THEOREMS CONCERNING EISENHART'S MODEL II[1]

By Franklin A. Graybill and Robert A. Hultquist

Colorado State University and Oklahoma State University

1. Introduction. Eisenhart's Model II has been discussed in many papers [1], [2], [3], and, since it has become quite important as a statistical model it seems worthwhile to investigate it in some generality. The purposes of this paper are (1) to study the covariance matrix of certain cases, (2) to give some theorems concerning minimal sufficient statistics, (3) to give some theorems concerning best quadratic unbiased estimation, (4) to give some theorems concerning analysis of variance.

2. Notation, Definitions, and Assumptions. In this paper we consider Eisenhart's Model II [4] which can be described as follows. An $n \times 1$ vector of observation $\mathbf{Y}$ is assumed to be a linear sum of $k + 2$ quantities,

$$\mathbf{Y} = \sum_{i=0}^{k+1} \mathbf{X}_i \boldsymbol{\beta}_i, \tag{1}$$

where $\boldsymbol{\beta}_0 = \mu$ is a fixed unknown constant, $\boldsymbol{\beta}_i$ $(i = 1, \cdots k)$ is a vector of p_i random variables, $\boldsymbol{\beta}_{k+1} = \mathbf{e}$ is an $n \times 1$ vector of random errors, $\mathbf{X}_0 = \mathbf{j}$ is an $n \times 1$ vector of 1's, $\mathbf{X}_i$ $(i = 1, \cdots k)$ is a matrix of known constants, and $\mathbf{X}_{k+1} = \mathbf{I}$ is the identity matrix.

Throughout this paper we assume all random variables in and between the vectors $\boldsymbol{\beta}_i$ are independent. $\mathbf{0}$ will denote the null matrix and $\boldsymbol{\beta}_i$ will be distributed with mean $\mathbf{0}$ and covariance matrix $\sigma_i^2 \mathbf{I}$. The covariance matrix of the vector $\mathbf{Y}$ will be denoted by $\mathbf{V}$ and $\mathbf{W}$ will denote $E(\mathbf{YY}')$. $\mathbf{Y}'$ denotes the transpose of $\mathbf{Y}$. Throughout the paper E is the operator denoting the expected value of what follows. $\mathbf{A}_i$ will denote $\mathbf{X}_i \mathbf{X}_i'$ and $\mathbf{A}_i$ $(i = 0, 1, \cdots k + 1)$ will be assumed linearly independent. $\mathbf{J}$ will denote the matrix $\mathbf{jj}'$.

Some of the following assumptions are made in certain sections of this paper.

(i) $\boldsymbol{\beta}_i$ $(i = 1, \cdots k + 1)$ have multivariate normal densities.

(ii) Finite third (fourth) moments exist for all random variables and third (fourth) moments are equal for all variables in a given vector $\boldsymbol{\beta}_i$.

(iii) $\mathbf{A}_i$ and $\mathbf{A}_j$ commute $(i, j = 0, 1, \cdots k + 1)$.

(iv) The matrix $\mathbf{X}_i$ is such that $\mathbf{j}_n' \mathbf{X}_i = r_i \mathbf{j}_{p_i}'$ and $\mathbf{X}_i \mathbf{j}_{p_i}' = \mathbf{j}_n$, where r_i is a positive integer and the subscripts n and p_i are the dimensions of the vectors $\mathbf{j}$.

Many of the commonly used models satisfy most of the above assumptions. For instance, the regression model is included in our discussion when assumptions (iii) and (iv) are deleted. The experimental design models with *equal* numbers in the subclasses satisfy the assumptions. These include the n way cross classifi-

Received April 27, 1959; revised June 27, 1960.

[1] Research sponsored by the National Science Foundation, Grant No. N.S.F. G-3970.

cation models with or without interaction, the n fold nested classification, the split-plot models, etc.

3. Characteristic Roots of the Covariance Matrix. The covariance matrix of $\mathbf{Y}$ is $\mathbf{V} = \sum_{i=1}^{k+1} \sigma_i^2 \mathbf{A}_i$. Since the characteristic roots of $\mathbf{V}$ play an important role in the sections to follow, we shall devote this section to a discussion of some of the properties of those characteristic roots. Throughout this section we shall assume that assumptions (iii) and (iv) hold.

