TRAINING & RESEARCH

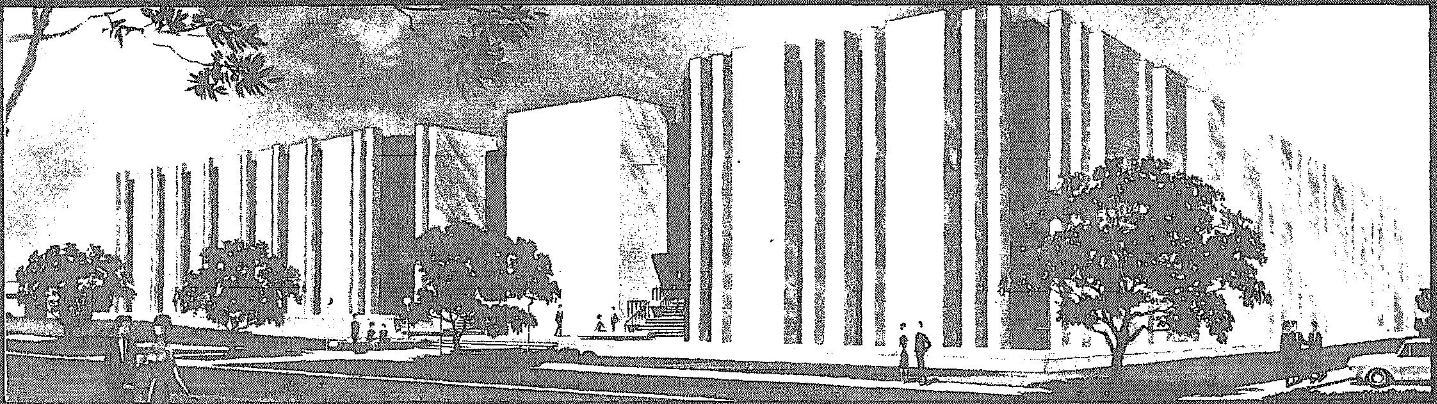
N71-35669

MAXIMUM LIKELIHOOD ANALYSIS OF BALANCED INCOMPLETE BLOCK MODELS

by

Michael Henry Kutner
Texas A&M University

GRADUATE INSTITUTE OF STATISTICS



TEXAS A&M UNIVERSITY · COLLEGE STATION

N71-35669

MAXIMUM LIKELIHOOD ANALYSIS OF BALANCED INCOMPLETE BLOCK MODELS

by

Michael Henry Kutner
Texas A&M University

Institute of Statistics
Texas A&M University

Technical Report #7
National Aeronautics and Space Administration
Research Grant NGR 44-001-095

July 1971

CHAPTER I

INTRODUCTION

1.1 Preliminaries

Presently a number of methods exist for obtaining point estimates of variance components with unbalanced data. Henderson [12] gives three unbiased methods for estimating variance components. Searle [26] critically reviews and reformulates, using matrix theory, Henderson's methods, as well as presenting a modified fourth method. Hartley and Rao [10] describe a procedure for obtaining maximum likelihood estimates of variance components and fixed effects in a mixed analysis of variance model. LaMotte [20] considers a class of estimators of variance components which are closely akin to maximum likelihood estimates. Symmetric sums is a method developed by Koch [17, 18], and Townsend [30] has developed a method based on best quadratic unbiased estimation. For a more complete listing see Crump [6] and Searle [27]. Any hope for a uniformly best estimation technique for unbalanced data appears to be futile. Evidence of this is found in the few comparative studies which have been made (Bush and Anderson [4], Anderson and Crump [1], as examples).

Citations follow the style of *The Journal of the American Statistical Association*.

Because maximum likelihood estimates yield asymptotically optimal properties and because there is evidence to indicate that maximum likelihood estimates also frequently have good small-sample properties (e.g., Klotz, et al. [17]), it is the intent of this dissertation to examine closely the likelihood equations. Unfortunately the likelihood equations generally require a numerical method of solution; therefore, it is extremely important to identify structure of the likelihood equations which might allow a less cumbersome computational task of maximizing the likelihood function. In Chapter III a structural form of the likelihood equations is developed. In order to investigate its computational efficiency and to pursue the investigation of small-sample properties of maximum likelihood estimates, it was decided to concentrate on the analysis of balanced incomplete block designs. There were several reasons for choosing the balanced incomplete block design. It was felt to be one of the easiest of the unbalanced designs to characterize and, further, that valuable information would be gained for later study of other unbalanced designs. Also, beyond unbiasedness very little is known about the properties of the estimates which are presently used in balanced incomplete block analyses.

Chapter II treats the matrix analysis of the fixed effects model primarily for establishing notation and completeness.

Chapter III is concerned with the analysis of the mixed model. Here, the structure of the likelihood equations is employed. This chapter completely spells out the relationship of maximum likelihood to analysis of variance. The sometimes misunderstood recovery of interblock information is considered in great detail. Chapter IV analyses the random model. Again, the structure developed in Chapter III is used to solve numerically the likelihood equations. Also included in this chapter is a numerical comparison of the maximum likelihood estimates with Henderson's fitting constants estimates and estimates based on a minimum variance combination of the sufficient statistics. Mention is made of the relationship between Rao's [25] estimation procedure and LaMotte's [21] estimation procedure. Chapter V is devoted to a summarization and recommendations for areas of future research.

CHAPTER II

MATRIX ANALYSIS OF THE FIXED EFFECTS MODEL (MODEL I)

2.1 Definition of the Model and Notation

Consider the model where each observation may be expressed as

$$y_{ijm} = \mu + \tau_i + \beta_j + e_{ijm} \tag{2.1}$$

$$i = 1, \ldots, t;\ j = 1, \ldots, b;\ \text{and } m = n_{ij} = \begin{cases} 1 \text{ if the ith treatment occurs in the jth block} \\ 0 \text{ otherwise.} \end{cases}$$

It should be noted that if $n_{ij} = 0$ there is no observation. Also, define $\sum\sum_{ij} n_{ij} = n$ (total number of observations). Further, each of the b blocks contains k experimental units, each of the t treatments is replicated r times, and each treatment occurs λ times in the same block with every other treatment. Equation (2.1) is the mathematical model for the balanced incomplete block design (BIB).

If we assume $E(y_{ijm}) = \mu + \tau_i + \beta_j$ and $e_{ijm} \sim NID(0, \sigma^2)$, then we have defined what is generally referred to as Eisenhart's Model I [7]. (Notationally, the capital letter E will always denote the expected value operator.)

In matrix notation we write equation (2.1) as

$$y = \mu J + X_1\beta + X_2\tau + e \tag{2.2}$$

where y is an $(n \times 1)$ vector of observations;

μ is a scalar constant;

J is an $(n \times 1)$ vector of ones;

X_1 is an $(n \times b)$ matrix of known numbers;

β is a $(b \times 1)$ vector of unknown constants;

X_2 is an $(n \times t)$ matrix of known numbers;

τ is an $(t \times 1)$ vector of unknown constants;

e is an $(n \times 1)$ vector of independent variables from $N(0, \sigma^2)$.

This description can be summarized as $y \sim N(\mu J + X_1\beta + X_2\tau, \Sigma)$

with $$\Sigma = E[y - E(y)][y - E(y)]' = \sigma^2 I ,$$

where the prime denotes transpose and I is an $(n \times n)$ identity matrix. As above we will adopt the convention of using Greek letters to represent parameters or matrices which involve parameters, with the exception being that we will not alter 'standard' notation.

The following is a list of helpful results:

$$\sum_j n_{ij} = r, \sum_i n_{ij} = k, \sum_i \sum_j n_{ij} = n = bk = rt , \tag{2.3}$$

$$\sum_j n_{ij} n_{i'j} (i \neq i') = \lambda = r(k - 1)/(t - 1) , \tag{2.4}$$

or $\lambda t + r - \lambda = rk$,

and

$$bk(k-1) = \lambda t(t-1),\ k(b-r) = r(t-k) = (r-\lambda)(t-1). \quad (2.5)$$

We define $f = bk - b - t + 1$, where $b \geq t$, $r \geq k$, $t > k$, and $b > r$. We will adopt the 'dot' notation to indicate a variable has been summed over its index or indices, e.g., $y_{i..} = \sum_j \sum_m y_{ijm}$.

For the matrix model (2.2), let $X_2'X_1 = N$, which is called the incidence matrix and define $A' = X_2' - k^{-1} N X_1'$. The following relationships hold for the matrix model:

$$X_1'X_1 = kI,\ X_2'X_2 = rI\ , \quad (2.6)$$

$J'X_1 = kJ'$, $J'X_1' = J'$, $J'X_2 = rJ'$, $J'X_2' = J'$ (where J' denotes a row vector of ones of proper dimension) , (2.7)

$$NN' = (r-\lambda)\ I + \lambda JJ'\ , \quad (2.8)$$

$$A'X_2 = \lambda k^{-1}(tI - JJ'),\ A'X_1 = 0\ ,\ A'J = 0\ , \quad (2.9)$$

$$tr(NN')^m = (rk)^m + (r-\lambda)^m(t-1)\ , \quad (2.10)$$

where tr denotes the trace operator and m is a positive integer.

