_q\underset{\sim}{Z}_q') \quad . \tag{19}$$

b. Estimation

It is well known for many cases of balanced data that BLUEs of estimated functions of parameters in fixed effects models are simple functions of observed means. For example, in the fixed effects form of the 2-factor model (1), the BLUE of $\alpha_i - \alpha_{i'}$ is BLUE$(\alpha_i - \alpha_{i'}) = \bar{y}_{i..} - \bar{y}_{i'..}$ for $i \neq i'$. The question of interest is "Is the BLUE of $\alpha_i - \alpha_{i'}$ also $\bar{y}_{i..} - \bar{y}_{i'..}$ in a mixed model form of (1) where the β_j and δ_{ij} are treated as random effects?" The answer is 'yes'; moreover, in all cases of balanced data (as defined in the preceding section) the BLUE in a mixed model is the same as the estimator yielded by using OLS. This we now prove, by showing that (13) is satisfied for $\underset{\sim}{V}$ of (19) and $\underset{\sim}{X} = \{\underset{\sim}{X}_d\}$, $d = 1,\cdots,f$ of (14) with $\underset{\sim}{X}_d$ being a KP of $\underset{\sim}{I}$-matrices and $\underset{\sim}{1}$-vectors.

Writing $\underset{\sim}{W}_q$ for $\underset{\sim}{Z}_q\underset{\sim}{Z}_q'$ of (19) we have

$$\underset{\sim}{W}_q = \underset{\sim}{Z}_q\underset{\sim}{Z}_q' = (\underset{\sim}{W}_{q1} * \underset{\sim}{W}_{q2} * \cdots * \underset{\sim}{W}_{qt} * \cdots * \underset{\sim}{W}_{q,m+1}) = \mathop{*}_{t=1}^{m+1} \underset{\sim}{W}_{qt} \quad , \tag{20}$$

and, similarly, for

$$\underset{\sim}{X} = [\underset{\sim}{X}_1 \quad \underset{\sim}{X}_2 \cdots \underset{\sim}{X}_d \cdots \underset{\sim}{X}_f] \text{ with } \underset{\sim}{X}_d = \mathop{*}_{t=1}^{m+1} \underset{\sim}{X}_{dt} \quad , \tag{21}$$

where each $\underset{\sim}{X}_{dt}$ is either $\underset{\sim}{I}_{N_t}$ or $\underset{\sim}{1}_{N_t}$. Then from (19)

$$\underset{\sim}{V}\underset{\sim}{X} = \left\{ \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{Z}_q\underset{\sim}{Z}_q'\underset{\sim}{X}_d \right\}_{d=1}^{d=f}$$

where, by the curly braces notation, we mean that $\underset{\sim}{V}\underset{\sim}{X}$ is partitioned into a row of f sub-matrices. Thus

$$\underset{\sim}{V}\underset{\sim}{X} = \left\{ \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{W}_q\underset{\sim}{X}_d \right\}_{d=1}^{d=f} \tag{22}$$

$$= \left\{ \sum_{q=1}^{r} \sigma_q^2 \mathop{*}_{t=1}^{m+1} \underset{\sim}{W}_{qt} \underset{\sim}{X}_{dt} \right\}_{d=1}^{d=f} . \qquad (23)$$

Now, from (19) and (20), $\underset{\sim}{W}_{qt}$ is either $\underset{\sim}{I}$ or $\underset{\sim}{J}$ and $\underset{\sim}{X}_{dt}$ is either $\underset{\sim}{I}$ or $\underset{\sim}{1}$, all of order $\underset{\sim}{N}_t$. Therefore the four possible values of the product $\underset{\sim}{W}_{qt}\underset{\sim}{X}_{dt}$, together with the definition of a matrix $\underset{\sim}{M}_{qdt}$ such that $\underset{\sim}{W}_{qt}\underset{\sim}{X}_{dt} = \underset{\sim}{X}_{dt}\underset{\sim}{M}_{qdt}$ in each case, are as follows:

$\underset{\sim}{W}_{qt}$	$\underset{\sim}{X}_{dt}$	$\underset{\sim}{W}_{qt}\underset{\sim}{X}_{dt} = \underset{\sim}{X}_{dt}\underset{\sim}{M}_{qdt}$	$\underset{\sim}{M}_{qdt}$
$\underset{\sim}{I}$	$\underset{\sim}{I}$	$\underset{\sim}{I} = \underset{\sim}{I}\underset{\sim}{I}$	$\underset{\sim}{I}$
$\underset{\sim}{I}$	$\underset{\sim}{1}$	$\underset{\sim}{1} = \underset{\sim}{1}1$	1
$\underset{\sim}{J}$	$\underset{\sim}{I}$	$\underset{\sim}{J} = \underset{\sim}{I}\underset{\sim}{J}$	$\underset{\sim}{J}$
$\underset{\sim}{J}$	$\underset{\sim}{1}$	$N_t\underset{\sim}{1} = \underset{\sim}{1}N_t$	N_t

Therefore from (23)

$$\underset{\sim}{V}\underset{\sim}{X} = \left\{ \sum_{q=1}^{r} \sigma_q^2 \mathop{*}_{t=1}^{m+1} \underset{\sim}{X}_{dt} \underset{\sim}{M}_{qdt} \right\}_{d=1}^{d=f} , \qquad (24)$$

$$= \left\{ \sum_{q=1}^{r} \sigma_q^2 \underset{\sim}{X}_d \underset{\sim}{M}_{qd} \right\}_{d=1}^{d=f} , \qquad (25)$$

for

$$\underset{\sim}{M}_{qd} = \underset{\sim}{M}_{qd1} * \underset{\sim}{M}_{qd2} * \cdots * \underset{\sim}{M}_{qdt} * \cdots * \underset{\sim}{M}_{q,d,m+1} . \qquad (26)$$

Derivation both of (23) from (22) and of (25) from (24) is based both on $\underset{\sim}{X}_d$ and $\underset{\sim}{M}_q$ each being a KP, and on the product rule for KP quoted earlier.

The conformability requirements of the regular products in (24) might seem to be lacking because, from the preceding table, two forms of $\underset{\sim}{M}_{qdt}$ are scalars. However, both regular and Kronecker products of matrices do exist when one or more of the matrices is a scalar; e.g., for scalar θ, both $\underset{\sim}{A}\theta$ and $(\underset{\sim}{A} * \underset{\sim}{B})(\theta * \underset{\sim}{L}) = \underset{\sim}{A}\theta * \underset{\sim}{B}\underset{\sim}{L}$ exist. Therefore (25)

exists. Hence, on writing
$$\mathbf{Q} = \operatorname{diag}\left\{\sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qd}\right\}_{d=1}^{d=f} ,$$

the block diagonal matrix of matrices $\sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qd}$, we get from (25)

$$\mathbf{V}\mathbf{X} = [\mathbf{X}_1 \cdots \mathbf{X}_d \cdots \mathbf{X}_f]\begin{bmatrix} \sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{q1} & & & & \mathbf{0} \\ & \ddots & & & \\ & & \sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qd} & & \\ & & & \ddots & \\ \mathbf{0} & & & & \sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qf} \end{bmatrix} \tag{27}$$

$$= \mathbf{X}\mathbf{Q} \quad .$$

Thus Zyskind's condition of (13) is satisfied. Hence, with balanced data as here defined, the BLUE of an estimable function of the fixed effects in any mixed model is the same as the estimator obtained using ordinary least squares.