Since $\mathbf{A}_0, \mathbf{A}_1, \cdots, \mathbf{A}_{k+1}$, is a set of real symmetric matrices which commute in pairs, there exists an orthogonal matrix $\mathbf{P}$ such that $\mathbf{PA}_i\mathbf{P}' = \mathbf{D}_i$ $(i = 0, 1, \cdots k + 1)$ where the $\mathbf{D}_i$ are diagonal matrices [11] (p. 189). It is clear from the relation of $\mathbf{V}$ to the $\mathbf{A}_i$ that $\mathbf{V}$ is also diagonalized by $\mathbf{P}$ and $\mathbf{PVP}' = \sum_{i=1}^{k+1} \sigma_i^2 \mathbf{D}_i^2$.

The following theorems concern bounds on the number of distinct characteristic roots of $\mathbf{V}$.

THEOREM 1. *The maximum number of distinct characteristic roots of* $\mathbf{V}$ *is* 1 *plus the rank of the matrix* $[\mathbf{X}_0, \cdots, \mathbf{X}_k]$.

PROOF: Let the rank of $[\mathbf{X}_0, \cdots, \mathbf{X}_k]$ be q. As a consequence of assumption (iv), $\sum_{i=1}^{k} \mathbf{X}_i \mathbf{X}_1'$ also has rank q. Hence the matrix $\sum_{i=1}^{k} \mathbf{D}_i = \sum_{i=1}^{k} \mathbf{PX}_i \mathbf{X}_i' \mathbf{P}'$ has q characteristic roots not equal to zero. Since the A_i are positive semidefinite, these q characteristic roots are positive which implies $\sum_{i=1}^{k} \sigma_i^2 \mathbf{D}_i$ has q positive characteristic roots and $n - q$ characteristic roots equal to zero. Now since $\mathbf{PVP}'$ equals $\sum_{i=1}^{k} \sigma_i^2 \mathbf{D}_i + \sigma^2\mathbf{I}$, $n - q$ of its n positive characteristic roots must be σ^2. Thus the maximum number of distinct characteristic roots of $\mathbf{V}$ is $q + 1$.

We shall at times use the following theorem which we state without proof.

THEOREM 2. *One row of the matrix* $\mathbf{P}$ *which diagonalizes* $\mathbf{A}_i$ $(i = 0, 1, \cdots k + 1)$ *is a row of equal elements either* $n^{-\frac{1}{2}}$ *or* $-n^{-\frac{1}{2}}$.

THEOREM 3. *The number of distinct characteristic roots of* $\mathbf{W}$ *is not less than* $k + 2$.

PROOF: $\mathbf{W} = \sum_{i=1}^{k+1} \sigma_i^2 \mathbf{A}_i$ where σ_0^2 is used to denote μ^2, and $\mathbf{PWP}' = \sum_{i=0}^{k+1} \sigma_i^2 \mathbf{D}_i$. Let $\mathbf{h}^{(i)}$ be the vector composed of the diagonal elements of $\mathbf{D}_i$. Suppose $\mathbf{W}$ has exactly s distinct characteristic roots $d_1, \cdots, d_s$, then

$$\sum_{i=0}^{k+1} \sigma_i^2 \mathbf{h}^{(i)} = [d_1 \mathbf{j}_1', \cdots, d_u \mathbf{j}_u', \cdots, d_s \mathbf{j}_s']'$$

