

CELL MEANS FORMULATIONS OF MIXED MODELS

Shayle R. Searle¹⁾

Biometrics Unit, Cornell University, Ithaca, N.Y., U.S.A.

and

Institut für Mathematik, Universität Augsburg,
D-8900 Augsburg, Federal Republic of Germany

BU-878-M

May 1985

Abstract

The cell means formulation of a mixed model has the fixed effects part of the model as cell means and the random effects part gives structure to the dispersion matrix. For balanced data, the best linear unbiased estimator (BLUE) of cell means are well known to be equal ordinary least squares estimators (OLSE). Conditions are considered under which this equality also holds for unbalanced data. Specific expressions are derived for unbalanced data from randomized complete blocks designs, of which balanced incomplete blocks are a special case.

¹⁾ Supported by a U.S. Senior Scientist Award at the University of Augsburg, from the Alexander von Humboldt-Stiftung, January - April, 1985 and 1986.

1. INTRODUCTION

a. Fixed effects models

Analysis of variance models are traditionally 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 horticultural experiment. A customary model for this is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \quad (1)$$

where y_{ijk} is the k 'th observation on treatment i and variety j , and μ is a general mean, α_i is the effect 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, with s of the ab cells containing data ($s \leq ab$), the number of parameters in (1) is $1 + a + b + s$, whereas the number of observed cell means available from the data is s . Thus there are more parameters in the model than there are cell means to estimate them from. Hence (1) is an example of 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

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

where the y_{ijk} for $k = 1, \dots, n_{ij}$ are deemed to be a random sample of n_{ij} 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, i.e., for which the n_{ij} are not all equal. Compared to (1), 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

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

which specify absence of interaction. We return to this in section 5.

Models like (1), where point and interval estimation of (and testing of hypothesis 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 \mathbf{V} of the vector of observations \mathbf{y} then has the form

$$\mathbf{V} = \sigma^2 \mathbf{I}, \quad (4)$$

\mathbf{I} being an identity matrix and σ^2 being the variance of every observation. An assumption about \mathbf{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 γ_{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),

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

that the data are from a randomized block experiment, with the β_j representing block effects. Then the β_j for $j = 1, \dots, b$, 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:

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

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

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) and (6) are together known as a mixed model. And the variances σ_β^2 and σ_e^2 of (6) are the variance components. The structure of (6) then leads to \mathbf{V} having elements that are either zero, $\sigma_\beta^2 + \sigma_e^2$, or σ_β^2 ; in general to elements that are either zero, or one of the variance components or a sum of them.

Example 1 Consider the case of 2 treatments and 3 blocks, with one observation on treatment 1 in blocks 1, 2 and 3, and on treatment 2 in just blocks 1 and 2. Then, where an element of a matrix that is zero is shown as a dot,

$$\mathbf{V} = \text{var} \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \end{bmatrix} = \begin{bmatrix} \sigma_\beta^2 + \sigma_e^2 & \cdot & \cdot & \sigma_\beta^2 & \cdot \\ \cdot & \sigma_\beta^2 + \sigma_e^2 & \cdot & \cdot & \sigma_\beta^2 \\ \cdot & \cdot & \sigma_\beta^2 + \sigma_e^2 & \cdot & \cdot \\ \sigma_\beta^2 & \cdot & \cdot & \sigma_\beta^2 + \sigma_e^2 & \cdot \\ \cdot & \sigma_\beta^2 & \cdot & \cdot & \sigma_\beta^2 + \sigma_e^2 \end{bmatrix}.$$

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. For balanced data we show why the cell means formulation always yields the same BLUEs as does an overparameterized model; we also give conditions under which this situation is true for unbalanced data; and for unbalanced randomized block designs, when the conditions are not satisfied, we give explicit expressions for the BLUEs of treatment means; and we show how these expressions simplify for balanced incomplete block designs, and are then consistent with results given in Scheffé.

2. A general formulation of cell means models

Consider the case of m factors, with the t 'th having N_t levels, for $t = 1, 2, \dots, m$. Then the k 'th observation in the cell defined by the i_t 'th level of the t 'th factor can be represented as y_{ik} for $\mathbf{i} = [i_1 \ i_2 \ \dots \ i_m]$ and with $k = 1, 2, \dots, n_{\mathbf{i}}$ where $n_{\mathbf{i}}$ is the number of observations in the \mathbf{i} 'th cell, where $\mathbf{i} = \mathbf{l}'_m, \dots, \mathbf{N}'$ for \mathbf{l}'_m being a row vector of m unities and \mathbf{N}' being a row vector $[N_1 \ N_2 \ \dots \ N_m]$. Then the cell means model (2) for y_{ijk} of the 2-factor case extends very naturally to y_{ik} :

$$y_{ik} = \mu_{\mathbf{i}} + e_{ik} \quad \text{with } E(y_{ik}) = \mu_{\mathbf{i}} .$$

For \mathbf{y} , $\boldsymbol{\mu}$ and \mathbf{e} being the vectors, respectively, of the y_{ik} , $\mu_{\mathbf{i}}$ and e_{ik} , arranged in lexicon order in each case, we write

$$\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \mathbf{e} . \tag{7}$$

Then \mathbf{X} is a direct sum of vectors $\mathbf{l}_{n_{\mathbf{i}}}$,

$$\mathbf{X} = \begin{matrix} i=\mathbf{N}' \\ (+) \mathbf{1} \\ i=\mathbf{1}'_m \end{matrix} \mathbf{1}_{n_i}, \quad (8)$$