Relationship (2.10) can be established directly by using (2.8) and (2.4) as follows:

$$(NN')^m = [(r - \lambda)\ I + \lambda JJ']^m$$

$$= \sum_{\ell=0}^{m} \binom{m}{\ell} (r - \lambda)^{m-\ell}\ \lambda^{\ell}\ I^{m-\ell}(JJ')^{\ell}$$

$$= (r - \lambda)^m\ I + \sum_{\ell=1}^{m} \binom{m}{\ell} (r - \ell)^{m-\ell}\ \lambda^{\ell}\ t^{\ell-1}(JJ')\ ,$$

since $(JJ') = t^{\ell-1}(JJ')$. Now

$$tr(NN')^m = (r - \lambda)^m\ t + \sum_{\ell=1}^{m} \binom{m}{\ell} (r - \lambda)^{m-\ell}(\lambda t)^{\ell}$$

$$= (r - \lambda)^m(t - 1) + [r - \lambda(t - 1)]^m$$

$$= (r - \lambda)^m(t - 1) + (rk)^m\ .$$

2.2 Solution of the Normal Equations

Form the logarithm of the likelihood function

$$\ell nL = -\frac{n}{2}\ \ell n(2\pi) - \frac{n}{2}\ \ell n\ \sigma^2$$

$$- \frac{1}{2\sigma^2}\ (y - \mu J - X_1\beta - X_2\tau)'(y - \mu J - X_1\beta - X_2\tau) \qquad (2.11)$$

and differentiate (2.11) with respect to μ, β, and τ. The resulting equations when set equal to zero constitute what are called the normal equations (N.E.). In matrix notation the N.E. are given by

$$J'J\mu + J'X_1\beta + J'X_2\tau = J'y \ , \tag{2.12}$$

$$X_1'J\mu + X_1'X_1\beta + X_1'X_2\tau = X_1'y \ , \tag{2.13}$$

and

$$X_2'J\mu + X_2'X_1\beta + X_2'X_2\tau = X_2'y \ . \tag{2.14}$$

To solve for τ, multiply equation (2.13) by $k^{-1}N$ and subtract this from equation (2.14), yielding

$$(X_2' - k^{-1}NX_1')J\mu + (X_2' - k^{-1}NX_1')X_2\tau +$$

$$+ (X_2' - k^{-1}NX_1')X_1\beta = (X_2' - k^{-1}NX_1')y$$

or

$$A'(J\mu + X_2\tau + X_1\beta] = A'y \ . \tag{2.15}$$

Since $A'J = A'X_1 = 0$, we have

$$A'X_2\tau = A'y \ . \tag{2.16}$$

Simplifying (2.16) we obtain

$$(X_2'X_2 - k^{-1}NN')\tau = (X_2' - k^{-1}NX_1')y$$

or

$$r\tau - k^{-1}[(r - \lambda)I + \lambda JJ']\tau = (X_2' - k^{-1}NX_1')y \qquad (2.17)$$

Imposing $J'\tau = 0$ yields

$$\hat{\tau} = \frac{k}{\lambda t} A'y \; . \qquad (2.18)$$

Imposing $J'\tau = J'\beta = 0$ in equation (2.12), we get

$$\hat{\mu} = \frac{J'y}{n} \; . \qquad (2.19)$$

Now from equation (2.13)

$$\hat{\beta} = k^{-1}[X_1'y - kJ\hat{\mu} - N'\hat{\tau}] \qquad (2.20)$$

Hence, (2.18), (2.19), and (2.20) are solutions to the N.E.

To estimate σ^2, differentiate (2.11) with respect to σ^2 and set the resulting equation equal to zero yielding

$$\tilde{\sigma}^2 = \frac{1}{n} (y - \hat{\mu}J - X_2\hat{\tau} - X_1\hat{\beta})'(y - \hat{\mu}J - X_2\hat{\tau} - X_1\hat{\beta})$$

$$= \frac{1}{n} y'[I - k^{-1}X_1X_1' - \frac{k}{\lambda t} AA']'[I - k^{-1}X_1X_1' - \frac{k}{\lambda t} AA']y \; . \tag{2.21}$$

2.3 Tests of Hypotheses

Suppose we write model (2.2) in partitioned form as

$$y = X\gamma + e \tag{2.22}$$

where $X = (J|X_1|X_2)$ and $\gamma' = (\mu|\beta|\tau)$. Now consider testing the hypothesis

$$H_o: K'\gamma = c \tag{2.23}$$

where K' has rows k_i', $i = 1, \ldots, t + b + 1$, c is a vector of constants, and $k_i'\gamma$ is estimable. The test statistic for testing (2.23) is given by

$$F = \frac{(K'\hat{\gamma}-c)'[K'(X'X)^{-}K]^{-1}(K'\hat{\gamma}-c)}{r(K)\ \hat{\sigma}^2} \tag{2.24}$$

which is a noncentral F distribution with $r(K)$ degrees of freedom in the numerator and $[n - r(K)]$ degrees of freedom in the denominator and with noncentrality parameter

$\lambda = \frac{1}{2\sigma^2} (K'\gamma - c)'(K'(X'X)^{-}K)^{-1}(K'\gamma - c)$. Note that $\hat{\gamma}$ is any solution to the N.E., $(X'X)^{-}$ denotes any generalized inverse of $X'X$, $r(\cdot)$ denotes rank, and

$$\hat{\sigma}^2 = \frac{(y-X\hat{\gamma})'(y-X\hat{\gamma})}{n-r(X)} = [y - XK(K'K)^{-1}c]'(I - X(X'X)^{-}X')$$

$$[y - XK(K'K)^{-1}c]/[n - r(X)]$$

$$= y'[I - X(X'X)^{-}X']y/[n - r(X)] \ .$$

For the general development see Searle [28, Chapter 5].

In particular, consider testing the equality of treatments, i.e.,

$$H_o\colon \tau_1 = \tau_2 = \dots = \tau_t \ . \tag{2.25}$$

From (2.23) let

$$K' = [0|0|J, - I] \quad \gamma' = (\mu|\beta|\tau_1, \tau_2, \dots, \tau_t), \text{ and } c = 0, \tag{2.26}$$

which is equivalent to (2.25). It can be shown that by testing (2.26) using (2.24) we obtain the same F statistic under H_o as in analysis of variance (AOV) (see Graybill, [9, p. 312]). Here, $r(K) = t - 1$ and $r(X) = t + b - 1$.

CHAPTER III

MATRIX ANALYSIS OF THE MIXED MODEL (MODEL III)

3.1 Definition of the Model

Recall equation (2.2) and set $\beta = B$, yielding

$$y = \mu J + X_1 B + X_2 \tau + e$$

where

$$B \sim NID(0, \sigma_B^2) \quad .$$

Alternately, we can write

$$y \sim N(\mu J + X_2 \tau, \Sigma)$$

where

$$\Sigma = \sigma^2 I + \sigma_B^2 X_1 X_1' \quad .$$

This description is commonly regarded as Eisenhart's Model III (treatments fixed, blocks random) [7].

3.2 Identification of Sufficient Statistics

In order to identify sufficient statistics, the following relationship is useful:

$$(I + CG)^{-1} = I - C(I + GC)^{-1}G \quad . \tag{3.1}$$

(See LaMotte [20] for proof.) To identify a minimal set of sufficient statistics, we write the exponent of the likelihood

$$(y - \mu J - X_2\tau)' \, \Sigma^{-1}(y - \mu J - X_2\tau)$$

$$= (y - \mu J - X_2\tau)' \frac{1}{\sigma^2} \left(I - \frac{\sigma_B^2}{\sigma^2 + k\sigma_B^2} X_1X_1'\right)(y - \mu J - X_2\tau) \tag{3.2}$$

utilizing (3.1) to obtain the inverse of Σ. In the exponent we expand (3.2) yielding

$$\frac{1}{\sigma^2} y'y - \frac{\sigma_B^2}{\sigma^2(\sigma^2+k\sigma_B^2)} y'X_1X_1'y - \frac{2\mu}{\sigma^2+k\sigma_B^2} J'y - \frac{2}{\sigma^2} \tau'[X_2' - \frac{\sigma_B^2 N X_1'}{\sigma^2+k\sigma_B^2}]y$$

$$+ \mu^2 J'\Sigma^{-1}J + \tau'X_2'\Sigma^{-1}X_2\tau + 2\mu J'\Sigma^{-1}X_2\tau \quad . \tag{3.3}$$

Since the last three terms in (3.3) do not involve the observations they can be ignored with regard to identifying sufficient statistics via the factorization theorem. (See Mood and Graybill [23, pp. 168-169].) Rewriting the first four terms of (3.3), we have in summation notation

$$\frac{1}{\sigma^2} \sum_{ijm} y_{ijm}^2 - \frac{\sigma_B^2}{\sigma^2(\sigma^2+k\sigma_B^2)} \sum_{j=1}^{b} y_{.j.}^2 - \frac{2\mu}{(\sigma^2+k\sigma_B^2)} y_{...}$$

$$- \frac{2}{\sigma^2} [\sum_{i=1}^{t} \tau_i y_{i..} - \frac{\sigma_B^2}{\sigma^2+k\sigma_B^2} \sum_{i=1}^{t} \tau_i \sum_{j=1}^{b} n_{ij} y_{.j.}] . \qquad (3.4)$$

For b > t we identify, after eliminating redundancies, the statistics

$$\sum_{ijm} y_{ijm}^2, \sum_j y_{.j.}^2, y_{i..} \; (i = 1, \ldots, t-1)$$

and $$\sum_j n_{ij} y_{.j.} \; (i = 1, \ldots, t) \qquad (3.5)$$

which agree with those previously reported by Hultquist and Graybill [14]. By using the factorization theorem we have a direct procedure for identifying sufficient statistics. However, we do require a closed form solution to the inverse of the variance-covariance matrix. When b = t, $\sum_j n_{ij} y_{.j.}$ is a set of t relations involving b $y_{.j.}$'s and because b = t we can determine each $y_{.j.}$ (j = 1, ..., b). Therefore, since each $y_{.j.}$ is known $\sum_j n_{ij}y_{.j.}$ (i = 1, ..., t) and $\sum_j y^2_{.j.}$ in (3.5) can be reduced to $y_{.j.}$ (j = 1, ..., b) when b = t.