where $\mathbf{j}_u$ has dimension n_u equal to the multiplicity of the characteristic root d_u. If we make the partition $\mathbf{h}^{(i)\prime} = [\mathbf{h}_1^{(i)\prime}, \cdots, \mathbf{h}_u^{(i)\prime}, \cdots, \mathbf{h}_s^{(i)\prime}]$, such that $\mathbf{h}_u^{(i)}$ has the dimension n_u for all i, then we can write $\sum_{i=0}^{k+1} \sigma_i^2 \mathbf{h}_u^{(i)} = d_u \mathbf{j}_u$. Let $h_{ur}^{(i)}$ be the rth and $h_{ut}^{(i)}$ be the tth element of $\mathbf{h}_u^{(i)}$. We then assert that $\sum_{i=0}^{k+1} \sigma^2 h_{ur}^{(i)} = d_u$ and $\sum_{i=0}^{k+1} \sigma_i^2 h_{ut}^{(i)} = d_u$. Subtracting we have $\sum_{i=0}^{k+1} \sigma_i^2(h_{ur}^{(i)} - h_{ut}^{(i)}) = 0$. The above equation implies $h_{ur}^{(i)} = h_{ut}^{(i)}$ for all r and t. Thus $\mathbf{h}^{(i)}$ can be written $h^{(i)\prime} = [a_1^{(i)}\mathbf{j}_1', \cdots, a_u^{(i)}\mathbf{j}_u', \cdots, a_s^{(i)}\mathbf{j}_s']$ where $a_u^{(i)}$ is a scalar. The $\mathbf{A}_i$ being linearly independent implies the $\mathbf{D}_i$ are linearly independent which in turn implies the $\mathbf{h}^{(i)}$ are linearly independent. Thus the $k + 2$ vectors

$$[a_1^{(i)}, \cdots a_s^{(i)}]' \qquad (i = 0, \cdots k + 1)$$

form a matrix of column rank $k + 2$. This matrix must also have row rank $k + 2$ which implies $s \geqq k + 2$.

Except for the first characteristic root d_1, the characteristic roots of $\mathbf{V}$ are identical with those of $\mathbf{W}$ hence the number of distinct characteristic roots of $\mathbf{V}$ is s or $s - 1$, the latter happening only when the characteristic root $\mathbf{P}_1 \mathbf{V} \mathbf{P}_1'$ is not equal to some other of the $s - 1$ roots.

4. Minimal Sufficient Statistics. In this section we exhibit minimal sets of sufficient statistics for the model defined in Section 2 under assumptions (i) and (iii) and we obtain their distribution. Conditions are also given for a set to be complete.

As in the previous section let the number of distinct characteristic roots of the matrix $\mathbf{W}$ be s. By the proper choice of $\mathbf{P}$ the matrix $\mathbf{PVP}'$ can be written Diag $[d_1^*, d_2 \mathbf{I}_2, \cdots, d_u \mathbf{I}_u, \cdots, d_s \mathbf{I}_s]$ where $d_1^* = d_1 - n\mu^2$ and $d_1, d_2, \cdots, d_s$ are the s distinct characteristic roots of $\mathbf{W}$. The dimension of $\mathbf{I}_u$ is equal to the multiplicity of the root d_u.

Consider now the joint distribution of $y_1, \cdots, y_n$. The quadratic form is $Q = (\mathbf{Y} - \mathbf{j}\mu)'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{j}\mu)$ which can be rewritten in the following manner.

$$Q = (\mathbf{PY} - \mathbf{Pj}\mu)'(\mathbf{PVP}')^{-1}(\mathbf{PY} - \mathbf{Pj}\mu). \tag{2}$$

Partition $\mathbf{P}$ as follows. $\mathbf{P}' = [\mathbf{P}_1', \mathbf{P}_2', \cdots, \mathbf{P}_u', \cdots, \mathbf{P}_s']$ where the dimension of $\mathbf{P}_u$ is $n_u \times n$. Then since $\mathbf{P}_1 \mathbf{j} = n^{\frac{1}{2}}$ and $\mathbf{P}_u \mathbf{j} = \mathbf{0}$ $(u \neq 1)$ Q can be written

$$\begin{bmatrix} \mathbf{P}_1\mathbf{Y} - n^{\frac{1}{2}}\mu \\ \mathbf{P}_2\mathbf{Y} \\ \vdots \\ \mathbf{P}_u\mathbf{Y} \\ \vdots \\ \mathbf{P}_s\mathbf{Y} \end{bmatrix}' \begin{bmatrix} 1/d_1^* & & & & & \\ & (1/d_2)\mathbf{I}_2 & & & \mathbf{0} & \\ & & \ddots & & & \\ & & & (1/d_u)\mathbf{I}_u & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & (1/d_s)\mathbf{I}_s \end{bmatrix} \begin{bmatrix} \mathbf{P}_1\mathbf{Y} - n^{\frac{1}{2}}\mu \\ \mathbf{P}_2\mathbf{Y} \\ \vdots \\ \mathbf{P}_u\mathbf{Y} \\ \vdots \\ \mathbf{P}_s\mathbf{Y} \end{bmatrix} \tag{3}$$