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

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

$$\mathbf{X} = \begin{matrix} i=[2 \ 2] \\ (+) \mathbf{1} \\ i=[1 \ 1] \end{matrix} \mathbf{1}_{n_i} = \begin{bmatrix} \mathbf{1}_{n_{11}} & . & . & . \\ . & \mathbf{1}_{n_{12}} & . & . \\ . & . & \mathbf{1}_{n_{21}} & . \\ . & . & . & \mathbf{1}_{n_{22}} \end{bmatrix}.$$

The OLS estimator of μ in (7) is

$$\text{OLSE}(\mu) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \bar{\mathbf{y}} \quad (9)$$

with, from (8), the matrix $\mathbf{X}'\mathbf{X}$ being $\mathbf{D}\{n_i\}$, the diagonal matrix of the n_i , and $\mathbf{X}'\mathbf{y}$ being the vector of cell totals $y_{i\cdot}$. Hence $\text{OLSE}(\mu) = \mathbf{D}\{1/n_i\}\{y_{i\cdot}\} = \{\bar{y}_{i\cdot}\} = \bar{\mathbf{y}}$, the vector of observed cell means, as in (9).

Adapting the cell means model to models where the dispersion matrix of \mathbf{y} is other than $\sigma^2 \mathbf{I}$, i.e., for a mixed model, involves using the cell means formulation for only the cells defined by the fixed effects. For example, with randomized complete blocks as in (5), where blocks are random, the cell means model is

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad (10)$$

where, in terms of (5), the μ_i of (10) is $\mu_i = \mu + \alpha_i$ for the fixed effects part of the model and $\varepsilon_{ij} = \beta_j + e_{ij}$. The difference is, though, that we do not formally identify ε_{ij} as $\beta_j + e_{ij}$, but merely attribute some form to the dispersion matrix of the observations, namely for (7)

$$\mathbf{V} = \text{var}(\mathbf{y}) = \text{var}(\mathbf{e}). \quad (11)$$

The \mathbf{V} following (6) is an example.

Estimation of fixed effects in mixed models using OLSE takes no account of the random effects part of the mixed model. It is as if the random effects were totally ignored. An alternative, that takes the random effects into account by way of their variances, is to use BLUEs. With $\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \mathbf{e}$ and $\mathbf{V} = \text{var}(\mathbf{e})$ of (1) and (11), respectively, and assuming \mathbf{V} is positive definite, we then have

$$\text{BLUE}(\boldsymbol{\mu}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}, \quad (12)$$

where $\boldsymbol{\mu}$ is estimable because \mathbf{X} of (8) has full column rank. And the sampling variances of these estimators are

$$\text{var}[\text{OLSE}(\boldsymbol{\mu})] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad \text{and} \quad \text{var}[\text{BLUE}(\boldsymbol{\mu})] = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \quad (13)$$

Thus as an alternative to any over-parameterized model, (1) and (8) represent a cell means model formulation, and for that formulation the BLUE($\boldsymbol{\mu}$) of (12) and (13) is a suitable method of estimation. We now consider certain aspects of that procedure.

3. Estimation from balanced data

Zyskind (1967) has shown for any linear model having $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ and $\text{var}(\mathbf{y}) = \mathbf{V}$, non-negative definite, that for any estimable function of elements of $\boldsymbol{\beta}$, the BLUE and the OLSE are the same if and only if

$$\mathbf{V}\mathbf{X} = \mathbf{X}\mathbf{Q} \quad (14)$$

for some \mathbf{Q} . This condition is directly applicable to cell means models. We consider balanced data first.

For over-parameterized models and a broad class of balanced data we know (Searle, 1984) that the BLUEs of estimable functions of the fixed effects are the same as the OLSEs. Furthermore, with balanced data, all population cell means (μ_{ij} s) are estimable functions of the parameters in an over-parameterized model. Hence, with balanced data, the BLUE($\boldsymbol{\mu}$) that we obtain from the cell means formulation of these models is the same as is obtained from the BLUE of appropriate estimable functions of the fixed

effects parameters in the comparable over-parameterized model. Thus for balanced data the cell means model gives the same estimation results as does the over-parameterized model.

4. Estimation from unbalanced data

a. The general case

The general estimation procedure is

$$\text{BLUE}(\boldsymbol{\mu}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

of (12). This is the procedure for estimation of fixed effects in mixed models whether $\boldsymbol{\mu}$ is a vector cell means with \mathbf{X} as in (8), or is a vector of fixed effects parameters in an over-parameterized model. Using BLUE rather than OLSE as a method of estimation is what takes account of the random effects.

One can rightly ask: when are BLUE and OLSE the same? (It might be thought, perhaps - and incorrectly so - that OLSE is what one would use for over-parameterized models.) The answer is (14): when \mathbf{Q} exists such that $\mathbf{VX} = \mathbf{XQ}$.