As an alternate to the above set of sufficient statistics for the case b > t we may consider the set

$$y_{i..} \ (i = 1, \ldots, t-1) \ , \quad \sum_j n_{ij}y_{.j.} \ (i = 1, \ldots, t)$$

and

$$s_2 = y'[\frac{1}{k} X_1X_1' - \frac{1}{k(r-\lambda)} X_1N'NX_1' + \frac{\lambda t \ JJ'}{bk(r-\lambda)}] \ y \quad ,$$

$$s_6 = y'[-\frac{1}{k} X_1X_1' - \frac{k}{\lambda t} X_2X_2' + I + \frac{1}{\lambda t} (X_1N'X_2' + X_2NX_1')$$

$$- \frac{1}{k\lambda t} X_1N'NX_1'] \ y \quad . \tag{3.6}$$

For convenience and for later reference we display the statistics s_2 and s_6 in tabular form (Table I) to illustrate their relation to quadratic forms. We also display statistics s_3, s_4, and s_5 in a similar way. These statistics will be useful in later analysis and will also appear in the analysis of the random model in Chapter V. The notation s_i, i = 2, ..., 6, is made to conform with that given later for the random model.

TABLE I. RELATIONSHIP OF AN s_i TO ITS QUADRATIC FORM

	X_1X_1'	X_2X_2'	I	$X_1N'X_2' + X_2NX_1'$	$X_1N'NX_1'$	JJ'
s_2	$\frac{1}{k}$				$-\frac{1}{k(r-\lambda)}$	$\frac{\lambda t}{bk(r-\lambda)}$
s_3					$\frac{1}{k(r-\lambda)}$	$-\frac{k}{t(r-\lambda)}$
s_4				$\frac{1}{2k}$	$-\frac{1}{k^2}$	
s_5		$-\frac{k}{\lambda t}$		$-\frac{1}{\lambda t}$	$\frac{1}{k\lambda t}$	
s_6	$-\frac{1}{k}$	$-\frac{k}{\lambda t}$	1	$\frac{1}{\lambda t}$	$-\frac{1}{k\lambda t}$	

To use Table I write

$$s_3 = y'[\frac{1}{k(r-\lambda)} X_1N'NX_1' - \frac{k}{t(r-\lambda)} JJ'] y \quad .$$

To obtain the results in Table I and a number of later results it is useful to construct a multiplication table (Table II).

TABLE II. MULTIPLICATION TABLE FOR V_j TIMES V_i

j \ i	V_1	V_2	V_3	V_4	V_5	V_6	V_7
V_1	kV_1	V_4	V_1	kV_4	V_6	kV_6	kV_7
V_2	V_5	rV_2	V_2	$(r-\lambda)V_2 + \lambda V_7$	rV_5	$(r-\lambda)V_5 + \lambda kV_7$	rV_7
V_3	V_1	V_2	V_3	V_4	V_5	V_6	V_7
V_4	V_6	rV_4	V_4	$(r-\lambda)V_4 + \lambda kV_7$	rV_6	$(r-\lambda)V_6 + \lambda k^2V_7$	rkV_7
V_5	kV_5	$(r-\lambda)V_2 + \lambda V_7$	V_5	$k(r-\lambda)V_2 + \lambda kV_7$	$(r-\lambda)V_5 + \lambda kV_7$	$k(r-\lambda)V_5 + \lambda k^2V_7$	rkV_7
V_6	kV_6	$(r-\lambda)V_4 + \lambda kV_7$	V_6	$k(r-\lambda)V_4 + \lambda k^2V_7$	$(r-\lambda)V_6 + \lambda k^2V_7$	$k(r-\lambda)V_6 + \lambda k^3V_7$	rk^2V_7
V_7	kV_7	rV_7	V_7	rkV_7	rkV_7	rk^2V_7	bkV_7

In Table II, $V_1 = X_1X_1'$, $V_2 = X_2X_2'$, $V_3 = I$, $V_4 = X_1N'X_2'$, $V_5 = X_2NX_1'$, $V_6 = X_1N'NX_1'$, and $V_7 = JJ'$.

For later reference, it is convenient to rewrite the exponent (3.2) in terms of quadratic forms as

$$\frac{1}{\sigma^2}\,[y'(I - r^{-1}\,X_2X_2')y + r^{-1}(y - \mu J - X_2\tau)'\,X_2X_2'(y - \mu J - X_2\tau)]$$

$$- \frac{\sigma_B^2}{\sigma^2(\sigma^2+k\sigma_B^2)}\;[(y - \mu J - X_2\tau)'\,X_1X_1'\,(y - \mu J - X_2\tau)]$$

$$= \frac{1}{\sigma^2}\,(S_1 + S_2) - \frac{\sigma_B^2}{\sigma^2(\sigma^2 + k\sigma_B^2)}\;S_3\;. \qquad (3.7)$$

Note that $(I - r^{-1}\,X_2X_2')\,J = (I - r^{-1}\,X_2X_2')\,X_2 = 0$.

3.3 Likelihood Equations (L.E.)

3.3.1 General structure of the L.E. for mixed models

Before we proceed to obtain the L.E. for the BIB design Model III, a general form of the L.E. will be developed. To develop this form the following rules for differentiation are given:

$$1. \quad \frac{\partial \ln|\Sigma|}{\partial \sigma_i^2} = tr(\Sigma^{-1} \frac{\partial \Sigma}{\partial \sigma_i^2})$$

$$2. \quad \frac{\partial \Sigma^{-1}}{\partial \sigma_i^2} = -\Sigma^{-1} \frac{\partial \Sigma}{\partial \sigma_i^2} \Sigma^{-1}. \qquad (3.8)$$

Define $V_i = \frac{\partial \Sigma}{\partial \sigma_i^2}$.

The general mixed model is written as

$$y = X_o \alpha + \sum_{i=1}^{p} X_i F_i \qquad (3.9)$$

where X_o is an (n x k) matrix of known fixed numbers, $k \leq n$;

X_i, i = 1, ..., p, is an (n x m_i) matrix of known numbers, $m_i \leq n$;

α is a (k x 1) vector of unknown constants;

F_i, i = 1, ..., p, is an (m_i x 1) vector of independent variables from $N(0, \sigma_i^2)$.

Note that $E(y) = X_o \alpha$ and $\Sigma = \sum_{i=1}^{p} \sigma_i^2 X_i X_i'$. Here, $V_i = X_i X_i'$, i = 1, ..., p, so that

$$\Sigma = \sum_{i=1}^{p} \sigma_i^2 V_i \quad . \tag{3.10}$$

Using (3.8) differentiate

$$\ell nL = -\frac{n}{2}\,\ell n(2\pi) - \tfrac{1}{2}\ell n\,|\Sigma| - \tfrac{1}{2}tr[\Sigma^{-1}(y - X_o\alpha)(y - X_o\alpha)'] \tag{3.11}$$

with respect to α and σ_i^2, $i = 1, \ldots, p$, and obtain the L.E.

$$X_o'\,\Sigma^{-1}\,X_o\alpha = X_o'\,\Sigma^{-1}y \ , \tag{3.12}$$

and

$$[\tfrac{1}{2}tr(\Sigma^{-1}\,V_i)] = [\tfrac{1}{2}(y - X_o\alpha)'\ \ \Sigma^{-1}\,V_i\,\Sigma^{-1}(y - X_o\alpha)] \ . \tag{3.13}$$

Note that the left hand side (l.h.s.) of (3.13) can be written

$$[\tfrac{1}{2}tr(\Sigma^{-1}\,V_i)] = [\tfrac{1}{2}tr(\Sigma^{-1}\,V_i\,\Sigma^{-1}\,\Sigma)]$$

$$= [\tfrac{1}{2}\sum_j tr(\Sigma^{-1}\,V_i\,\Sigma^{-1}\,V_j\,\sigma_j^2)]$$

or in matrix form

$$[\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1} V_i)] = \Omega \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_p^2 \end{pmatrix} = \Omega\ \Theta\ , \qquad (3.14)$$

where $\Omega = (\omega_{ij}) = [\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1} V_i\ \Sigma^{-1} V_j)]$. Therefore, we can write (3.13) as

$$\Omega\ \theta = [\tfrac{1}{2}(y - X_o\alpha)'\ \Sigma^{-1} V_i\ \Sigma^{-1}(y - X_o\alpha)]\ . \qquad (3.15)$$

Note that Ω is the information matrix, i.e.,

$$\Omega = \left[E\left(\frac{-\partial^2 \ell n\ L}{\partial\sigma_i^2\ \partial\sigma_j^2} \right) \right]_{p \times p}\ .$$

(See LaMotte [19].) This suggests an iterative technique for simultaneously solving (3.12) and (3.15). Furthermore, the right hand side (r.h.s.) of (3.15) can be written as a matrix, say Δ, dependent upon unknown parameters, times a vector of sufficient statistics and fixed parameters, say S. Hence, (3.15) can be written in matrix form as

$$\Omega\ \theta = \Delta\ S\ . \qquad (3.16)$$

This will be demonstrated for the BIB mixed and random design.