A final note: each sum $\sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qd}$ in (27) does exist because, as a result of (26), the order of $\mathbf{M}_{qd}$ is the product of the orders of $\mathbf{M}_{qdt}$ for $t = 1, \cdots, m+1$; and (from the Table) each $\mathbf{M}_{qdt}$ is square of order either N_t or 1. Furthermore, that order is N_t only when $\mathbf{X}_{dt} = \mathbf{I}$; and this is so only when the t'th main effects factor is involved in defining the d'th fixed effects factor. Hence the order of $\mathbf{M}_{qd}$ is the product of such N_t values, and this is h_d; thus $\mathbf{M}_{qd}$ has order h_d for all q and so $\sum_{q=1}^{r} \sigma_q^2 \mathbf{M}_{qd}$ exists.

Example Suppose in (1) and (17) that the βs and γs are random effects. Then

$$\mathbf{X} = [\mathbf{1}_a * \mathbf{1}_b * \mathbf{1}_n \quad \mathbf{I}_a * \mathbf{1}_b * \mathbf{1}_n]$$

and

$$\mathbf{V} = \sigma_\beta^2(\mathbf{J}_a * \mathbf{I}_b * \mathbf{J}_n) + \sigma_\gamma^2(\mathbf{I}_a * \mathbf{I}_b * \mathbf{J}_n) + \sigma_e^2(\mathbf{I}_a * \mathbf{I}_b * \mathbf{I}_n) \quad .$$

Hence in $\underset{\sim}{V}\underset{\sim}{X}$ the first sub-matrix is

$$\underset{\sim}{V}(\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n) = \sigma^2_\beta(a\underset{\sim}{1}_a * \underset{\sim}{1}_b * n\underset{\sim}{1}_n) + \sigma^2_\gamma(\underset{\sim}{1}_a * \underset{\sim}{1}_b * n\underset{\sim}{1}_n) + \sigma^2_e(\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)$$

$$= (\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)[\sigma^2_\beta(a * 1 * n) + \sigma^2_\gamma(1 * 1 * n) + \sigma^2_e(1 * 1 * 1)] \quad . \qquad (28)$$

Similarly, the second sub-matrix of $\underset{\sim}{V}\underset{\sim}{X}$ is

$$\underset{\sim}{V}(\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n) = \sigma^2_\beta(\underset{\sim}{J}_a * \underset{\sim}{1}_b * n\underset{\sim}{1}_n) + \sigma^2_\gamma(\underset{\sim}{I}_a * \underset{\sim}{1}_b * n\underset{\sim}{1}_n) + \sigma^2_e(\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)$$

$$= (\underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n)[\sigma^2_\beta(\underset{\sim}{J}_a * 1 * n) + \sigma^2_\gamma(\underset{\sim}{I}_a * 1 * n) + \sigma^2_e(\underset{\sim}{I}_a * 1 * 1)] \; . \qquad (29)$$

Hence

$$\underset{\sim}{V}\underset{\sim}{X} = [\underset{\sim}{1}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n \quad \underset{\sim}{I}_a * \underset{\sim}{1}_b * \underset{\sim}{1}_n]\begin{bmatrix}\underset{\sim}{M}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{M}_2\end{bmatrix} = \underset{\sim}{X}\begin{bmatrix}\underset{\sim}{M}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{M}_2\end{bmatrix}$$

for $\underset{\sim}{M}_1$ and $\underset{\sim}{M}_2$ being the matrices in square braces in (28) and (29), respectively, namely

$$\underset{\sim}{M}_1 = an\sigma^2_\beta + n\sigma^2_\gamma + \sigma^2_e \qquad \text{and} \qquad \underset{\sim}{M}_2 = n\sigma^2_\beta\underset{\sim}{J}_a + n\sigma^2_\gamma\underset{\sim}{I}_a + \sigma^2_e\underset{\sim}{I}_a \quad .$$

3. CELL MEANS MODELS

a. A general formulation

The cell means model (2) for y_{ijk} in the 2-factor case extends very naturally to $y_{\underset{\sim}{i}k}$ for any number of factors:

$$y_{\underset{\sim}{i}k} = \mu_{\underset{\sim}{i}} + e_{\underset{\sim}{i}k} \qquad \text{with } E(y_{\underset{\sim}{i}k}) = \mu_{\underset{\sim}{i}}$$

for $\underset{\sim}{i} = \underset{\sim}{1}'_m, \cdots, \underset{\sim}{N}'$ and $k = 1, 2, \cdots, n_{\underset{\sim}{i}}$. For $\underset{\sim}{y}$, $\underset{\sim}{\mu}$ and $\underset{\sim}{e}$ being the vectors, respectively, of the $y_{\underset{\sim}{i}k}$, $\mu_{\underset{\sim}{i}}$ and $e_{\underset{\sim}{i}k}$, arranged in lexicon order in each case, we write

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{e} \quad . \qquad (30)$$

Then $\underset{\sim}{X}$ is a direct sum of vectors $\underset{\sim}{1}_{n_{\underset{\sim}{i}}}$,

$$\underset{\sim}{X} = \overset{i=\underset{\sim}{N}'}{\underset{i=\underset{\sim m}{1}'}{(+)}} \underset{\sim}{1}_{n_i} \quad . \tag{31}$$

where (+) represents the direct sum operation; and $\underset{\sim}{X}$ has full column rank.

Example For $m = 2$ and $N_1 = 2$ and $N_2 = 3$

$$\underset{\sim}{X} = \overset{i=[2\ 3]}{\underset{i=[1\ 1]}{(+)}} \underset{\sim}{1}_{n_{\underset{\sim}{i}}} = \begin{bmatrix} \underset{\sim}{1}_{n_{11}} & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \underset{\sim}{1}_{n_{12}} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \underset{\sim}{1}_{n_{13}} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \underset{\sim}{1}_{n_{21}} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \underset{\sim}{1}_{n_{22}} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \underset{\sim}{1}_{n_{23}} \end{bmatrix} .$$

The OLS estimator of $\underset{\sim}{\mu}$ in (31) is

$$\tilde{\underset{\sim}{\mu}} = \text{OLS}(\underset{\sim}{\mu}) = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \bar{\underset{\sim}{y}} \quad . \tag{32}$$

with, from the nature of $\underset{\sim}{X}$ in (31), $\underset{\sim}{X}'\underset{\sim}{X}$ being $D\{n_{\underset{\sim}{i}}\}$, the diagonal matrix of the $n_{\underset{\sim}{i}}$, and $\underset{\sim}{X}'\underset{\sim}{y}$ being the vector of cell totals $y_{\underset{\sim}{i}.}$. Hence $\tilde{\underset{\sim}{\mu}} = D\{1/n_{\underset{\sim}{i}}\}\{y_{\underset{\sim}{i}.}\} = \{\bar{y}_{\underset{\sim}{i}.}\} = \bar{\underset{\sim}{y}}$, the vector of observed cell means, as in (32).