or $Q = d_1^{*-1}(\mathbf{P}_1 \mathbf{Y} - n^{\frac{1}{2}}\mu)^2 + \sum_{u=2}^{s} d_u^{-1}\mathbf{Y}'\mathbf{P}_u' \mathbf{P}_u \mathbf{Y}$. This last form of Q exhibits according to Koopman [5], a set of s sufficient statistics namely

$$\mathbf{Y}'\mathbf{P}_u' \mathbf{P}_u \mathbf{Y} \qquad (u = 2, \cdots s)$$

and $\mathbf{P}_1 \mathbf{Y}$.

$\mathbf{P}_1 \mathbf{Y}$ is distributed as a univariate normal with mean $\mathbf{P}_1 \mathbf{j}\mu = n^{\frac{1}{2}}\mu$ and variance $\mathbf{P}_1 \mathbf{V}\mathbf{P}_1' = d_1^*$. In order to obtain the distribution of the remaining statistics we note the following: (a) $\mathbf{P}_u' \mathbf{P}_u \mathbf{V}/d_u$ is idempotent. (b) The non-centrality parameter $\lambda = \frac{1}{2}\mu\mathbf{j}'\mathbf{P}_u' \mathbf{P}_u \mathbf{j}\mu = 0$ $(u \neq 1)$. (c) The rank of $\mathbf{P}_u' \mathbf{P}_u \mathbf{V}$ is n_u. These conditions according to Theorem 5, Section 3 of [6], are sufficient for $\mathbf{Y}'\mathbf{P}_u' \mathbf{P}_u \mathbf{Y}/d_u$ to be distributed as a central chi square variable with n_u degrees of freedom. Since for $u \neq v$, $\mathbf{P}_u \mathbf{V}\mathbf{P}_v' = \mathbf{0}$, we have $\mathbf{P}_u' \mathbf{P}_u \mathbf{V}\mathbf{P}_v' \mathbf{P}_v = \mathbf{0}$, which is sufficient [6] to imply the independence of $\mathbf{Y}'\mathbf{P}_u' \mathbf{P}_u \mathbf{Y}$ and $\mathbf{Y}'\mathbf{P}_v' \mathbf{P}_v \mathbf{Y}$ and the independence of $\mathbf{P}_1 \mathbf{Y}$ and $\mathbf{Y}'\mathbf{P}_u' \mathbf{P}_u \mathbf{Y}$.

The following theorem establishes the minimal property of this set of statistics.

THEOREM 4. *If* $\mathbf{W}$ *has s distinct characteristic roots, then the s statistics,*

$$\mathbf{Y}'\mathbf{P}_u'\mathbf{P}_u\mathbf{Y} \qquad (u = 2, \cdots s)$$

and $\mathbf{P}_1\mathbf{Y}$ *form a minimal sufficient set.*

PROOF: Two cases must be examined: (i) d_1^* is not equal to some other of the $s - 1$ roots; (ii) d_1^* is equal to d_2.

If f is the joint frequency distribution function of $y_1, \cdots, y_n$, then for case (i) a straight forward application of the procedure of Lehmann and Scheffé [7] (pp. 327–329) to $K(\mathbf{Y}, \mathbf{Y}_0) = f(\mathbf{Y})/f(\mathbf{Y}_0)$ establishes the theorem.

Case (ii) differs from case (i) only in that $d_1^* = d_2$. However, Lemma 1, the proof of which follows, implies $(\mathbf{P}_1\mathbf{Y} - \mathbf{P}_1\mathbf{j}\mu)^2 - (\mathbf{P}_1\mathbf{Y}_0 - \mathbf{P}_1\mathbf{j}\mu)^2 + \mathbf{Y}'\mathbf{P}_2'\mathbf{P}_2\mathbf{Y} - \mathbf{Y}_0'\mathbf{P}_2'\mathbf{P}_2\mathbf{Y}_0 \equiv 0$. However, since this is an identity in μ we have $\mathbf{P}_1\mathbf{Y} = \mathbf{P}_1\mathbf{Y}_0$ and $\mathbf{Y}'\mathbf{P}_2'\mathbf{P}_2\mathbf{Y} = \mathbf{Y}_0'\mathbf{P}_2'\mathbf{P}_2\mathbf{Y}_0$. Thus in this case also the set described is a minimal sufficient set of statistics.