5. Some fixed effects interactions omitted

a. Unbalanced data

The formulation $\mathbf{X}\boldsymbol{\mu}$ in (7), with \mathbf{X} of (8), 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 over-parameterized 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} s. This is done by using an appropriate number of equations of the form

$$\mu_{ij} - \mu_{i',j} - \mu_{ij'} - \mu_{i',j'} = 0 \quad (15)$$

for $i \neq i'$ and $j \neq j'$. This is tantamount to imposing restrictions on the elements of μ , which we now do by the representation

$$H\mu = 0 . \quad (16)$$

H is of full row rank and every element of any $H\mu$ is estimable, so that, following Searle, (1971, p. 206) the OLSE of μ for the restricted model $E(y) = X\mu$ and $H\mu = 0$ is

$$\text{OLSE}(\mu_r) = (X'X)^{-1}X'y - (X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X'y \quad (17)$$

$$= \bar{y} - (X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H\bar{y} , \quad (18)$$

after using (9). Similarly the BLUE is

$$\begin{aligned} \text{BLUE}(\mu_r) &= (X'V^{-1}X)^{-1}X'V^{-1}y \\ &\quad - (X'V^{-1}X)^{-1}H'[H(X'V^{-1}X)^{-1}H']^{-1}H(X'V^{-1}X)^{-1}X'V^{-1}y . \end{aligned} \quad (19)$$

Let us now consider when $\text{BLUE}(\mu_r)$ and $\text{OLSE}(\mu_r)$ can be equal. As a first condition we impose (14), and in so doing confine attention to situations in which $VX = XQ$ for some Q . Then (19) reduces to

$$\text{BLUE}(\mu_r) = \bar{y} - (X'V^{-1}X)^{-1}H'[H(X'V^{-1}X)^{-1}H']^{-1}H\bar{y} . \quad (20)$$

Then, on using (14) to derive $(X'V^{-1}X)^{-1} = Q(X'X)^{-1} = (X'X)^{-1}Q'$, the latter equality arising from symmetry, we find that (20) equals (18) if and only if

$$(X'X)^{-1}Q'H'[H(X'X)^{-1}Q'H']^{-1}H = (X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H , \quad (21)$$

i.e., if and only if, in using $VX = XQ$ and the full row rank property of H ,

$$\mathbf{Q}'\mathbf{H}'[\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Q}'\mathbf{H}']^{-1} = \mathbf{H}'[\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1}.$$

A sufficient condition for this equality to hold is

$$\mathbf{H}\mathbf{Q} = \mathbf{P}\mathbf{H} \quad \text{for some non-singular } \mathbf{P}. \quad (22)$$

Thus (22) is a condition for mixed models $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\mu}$ with $\text{var}(\mathbf{y}) = \mathbf{V}$, and restrictions $\mathbf{H}\boldsymbol{\mu} = \mathbf{0}$ under which with $\mathbf{V}\mathbf{X} = \mathbf{X}\mathbf{Q}$, the BLUE of $\boldsymbol{\mu}$ is the same as the OLSE. Two situations when (22) is trivially true are as follows: (i) models that include all interactions among their fixed, main effects factors, because then \mathbf{H} is null and so (22) is obviously satisfied; and (ii) models in which $\mathbf{V} = \sigma^2 \mathbf{I}$, for then \mathbf{Q} and \mathbf{P} can both be taken as $\sigma^2 \mathbf{I}$ and (22) is satisfied. In general, though, (22) is a sufficient condition, along with $\mathbf{V}\mathbf{X} = \mathbf{X}\mathbf{Q}$, for cell means models with some interactions omitted (represented by $\mathbf{H}\boldsymbol{\mu} = \mathbf{0}$) to have the BLUE and OLSE of $\boldsymbol{\mu}$ be the same.

b. Balanced data

Section 3 describes why estimation using BLUE gives the same results for a cell means model as does its over-parameterized equivalent. Nevertheless, for the case of some interactions omitted it is convincing to see that (22) is satisfied.

We begin with an example.

Example Consider 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 over-parameterized model could be

$$y_{ijklv} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{ik} + \delta_\ell + e_{ijklv}$$

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

$$y_{ijklv} = \mu_{ijk} + e_{ijk} \quad (23)$$

with restrictions of the form

$$\mu_{i \cdot k} - \mu_{i' \cdot k} - \mu_{i \cdot k'} + \mu_{i' \cdot k'} = 0 \quad (24)$$

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

$$\mu_{ijk} - \mu_{i'jk} - \mu_{ij'k} + \mu_{i'j'k} - (\mu_{ijk'} - \mu_{i'jk'} - \mu_{ij'k'} + \mu_{i'j'k'}) = 0, \quad (25)$$

for $i \neq i'$, $j \neq j'$ and $k \neq k'$. In writing (23) as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\mu} + \mathbf{e},$$

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

$$\mathbf{X} = \mathbf{I}_a * \mathbf{I}_b * \mathbf{I}_c * \mathbf{I}_d * \mathbf{I}_n \quad (26)$$

and

$$\mathbf{V} = (\mathbf{J}_a * \mathbf{J}_b * \mathbf{J}_c * \mathbf{I}_d * \mathbf{J}_n) \sigma_\delta^2 + \mathbf{I}_{abcdn} \sigma_e^2, \quad (27)$$

where $*$ represents the Kronecker product (KP) operator. Then, on defining \mathbf{T}_a as the $(a-1) \times a$ matrix

$$\mathbf{T}_a = [\mathbf{I}_{a-1} \quad -\mathbf{I}_{a-1}] \quad \text{with} \quad \mathbf{T}_a \mathbf{J}_a = \mathbf{0}, \quad (28)$$