3.3.2 L.E. for the fixed effects

The logarithm of the likelihood is given by

$$\ell n\ L = -\frac{n}{2}\ell n(2\pi) - \tfrac{1}{2}\ell n\ |\Sigma|$$

$$- \tfrac{1}{2}\mathrm{tr}[\Sigma^{-1}(y - \mu J - X_2\tau)(y - \mu J - X_2\tau)'] \ . \qquad (3.17)$$

Differentiating (3.17) with respect to μ and τ and setting the resulting equations equal to zero yields

$$\begin{bmatrix} J' \\ X_2' \end{bmatrix} \Sigma^{-1} [J\ X_2] \begin{pmatrix} \mu \\ \tau \end{pmatrix} = \begin{bmatrix} J' \\ X_2' \end{bmatrix} \Sigma^{-1} y \ . \qquad (3.18)$$

To simplify (3.18), we write the l.h.s. as

$$\frac{1}{\sigma^2}\begin{bmatrix} J' \\ X_2' \end{bmatrix}[I - \frac{\sigma_B^2}{\sigma^2+k\sigma_B^2}\ X_1X_1'] [J\ X_2] \begin{pmatrix} \mu \\ \tau \end{pmatrix}$$

$$= \frac{1}{\sigma^2}\left\{\begin{bmatrix} n & rJ' \\ rJ & rI \end{bmatrix} - \frac{k\sigma_B^2}{\sigma^2+k\sigma_B^2}\begin{bmatrix} n & rJ' \\ rJ & k^{-1}NN' \end{bmatrix}\right\} \begin{pmatrix} \mu \\ \tau \end{pmatrix} \ , \qquad (3.19)$$

and the r.h.s. as

$$\frac{1}{\sigma^2}\begin{bmatrix} J' \\ X_2' \end{bmatrix} \left[I - \frac{\sigma_B^2}{\sigma^2+k\sigma_B^2} X_1X_1'\right] y$$

$$= \frac{1}{\sigma^2}\begin{bmatrix} \dfrac{\sigma^2}{\sigma^2+k\sigma_B^2} \; J'y \\ \\ \dfrac{\sigma^2}{\sigma^2+k\sigma_B^2} \; X_2'y + \dfrac{k\sigma_B^2}{\sigma^2+k\sigma_B^2} \; (X_2' - k^{-1} \; NX_1') \; y \end{bmatrix} . \qquad (3.20)$$

3.3.3 L.E. for the variance components

The L.E. for the variance components of the BIB design Model III are obtained by finding those quantities listed in equation (3.15). The information matrix is found using $[\tfrac{1}{2}tr(\Sigma^{-1} \; V_i \; \Sigma^{-1} \; V_j)]$ and is given by

$$\Omega = \tfrac{1}{2} \; \frac{b}{\sigma^4(\sigma^2+k\sigma_B^2)^2} \begin{bmatrix} \sigma^4 + (k-1)(\sigma^2 + k\sigma_B^2)^2 & k\sigma^4 \\ \\ k\sigma^4 & k^2\sigma^4 \end{bmatrix}$$

$$= \tfrac{1}{2}\left\{\frac{b}{(\sigma^2+k\sigma_B^2)^2} \begin{bmatrix} 1 & k \\ k & k^2 \end{bmatrix} + \frac{b(k-1)}{\sigma^4} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\} .$$

The r.h.s. of (3.15) is found by expanding $[\frac{1}{2}(y - \mu J - X_2\tau)' \cdot \Sigma^{-1} V_i \Sigma^{-1}(y - \mu J - X_2\tau)]$, the result of which is

$$\Delta S \tag{3.21}$$

where

$$\Delta = \frac{1}{2}\begin{bmatrix} \frac{1}{\sigma^4} & \frac{1}{\sigma^4} & \frac{1}{\sigma^4 k}\left(\left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)^2 - 1\right) \\ 0 & 0 & \frac{1}{(\sigma^2+k\sigma_B^2)^2} \end{bmatrix},$$

and

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix},$$

and S_i is defined by (3.7).

Letting
$$\theta = \begin{pmatrix} \sigma^2 \\ \sigma_B^2 \end{pmatrix}$$

we obtain the equation in the form given in (3.16).

3.4 Discussion of the Estimation of μ and τ

In most of the standard references where BIB analysis is discussed, there frequently appears a section dealing with intrablock analysis and the recovery of interblock information, (e.g., see Kempthorne [15], Graybill [9], Cochran and Cox [5]). The recovery of interblock information was first considered by Yates [32] and subsequently by a number of other writers (e.g., Zelen [33], Rao, C. R. [24], and Shah [29]). The basic idea employed in the recovery of interblock information can be summarized as follows: If one performs only the intrablock analysis then he has 'lost' some treatment information which is contained in the blocks, i.e., there exists another unbiased estimate of treatments independent of the intrablock estimates and it is a function of the block totals. (See Graybill [9, p. 408].) To pursue this, recall equation (3.18) rewritten

$$\left[\left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)\begin{pmatrix} n & rJ' \\ rJ & k^{-1}NN' \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & rI - k^{-1}NN' \end{pmatrix}\right]\begin{pmatrix} \mu \\ \tau \end{pmatrix}$$

$$= \begin{bmatrix} 0 \\ X_2'y - k^{-1}NX_1'y \end{bmatrix} + \left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)\begin{bmatrix} J'y \\ k^{-1}NX_1'y \end{bmatrix} . \qquad (3.22)$$

Imposing $J'\tau = 0$ yields

$$\tilde{\tilde{\mu}} = \frac{1}{n} J'y \tag{3.23}$$

and

$$\left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)\left(r\mu J + \frac{(r-\lambda)}{k}\tau\right) + \frac{\lambda t}{k}\tau = A'y + \frac{\sigma^2}{\sigma^2+k\sigma_B^2} k^{-1} NX_1'y \; . \tag{3.24}$$

(See Bargmann [2] for equivalent result.)

Investigation of (3.24) shows that it is a sum of two sets of equations

$$\frac{\lambda t}{k}\tau = A'y \tag{3.24a}$$

and

$$\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\left(r\mu J + \frac{r-\lambda}{k}\tau\right) = \left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right) k^{-1}NX_1'y \; . \tag{3.24b}$$

Recalling the estimates which were obtained for the fixed model, equations (2.18) and (2.19), we see that these agree with (3.23) and the solution of (3.24a). The solution of (3.24a) is

$$\hat{\tau} = \frac{k}{\lambda t} A'y \tag{3.24c}$$

and is referred to as the 'intrablock' estimate of τ. Now solving (3.24b) with (3.23) we obtain

$$\tilde{\tau} = \frac{k}{r-\lambda} [k^{-1} NX_1'y - r\tilde{\tilde{\mu}}J]$$

$$= \frac{1}{r-\lambda} [NX_1'y - \frac{1}{t} J'NX_1'yJ] \tag{3.24d}$$

since $kJ'y = J'NX_1'y$. The estimator $\tilde{\tau}$ is known as the 'interblock' estimator of τ (Graybill [9, p. 408]). It is shown in Graybill that $\hat{\tau}$ and $\tilde{\tau}$ are unbiased with variances $\frac{k(t-1)}{\lambda t^2} \sigma^2$ and $\frac{(t-1)k}{t(r-\lambda)} (\sigma^2 + k\sigma_B^2)$, respectively. Hence the problem of 'recovering' the interblock information is just that of improving on the estimator $\hat{\tau}$ by considering a minimum variance combination of (3.24c) and (3.24d), namely

$$\tilde{\tilde{\tau}} = \left(\frac{\sigma^2 + k\sigma_B^2}{r\sigma^2 + \lambda t\sigma_B^2}\right) \left[\frac{\lambda t}{k} \hat{\tau} + \left(\frac{\sigma^2}{\sigma^2 + k\sigma_B^2}\right) \frac{(r-\lambda)}{k} \tilde{\tau}\right] . \tag{3.24e}$$

Note that (3.24) can be rewritten as

$$\left[\frac{\lambda t}{k} + \frac{r-\lambda}{k} \left(\frac{\sigma^2}{\sigma^2 + k\sigma_B^2}\right)\right] \tau = A'y + \left(\frac{\sigma^2}{\sigma^2 + k\sigma_B^2}\right)(k^{-1}NX'y - r\tilde{\tilde{\mu}}J)$$

$$= \frac{\lambda t}{k} \hat{\tau} + \left(\frac{\sigma^2}{\sigma^2 + k\sigma_B^2}\right)\left(\frac{r-\lambda}{k}\right) \tilde{\tau} .$$

Therefore, the solution of (3.24) is identical with the minimum variance combination (3.24c). Hence, the often confusing 'recovery' is nothing more than the solution of the L.E. Of course, the estimate depends on the unknown ratio $\frac{\sigma^2}{\sigma^2+k\sigma_B^2}$ and since we generally do not know this ratio, we must estimate it. Much has been written in regard to this problem and the solutions depend upon the manner in which one estimates this ratio. (See Shah [29] for references and some comparisons of methods.)

Yates [32] suggests a procedure based on the AOV estimates given in Table III.

The estimates $\hat{\sigma}^2 = E_e$ and $\hat{\sigma}_B^2 = \frac{(b-1)}{t(r-1)}(E_b - E_e)$ are used to form $\frac{\sigma^2}{\sigma^2+k\sigma_B^2}$ in Yates' method. We suggest σ^2 and σ_B^2 be estimated by solving the L.E. However, the equations, given by (3.34) and (3.36), generally require an iterative solution.

TABLE III. AOV TABLE FOR THE MIXED MODEL

Source	df	SS	MS	EMS
Treatments	t-1	$\sum_{i=1}^{t} \frac{y_{i..}^2}{r} - \frac{y_{...}^2}{bk}$		
Blocks (adj.)	b-1	$\frac{k}{\lambda t} \sum_{j=1}^{t} (y_{i..} - \frac{1}{k} \sum_{j=1}^{b} n_{ij} y_{.j.})^2 + \sum_{j=1}^{b} \frac{y_{.j.}^2}{k} - \sum_{i=1}^{t} \frac{y_{i..}^2}{r}$	E_b	$\sigma^2 + \frac{t(r-1)}{b-1} \sigma_B^2$
Interblock Error	bk-t-b+1	Subtraction	E_e	σ^2
Total	bk-1	$\sum_i \sum_j \sum_m \dot{y}_{ijm}^2 - \frac{y_{...}^2}{bk}$		

3.5 Solution of the L.E.

3.5.1 Approximate solution using the fixed model estimates

Let us now consider the task of approximately solving the L.E. First, we find

$$\Omega^{-1} = \frac{2\sigma^4}{bk^2(k-1)} \begin{bmatrix} k^2 & -k \\ -k & 1 \end{bmatrix} + \frac{2(\sigma^2+k\sigma_B^2)^2}{bk^2} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} .$$

Next, pre-multiply $\Omega\theta = \Delta S$(found in Section 3.3.3) by Ω^{-1} yielding

$$\begin{pmatrix} \sigma^2 \\ \sigma_B^2 \end{pmatrix} = \left\{ \frac{1}{bk^2(k-1)} \begin{bmatrix} k^2 & k^2 & -k \\ -k & -k & 1 \end{bmatrix} + \frac{1}{bk^2} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

$$= \frac{1}{bk^2(k-1)} \begin{bmatrix} k^2 & k^2 & -k \\ -k & -k & k \end{bmatrix} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} . \qquad (3.25)$$

Since S_2 and S_3 depend on μ and τ we must estimate them. As a first approximation consider $\hat{\mu} = \frac{J'y}{n}$ and $\hat{\tau} = \frac{k}{\lambda t} A'y$. Define $\hat{S}_i = S_i(\hat{\mu}, \hat{\tau})$. We then obtain from the σ^2 equation in (3.25)

$$\tilde{\tilde{\sigma}}^2 = \frac{s_6}{b(k-1)} \quad , \tag{3.26}$$

where the $\hat{S}_i$'s are given in Table IV. Note that s_6 is defined in Table I.