Adapting the cell means model to models where the dispersion matrix of $\underset{\sim}{y}$ is other than $\sigma^2\underset{\sim}{I}$, i.e., for a mixed model, involves using the cell means formulation for only the sub-most cells as defined by the fixed effects. For example, in a randomized complete blocks where blocks are random the cell mean model is

$$y_{ij} = \mu_i + e_{ij}$$

where, in terms of an overparameterized model $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$, the μ_i is $\mu_i = \mu + \alpha_i$ for the fixed effects part of the model and $e_{ij} =$

$\beta_j + \varepsilon_{ij}$. The difference is, though, that we do not formally identify e_{ij} as $\beta_j + \varepsilon_{ij}$, but merely attribute some form to the dispersion matrix of the e_{ij}, namely

$$\underset{\sim}{V} = \mathrm{var}(\underset{\sim}{y}) = \mathrm{var}(\underset{\sim}{e}) \tag{33}$$

in this case

$$\underset{\sim}{V} = \sigma^2_\beta(\underset{\sim}{J}_a * \underset{\sim}{I}_b) + \sigma^2_e(\underset{\sim}{I}_a * \underset{\sim}{I}_b) \quad . \tag{34}$$

In general we use $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{e}$ and $\underset{\sim}{V} = \mathrm{var}(\underset{\sim}{e})$ of (30) and (33), respectively, and then the BLUE of $\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\mu}$ is

$$\mathrm{BLUE}(\underset{\sim}{\lambda}'\underset{\sim}{X}\underset{\sim}{\mu}) = \underset{\sim}{\lambda}'\underset{\sim}{X}\hat{\underset{\sim}{\mu}} \quad \text{for} \quad \hat{\underset{\sim}{\mu}} = \mathrm{BLUE}(\underset{\sim}{\mu}) = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}')^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \tag{35}$$

where $\underset{\sim}{\mu}$ is estimable because in the cell means model $\underset{\sim}{X}$ of (31) has full column rank. And the sampling variances of these estimators are

tors are

$$\mathrm{var}(\tilde{\underset{\sim}{\mu}}) = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} \quad \text{and} \ \mathrm{var}(\hat{\underset{\sim}{\mu}}) = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1} \quad . \tag{36}$$

We can note in passing, due to the non-singularity of $\underset{\sim}{X}'\underset{\sim}{X}$ and $\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}$ that it is not difficult to show that

$$\hat{\underset{\sim}{\mu}} = \tilde{\underset{\sim}{\mu}} = \bar{\underset{\sim}{y}} \quad \text{when} \quad \underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q} \ \text{for some} \ \underset{\sim}{Q} \quad , \tag{37}$$

i.e., when the Zyskind condition is satisfied; whereupon, of course, the sampling variances in (36) are also equal.

b. Some interactions zero

The formulation $\underset{\sim}{X}\underset{\sim}{\mu}$ in (30) for the fixed effects part of a mixed model implicitly includes interactions; e.g., for two fixed effects factors, μ_{ij} in terms of the overparameterized model implicitly includes interaction between the two factors. To use a cell means formulation for the no-interaction model requires defining an absence of interactions among the μ_{ij}. This is done by using an appropriate number of equations of the form

$$\mu_{ij} - \mu_{i'j} - \mu_{ij'} - \mu_{i'j'} = 0 \tag{38}$$

for $i \neq i'$ and $j \neq j'$. This is tantamount to imposing restrictions on the elements of μ, which in general we will represent as

$$\underset{\sim}{H}\underset{\sim}{\mu} = \underset{\sim}{0} \quad . \tag{39}$$

$\underset{\sim}{H}$ is of full row rank and, although every element of any $\underset{\sim}{H}\underset{\sim}{\mu}$ is estimable, because $\underset{\sim}{\mu}$ is estimable (since $\underset{\sim}{X}$ has full column rank), we can also invoke the principles of estimability to note that

$$\underset{\sim}{H} = \underset{\sim}{L}\underset{\sim}{X} \qquad \text{for some } \underset{\sim}{L} \quad . \tag{40}$$

Then, following Searle (1971, p. 206), for example, the OLS estimator of $\underset{\sim}{\mu}$ for the restricted model $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{\mu}$ and $\underset{\sim}{H}\underset{\sim}{\mu} = \underset{\sim}{0}$ is

$$\begin{aligned}\tilde{\underset{\sim}{\mu}}_r &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} - (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{H}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{H}']^{-1}\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} \\ &= \bar{\underset{\sim}{y}} - (\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{H}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{H}']^{-1}\underset{\sim}{H}\bar{\underset{\sim}{y}} \quad ,\end{aligned} \tag{41}$$

after using (32). Similarly the BLUE is

$$\hat{\underset{\sim}{\mu}}_r = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} - (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{H}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{H}']^{-1}\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \quad .$$

On invoking the Zyskind condition this reduces to

$$\hat{\underset{\sim}{\mu}}_r = \bar{\underset{\sim}{y}} - (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{H}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{H}']^{-1}\underset{\sim}{H}\bar{\underset{\sim}{y}} \quad . \tag{42}$$

Then, in association with $\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q}$ for some $\underset{\sim}{Q}$, the question now is under what condition is the BLUE the same as the OLS estimator, i.e., when does $\hat{\underset{\sim}{\mu}}_r = \tilde{\underset{\sim}{\mu}}_r$? Since $\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q}$ implies $(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1} = \underset{\sim}{Q}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{Q}'$, the latter equality arising from symmetry, and because $\underset{\sim}{H} = \underset{\sim}{L}\underset{\sim}{X}$ for some $\underset{\sim}{L}$, we find from (41) and (42) that $\hat{\underset{\sim}{\mu}}_r = \tilde{\underset{\sim}{\mu}}_r$ if and only if

$$(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{Q}'\underset{\sim}{X}'\underset{\sim}{L}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{Q}'\underset{\sim}{X}'\underset{\sim}{L}']^{-1}\underset{\sim}{H} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{L}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{L}']^{-1}\underset{\sim}{H} \quad ,$$

i.e., if and only if, in using $\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{Q}$ and the full row rank property of $\underset{\sim}{H}$,

$$\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{L}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{L}']^{-1} = \underset{\sim}{X}'\underset{\sim}{L}'[\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{L}']^{-1} \quad . \tag{43}$$

A necessary and sufficient condition for this equality to hold is $\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{L}' =$

$\underset{\sim}{X}'\underset{\sim}{L}'\underset{\sim}{K}'$ for some non-singular $\underset{\sim}{K}$, where, in the necessity condition $\underset{\sim}{K}' = [\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{L}']^{-1}\underset{\sim}{H}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{L}'$. A simpler sufficient condition is $\underset{\sim}{V}\underset{\sim}{L}' = \underset{\sim}{L}'\underset{\sim}{P}'$ for some non-singular $\underset{\sim}{P}$; i.e.,

$$\underset{\sim}{L}\underset{\sim}{V} = \underset{\sim}{P}\underset{\sim}{L} \qquad \text{for some non-singular } \underset{\sim}{P} \quad . \tag{44}$$

Thus (44) is a condition for mixed models $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{\mu}$ with $\text{var}(\underset{\sim}{y}) = \underset{\sim}{V}$, and restrictions $\underset{\sim}{H}\underset{\sim}{\mu} = \underset{\sim}{0}$ for $\underset{\sim}{H} = \underset{\sim}{L}\underset{\sim}{X}$, under which the BLUE of $\underset{\sim}{\mu}$ is the same as the estimator obtained from OLS. Two situations when (44) is trivially true are as follows: (i) models that include all interactions among their fixed, main effects factors, because then in terms of (40) $\underset{\sim}{L}$ is null and so (44) is obviously satisfied; and (ii) models in which $\underset{\sim}{V} = \sigma^2\underset{\sim}{I}$, for then with $\underset{\sim}{P} = \underset{\sim}{V}$ (44) is also satisfied. It remains for us to consider mixed models, with $\underset{\sim}{V}$ having some form other than $\sigma^2\underset{\sim}{I}$ and in which some interactions among the fixed, main effects factors are assumed to be non-existent. We do so for balanced data only.