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

$$\mathbf{H}\boldsymbol{\mu} = \mathbf{0} \quad \text{for} \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} \quad (29)$$

and

$$\mathbf{H}_1 = [\mathbf{T}_a * \mathbf{I}_b' * \mathbf{T}_c] \quad \text{and} \quad [\mathbf{H}_2 = \mathbf{T}_a * \mathbf{T}_b * \mathbf{T}_c]. \quad (30)$$

Then from (26) and (27), $\mathbf{VX} = \mathbf{XQ}$ for

$$\mathbf{Q} = (\mathbf{J}_a * \mathbf{J}_b * \mathbf{J}_c * \mathbf{I}_d * n\mathbf{I}_n) \sigma_\delta^2 + \mathbf{I}_{abcdn} \sigma_e^2. \quad (31)$$

Hence in both $H_1 Q$ and $H_2 Q$ we get, from (30) and (31), products $T_a J_a = 0$; hence $HQ = I_{abcdn} \sigma_e^2 H$, which satisfies (22).

The result just obtained for the example is true in general. X , like (26), is always a KP of matrices for which those corresponding to the main effects that define the fixed effects are identity matrices. V , like (23), is always $\sigma_e^2 I_N$ plus a weighted sum (using variance components as weights) of KPs, in which the matrices corresponding to the main effects that define fixed effects are J -matrices (with two exceptions that shall be considered shortly). Hence VX equals XQ for Q being $\sigma_e^2 I$ plus a weighted sum of KPs each of which has J -matrices corresponding to the main effects that define fixed effects. Also, H can always, as in (29) and (30), be partitioned into subsets of rows, each subset being a KP of T s and $(1')$ s. Hence HQ involves products TJ which are null, plus $\sigma_e^2 H$. Thus (22) is always satisfied.

The two exceptions are for nested random factors, and for random factors that are interactions between fixed and random factors. Each of these affect V by changing some of the J s corresponding to main effects that define fixed effects to be I s. This affects Q by replacing its term $\sigma_e^2 I$ by λI where λ is a linear combination of variance components. Thus, by the same argument as previously, HQ involves null products of the form TJ , plus λH ; and so again, (22) is satisfied.

6. 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 $i = 1, \dots, a$ and $j = 1, \dots, b$. The cell means formulation for the k 'th observation ($k = 1, 2, \dots, n_{ij}$) on treatment i in block j is

$$E(y_{ijk}) = \mu_i \quad (32)$$

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

$$\begin{aligned}
v(y_{ijk}) &= \sigma_e^2 + \sigma_\beta^2, \\
\text{cov}(y_{ijk}, y_{ijk'}) &= \sigma_\beta^2 \quad \text{for } k \neq k' = 1, 2, \dots, n_{ij}, \\
\text{cov}(y_{ijk}, y_{i'jk'}) &= \sigma_\beta^2 \quad \text{for } i \neq i', k = 1, \dots, n_{ij} \text{ and } k' = 1, \dots, n_{i'j},
\end{aligned} \tag{33}$$

and

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

The consequence of this is that for

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_a \end{bmatrix} \quad \text{with} \quad Z_i = \begin{matrix} b \\ (+)^* \mathbf{1}_{n_{ij}} \\ j=1 \end{matrix} \tag{34}$$

and

$$V = \sigma_\beta^2 ZZ' + \sigma_e^2 I_N. \tag{35}$$

In (34) the symbol $(+)^*$ represents the direct sum operator with the adaptation that every Z_i has b columns, and for every $n_{ij} \neq 0$, $\mathbf{1}_{n_{ij}}$ is always in column j of Z_i for $j = 1, \dots, b$. Thus, for example, with $b = 3$ and $n_{i1} = 4$, $n_{i2} = 0$ and $n_{i3} = 5$,

$$Z_i = \begin{bmatrix} \mathbf{1}_4 & 0 & 0 \\ 0 & 0 & \mathbf{1}_5 \end{bmatrix}.$$

Furthermore, from (32)

$$X = \begin{matrix} a \\ (+) \mathbf{1}_{n_i} \\ i=1 \end{matrix}. \tag{36}$$

Applying to (35) the general result

$$(D - CA^{-1}B)^{-1} = D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1}$$

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

$$\mathbf{V}^{-1} = [\mathbf{I} - \mathbf{Z} \left\{ \begin{smallmatrix} b \\ (+) \end{smallmatrix} \frac{\sigma_{\beta}^2}{\sigma_e^2 + n_{\cdot j} \sigma_{\beta}^2} \right\} \mathbf{Z}'] / \sigma_e^2 \quad . \quad (37)$$

Then $\mathbf{X}'\mathbf{V}^{-1}$ utilizes $\mathbf{X}'\mathbf{Z}$ which from (34) and (36) is

$$\begin{aligned} \mathbf{X}'\mathbf{Z} &= \{n_{ij}\} \quad \text{for } i = 1, \dots, a \text{ and } j = 1, \dots, b, \\ &= \{\mathbf{c}_j\} \quad \text{for } j = 1, \dots, b \text{ on defining } \mathbf{c}_j = [n_{1j} \ n_{2j} \dots n_{aj}]' \end{aligned} \quad (38)$$