TABLE IV. RELATIONSHIP BETWEEN $\hat{S}_i$ AND ITS QUADRATIC FORM

	X_1X_1'	X_2X_2'	I	$X_1N'X_2' + X_2NX_1'$	$X_1N'NX_1'$	JJ'
$\hat{S}_1$		$-\frac{1}{r}$	1			
$\hat{S}_2$		$\frac{1}{r}(\frac{r-\lambda}{\lambda t})^2$		$-\frac{(r-\lambda)}{(\lambda t)^2}$	$\frac{r}{(\lambda t)^2}$	$-\frac{1}{rt}$
$\hat{S}_3$	1	$(r-\lambda)(\frac{k}{\lambda t})^2$		$-\frac{rk^2}{(\lambda t)^2}$	$\frac{(rk + \lambda t)}{(\lambda t)^2}$	$-\frac{1}{b}$

To use Table IV, we take for example

$$\hat{S}_3 = y'[\frac{1}{r}\ (\frac{r-\lambda}{\lambda t})^2\ X_2X_2' - \frac{(r-\lambda)}{(\lambda t)^2}\ (X_1N'X_2' + X_2NX_1') + \frac{r}{(\lambda t)^2}\ X_1N'NX_1'$$

$$- \frac{1}{rt}\ JJ']\ y \quad .$$

By writing

$$\hat{S}_3 = ks_7 + (r-\lambda)\,\frac{k}{\lambda t}\,w \quad ,$$

where

$$s_7 = s_2 + \left(\frac{r-\lambda}{rk}\right) s_5 - \frac{2s_4}{r} + \frac{\lambda t}{rk}\, s_3 = w + s_2$$

and the s_i's are defined in Table I, we obtain

$$\widetilde{\widetilde{bk(\sigma^2+k\sigma_B^2)}} = \hat{S}_3 = ks_7 + \frac{(r-\lambda)}{\lambda t}\, kw = ks_2 + \frac{rk^2}{\lambda t}\, w \ . \qquad (3.27)$$

Solving for $\tilde{\tilde{\sigma}}_B^2$, we obtain

$$\tilde{\tilde{\sigma}}_B^2 = \frac{s_2}{bk} - \frac{s_6}{bk(k-1)} + \frac{r}{\lambda tb}\, w \ . \qquad (3.28)$$

It is interesting to compare estimates (3.26) and (3.27) with those obtained by using Yates' method given in the previous section. (These are also Henderson's method 3 estimates.) Apart from degrees of freedom the estimates are the same, i.e., if $b(k-1) \doteq b(k-1) - (t-1)$ and if we eliminate some of the -1's from the degrees of freedom. To see this, look at (3.27) and observe

$$\widetilde{\widetilde{bk(\sigma^2+k\sigma_B^2)}} = ks_2 + \frac{rk^2}{\lambda t} w = ks_2 + \frac{rk^2(t-1)}{rt(k-1)} w$$

$$\doteq k(s_2+ w) = ks_7 \quad .$$

Hence

$$\tilde{\tilde{\sigma}}_B^2 = \frac{s_7}{bk} - \frac{s_6}{bk(k-1)} \quad . \tag{3.29}$$

Yates' estimates are given as

$$\dot{\sigma}^2 = \frac{s_6}{bk-b-t+1} \qquad \text{and}$$

$$\dot{\sigma}_B^2 = \frac{(b-1)}{t(r-1)} \left[\frac{s_7}{(b-1)} - \frac{s_6}{bk-b-t+1}\right] \quad , \tag{3.30}$$

and the comparisons are clearly as noted above. Therefore, we conclude that if the fixed model estimates are used in the L.E. then, apart from small discrepancies in degrees of freedom, the estimates of the variance components are equal to those obtained using AOV techniques.

3.5.2 Direct solution using the maximum likelihood estimates

Suppose we substitute $\tilde{\tilde{\mu}} = \frac{J'y}{n}$ and

$$\tilde{\tilde{\tau}} = \left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right) \left[\frac{\lambda t}{k}\hat{\tau} + \left(\frac{\sigma_B^2}{\sigma^2+k\sigma_B^2}\right)\left(\frac{r-\lambda}{k}\right)\tilde{\tau}\right] \text{ , i.e.,}$$

the maximum likelihood estimates (3.23) and (3.24e), into (3.21). Let

$$\tilde{\tilde{h}} = (y - \tilde{\tilde{\mu}}J - X_2\tilde{\tilde{\tau}}) = \left[I - \frac{\lambda}{r}\left(\frac{\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right) JJ' - \frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2} X_2X_2' \right.$$

$$\left. + \frac{\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2} X_2NX_1'\right] y \quad .$$

Now compute

$$X_2X_2'\tilde{\tilde{h}} = \left(\frac{\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right) [-(r-\lambda)\ X_2X_2' - \lambda JJ' + rX_2NX_1']\ y \ .$$

Therefore

$$\tilde{\tilde{S}}_2 = \frac{1}{r}\tilde{\tilde{h}}'X_2X_2'\tilde{\tilde{h}}$$

$$= \frac{1}{r}\left(\frac{\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 y'[(r-\lambda)^2\ X_2X_2' - r(r-\lambda)(X_1N'X_2' + X_2NX_1')$$

$$+ r^2X_1N'NX_1' - \lambda^2tJJ']\ y \quad . \tag{3.31}$$

These results were obtained with the aid of multiplication Table II and results (2.3) through (2.10).

Similarly, we obtain

$$\tilde{\tilde{h}}'X_1X_1'\tilde{\tilde{h}} = \tilde{\tilde{S}}_3 = y'\Bigg[X_1X_1' + (r-\lambda)\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 X_2X_2' - r\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 (X_1N'X_2'$$

$$+ X_2NX_1') + \left(\frac{\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)\left(1 + r\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right) X_1N'NX_1'\right)$$

$$+ \left(\lambda\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 - \frac{\lambda k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\left(\frac{1}{r} + \frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)\right) JJ'\Bigg] y ,$$

or to express this in terms of the s_i's

$$\tilde{\tilde{S}}_3 = ks_2 + \lambda t\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 \left(\frac{r-\lambda}{k} s_5 - 2s_4 + \frac{\lambda t}{k} s_3\right)$$

$$= ks_2 + r\lambda t\left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 w . \qquad (3.32)$$

Now consider the equation for σ^2 from (3.25), which can be written

$$bk(k-1)\ \sigma^2 = k\tilde{\tilde{S}}_1 + k\tilde{\tilde{S}}_2 - \tilde{\tilde{S}}_3 \quad . \tag{3.33}$$

Again as above, equation (3.33) can be simplified to

$$bk(k-1)\ \sigma^2 = ks_6 + \left(\frac{\sigma^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 r(r-\lambda)\ w \ . \tag{3.34}$$

The equation for σ_B^2 from (3.25) can be written as

$$bk^2(k-1)\ \sigma_B^2 = -bk(k-1)\ \sigma^2 + (k-1)\ \tilde{\tilde{S}}_3$$

or

$$bk(\sigma^2 + k\sigma_B^2) = \tilde{\tilde{S}}_3 \ . \tag{3.35}$$

Thus (3.35) results in

$$bk(\sigma^2 + k\sigma_B^2) = ks_2 + r\lambda t \left(\frac{\sigma^2+k\sigma_B^2}{r\sigma^2+\lambda t\sigma_B^2}\right)^2 w \ . \tag{3.36}$$

As done previously, let us consider the simultaneous solution of (3.34) and (3.36) and relate this to Yates' AOV estimates. If $\frac{\sigma^2}{\sigma^2+k\sigma_B^2}$ is small and if we eliminate some of the -1's in the degrees of freedom, we can show that maximum likelihood estimates and AOV estimates are equivalent. To exhibit this result, note that (3.34) can be written

$$bk(k-1)\ \sigma^2 = ks_6 + \left[\frac{\sigma^2}{\sigma^2 + \frac{\lambda t}{r}\sigma_B^2}\right]^2 \left(\frac{r-\lambda}{r}\right) w$$

$$\doteq ks_6 + \left[\frac{\sigma^2}{\sigma^2 + k\sigma_B^2}\right]^2 \left(\frac{r-\lambda}{r}\right) w\ , \qquad (3.37)$$

using $\lambda t = rk$ instead of $\lambda(t-1) = r(k-1)$.

It follows that (3.37) is approximately

$$\approx{\sigma}^2 \doteq \frac{s_6}{b(k-1)}$$

versus Yates' estimate

$$\dot{\sigma}^2 = \frac{s_6}{bk-b-t+1}\quad .$$

Also, (3.36) can be written

$$bk\widetilde{\widetilde{(\sigma^2+k\sigma_B^2)}} = ks_2 + \frac{\lambda t}{r}\left[\frac{\sigma^2+k\sigma_B^2}{\sigma^2 + \frac{\lambda t}{r}\sigma_B^2}\right] w \doteq ks_2 + kw = ks_7$$

using the relationship that yielded (3.37). Therefore,

$$\approx\!\!\sigma_B^2 \doteq \left[\frac{s_7}{bk} - \frac{s_6}{bk(k-1)}\right]$$

versus Yates' estimate

$$\dot{\sigma}_B^2 = \frac{(b-1)}{t(r-1)}\left[\frac{s_7}{b-1} - \frac{s_6}{bk-b-t+1}\right] .$$

It also follows that if $\left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)$ is small, then applying the same degree of freedom adjustment as above results in $\approx\!\!\tau = \hat{\tau}$.