c. Balanced data, mixed models, some fixed effects interactions missing

Example We begin with the example of a four-way crossed classification, with one factor random and with the third order and one set of second order interactions among fixed effects being zero. Thus the overparameterized model could be

$$E(y_{ijk\ell v}) = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{jk} + \delta_\ell + e_{ijk\ell v}$$

for a, b, c, and d levels of the four main effects factors, respectively, and n observations per cell. For the α_i, β_j and γ_k effects taken as fixed, and the δ_ℓ effects as random, the cell means formulation would be

$$y_{ijk\ell v} = \mu_{ijk} + \varepsilon_{ijk} \tag{45}$$

with restrictions of the form

$$\mu_{i\cdot k} - \mu_{i'\cdot k} - \mu_{i\cdot k'} + \mu_{i'\cdot k'} = 0 \tag{46}$$

for $i \neq i'$ and $j \neq j'$; and

$$\mu_{ijk} - \mu_{i'jk} - \mu_{ij'k} + \mu_{i'j'k} - (\mu_{ijk'} - \mu_{i'jk'} - \mu_{ij'k'} + \mu_{i'j'k'}) = 0 \quad , \qquad (47)$$

for $i \neq i'$, $j \neq j'$ and $k \neq k'$. In writing (45) as

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{e} \quad ,$$

with elements of $\underset{\sim}{y}$, $\underset{\sim}{\mu}$ and $\underset{\sim}{e}$ in lexicon order, we have

$$\underset{\sim}{X} = \underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_c * \underset{\sim}{1}_d * \underset{\sim}{1}_n \qquad (48)$$

and

$$\underset{\sim}{V} = (\underset{\sim}{J}_a * \underset{\sim}{J}_b * \underset{\sim}{J}_c * \underset{\sim}{I}_d * \underset{\sim}{J}_n)\sigma^2_\delta + \underset{\sim}{I}_{abcdn}\sigma^2_e \ . \qquad (49)$$

Then, on defining $\underset{\sim}{T}_a$ as the $(a-1) \times a$ matrix

$$\underset{\sim}{T}_a = [\underset{\sim}{1}_{a-1} \quad -\underset{\sim}{I}_{a-1}] \quad \text{with} \quad \underset{\sim}{T}_a\underset{\sim}{J}_a = 0 \quad , \qquad (50)$$

the absence of the $(\alpha\gamma)$ and $(\alpha\beta\gamma)$ interactions can be written as

$$\underset{\sim}{H}\underset{\sim}{\mu} = \underset{\sim}{0} \quad \text{for} \quad \underset{\sim}{H} = \begin{bmatrix} \underset{\sim}{H}_1 \\ \underset{\sim}{H}_2 \end{bmatrix} \qquad (51)$$

and

$$\underset{\sim}{H}_1 = [\underset{\sim}{T}_a * \underset{\sim}{1}'_b * \underset{\sim}{T}_c] \quad \text{and} \quad \underset{\sim}{H}_2 = [\underset{\sim}{T}_a * \underset{\sim}{T}_b * \underset{\sim}{T}_c] \ . \qquad (52)$$

Thus on comparing $\underset{\sim}{H}_1$ and $\underset{\sim}{H}_2$ with $\underset{\sim}{X}$ of (48), it can be seen that

$$\underset{\sim}{H} = \underset{\sim}{L}\underset{\sim}{X} \quad \text{for} \quad \underset{\sim}{L} = \underset{\sim}{H} * (\underset{\sim}{1}'_d/d) * (\underset{\sim}{1}'_n/n) \ . \qquad (53)$$

Then, on using $\underset{\sim}{T}_a\underset{\sim}{J}_a = \underset{\sim}{0}$ of (50), it is evident from (49) and (53) that

$$\underset{\sim}{L}\underset{\sim}{V} = \underset{\sim}{L}\underset{\sim}{I}\sigma^2_e = (\sigma^2_e\underset{\sim}{I})\underset{\sim}{L} \qquad (54)$$

and so (44) is satisfied. Hence in this example of balanced data, from a mixed model with some of the interactions between fixed main effects assumed as being zero, the Zyskind condition is upheld and so in the restricted cell means model the BLUE of $\underset{\sim}{\mu}$ is the same as the OLS estimator.

The result just obtained for the example is true in general. $\underset{\sim}{X}$, like (48), is always a KP of $\underset{\sim}{I}$-matrices corresponding to the main effects that define the fixed effects, of $\underset{\sim}{1}$-vectors corresponding to the main effects that define the random effects, and of $\underset{\sim}{1}_n$ for n observations per cell. $\underset{\sim}{V}$, like (19) and (49) is always $\sigma_e^2\underset{\sim}{I}_N$ plus a weighted sum (using variance components as weights) of KP's of m+1 $\underset{\sim}{I}$ and $\underset{\sim}{J}$ matrices, with the matrices corresponding to the main effects that define fixed effects being $\underset{\sim}{J}$-matrices — with two exceptions that shall be considered shortly. $\underset{\sim}{H}$ can always, as in (51) and (52), be partitioned into subsets of rows, each subset being a KP of $\underset{\sim}{T}$s and $(\underset{\sim}{1}')$s, and $\underset{\sim}{L}$ is then the KP of $\underset{\sim}{H}$ and a KP of vectors $\underset{\sim}{1}'_{N_t}/N_t$ and $\underset{\sim}{1}'_n/n$, as in (53). Hence in the product $\underset{\sim}{L}\underset{\sim}{V}$ every term except $\underset{\sim}{L}\sigma_e^2\underset{\sim}{I}$ has a product $\underset{\sim}{T}\underset{\sim}{J}$ in it, which by (50) is null; and so $\underset{\sim}{L}\underset{\sim}{V} = (\sigma_e^2\underset{\sim}{I})\underset{\sim}{L}$, which satisfies (44).

The two exceptions are for nested random factors, and for random factors that are interactions between fixed and random factors. Each of these affect $\underset{\sim}{V}$ by changing some of the $\underset{\sim}{J}$s corresponding to main effects that define fixed effects to be $\underset{\sim}{I}$s. The only occasion that this affects a term in $\underset{\sim}{L}\underset{\sim}{V}$ is if, for every $\underset{\sim}{T}$ in $\underset{\sim}{H}$ (and $\underset{\sim}{L}$), the corresponding $\underset{\sim}{J}$ in a term in $\underset{\sim}{V}$ becomes $\underset{\sim}{I}$ on the occurrence of either of these exceptions, then the resulting term in $\underset{\sim}{L}\underset{\sim}{V}$ that was null will become a multiple of $\underset{\sim}{L}$. Hence $\underset{\sim}{L}\underset{\sim}{V} = \underset{\sim}{P}\underset{\sim}{L}$ is still upheld, for $\underset{\sim}{P}$ being a scalar matrix, although different from $\sigma_e^2\underset{\sim}{I}$ of (54).