LEMMA 1. *If the distinct positive quantities* d_u $(u = 1, \cdots, k)$, *are of the form* $d_u = l_u + a \neq 0$ *and a is functionally independent of each* l_u, *then the quantities* d_u^{-1}; $(u = 1, \cdots, k)$, *are linearly independent.*

PROOF: Consider the set of constants c_u; $(u = 1, \cdots, k)$, such that $\sum_{u=1}^{k} c_u^{-1} d_u = 0$. It follows then that $\sum_{u=1}^{k} (c_u \prod_{v \neq u} d_v) = 0$ or equivalently $\sum_{u=1}^{k} [c_u \prod_{v \neq u} (l_v + a)] = 0$. Expanding and collecting coefficients of powers of a we have a system of k equations which can be written as $\mathbf{BC} = \mathbf{0}$ where $\mathbf{C}' = (c_1, \cdots, c_k)$ and

$$\mathbf{B} = \begin{bmatrix} 1 & \cdot & 1 & \cdot & 1 \\ \sum_{v \neq 1} l_v & \vdots & \sum_{v \neq u} l_v & \vdots & \sum_{v \neq k} l_v \\ \sum_{\substack{v \neq 1 \\ v_1 < v_2}} l_{v_1} l_{v_2} & \vdots & \sum_{\substack{v \neq u \\ v_1 < v_2}} l_{v_1} l_{v_2} & \vdots & \sum_{\substack{v \neq k \\ v_1 < v_2}} l_{v_1} l_{v_2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ l_2 l_3 \cdots l_k & \cdot & l_1 \cdots \underset{v \neq u}{l_v} \cdots l_k & \cdot & l_1 l_2 \cdots l_{k-1} \end{bmatrix} \tag{4}$$

If $k = 2$, then $|\mathbf{B}| = (l_1 - l_2)$. Assuming for $k = m$ that $|\mathbf{B}| = \prod_{u<j}^{m} (l_u - l_j)$ it readily follows that for $k = m + 1$, $|\mathbf{B}| = \prod_{u<j}^{m+1} (l_u - l_j)$. Since the d_u are distinct the l_u are also distinct. Thus by induction $|\mathbf{B}| \neq 0$. This implies $\mathbf{C} = \mathbf{0}$ which asserts that the quantities d_u^{-1}; $(u = 1, \cdots k)$, are linearly independent.

In order to prove a result concerning completeness we prove the following.

LEMMA 2. *If the number of distinct characteristic roots of* $\mathbf{W}$ *is* $k + 2$ *then the distinct characteristic roots* $d_1 \cdots d_{k+2}$ *are functionally independent.*

PROOF: Consider the equation $\mathbf{PWP}' = \sum_{i=0}^{k+1} \sigma_i^2 \mathbf{PA}_i\mathbf{P}'$. Let $\mathbf{D}^*$ and $\mathbf{D}_i^*$ be the vectors of the diagonal elements of the diagonal matrices $\mathbf{PWP}'$ and $\mathbf{PA}_i\mathbf{P}'$ respectively. Then $\mathbf{D}^* = \sum_{i=0}^{k+1} \sigma_i^2 \mathbf{D}_i^* = (\mathbf{D}_0^*, \mathbf{D}_1^*, \cdots \mathbf{D}_{k+1}^*)\, \mathbf{\Sigma}$ where $\mathbf{\Sigma}' = (\sigma_0^2, \cdots, \sigma_{k+1}^2)$. Since the $\mathbf{A}_i$ are linearly independent matrices the $\mathbf{D}_i^*$ are linearly independent vectors which implies the matrix $(\mathbf{D}_0^*, \mathbf{D}_1^*, \cdots, \mathbf{D}_{k+1}^*)$ has rank $k + 2$. This together with the fact that $\mathbf{\Sigma}$ has $k + 2$ functionally independent elements implies $\mathbf{D}^*$ has $k + 2$ functionally independent elements. These clearly are the $k + 2$ distinct elements $d_1, \cdots, d_{k+2}$.