What if $\left(\frac{\sigma^2}{\sigma^2+k\sigma_B^2}\right)$ is not small? Then maximum likelihood estimates and AOV estimates can differ appreciably. To see an example of this, consider Graybill's [9] AOV on page 418, Table 18.10. Yates' method produces the estimates

$$\dot{\sigma}^2 = .2424$$

$$\dot{\sigma}_B^2 = .00596 \quad .$$

Iterating on (3.34) and (3.36) yields

$$\approx\!\!\sigma^2 = .1587$$

$$\approx\!\!\sigma_B^2 = .00713 \quad ,$$

which, as indicated above, differ considerably from the AOV estimates. Note that in this example the estimate of $\frac{\sigma^2}{\sigma^2+k\sigma_B^2}$ is approximately 1. It is important to realize that (3.34) and (3.36) must in most cases be solved iteratively, i.e., we must estimate σ^2 and σ_B^2 on the r.h.s. of (3.34) and (3.36) and obtain the first iteration estimates for σ^2 and $\sigma^2+k\sigma_B^2$ on the l.h.s.; then put these first iteration estimates into the r.h.s. of (3.34) and (3.36) and continue this process until convergence is obtained.

3.6 Large Sample Properties of the Estimators

Maximum likelihood estimates possess a number of desirable large sample properties. These include: (1) that the vector of estimates is consistent; (2) that the joint maximum likelihood estimates tend to a multivariate normal distribution (under regularity conditions); (3) that the large sample variance-covariance matrix is the inverse of the information matrix and; (4) that the estimates are asymptotically unbiased and asymptotically efficient. (See Kendall and Stuart [16].)

In reference to the previous section, if the ratios of the parameters on the r.h.s. of (3.34) and (3.36) were known, then we would have the exact solution of the L.E. and could claim the properties of maximum likelihood. However, we practically never know these ratios and consequently are hopeful that our iterative technique converges to the maximum of the likelihood function.

CHAPTER IV

MATRIX ANALYSIS OF THE RANDOM MODEL (MODEL II)

4.1 Definition of the Model

Consider equation (2.2) and set $\beta = B$, $\tau = T$ resulting in

$$y = \mu J + X_1 B + X_2 T + e \quad \text{where} \quad B \sim N(0,\ \sigma_B^2 I)$$

$$\text{and} \quad T \sim N(0,\ \sigma_T^2 I)\ .$$

Alternately, we can write

$$y \sim N(\mu J,\ \Sigma) \text{ a and } \quad \Sigma = \sigma^2 I + \sigma_B^2 X_1 X_1' + \sigma_T^2 X_2 X_2'\ .$$

With this definition of the model we have what is commonly called Eisenhart's Model II [7].

4.2 Identification of Sufficient Statistics

Write

$$\Sigma = (\sigma^2 I + \sigma_B^2 X_1 X_1')[I + (\sigma^2 I + \sigma_B^2 X_1 X_1')^{-1}\ \sigma_T^2 X_2 X_2']\ ,$$

then

$$\Sigma^{-1} = [I + (\sigma^2 I + \sigma_B^2 X_1 X_1')^{-1}\ \sigma_T^2 X_2 X_2']^{-1} (\sigma^2 I + \sigma_B^2 X_1 X_1')^{-1}\ , \quad (4.1)$$

where each inverse on the r.h.s. of (4.1) can be found by applying (3.1). The resulting inverse can be shown to be

$$\Sigma^{-1} = \sigma^{-2}\left[-\frac{\sigma_B^2}{\eta} X_1X_1' - \frac{\sigma_T^2\eta}{\phi} X_2X_2' + I + \frac{\sigma_B^2\sigma_T^2}{\phi}(X_1N'X_2' + X_2NX_1') - \frac{\sigma_B^4\sigma_T^2}{\eta\phi} X_1N'NX_1' - \frac{\lambda\sigma^2\sigma_B^2\sigma_T^2}{\phi\nu\eta} JJ'\right], \tag{4.2}$$

where

$$\eta = \sigma^2 + k\sigma_B^2 , \quad \nu = \eta + r\sigma_T^2 , \quad \text{and} \quad \phi = \sigma^2\nu + \lambda t\sigma_B^2\sigma_T^2 .$$

To identify sufficient statistics (b > t), write the exponent

$$(y - \mu J)' \Sigma^{-1}(y - \mu J) = \sum_{j=1}^{6} \rho_j (y - \mu J)' U_j(y - \mu J) \tag{4.3}$$

where the U_j's are linear combinations of the matrices found in (4.2) such that these quadratic forms reveal the sufficient statistics and the ρ_j's involve the parameters. Table IV is given to identify the ρ_j's and U_j's and further, this set of sufficient statistics was chosen subject to $U_jJ = 0$, $j = 1, \ldots, 5$. This choice was made in order to obtain the sufficient statistics s_i, $i = 1, \ldots, 6$, previously obtained by Weeks and Graybill [31].

TABLE V. RELATION OF ρ_j AND U_j TO s_j

	ρ_j	X_1X_1'	X_2X_2'	I	$X_1N'X_2' + X_2NX_1'$	$X_1N'NX_1'$	JJ'
s_2	$\frac{1}{\eta}$	$\frac{1}{k}$				$-\frac{1}{k(r-\lambda)}$	$\frac{\lambda t}{bk(r-\lambda)}$
s_3	$\frac{\left(\sigma^2 + \frac{\lambda t}{k}\sigma_T^2\right)}{\phi}$					$\frac{1}{k(r-\lambda)}$	$-\frac{k}{t(r-\lambda)}$
s_4	$\frac{-2\sigma_T^2}{\phi}$				$\frac{1}{2k}$	$-\frac{1}{k^2}$	
s_5	$\frac{(\eta + \frac{r-\lambda}{k}\sigma_T^2}{\phi}$		$\frac{k}{\lambda t}$		$-\frac{1}{\lambda t}$	$\frac{1}{k\lambda t}$	
s_6	$\frac{1}{\sigma^2}$	$-\frac{1}{k}$	$\frac{k}{\lambda t}$	1	$\frac{1}{\lambda t}$	$-\frac{1}{k\lambda t}$	

To illustrate the use of Table V, consider for example

$$s_5 = y'[\frac{k}{\lambda t} X_2X_2' - \frac{1}{\lambda t}(X_1N'X_2' + X_2NX_1') + \frac{1}{k\lambda t} X_1N'NX_1']y$$

$$= y'U_5y$$

where $\rho_5 = \frac{(\eta + \frac{r-\lambda}{k}\sigma_T^2)}{\phi}$ is substituted into the r.h.s. of equation

(4.3) when j = 5. The remaining parts on the r.h.s. of (4.3) are obtained analogously. When b = t, s_2 is not defined.

4.3 L. E. for μ and the Variance Components

The logarithm of the likelihood function is, apart from an additive constant,

$$\ell n\ L = -\frac{n}{2}\ \ell n\ |\Sigma| - \frac{1}{2}(y - \mu J)'\ \Sigma^{-1}(y - \mu J)\ . \tag{4.4}$$

Differentiating (4.4) with respect to μ, σ^2, σ_B^2, and σ_T^2, yields the L.E., which can be written

$$J'\Sigma^{-1}\ J\mu = J'\Sigma^{-1}\ y\ , \tag{4.5}$$

and

$$\Omega \begin{pmatrix} \sigma^2 \\ \sigma_B^2 \\ \sigma_T^2 \end{pmatrix} = [\tfrac{1}{2}(y - \mu J)'\ \Sigma^{-1}\ V_i\ \Sigma^{-1}(y - \mu J)],\quad i = 1, 2, 3, \tag{4.6}$$

where $V_1 = \dfrac{\partial\Sigma}{\partial\sigma^2}$, $V_2 = \dfrac{\partial\Sigma}{\partial\sigma_B^2}$, $V_3 = \dfrac{\partial\Sigma}{\partial\sigma_T^2}$ and Ω is the information matrix. The reader is referred to equation (3.15) for the development.

4.4 Solution of the L. E.

Equation (4.5) is easily shown to yield

$$\hat{\mu} = \frac{J'y}{n} = \bar{y}_{...} \; . \qquad (4.7)$$

Equation (4.6) is difficult to solve. In order to develop the comparison between maximum likelihood and a minimum variance combination of the sufficient statistics, we decided to work with a different set of sufficient statistics from those given by Weeks and Graybill [31]. Table VI gives the relationship between the two sets of sufficient statistics and the expected values of this new set, denoted by P_1, P_2, P_3, Z_1, and Z_2. Further, it can be shown that P_1, P_2, and P_3 are independent of Z_1 and Z_2. Also, the expected values in Table VI suggested the reparameterization $\sigma_B^2{*} = \sigma^2 + k\sigma_B^2$.