Example (continued) Suppose in the preceding example that the covariance structure includes the assumption that all observations in the same i,j,k cell of the three fixed effects factors have a common variance. Then, instead of $\underset{\sim}{V}$ of (49) the variance-covariance matrix of $\underset{\sim}{y}$ will be

$\underset{\sim}{V}_1$ with

$$\underset{\sim}{V}_1 = \underset{\sim}{V} + (\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_c * \underset{\sim}{J}_d * \underset{\sim}{J}_n)\sigma^2_{\alpha\beta\gamma} \quad .$$

Then, for $\underset{\sim}{L}$ of (53), using (54),

$$\begin{aligned}\underset{\sim}{L}\underset{\sim}{V}_1 &= \underset{\sim}{L}\underset{\sim}{V} + \{[\underset{\sim}{H}(\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_c)] * (\underset{\sim}{1}'_d/d)\underset{\sim}{J}_d * (\underset{\sim}{1}'_n/n)\underset{\sim}{J}_n\}\sigma^2_{\alpha\beta\gamma} \\ &= \sigma^2_e\underset{\sim}{L} + \{\underset{\sim}{H}(\underset{\sim}{I}_a * \underset{\sim}{I}_b * \underset{\sim}{I}_c) * \underset{\sim}{1}'_d * \underset{\sim}{1}'_n\}\sigma^2_{\alpha\beta\gamma} \\ &= \sigma^2_e\underset{\sim}{L} + dn\sigma^2_{\alpha\beta\gamma}[\underset{\sim}{H} * (\underset{\sim}{1}'_d/d) * (\underset{\sim}{1}'_n/n)] \\ &= (\sigma^2_e + dn\sigma^2_{\alpha\beta\gamma})\underset{\sim}{L} \quad .\end{aligned}$$

Thus we conclude for balanced data in general, from mixed models with some (or all) of the interactions among fixed effects being assumed non-existent, that the BLUE of $\underset{\sim}{\mu}$ is the same as the OLS estimator. Moreover, this result holds for all cases of balanced data from mixed models be some, all or none of the interactions be assumed zero. This would seem to satisfactorily refute the suggestion made by Steinhorst (1982), quoted near the end of Section 1(b) of this paper, that the cell means model is inapplicable to mixed models — at least for balanced data as have been defined in Section 2. We now turn to a particular example of unbalanced data, and a special case thereof, the balanced incomplete blocks design.

4. RANDOMIZED BLOCKS WITH UNBALANCED DATA

We consider the case of testing a treatments in b blocks with n_{ij} observations on treatment i in block j for $i = 1, \cdots, a$ and $j = 1, \cdots, b$. The cell means formulation for the k'th observation ($k = 1, 2, \cdots, n_{ij}$) on treatment i in block j is

$$E(y_{ijk}) = \mu_i \quad . \tag{55}$$

We assume that all observations in the same block have a common covariance, σ_β^2 say, and more specifically that the variance-covariance structure among the observations is

$$v(y_{ijk}) = \sigma_e^2 + \sigma_\beta^2 \quad ,$$

$$\begin{aligned} cov(y_{ijk}, y_{ijk'}) &= \sigma_\beta^2 \quad \text{for } k \neq k' = 1, 2, \cdots, n_{ij} \quad , \\ cov(y_{ijk}, y_{i'jk'}) &= \sigma_\beta^2 \quad \text{for } i \neq i', \ k = 1, \cdots, n_{ij} \text{ and } k' = 1, \cdots, n_{i'j} \quad , \end{aligned} \tag{56}$$

and

$$cov(y_{ijk}, y_{i'j'k'}) = 0 \quad \text{for } j \neq j' \quad .$$

The consequence of this is that for

$$\underset{\sim}{Z} = \begin{bmatrix} \underset{\sim}{Z}_1 \\ \underset{\sim}{Z}_2 \\ \vdots \\ \underset{\sim}{Z}_a \end{bmatrix} \quad \text{with} \quad \underset{\sim}{Z}_i = \overset{b}{\underset{j=1}{(+)}} \underset{\sim}{1}_{n_{ij}} \tag{57}$$

$$\underset{\sim}{V} = \sigma_\beta^2 \underset{\sim}{Z}\underset{\sim}{Z}' + \sigma_e^2 \underset{\sim}{I}_N \quad . \tag{58}$$

Furthermore, from (55)

$$\underset{\sim}{X} = \overset{a}{\underset{i=1}{(+)}} \underset{\sim}{1}_{n_{i\cdot}} \quad . \tag{59}$$

Applying to (58) the general result

$$(\underset{\sim}{D} - \underset{\sim}{C}\underset{\sim}{A}^{-1}\underset{\sim}{B})^{-1} = \underset{\sim}{D}^{-1} + \underset{\sim}{D}^{-1}\underset{\sim}{C}(\underset{\sim}{A} - \underset{\sim}{B}\underset{\sim}{D}^{-1}\underset{\sim}{C})^{-1}\underset{\sim}{B}\underset{\sim}{D}^{-1}$$

from, for example, Searle (1982, p. 261) gives, after a little simplification

$$\underset{\sim}{V}^{-1} = \left[\underset{\sim}{I} - \underset{\sim}{Z}\left(\overset{b}{\underset{j=1}{(+)}} \frac{\sigma_\beta^2}{\sigma_e^2 + n_{\cdot j}\sigma_\beta^2}\right)\underset{\sim}{Z}'\right] / \sigma_e^2 \quad . \tag{60}$$

Then $\underset{\sim}{X}'\underset{\sim}{V}^{-1}$ utilizes $\underset{\sim}{X}'\underset{\sim}{Z}$ which from (57) and (59) is

$$\underset{\sim}{X}'\underset{\sim}{Z} = \{n_{ij}\} \quad \text{for } i = 1,\cdots,a \text{ and } j = 1,\cdots,b$$

$$= \{\underset{\sim}{c}_j\} \quad \text{for } j = 1,\cdots,b \text{ on defining } \underset{\sim}{c}_j = [n_{ij}\ n_{2j}\cdots n_{aj}]' \ . \qquad (61)$$

Thus we find that

$$\hat{\underset{\sim}{\mu}} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

$$= \left[\underset{\sim}{X}'\underset{\sim}{X} - \{\underset{\sim}{c}_j\} \overset{b}{\underset{j=1}{(+)}} \frac{\sigma^2_\beta}{\sigma^2_e+n_{\cdot j}\sigma^2_\beta} \{\underset{\sim}{c}_j\}'\right]^{-1}\left[\underset{\sim}{X}'\underset{\sim}{y} - \{\underset{\sim}{c}_j\} \overset{b}{\underset{j=1}{(+)}} \frac{\sigma^2_\beta}{\sigma^2_e+n_{\cdot j}\sigma^2_\beta} \underset{\sim}{Z}'\underset{\sim}{y}\right]$$

$$= \left[\overset{a}{\underset{i=1}{+}} n_{i\cdot} - \sum_{j=1}^{b} \frac{\sigma^2_\beta}{\sigma^2_e+n_{\cdot j}\sigma^2_\beta} \underset{\sim}{c}_j\underset{\sim}{c}_j'\right]^{-1}\left[\left\{y_{i\cdot\cdot} - \sum_{j=1}^{b} \frac{n_{ij}\sigma^2_\beta y_{\cdot j\cdot}}{\sigma^2_e+n_{\cdot j}\sigma^2_\beta}\right\}\right] \ . \qquad (62)$$

This is a general result for estimating treatment effects in a randomized blocks when the treatments have different numbers of observations within a block, and also from block to block. And, of course

$$\text{var}(\hat{\underset{\sim}{\mu}}) = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1} = \sigma^2_e\left[\overset{a}{\underset{i=1}{(+)}} n_{i\cdot} - \sum_{j=1}^{b} \frac{\sigma^2_\beta}{\sigma^2_e+n_{\cdot j}\sigma^2_\beta} \underset{\sim}{c}_j\underset{\sim}{c}_j'\right]^{-1} \ . \qquad (63)$$