Elements of $\mathbf{X}'\mathbf{Z}$ and \mathbf{c}_j do, of course, include values $n_{ij} = 0$ when they exist. Thus we find that

$$\begin{aligned} \text{BLUE}(\mu) &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ &= [\mathbf{X}'\mathbf{X} - \{\mathbf{c}_j\} \left\{ \begin{smallmatrix} b \\ (+) \end{smallmatrix} \frac{\sigma_{\beta}^2}{\sigma_e^2 + n_{\cdot j} \sigma_{\beta}^2} \right\} \{\mathbf{c}_j\}']^{-1} [\mathbf{X}'\mathbf{y} - \{\mathbf{c}_j\} \left\{ \begin{smallmatrix} b \\ (+) \end{smallmatrix} \frac{\sigma_{\beta}^2}{\sigma_e^2 + n_{\cdot j} \sigma_{\beta}^2} \right\} \mathbf{Z}'\mathbf{y}] \\ &= \left[\begin{smallmatrix} a \\ (+) \end{smallmatrix} n_{i\cdot} - \sum_{j=1}^b \frac{\sigma_{\beta}^2}{\sigma_e^2 + n_{\cdot j} \sigma_{\beta}^2} \mathbf{c}_j \mathbf{c}_j' \right]^{-1} \left[\{y_{i\cdot}\} - \sum_{j=1}^b \frac{n_{ij} \sigma_{\beta}^2 y_{\cdot j}}{\sigma_e^2 + n_{\cdot j} \sigma_{\beta}^2} \right] \quad . \end{aligned} \quad (39)$$

and for $\rho = \sigma_e^2 / \sigma_{\beta}^2$ this is

$$\begin{aligned} \text{BLUE}(\mu) &= \left\{ D\{n_{i\cdot}\}_{i=1}^a - \sum_{j=1}^b \frac{1}{\rho + n_{\cdot j}} \begin{bmatrix} n_{1j}^2 & n_{1j}n_{2j} & \dots & n_{1j}n_{aj} \\ n_{2j}n_{1j} & n_{2j}^2 & \dots & n_{2j}n_{aj} \\ \vdots & \vdots & \ddots & \vdots \\ n_{aj}n_{1j} & n_{aj}n_{2j} & \dots & n_{aj}^2 \end{bmatrix} \right\}^{-1} \\ &\quad \times \left\{ \begin{bmatrix} y_{1\cdot} \\ y_{2\cdot} \\ \vdots \\ y_{a\cdot} \end{bmatrix} - \sum_{j=1}^b \frac{y_{\cdot j}}{\rho + n_{\cdot j}} \begin{bmatrix} n_{1j} \\ n_{2j} \\ \vdots \\ n_{aj} \end{bmatrix} \right\} \quad . \end{aligned} \quad (40)$$

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

$$\text{var}[\text{BLUE}(\boldsymbol{\mu})] = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} = \sigma_e^2 \left[\sum_{i=1}^a (+)n_{i.} - \sum_{j=1}^b \frac{\sigma_\beta}{\sigma_e^2 + n_{.j}\sigma_\beta^2} \mathbf{c}_j \mathbf{c}_j' \right]^{-1}. \quad (41)$$

which is σ_e^2 multiplying the inverted matrix in (39) and (40).

Several features of $\text{BLUE}(\boldsymbol{\mu})$ are worth noting. (i) Calculating an estimate from $\text{BLUE}(\boldsymbol{\mu})$ requires estimates of σ_β^2 and σ_e^2 , or at least of $\rho = \sigma_e^2/\sigma_\beta^2$. (ii) Using any one of a wide class of estimators of σ_e^2 and σ_β^2 in place of σ_e^2 and σ_β^2 in $\text{BLUE}(\boldsymbol{\mu})$ gives, as Kackar and Harville (1981, 1984) have shown, an estimate that is unbiased. (iii) When $n_{ij} = n \forall i$ and j (corresponding to balanced data), $\text{BLUE}(\mu_i)$ from (40) is $\bar{y}_{i..}$, as one would expect. (iv) When $\sigma_\beta^2 = 0$ (corresponding to the model $y_{ijk} = \mu + \alpha_i + e_{ijk}$), we also have $\text{BLUE}(\mu_i)$ from (40) as $\bar{y}_{i..}$. (v) When $\sigma_\beta^2 \rightarrow \infty$ (corresponding to a fixed effects no-interaction model), $\text{BLUE}(\mu_i - \mu_h)$ obtained from (40) is the same as obtained directly (e.g., Searle, 1971, Chapter 7). Note in that case, though, that the matrix to be inverted in (40) is then singular - because its row sums are zero: $n_{i.} - \sum_{j=1}^b (1/n_{.j}) n_{ij} \sum_{i=1}^a n_{ij} = 0$. Instead of the regular inverse, a generalized inverse must be used.

An extension of these results would be to include in the variance-covariance structure of (33) a covariance among observations in the same cell so that $v(y_{ijk}) = \sigma_e^2 + \sigma_\beta^2$ of (33) would become $\sigma_e^2 + \sigma_\beta^2 + \sigma_\gamma^2$; and $\text{cov}(y_{ijk}, y_{ijk'}) = \sigma_\beta^2$ for $k \neq k' = 1, \dots, n_{ij}$ would become $\sigma_\beta^2 + \sigma_\gamma^2$. The other terms in (33) would remain unaltered.

7. 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 (39).