TABLE VI. RELATION OF P_1, P_2, P_3, Z_1, Z_2 TO s_2, s_3, s_4, s_5, s_6

	s_2	s_3	s_4	s_5	s_6	Expected Value
P_1			$\frac{-k}{(t-1)(r-\lambda)}$	$\frac{1}{(t-1)}$		σ^2
P_2		$\frac{1}{(t-1)}$	$\frac{-k}{\lambda t(t-1)}$			$\sigma^2 + k\sigma_B^2$
P_3			$\frac{k^2}{\lambda t(t-1)(r-\lambda)}$			σ_T^2
Z_1					$\frac{1}{f}$	σ^2
Z_2	$\frac{1}{b-t}$					$\sigma^2 + k\sigma_B^2$

To use Table VI, write for example

$$P_1 = \frac{-k}{(t-1)(r-\lambda)} \ s_4 + \frac{1}{(t-1)} s_5 \ .$$

Low [22] suggests a linear combination of the sufficient statistics with constant coefficients as estimates of the variance components. She then compares the variances of this set of estimates. As alternatives consider maximum likelihood estimates or a minimum variance combination of these sufficient statistics. To pursue the connection between the latter two approaches, note that since the P's and Z's are independent, we have

$$\Omega = \Omega_P + D'\Omega_Z D = \Omega_P[I + \Omega_P^{-1}D'\Omega_Z D] \ , \tag{4.8}$$

where $D = \begin{bmatrix} 100 \\ 010 \end{bmatrix}$, and Ω_P and Ω_Z are the information matrices associated with the P's and Z's, respectively. In order to find Ω^{-1}, we apply (3.1) to (4.8) and find

$$\Omega^{-1} = (I + \Phi D)\Omega_P^{-1} \tag{4.9}$$

where $\Phi = -\Omega_P^{-1}D'[D\Omega_P^{-1}D' + \Omega_Z^{-1}]^{-1}$. Therefore, one must find Ω_p, Ω_Z, and their inverses.

To form Ω_P and Ω_Z, we apply

$$\left[\frac{1}{2}\,\mathrm{tr}(\Sigma^{-1}V_i\Sigma^{-1}V_j)\right] = \left[E\left(\frac{-\partial^2 \ell n\ L}{\partial\sigma_i^2\partial\sigma_j^2}\right)\right]_{3\times 3} ,$$

where Σ^{-1} is given by (4.2). The algebra involved is quite tedious and will not be given. However, one makes extensive use of the relationships (2.6) through (2.10)$'$ and multiplication Table II. After simplification, we obtain

$$\Omega_P = \frac{(t-1)}{2\phi^2}\begin{bmatrix} (\sigma_{B*}^2 + \frac{r-\lambda}{k}\sigma_T^2)^2 & \frac{\lambda t(r-\lambda)}{k^2}\sigma_T^4 & \frac{\lambda t}{k}\sigma_{B*}^4 \\ & (\sigma^2 + \frac{\lambda t}{k}\sigma_T^2)^2 & \frac{r-\lambda}{k}\sigma^4 \\ & & (\frac{\lambda t}{k}\sigma_{B*}^2 + \frac{r-\lambda}{k}\sigma_T^2)^2 \end{bmatrix}$$

$$+ \frac{1}{2\nu^2}\begin{pmatrix}0\\1\\r\end{pmatrix}(0,\ 1,\ r), \tag{4.10}$$

where ν and ϕ are defined in (4.2), and

$$\Omega_Z = \frac{1}{2}\begin{bmatrix} \frac{f}{\sigma^4} & 0 \\ 0 & \frac{b-t}{\sigma_{B*}^4} \end{bmatrix} . \tag{4.11}$$

Furthermore, the r.h.s. of (4.6) can be written after substituting

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

$$\left[\Omega_P - \frac{1}{2\nu^2}\begin{pmatrix}0\\1\\r\end{pmatrix}(0,\ 1,\ r) \;\vdots\; D'\Omega_Z\right] \begin{bmatrix}P\\Z\end{bmatrix} \quad , \tag{4.12}$$

which displays the form given in (3.16). To obtain (4.12), we can use the orthogonal matrix given by Weeks and Graybill [31] to form the decomposition of the r.h.s. into a matrix times the vector of sufficient statistics and then transform these s_i's, $i = 1, \ldots, 6$, to P_1, P_2, P_3, Z_1, Z_2 via Table IV. However, one can obtain the result more directly by simply expanding $[(y - \mu J)'\Sigma^{-1}V_i\Sigma^{-1}(y - \mu J)]$ using (4.2) for Σ^{-1}, and utilizing multiplication Table II and Table I.

Now let us look at the L. E. for the variance components, i.e.,

$$\begin{pmatrix}\sigma^2\\ \sigma^2_{B*}\\ \sigma^2_T\end{pmatrix} = \Omega^{-1}\left[\Omega_P - \frac{1}{2\nu^2}\begin{pmatrix}0\\1\\r\end{pmatrix}(0,\ 1,\ r) \;\vdots\; D'\Omega_Z\right] \begin{bmatrix}P\\Z\end{bmatrix} . \tag{4.13}$$

Note that

$$\Omega^{-1}D'\Omega_Z = (I + \Phi D)\Omega_P^{-1}\, D'\Omega_Z \tag{4.14}$$

and further recall that

$$\Phi = \Omega_P^{-1}D'[D\Omega_P^{-1}D' + \Omega_Z^{-1}]^{-1} .$$

Also, $\Phi[D\Omega_P^{-1}D' + \Omega_Z^{-1}] = -\Omega_P^{-1}D'$, and rearranging, we get

$$(I + \Phi D)\Omega_P^{-1}D' = -\Phi\Omega_Z^{-1} . \qquad (4.15)$$

Therefore,

$$\Omega^{-1}D'\Omega_Z = -\Phi . \qquad (4.16)$$

Hence,

$$\begin{pmatrix} \sigma^2 \\ \sigma^2_{B*} \\ \sigma^2_T \end{pmatrix} = P + \Phi(DP - Z) - \frac{\Omega^{-1}}{2\nu^2}\begin{pmatrix} 0 \\ 1 \\ r \end{pmatrix} (0,\ 1,\ r)P . \qquad (4.17)$$

The first two terms on the r.h.s. of (4.17) are a linear combination of the sufficient statistics where Φ is chosen to minimize the variance of this linear combination (See Hocking, et al., [13]). Thus, we observe from (4.17) that maximum likelihood and a minimum variance combination of the sufficient statistics differ by the last term on the r.h.s. of (4.17). This difference will be discussed in the next section. Note that (4.17) is now in a good form for iteration.

4.5 Numerical Comparisons

As noted in the previous section, one must iterate on (4.17) to estimate the variance components. Therefore, a computer program was written to solve (4.17). In addition, a simulation study was made comparing (1) maximum likelihood, (2) the minimum variance combination of the sufficient statistics, and (3) Henderson's fitting constants estimates. The fitting constants estimates were chosen because Low [22] states that these estimates "...tend to have smaller variance when the σ^2's are such that a BIB might be used."

Without loss of generality μ was assumed equal to zero. To generate the multivariate normal data, let $x \sim N(0, \Sigma)$ and $y \sim N(0, I)$. Let M be a matrix whose columns are the eigenvectors of Σ, then $M\Sigma M' = \text{Diag}(\lambda_i)$, where $\text{Diag}(\lambda_i)$ is a diagonal matrix with λ_i the characteristic roots of Σ. Choose $H = \text{Diag}(\frac{1}{\sqrt{\lambda_i}})M$, then $x = H^{-1}y$ is the transformation needed such that $\Sigma = H^{-1}(H')^{-1}$, i.e., the variance-covariance matrix can be written as a product of a triangular matrix and its transposed matrix. To find H^{-1} we used the square-root method. (See Faddeev and Faddeeva [8 , pp. 144-147].) Therefore, we generated N(0, 1)'s and applied the above transformation to obtain $N(0, \Sigma)$. Incidentally, the N(0, 1)'s are generated using the Box-Muller [3] equations.

Two hundred samples were generated for each of the BIB designs under study. BIB designs with n = 12, 30, and 42 observations were generated for numerous sets of parameters. The starting values chosen for both the L. E. and the minimum variance combination of the sufficient statistics were the fitting constants estimators. However, this choice was for convenience and other starting values were investigated later. The iterative technique was performed using the unconstrained L. E. The results are summarized as follows. All three methods on the average produce the same estimates when either $\frac{\sigma_B^2}{\sigma^2}$ or $\frac{\sigma_T^2}{\sigma^2}$ is greater than 1. Also, their sample variances are almost identical and are quite close to the large-sample maximum likelihood variances. If both $\frac{\sigma_B^2}{\sigma^2}$ and $\frac{\sigma_T^2}{\sigma^2}$ are considerably smaller than 1, then one encounters some convergence difficulties solving the L. E. However, even in this case if we eliminate the divergent results and compare the remaining estimates we find that on the average they are not different. It is noted that for individual cases one can find the fitting constants estimates and the maximum likelihood estimates to be quite different. It is very uncommon to find the maximum likelihood estimates and the minimum variance combination estimates differing significantly. For all the cases considered it appears that the starting values can be vastly different and have little effect on convergence when either $\frac{\sigma_B^2}{\sigma^2}$ or $\frac{\sigma_T^2}{\sigma^2}$ is near or greater than 1, but when divergent cases are present starting values do not seem to be the cause for divergence. Since we are iterating on the L. E. we cannot guarantee a global maximum and hence can only hope that the apparent indifference to starting values indicates the desired maximum.

Rao [25] recently has proposed a procedure based on MInimum Norm Quadratic Unbiased Estimation (MINQUE). LaMotte's [21] procedure is an iterative procedure which is essentially maximum likelihood corrected for bias. If one takes as starting values

$\sigma^2 = 1$, $\sigma_B^2 = \sigma_T^2 = 0$ using LaMotte's technique, we obtain after the first iteration estimates based on the unadjusted BIB analysis, i.e., one simultaneously equates the sample mean squares to the expected mean squares and solves for the individual variance components from the AOV Table VII.