Two minor features of these results are worth commenting on. One is that estimates of σ^2_e and σ^2_β are required for calculating an estimate from $\hat{\underset{\sim}{\mu}}$; and second, for balanced data, i.e., $n_{ij} = n$ for all i and j, $\hat{\underset{\sim}{\mu}}$ of (62) simplifies to being $\hat{\mu}_i = \bar{y}_{i\cdot\cdot}$, as one would expect. An extension would be to include in the variance-covariance structure of (56) a covariance among observations in the same cell so that $v(y_{ijk}) = \sigma^2_e + \sigma^2_\beta$ of (56) would become $\sigma^2_e + \sigma^2_\beta + \sigma^2_\gamma$; and $\text{cov}(y_{ijk}, y_{ijk'}) = \sigma^2_\beta$ for $k \neq k' = 1, \cdots, n_{ij}$ would become $\sigma^2_\beta + \sigma^2_\gamma$. The other terms in (56) would remain unaltered.

5. BALANCED INCOMPLETE BLOCKS (BIB)

Data from a balanced incomplete blocks experiment can be arrayed as a 2-way crossed classification with values of n_{ij} being 0 and 1 in a patterned manner determined by the nature of the experiment. The estimation of treatment effects in a BIB experiment is therefore a special case of (62).

Example Consider four treatments (a=4) used in a BIB experiment of six blocks (b=6) with two treatments in each block. The pattern of n_{ij}-values can be arrayed as in Table 1, where a dash represents no observation.

Table 1

	Block						
Treatment	1	2	3	4	5	6	$n_{i.} = r$
I	1	1	1	–	–	–	3
II	1	–	–	1	1	–	3
III	–	1	–	1	–	1	3
IV	–	–	1	–	1	1	3
$n_{.j} = k$	2	2	2	2	2	2	$12 = n_{..} = ar = kb$

Characteristics of a BIB experiment, with values for the example, are as follows:

Number of blocks: $b = 6$.

Number of different treatments used in each block: $k = n_{.j} = 2$.

Number of treatments: $a = t = 4$.

Number of blocks containing each particular treatment: $r = n_{i.} = 3$.

Number of times each treatment pair occurs in the same block: $\lambda = 1$.

Total number of observations: $n_{..} = ar = bk = 12$.

Total number of treatment pairs in the same block that contain a particular treatment: $\lambda(a-1) = r(k-1) = 3$.

To simplify (62) first note that any cell containing data has only one observation (BIB designs with more than one can be considered, but are not dealt with here), and so we denote it by y_{ij}. Then (62) is

$$\left\{\hat{\mu}_i\right\}_{i=1}^{i=a} = \left[r\underset{\sim}{I}_a - \frac{\beta}{e + k\beta}\sum_{j=1}^{b}\underset{\sim}{c}_j\underset{\sim}{c}_j'\right]^{-1}\left\{y_{i.} - \frac{\beta}{e + k\beta}\sum_{j=1}^{b}n_{ij}y_{.j}\right\}_{i=1}^{i=a} \quad . \tag{64}$$

where, for notational convenience we write

$$\beta \text{ for } \sigma_\beta^2 \text{ and } e \text{ for } \sigma_e^2 \quad . \tag{65}$$

Simplifying (64) involves two summation terms. For the first we get assistance from the example.

Example (continued) Using the columns of unities and zeros in Table 1 as the columns $\underset{\sim}{c}_j$,

$$\sum_{j=1}^{b}\underset{\sim}{c}_j\underset{\sim}{c}_j' = \begin{bmatrix}1&1&\cdot&\cdot\\1&1&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\end{bmatrix} + \begin{bmatrix}1&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot\end{bmatrix} + \begin{bmatrix}1&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&1\end{bmatrix} + \begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\\cdot&1&1&\cdot\\\cdot&1&1&\cdot\\\cdot&\cdot&\cdot&\cdot\end{bmatrix} + \begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&1\end{bmatrix} + \begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&1&1\\\cdot&\cdot&1&1\end{bmatrix}$$

$$= \begin{bmatrix}3&1&1&1\\1&3&1&1\\1&1&3&1\\&1&1&3\end{bmatrix} = (3 - 1)\underset{\sim}{I}_4 + \underset{\sim}{J}_4 \quad .$$

Generalization for any BIB is that

$$\sum_{j=1}^{b}\underset{\sim}{c}_j\underset{\sim}{c}_j' = (r - \lambda)\underset{\sim}{I}_a + \lambda\underset{\sim}{J}_a \quad . \tag{66}$$

The second summation for (64) is

$$\sum_{j=1}^{b}n_{ij}y_{.j} = \sum_{j=1}^{b}n_{ij}k\bar{y}_{.j} = kr\left(\sum_{j=1}^{b}n_{ij}\bar{y}_{.j}\right)/r = kr\bar{y}_{i(j)} \tag{67}$$

where

$$\bar{y}_{i(j)} = \sum_{j=1}^{b} n_{ij}\bar{y}_{.j.}/r = \text{Mean of block means } \bar{y}_{.j.} \text{ for the blocks that contain treatment } i.$$

Substituting (66) and (67) into (64) gives

$$\left\{\hat{\mu}_i\right\}_{i=1}^{i=a} = \left[r\underset{\sim}{I}_a - \frac{\beta(r-\lambda)}{e+k\beta}\underset{\sim}{I}_a - \frac{\beta\lambda}{e+k\beta}\underset{\sim}{J}_a\right]^{-1}\left\{y_{i.} - \frac{\beta kr}{e+k\beta}\bar{y}_{i(j)}\right\}_{i=1}^{i=a} \tag{68}$$

$$= (e+k\beta)\Big([re + (rk - r + \lambda)\beta]\underset{\sim}{I}_a - \beta\lambda\underset{\sim}{J}_a\Big)^{-1}\left\{y_{i.} - \frac{kr\beta}{e+k\beta}\bar{y}_{i(j)}\right\}_{i=1}^{i=a} .$$

But $\lambda(a-1) = r(k-1)$, so that

$$\left\{\hat{\mu}_i\right\}_{i=1}^{i=a} = (e+k\beta)[(re+\lambda a\beta)\underset{\sim}{I}_a - \beta\lambda\underset{\sim}{J}_a]^{-1}\left\{y_{i.} - \frac{kr\beta}{e+k\beta}\bar{y}_{i(j)}\right\}_{i=1}^{i=a}$$

$$= \frac{e+k\beta}{re+\lambda a\beta}\left(\underset{\sim}{I}_a + \frac{\lambda\beta}{re}\underset{\sim}{J}_a\right)\left\{y_{i.} - \frac{kr\beta}{e+k\beta}\bar{y}_{i(j)}\right\}_{i=1}^{i=a} . \tag{69}$$

Hence

$$\hat{\mu}_i = \frac{e+k\beta}{re+\lambda a\beta}\left[y_{i.} + \frac{\lambda\beta}{re}y_{..} - \frac{kr\beta}{e+k\beta}\bar{y}_{i(j)} - \frac{\lambda\beta}{re}\frac{kr\beta}{e+k\beta}\sum_{i=1}^{a}\bar{y}_{i(j)}\right] .$$

But from (67)

$$\sum_{i=1}^{a}\bar{y}_{i(j)} = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\bar{y}_{.j}/r = \sum_{j=1}^{b} n_{.j}\bar{y}_{.j}/r = y_{..}/r .$$

Therefore

$$\hat{\mu}_i = \frac{e+k\beta}{re+\lambda a\beta}\left[y_{i.} - \frac{kr\beta}{e+k\beta}\bar{y}_{i(j)} + \frac{\lambda\beta}{re}\left(1 - \frac{k\beta}{e+k\beta}\right)y_{..}\right]$$

$$= \frac{r(e+k\beta)}{re+\lambda a\beta}\left[\bar{y}_{i.} - \frac{k\beta}{e+k\beta}\bar{y}_{i(j)} + \frac{\lambda a\beta}{r(e+k\beta)}\bar{y}_{..}\right] . \tag{70}$$

As shown in the appendix, this result is consistent with results given in Scheffé (1959).