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

Treatment	B l o c k						$n_{i.} = r$
	1	2	3	4	5	6	
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 within-block treatment pairs that contain a particular treatment: $\lambda(a-1) = r(k-1) = 3$. (42)

To simplify (34) 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 (39) is

$$\{\hat{\mu}_i\}_{i=1}^{i=a} = [rI_a - \frac{1}{\rho+k} \sum_{j=1}^b c_j c_j']^{-1} \{y_{i.} - \frac{1}{\rho+k} \sum_{j=1}^b n_{ij} y_{.j}\}_{i=1}^{i=a}. \quad (43)$$

The notation used in (43) is that $\hat{\mu}_i \equiv \text{BLUE}(\mu_i)$ an element of $\text{BLUE}(\mu)$

of (39), and the notation on the left-hand side indicates a column vector of $\hat{\mu}_i$ s, for $i = 1, 2, \dots, a$; and analogously so, on the right-hand side, also. (43) also involves $\rho = \sigma_e^2 / \sigma_\beta^2$.

Simplifying (43) involves two summations, the nature of which are best developed from (40). The first is

$$S = \frac{1}{\rho+k} \sum_{j=1}^b c_j c_j' = \frac{1}{\rho+k} \sum_{j=1}^b \begin{bmatrix} n_{1j}^2 & n_{1j}n_{2j} & \dots & n_{1j}n_{aj} \\ n_{2j}n_{1j} & n_{2j}^2 & \dots & n_{2j}n_{aj} \\ \vdots & \vdots & \ddots & \vdots \\ n_{aj}n_{1j} & n_{aj}n_{2j} & \dots & n_{aj}^2 \end{bmatrix}$$

and from the definitions of λ and r this is seen to be

$$S = \frac{1}{\rho+k} \begin{bmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \lambda & \dots & r \end{bmatrix} = \frac{1}{\rho+k} [(r-\lambda)I_a + \lambda J_a] \quad (44)$$

Using this in the first term of (43) gives that term as

$$\begin{aligned} [rI_a - \frac{r-\lambda}{\rho+k}I_a - \frac{\lambda}{\rho+k}J_a]^{-1} &= (\rho+k)[(r\rho + rk - r + \lambda)I_a - \lambda J_a]^{-1} \\ &= (\rho+k)[(r\rho + \lambda a)I_a - \lambda J_a]^{-1}, \text{ from (42),} \\ &= \frac{\rho+k}{r\rho + \lambda a} (I_a + \frac{\lambda}{r\rho} J_a) \quad (45) \end{aligned}$$

The second summation for (43) is

$$\begin{aligned} t_i &= \frac{1}{\rho+k} \sum_{j=1}^b n_{ij} y_{.j} = \frac{1}{\rho+k} \sum_{j=1}^b n_{ij} k \bar{y}_{.j} = \frac{kr}{\rho+k} \sum_{j=1}^b n_{ij} \bar{y}_{.j} / r \\ &= \frac{kr}{\rho+k} \bar{y}_{i(j)} \quad (46) \end{aligned}$$

where

$$\bar{y}_{i(j)} = \sum_{j=1}^b n_{ij} \bar{y}_{.j} / r = \text{mean of block means } \bar{y}_{.j} \text{ for the blocks that contain treatment } i. \quad (47)$$

Using this in the second term of (43), along with (45) for the first gives (43) as

$$\{\hat{\mu}_i\}_{i=1}^a = \frac{\rho+k}{r\rho+\lambda a} \left(\mathbf{I}_a + \frac{\lambda}{r\rho} \mathbf{J}_a \right) \{y_{i.} - \frac{kr}{\rho+k} \bar{y}_{i(j)}\}_{i=1}^{i=a}.$$

Hence, using $\sum_{i=1}^a \bar{y}_{i(j)} = \sum_{j=1}^b n_{.j} \bar{y}_{.j} / r = y_{..} / r$ derived from (47),

$$\begin{aligned} \hat{\mu}_i &= \frac{\rho+k}{r\rho+\lambda a} \left[y_{i.} + \frac{\lambda}{r\rho} y_{..} - \frac{kr}{\rho+k} \bar{y}_{i(j)} - \frac{\lambda}{r\rho} \frac{kr}{\rho+k} y_{..} / r \right] \\ &= \frac{r(\rho+k)}{r\rho+\lambda a} \left[\bar{y}_{i.} - \frac{k}{\rho+k} \bar{y}_{i(j)} + \frac{\lambda a}{r(\rho+k)} \bar{y}_{..} \right] \end{aligned} \quad (48)$$

$$= \bar{y}_{..} + \frac{r\rho}{r\rho+\lambda a} [\bar{y}_{i(j)} - \bar{y}_{..}] + \frac{r(\rho+k)}{r\rho+\lambda a} [\bar{y}_{i.} - \bar{y}_{i(j)}]. \quad (49)$$

As shown in the Appendix, this result is consistent with Scheffé (1959, pp 161-175).