TABLE VII. UNADJUSTED AOV TABLE FOR THE RANDOM MODEL

S.V.	d.f.	SS	E(MS)
Blocks	b - 1	$\frac{1}{k}\sum_j (y_{.j.} - \frac{y_{...}}{b})^2$	$\sigma^2 + k\sigma_B^2 + \left(\frac{b-r}{b-1}\right)\sigma_T^2$
Treatments	t - 1	$\frac{1}{r}\sum_i (y_{i..} - \frac{y_{...}}{t})^2$	$\sigma^2 + \left(\frac{t-k}{t-1}\right)\sigma_B^2 + r\sigma_T^2$
Remainder	bk-b-t+1	Subtraction	
Total	bk - 1	$\sum_i\sum_j\sum_m y_{ijm}^2 - \frac{y_{...}^2}{bk}$	$\sigma^2 + \frac{k(b-1)}{(bk-1)}\sigma^2 + \frac{r(t-1)}{(bk-1)}\sigma_T^2$

It has been shown by Low [22] that these estimates tend to have larger variance than the fitting constants estimates when the σ^2's are such that a BIB might be used. It can be shown that if one takes $\sigma^2 = \sigma_T^2 = \sigma_B^2 = 1$ as initial estimates using LaMotte's procedure then the resulting estimates after one iteration are Rao's MINQUE estimates. Consequently, one is skeptical about the variance of Rao's estimates.

In conclusion, we state that generally the fitting constants estimators compare quite favorably with both the maximum likelihood estimates and the minimum variance combination estimates. Occasionally, one of the latter two estimates is uniformly better than the others when the criterion for comparison is sample variance. Never did we find a case where the fitting constants estimates were uniformly better.

4.6 Negative Estimates of Variance Components

In the previous section and also in Section 5 of Chapter 3, we have considered the unconstrained solution of the L. E. Since reporting negative estimates for known nonnegative parameters is not generally well received, we present in this section the necessary steps required to modify and eliminate negative estimates.

It is relatively easy to impose nonnegativity constraints on the estimates of the variance components. If there is one negative estimate, then this component is set equal to zero and we delete the row and column which corresponds to the differentiation of the L. E. with respect to that component. If more than one negative estimate is obtained, then one considers the solutions to the L. E. by forcing each one of the negative components to zero separately. The analogous deletion procedure described above is used. If no solution is found in which all components are nonnegative, then force all pairs to be zero. Continue with three at a time, etc., until for some n, having examined all n-tuples, at least one solution obtained by forcing an n-tuple to be zero has all components nonnegative . Now the best of these solutions is the one that yields the largest value of the likelihood.

An alternate solution to the problem of negative estimates is to consider the quadratic programming problem

$$\max \ (\Delta S)' \theta - \frac{1}{2} \theta' \Omega \theta$$

$$\text{subject to } \theta \geq 0 ,$$

where ΔS and Ω are currently fixed.

Let $F = (\Delta S)' \theta - \frac{1}{2} \theta' \Omega \theta + \lambda' \theta$, then the necessary and sufficient conditions for optimality (Kuhn-Tucker Conditions) are

$$\nabla_\theta F = \Delta S - \Omega \theta + \lambda = 0$$

$$\lambda' \theta = 0$$

$$\lambda \geq 0 , \quad \theta \geq 0 .$$

Letting ∇F_i denote the i^{th} component of ∇F, we have

$$\nabla F_i = (\Delta S - \Omega \theta)_i = 0 \quad \text{if} \quad \lambda_i = 0$$

$$\leq 0 \quad \text{if} \quad \lambda_i > 0 , \quad \text{i.e.,} \ \theta_i = 0 .$$

Hence the quadratic programming solution will yield a solution to

$$\Omega\theta = \Delta S$$

satisfying $\theta \geq 0$. Note that this must be done at each iteration since we get a new Ω and ΔS. However, one would no doubt wait to see if the unrestrained solution of $\Omega\theta = \Delta S$ is going to yield negatives before initiating the quadratic program. (See Hadley [10] for the details.)

CHAPTER V

CONCLUSION

5.1 Summary

A general iterative technique for obtaining point estimates of variance components based on a particular form of the L.E. has been developed. The procedure has been implemented for the balanced incomplete block mixed and random models.

For the mixed model, the recovery of interblock information is discussed and the relationship between maximum likelihood estimates and analysis of variance estimates is uncovered. Also a direct method, dependent upon a closed form solution to the inverse of the variance-covariance matrix, is given for identifying sufficient statistics.

For the random model, the numerical solution of the L.E. using the form mentioned above is obtained. Numerical comparisons are made among (1) maximum likelihood, (2) a minimum variance combination of the sufficient statistics, and (3) Henderson's fitting constants estimates. The results indicate that Henderson's fitting constants estimates compare quite favorably to the other estimates in both average value and sample variance. The bias for the maximum likelihood estimates is extremely small if indeed they are biased. Since we solve the unconstrained L.E. and therefore, allow negative estimates for the variance components, there is a section devoted to considering only non-negative estimates. Also included is a direct method for identifying sufficient statistics similar to that presented for the mixed model.

5.2 Future Research

It appears that one cannot generally solve the L.E. in closed form. Therefore, one should investigate other forms of the L.E. which might further ease the computational task of maximizing the likelihood function. Also needed is a workable criterion which will allow us to establish convergence properties for the iterative technique which we present.

As far as unbalanced designs go we have admittedly just scratched the surface by considering BIB models. However, we are hopeful that the conclusions that were drawn for the BIB models will hold for other unbalanced designs and to pursue this we plan to investigate the properties of maximum likelihood estimation for other unbalanced design models.

REFERENCES

[1] Anderson, R. L. and Crump, P. P., "Comparisons of Designs and Estimation Procedures for Estimating Parameters in a Two-Stage Nested Process", Technometrics, 9, (1967), 499-516.

[2] Bargmann, R. C. and Mah, V., "Theory of Least Squares", Parts 1 and 2, Lecture Notes, Virginia Polytechnic Institute and State University, Blacksburg, Virginia (1961).

[3] Box, G. E. P. and Muller, M. E., "A Note on the Generation of Random Normal Deviates", The Annals of Mathematical Statistics, 29, (1958), 610-611.

[4] Bush, N. and Anderson, R. L., "A Comparison of Three Different Procedures for Estimating Variance Components", Technometrics, 5, (1963), 421-440.

[5] Cochran, W. G. and Cox, G. M., Experimental Designs, 2nd Edition, New York: John Wiley and Sons, Inc., 1960.

[6] Crump, S. L., "Present Status of Variance Component Analysis", Biometrics, 7, (1951), 1-16.

[7] Eisenhart, C., "The Assumptions Underlying the Analysis of Variance", Biometrics, 3, (1947), 1-21.

[8] Faddeev, D. K. and Faddeeva, V. N., Computational Methods of Linear Algebra, San Francisco, California: W. H. Freeman and Company, 1963.

[9] Graybill, F. A., An Introduction to Linear Statistical Models, Vol. I., New York: McGraw-Hill Book Company, 1961.

[10] Hadley, G., Nonlinear and Dynamic Programming, Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1964.

[11] Hartley, H. O. and Rao, J. N. K., "Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model", Biometrika, 54, (1967), 93-108.

[12] Henderson, C. R., "Estimation of Variance and Covariance Components", Biometrics, 9, (1953), 226-252.

[13] Hocking, R. R., Oxspring, H. H. and Waldron, B. R., "Maximum Likelihood Estimation with Incomplete Normal Data: Parts I and II", submitted to the Journal of the American Statistical Association, 1969.

[14] Hultquist, R. A. and Graybill, F. A., "Minimal Sufficient Statistics for the Two-Way Classification Mixed Model Design", Journal of the American Statistical Association, 60, (1965), 182-192.

[15] Kempthorne, O., The Design and Analysis of Experiments, New York: John Wiley and Sons, Inc., 1960.

[16] Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, Vol. II, London: Charles Griffin and Company Limited, 1961.

[17] Klotz, J. H., Milton, R. C. and Zacks, S., "Mean Square Efficiency of Estimators of Variance Components", Journal of the American Statistical Association, 64, (1968), 1383-1402.

[18] Koch, G. G., "A General Approach to the Estimation of Variance Components", Technometrics, 9, (1967), 93-118.

[19] ________, "Some Further Remarks on 'A General Approach to the Estimation of Variance Components'", Technometrics, 10, (1968), 551-558.

[20] LaMotte, L. R., "Selection of Regression Variables and Contributions to the Estimation of Variance Components", Ph.D. Thesis, Texas A&M University, College Station, Texas, 1969.

[21] ________, "A Class of Estimators of Variance Components", University of Kentucky, Department of Statistics, Technical Report No. 10, 1970.

[22] Low, L. Y., "Estimators of Variance Components in the Balanced Incomplete Block", Journal of the American Statistical Association, 64, (1969), 1014-1030.

[23] Mood, A. M. and Graybill, F. A., Introduction to the Theory of Statistics, 2nd Edition, New York: McGraw-Hill Book Company, Inc., 1963.

[24] Rao, C. R., "General Methods of Analysis for Incomplete Block Designs", Journal of the American Statistical Association, 42, (1947), 541-561.

[25] Rao, C. R., "Estimation of Variance and Covariance Components in Linear Models", Submitted to Journal of the American Statistical Association, 1971.

[26] Searle, S. R., "Another Look at Henderson's Methods of Estimating Variance Components", Biometrics, 24, (1968), 749-788.

[27] ________, "Topics in Variance Component Estimation", Biometrics, 27, (1971), 1-76.

[28] ________, Linear Models, New York: John Wiley and Sons, Inc., 1971.

[29] Shah, K. R., "Loss of Information in Combined Inter- and Intra-Block Estimation", Journal of the American Statistical Association, 65, (1970), 1562-1564.

[30] Townsend, E. C., "Unbiased Estimators of Variance Components in Simple Unbalanced Designs", Ph.D. Thesis, Cornell University, Ithaca, New York, 1968.

[31] Weeks, D. L., and Graybill, F. A., "Minimal Sufficient Statistics for the Balanced Incomplete Block Design under an Eisenhart Model II", Sankyā, Series A, 23, (1961), 261-298.

[32] Yates, F., "The Recovery of Inter-Block Information in Balanced Incomplete Block Designs", Annals of Eugenics, 10, (1940), 317-325.

[33] Zelen, M., "The Analysis of Incomplete Block Designs", Journal of the American Statistical Association, 52, (1951), 204-217.