Furthermore, from (63) and (68),

$$\text{var}(\hat{\underset{\sim}{\mu}}) = e\left[r\underset{\sim}{I}_a - \frac{\beta(r - \lambda)}{e + k\beta}\underset{\sim}{I}_a - \frac{\beta\lambda}{e + k\beta}\underset{\sim}{J}_a\right]^{-1}$$

and from (69) this is

$$\text{var}(\hat{\underset{\sim}{\mu}}) = \frac{e(e + k\beta)}{re + \lambda a\beta}\left(\underset{\sim}{I}_a + \frac{\lambda\beta}{re}\underset{\sim}{J}_a\right) .$$

Hence

$$v(\hat{\mu}_i) = \frac{(e + k\beta)(re + \lambda\beta)}{r(re + \lambda a\beta)} \tag{71}$$

and

$$\text{cov}(\hat{\mu}_i, \hat{\mu}_{i'}) = \frac{\lambda\beta(e + k\beta)}{r(re + \lambda a\beta)} \quad \text{for } i \neq i' . \tag{72}$$

Thus the estimated difference between treatments h and i by this method is, from (70),

$$\hat{\mu}_h - \hat{\mu}_i = \frac{r(e + k\beta)}{re + \lambda a\beta}\left\{\bar{y}_{h\cdot} - \bar{y}_{i\cdot} - \frac{k\beta}{e + k\beta}\left[\bar{y}_{h(j)} - \bar{y}_{i(j)}\right]\right\}$$

with, from (71) and (72)

$$v(\hat{\mu}_h - \hat{\mu}_i) = \frac{e + k\beta}{r(re + \lambda a\beta)}[2(re + \lambda\beta) + 2\lambda\beta]$$

$$= \frac{2(e + k\beta)(re + 2\lambda\beta)}{r(re + \lambda a\beta)} ,$$

where, as in (65), $\beta \equiv \sigma^2_\beta$ and $e \equiv \sigma^2_e$.

References

Anderson, R. D., Henderson, H. V., Pukelsheim, F. and Searle, S. R. (1984). Best estimation of variance components from balanced data with arbitrary kurtosis. *Mathematische Operationsforschung und Statistik. Series Statistik* 15, 163-176.

Cornfield, J. and Tukey, J. W. (1956). Average values of mean squares in factorials. *Ann. Math. Statist.* 27, 907-949.

Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Duxbury, North Scituate, Massachusetts.

Hocking, R. R. and Speed, F. M. (1975). A full rank analysis of some linear model problems. *J. American Statistical Association* 70, 706-712.

Khuri, A. I. (1981). Anova sums of squares for balanced mixed-effects models revisited. Technical Report No. 162, University of Florida.

McElroy, F. W. (1967). A necessary and sufficient condition that ordinary least squares estimators be best linear unbiased. *J. American Statistical Association* 62, 1302-1304.

Scheffé, Henry (1959). *The Analysis of Variance.* John Wiley & Sons, New York.

Searle, S. R. (1982). *Matrix Algebra Useful for Statistics.* John Wiley & Sons, New York.

Searle, S. R. and Henderson, H. V. (1979). Dispersion matrices in variance components models. *J. American Statistical Association* 74, 465-470.

Seifert, B. (1979). Optimal testing for fixed effects in general balanced mixed classification models. *Mathematische Operationsforschung und Statistik. Series Statistics* 10, 237-255.

Smith, D. W. and Hocking, R. R. (1978). Maximum likelihood analysis of the mixed model: the balanced case. *Communications in Statistics — Theory and Methods* A7, 1253-1266.

Speed, F. M. (1969). A new approach to the analysis of linear models. Ph.D. Thesis, Texas A & M University, College Station, Texas.

Speed, F. M. and Hocking, R. R. (1976). The use of the R( ) notation with unbalanced data. *The American Statistician* 30, 30-34.

Speed, F. M., Hocking, R. R. and Hackney, O. P. (1978). Methods of analysis of linear models with unbalanced data. *J. American Statistical Association* 73, 105-112.

Steinhorst, R. K. (1982). Resolving current controversies in analysis of variance. *The American Statistician* 36, 138-139.

Urquhart, N. S. and Weeks, D. L. (1978). Linear models in messy data: some problems and alternatives. *Biometrics* 34, 696-705.

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statistics* 38, 1092-1109.

APPENDIX: Analysis of BIB Data

a. Reconciliation of $\hat{\mu}_i$ with Scheffé.

One of the few places where the randomness of the blocks in a BIB design has been taken into account in estimating treatment effects is in Scheffé (1959) at pages 165-178. We show that the result given there, for estimation using recovery of interblock information, is consistent with $\hat{\mu}_i$ of (70). We begin with laying out equivalent notation.

Scheffé			This paper
p. 161:	# of treatments	I	$a = t$
	# of blocks	J	b
	# of replications	r	r
	block size	k	k
p. 162: (line 3 up)	# of occurrences of treatment i in block j	$K_{ij} = 0$ or 1	n_{ij}
p. 164: (lines 8-9)	i'th treatment total	g_i	$y_{i.}$
	j'th block total	h_j	$y_{.j}$
	i'th adjusted treatment total	$_i$	
(after 5.2.9):	$_i = g_i - k^{-1}\Sigma_j K_{ij} h_j$		$y_{i.} - \Sigma_j n_{ij}\bar{y}_{.j.}$ $= y_{i.} - r\bar{y}_{i(j)}$
	sum of block totals in which treatment i occurs	T_i	
(5.2.10):	$T_i = \Sigma_j n_{ij} h_j$		$kr\bar{y}_{i(j)}$
p. 166: (5.2.17)	efficiency factor	δ	
	$\delta = \dfrac{rk - r + \lambda}{rk} = \dfrac{(k-1)I}{k(I-1)}$		$\dfrac{\lambda a}{rk} = \dfrac{(k-1)a}{k(a-1)}$

p. 165: (last line)	$r\delta\hat{\alpha}_i = G_i$ $\hat{\alpha}_i = G_i/r\delta$	$\dfrac{y_{i.} - r\bar{y}_{i(j)}}{r\delta}$ $= rk(\bar{y}_{i.} - \bar{y}_{i(j)})/\lambda a$
p. 172: (5.2.33)	$\hat{\alpha}'_i = \dfrac{T_i - rJ^{-1}\Sigma_j h_j}{r - \lambda}$	$\dfrac{kr\bar{y}_{i(j)} - r\Sigma_j y_{.j}/b}{r - \lambda}$ $= \dfrac{kr\bar{y}_{i(j)} - ry_{..}/b}{r - \lambda}$ $= \dfrac{kr(\bar{y}_{i(j)} - \bar{y}_{..})}{r - \lambda}$
(5.2.32b):	$\sigma_f^2 = k^2\sigma_B^2 + k\sigma_e^2$	$k(e + k\beta)$
(line 5 up):	$\psi = \Sigma_i c_i \alpha_i$ $\Sigma_i c_i = 0$ $\hat{\psi}' = \Sigma_i c_i \hat{\alpha}'_i$	
p. 174: (5.2.41)	$w = r\delta/\sigma_e^2$ $w' = (r - \lambda)/\sigma_f^2$	$\lambda a/ke$ $(r - \lambda)/k(e + k\beta)$
p. 175: (5.2.42)	$\psi^* = \dfrac{w\hat{\psi} + w'\hat{\psi}'}{w + w'}$	