Furthermore, from (41) and (45), using

$$\begin{aligned} \text{var}(\hat{\mu}) &= \sigma_e^2 \left[r\mathbf{I}_a - \frac{r-\lambda}{\rho+k} \mathbf{I}_a - \frac{\lambda}{\rho+k} \mathbf{J}_a \right]^{-1} \\ &= \frac{\sigma_e^2 (\rho+k)}{(r\rho+\lambda a)} \left(\mathbf{I}_a + \frac{\lambda}{r\rho} \mathbf{J}_a \right). \end{aligned}$$

Hence

$$v(\hat{\mu}_i) = \frac{r\rho+\lambda}{r(r\rho+\lambda a)} (\sigma_e^2 + k\sigma_\beta^2) \quad (50)$$

and

$$\text{cov}(\hat{\mu}_i, \hat{\mu}_h) = \frac{\lambda}{r(r\rho+\lambda a)} (\sigma_e^2 + k\sigma_\beta^2) \quad \text{for } i \neq h. \quad (51)$$

Thus the estimated difference between treatments i and h is, from (48)

$$\hat{\mu}_i - \hat{\mu}_h = \frac{r(\rho+k)}{r\rho+\lambda a} (\bar{y}_{i.} - \bar{y}_{h.}) - \frac{kr}{r\rho+\lambda a} [\bar{y}_{i(j)} - \bar{y}_{h(j)}]$$

with, from (50) and (51)

$$v(\hat{\mu}_i - \hat{\mu}_n) = \frac{2(\sigma_e^2 + k\sigma_\beta^2)}{r(rp+ka)} (rp+\lambda+\lambda) = \frac{2(rp+2\lambda)}{r(rp+ka)} (\sigma_e^2 + k\sigma_\beta^2).$$

REFERENCES

- Hocking, R.R. and Speed, F.M. (1975). *A full rank analysis of some linear model problems*. J. of the American Statistical Assoc., 70, 706-712.
- Kackar, R.N. and Harville, D.A. (1981). *Unbiasedness of two-stage estimation and prediction procedures for mixed linear models*. Communications in Statistics - Theory and Methods, Ser. A., 10, 1249-1261.
- Kackar, R.N. and Harville, D.A. (1984). *Approximations for standard errors of estimators of fixed and random effects in mixed linear models*. J. of the American Statistical Assoc., 79, 853-861.
- Scheffé, H. (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. (1984). *Best linear unbiased estimation in mixed models of the analysis of variance*. Technical Report BU-864-M, Biometrics Unit, Cornell University.
- 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 Assoc. 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 a 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 displaying equivalent notation.

	<u>Scheffé</u>		<u>This paper</u>
p. 161 :	# of treatments	I	$a = t$
	# of blocks	J	b
	# of replications	r	r
	block size	k	k
p. 162 :	# of occurrences of	$K_{ij} = 0 \text{ or } 1$	n_{ij}
(line 3 up)	treatment i in block j		
p. 164 :	i'th treatment total	g_i	$y_{i.}$
(lines 8-9)	j'th block total	h_j	$y_{.j}$
	i'th adjusted treatment total		
(after 5.2.9):	$G_i = g_i - k^{-1} \sum_j K_{ij} h_j$		$y_{i.} - \sum_j n_{ij} \bar{y}_{.j}$ $= y_{i.} - r \bar{y}_{i(j)}$
	sum of block totals in which treatment i occurs	T_i	
(5.2.10):	$T_i = \sum_j n_{ij} h_j$		$kr \bar{y}_{i(j)}$
p. 166 :	efficiency factor	δ	
(5.2.17)	$\delta = \frac{rk - r + \lambda}{rk} = \frac{(k-1)I}{k(I-1)}$		$\frac{\lambda a}{rk} = \frac{(k-1)a}{k(a-1)}$
p. 165 :	$r\delta \hat{\alpha}_i = G_i$		$\frac{y_{i.} - r \bar{y}_{i(j)}}{r\delta}$
(last line)	$\hat{\alpha}_i = G_i / r\delta$		$= rk(\bar{y}_{i.} - \bar{y}_{i(j)}) / \lambda a$

p. 172 : $\hat{\alpha}'_i = \frac{T_i - rJ^{-1}\sum_j h_j}{r - \lambda}$ $\frac{kr\bar{y}_i(j) - r\sum_j y_{..j}/b}{r - \lambda}$
 (5.2.33) $= \frac{kr\bar{y}_i(j) - ry_{..}/b}{r - \lambda}$
 $= \frac{kr(\bar{y}_i(j) - \bar{y}_{..})}{r - \lambda}$

(5.2.32b): $\sigma_f^2 = k^2 \sigma_B^2 + k\sigma_e^2$ $k(e + k\beta)$, with
 $e \equiv \sigma_e^2$ and $\beta \equiv \sigma_\beta^2$

(line 5 up): $\psi = \sum_i c_i \alpha_i$ $\sum_i c_i = 0$

$$\hat{\psi}' = \sum_i c_i \hat{\alpha}'_i$$

p. 174 : $w = r\delta/\sigma_e^2$ $\lambda a/ke$
 (5.2.41) $w' = (r - \lambda)/\sigma_f^2$ $(r - \lambda)/k(e + k\beta)$

p. 175 : $\psi^* = \frac{w\hat{\psi} + w'\hat{\psi}'}{w + w'}$
 (5.2.42)

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

$$\hat{\mu}' = k \sum_j h_j / k^2 J = \sum_j y_{.j} / ka = y_{..} .$$

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

$$\begin{aligned} \mu_i^* &= \bar{y} + \frac{\frac{\lambda a}{ke} \frac{kr}{\lambda a} (\bar{y}_{i.} - \bar{y}_{i(j)}) + \frac{r - \lambda}{k(e + k\beta)} \frac{kr(\bar{y}_{i(j)} - \bar{y}_{..})}{r - \lambda}}{\frac{\lambda a}{ke} + \frac{r - \lambda}{k(e + k\beta)}} \\ &= \bar{y}_{..} + \frac{r[(\bar{y}_{i.} - \bar{y}_{i(j)})/e + (\bar{y}_{i(j)} - \bar{y}_{..})/(e + k\beta)]}{[\lambda a(e + k\beta) + (r - \lambda)e]/ke(e + k\beta)} \\ &= \bar{y}_{..} + \frac{rk[(e + k\beta)(\bar{y}_{i.} - \bar{y}_{i(j)}) + e(\bar{y}_{i(j)} - \bar{y}_{..})]}{\lambda ak\beta + rke}, \\ &\quad \text{because } \lambda a + r - \lambda = rk \\ &= \bar{y}_{..} + \frac{r(e + k\beta)}{re + a\lambda\beta} [\bar{y}_{i.} - \frac{k\beta}{e + k\beta} \bar{y}_{i(j)} - \frac{e}{e + k\beta} \bar{y}_{..}] \\ &= \frac{r(e + k\beta)}{re + a\lambda\beta} [\bar{y}_{i.} - \frac{k\beta}{e + k\beta} \bar{y}_{i(j)} + \frac{a\lambda\beta}{r(e + k\beta)} \bar{y}_{..}] \\ &= \frac{r(\rho + k)}{r\rho + \lambda a} [\bar{y}_{i.} - \frac{k}{\rho + k} \bar{y}_{i(j)} + \frac{\lambda a}{r(\rho + k)} \bar{y}_{..}] = \hat{\mu}_i \text{ of (48).} \end{aligned}$$

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

From (48)

$$\begin{aligned} v(\hat{\mu}_i) &= v\left\{\frac{r(e + k\beta)}{re + \lambda a\beta} \left[\bar{y}_{i.} - \frac{k\beta}{e + k\beta} \bar{y}_{i(j)} + \frac{\lambda a\beta}{r(e + k\beta)} \bar{y}_{..}\right]\right\} \\ &= \frac{r^2(e + k\beta)^2}{(re + \lambda a\beta)^2} \left\{v(\bar{y}_{i.}) + \frac{k^2\beta^2}{(e + k\beta)^2} v(\bar{y}_{i(j)}) + \frac{\lambda^2 a^2 \beta^2}{r^2(e + k\beta)^2} v(\bar{y}_{..})\right. \\ &\quad \left.+ 2\left[-\frac{k\beta}{e + k\beta} \text{cov}(\bar{y}_{i.}, \bar{y}_{i(j)}) - \frac{k\beta}{e + k\beta} \frac{\lambda a\beta}{r(e + k\beta)} \text{cov}(\bar{y}_{i(j)}, \bar{y}_{..})\right.\right. \\ &\quad \left.\left.+ \frac{\lambda a\beta}{r(e + k\beta)} \text{cov}(\bar{y}_{i.}, \bar{y}_{..})\right]\right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{r^2 (e + k\beta)^2}{(re + \lambda a\beta)^2} \left\{ \frac{r(e + \beta)}{r^2} + \frac{k^2 \beta^2 rk(e + k\beta)}{(e + k\beta)^2 r^2 k^2} + \frac{\lambda^2 a^2 \beta^2 ar(e + k\beta)}{r^2 (e + k\beta)^2 a^2 r^2} \right. \\
&\quad \left. + 2 \left[\frac{-k\beta}{e + k\beta} \frac{r(e + k\beta)}{rrk} - \frac{\lambda ka\beta^2}{r(e + k\beta)^2} \frac{kr(e + k\beta)}{krar} + \frac{\lambda a\beta r(e + k\beta)}{r(e + k\beta)rar} \right] \right\} \\
&= \frac{e + k\beta}{(re + \lambda a\beta)^2} \{ r(e+\beta)(e+k\beta) + rk\beta^2 + \lambda^2 a\beta^2 / r + 2[(-r\beta+\lambda\beta)(e+k\beta) - k\lambda\beta^2] \} \\
&= \frac{(e + k\beta)}{(re + \lambda a\beta)^2} \{ re^2 + \beta^2 (rk+rk+\lambda^2 a/r-2rk+2\lambda k-2k\lambda) + \beta e(r+rk-2r+2\lambda) \} \\
&= \frac{(e + k\beta)}{(re + \lambda a\beta)^2} \left[re^2 + \frac{\lambda^2 a}{r} \beta^2 + \beta e(rk - r + 2\lambda) \right] \\
&= \frac{(e + k\beta)}{(re + \lambda a\beta)^2} [r^2 e^2 + r\lambda(a + 1)\beta e + \lambda^2 a\beta^2] / r, \text{ because } rk-r+2\lambda = \lambda(a + 1) \\
&= \frac{(e + k\beta)}{r(re + \lambda a\beta)^2} (re + \lambda a\beta)(re + \lambda\beta) \\
&= \frac{(e + k\beta)(re + \lambda\beta)}{r(re + \lambda a\beta)} \\
&= \frac{(r\rho+\lambda)}{r(r\rho+\lambda a)} (\sigma_e^2 + k\sigma_\beta^2), \text{ which is (50).}
\end{aligned}$$