ψ^* is described by Scheffé as being unbiased and having minimum variance. It therefore corresponds to an element in our $\hat{\mu}$. Since ψ is a contrast of α_i terms it is also a contrast of $(\mu + \alpha_i)$ terms. The consistency of ψ^* with $\hat{\mu}$ will therefore be shown by adapting the i'th element ψ^* to be

$$\mu_i^* = \frac{w(\hat{\mu} + \hat{\alpha}_i) + w'(\hat{\mu}' + \hat{\alpha}'_i)}{w + w'}$$

and showing that $\mu_i^* = \hat{\mu}_i$.

Scheffé gives $\hat{\alpha}_i$ on page 165 — as shown above. Nowhere there does he show the corresponding $\hat{\mu}$. But in the last line of page 164 he mentions the "correction term for the grand mean". From that we infer that

$$\hat{\mu} = \bar{y}_{..}\ .$$

The expression for $\hat{\alpha}_i'$ is given at (5.2.34) on page 172. From (5.2.33) we get the corresponding

$$\hat{\mu}' = k\Sigma_j h_j/k^2 J = \Sigma_j y_{.j}/ka = \bar{y}_{..}\ .$$

Thus, using $\hat{\mu} = \hat{\mu}' = \bar{y}_{..}$ and w, w', $\hat{\alpha}$, $\hat{\alpha}'$ as above we have, from Scheffé's methodology,

$$\mu_i^* = \bar{y}_{..} + \frac{\frac{\lambda a}{ke}\frac{kr}{\lambda a}\left(\bar{y}_{i.} - \bar{y}_{i(j)}\right) + \frac{r-\lambda}{k(e+k\beta)}\frac{kr(\bar{y}_{i(j)} - \bar{y}_{..})}{r-\lambda}}{\frac{\lambda a}{ke} + \frac{r-\lambda}{k(e+k\beta)}}$$

$$= \bar{y}_{..} + \frac{r\left[\left(\bar{y}_{i.} - \bar{y}_{i(j)}\right)/e + \left(\bar{y}_{i(j)} - \bar{y}_{..}\right)/(e+k\beta)\right]}{[\lambda a(e+k\beta) + (r-\lambda)e]/ke(e+k\beta)}$$

$$= \bar{y}_{..} + \frac{rk\left[(e+k\beta)\left(\bar{y}_{i.} - \bar{y}_{i(j)}\right) + e\left(\bar{y}_{i(j)} - \bar{y}_{..}\right)\right]}{\lambda ak\beta + rke},$$

because $\lambda a + r - \lambda = rk$

$$= \bar{y}_{..} + \frac{r(e+k\beta)}{re + a\lambda\beta}\left[\bar{y}_{i.} - \frac{k\beta}{e+k\beta}\bar{y}_{i(j)} - \frac{e}{e+k\beta}\bar{y}_{..}\right]$$

$$= \frac{r(e+k\beta)}{re + a\lambda\beta}\left[\bar{y}_{i.} - \frac{k\beta}{e+k\beta}\bar{y}_{i(j)} + \frac{a\lambda\beta}{r(e+k\beta)}\bar{y}_{..}\right],$$

$= \hat{\mu}_i$ of (70).

b. The variance of $\hat{\mu}_i$

From (70)

$$v(\hat{\mu}_i) = v\left\{\frac{r(e + k\beta)}{re + \lambda a\beta}\left[\bar{y}_{i\cdot} - \frac{k\beta}{e + k\beta}\bar{y}_{i(j)} + \frac{\lambda a\beta}{r(e + k\beta)}\bar{y}_{\cdot\cdot}\right]\right\}$$

$$= \frac{r^2(e + k\beta)^2}{(re + \lambda a\beta)^2}\left\{v(\bar{y}_{i\cdot}) + \frac{k^2\beta^2}{(e + k\beta)^2}v(\bar{y}_{i(j)}) + \frac{\lambda^2 a^2\beta^2}{r^2(e + k\beta)^2}v(\bar{y}_{\cdot\cdot})\right.$$

$$+ 2\left[-\frac{k\beta}{e + k\beta}\operatorname{cov}(\bar{y}_{i\cdot},\bar{y}_{i(j)}) - \frac{k\beta}{e + k\beta}\frac{\lambda a\beta}{r(e + k\beta)}\operatorname{cov}(\bar{y}_{i(j)},\bar{y}_{\cdot\cdot})\right.$$

$$\left.\left. + \frac{\lambda a\beta}{r(e + k\beta)}\operatorname{cov}(\bar{y}_{i\cdot},\bar{y}_{\cdot\cdot})\right]\right\}$$

$$= \frac{r^2(e + k\beta)^2}{(re + \lambda a\beta)^2}\left\{\frac{r(e + \beta)}{r^2} + \frac{k^2\beta^2 rk(e + k\beta)}{(e + k\beta)^2 r^2 k^2} + \frac{\lambda^2 a^2\beta^2 ar(e + k\beta)}{r^2(e + k\beta)^2 a^2 r^2}\right.$$

$$\left. + 2\left[\frac{-k\beta}{e + k\beta}\frac{r(e + k\beta)}{rrk} - \frac{\lambda ka\beta^2}{r(e + k\beta)^2}\frac{kr(e + k\beta)}{krar} + \frac{\lambda a\beta r(e + k\beta)}{r(e + k\beta)rar}\right]\right\}$$

$$= \frac{e + k\beta}{(re + \lambda a\beta)^2}\left\{r(e+\beta)(e+k\beta) + rk\beta^2 + \lambda^2 a\beta^2/r + 2[(-r\beta+\lambda\beta)(e+k\beta) - k\lambda\beta^2]\right\}$$

$$= \frac{(e + k\beta)}{(re + \lambda a\beta)^2}\left\{re^2 + \beta^2(rk+rk+\lambda^2 a/r-2rk+2\lambda k-2k\lambda) + \beta e(r+rk-2r+2\lambda)\right\}$$

$$= \frac{(e + k\beta)}{(re + \lambda a\beta)^2}\left[re^2 + \frac{\lambda^2 a}{r}\beta^2 + \beta e(rk - r + 2\lambda)\right]$$

$$= \frac{(e + k\beta)}{(re + \lambda a\beta)^2}[r^2e^2 + r\lambda(a + 1)\beta e + \lambda^2 a\beta^2]/r, \text{ because } rk-r+2\lambda = \lambda(a + 1)$$

$$= \frac{(e + k\beta)}{r(re + \lambda a\beta)^2}(re + \lambda a\beta)(re + \lambda\beta)$$

$$= \frac{(e + k\beta)(re + \lambda\beta)}{r(re + \lambda a\beta)}, \text{ which is (71).}$$