THEOREM 5. *If* $\mathbf{W}$ *has* $k + 2$ *distinct characteristic roots then the* $k + 2$ *statistics* $\mathbf{Y}'\mathbf{P}_u'\mathbf{P}_u\mathbf{Y}$; $(u = 2, \cdots, k + 2)$ *and* $\mathbf{P}_1\mathbf{Y}$ *form a complete sufficient set.*

PROOF: By applying the result of Lemma 2 to a theorem due to Gautschi [8] the result follows.

5. An Example. Consider the model $y_{ij} = \mu + \beta_i + \tau_j + e_{ij}$; $(i = 1, 2)$ $(j = 1, 2)$; $(i = 3, 4)(j = 3, 4)$. In matrix notation $\mathbf{Y} = \mu\mathbf{j} + \mathbf{X}_1\beta + \mathbf{X}_2\tau + \mathbf{e}$. Suppose $E(\boldsymbol{\beta}) = \mathbf{0}$, $E(\boldsymbol{\beta}\boldsymbol{\beta}') = \sigma_1^2\mathbf{I}$, $E(\tau) = \mathbf{0}$, $E(\tau\tau') = \sigma_2^2\mathbf{I}$, $E(\mathbf{e}) = \mathbf{0}$, $E(\mathbf{ee}') = \sigma_3^2\mathbf{I}$. The observation vector and the matrices can be written

$$\mathbf{Y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{33} \\ y_{34} \\ y_{43} \\ y_{44} \end{bmatrix}; \quad \mathbf{X} = [\mathbf{j}, \mathbf{X}_1, \mathbf{X}_2] = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$\mathbf{A}_0 = \mathbf{J}\ (8 \times 8)$; $\mathbf{A}_1 =$ Diag. $[\mathbf{J}, \mathbf{J}, \mathbf{J}, \mathbf{J}]$ where $\mathbf{J}$ here is (2×2); and $\mathbf{A}_2$ is Diag. $(\mathbf{M}, \mathbf{M})$ where

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Matrix multiplication will verify that $\mathbf{A}_0$, $\mathbf{A}_1$, and $\mathbf{A}_2$ commute in pairs. If we choose $\mathbf{P}$ to be

$$8^{\frac{1}{2}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

then $\mathbf{PA}_1\mathbf{P}'$ and $\mathbf{PA}_2\mathbf{P}'$ are diagonal and $\mathbf{PVP}'$ has the following characteristic roots each of multiplicity two: $2\sigma_1^2 + 2\sigma_2^2 + \sigma_3^2$, $2\sigma_2^2 + \sigma_3^2$, $2\sigma_1^2 + \sigma_3^2$, σ_3^2. Hence there are five statistics in a minimal set of sufficient statistics. Since we have four parameters it follows that these statistics are not complete.

6. A Theorem on the Analysis of Variance. It is well known that for Eisenhart's Model I [4] with n observations the total sum of squares can be partitioned into n sums of squares, each sum of squares being independently distributed as noncentral or central chi-square. For the model II case in which there appear $k + 2$ unknown parameters σ_i^2 $(i = 0, \cdots, k + 1)$ we make the definition:

DEFINITION 1. *An analysis of variance will be said to exist under assumption* (i) *if matrices* $\mathbf{B}_i$ *of known constants exist, such that*

(1) $\mathbf{Y}'\mathbf{Y} = \sum_{i=0}^{k+1} \mathbf{Y}'\mathbf{B}_i\mathbf{Y}$

(2) $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}/c_i$ $(i = 0, 1, \cdots, k + 1)$ is distributed as a noncentral chi-square variate with p_i degrees of freedom and noncentrality parameter λ_i.

(3) $\mathbf{B}_0 = \mathbf{J}/n$ and $p_0 = 1$.

(4) $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}$ $(i = 0, 1, \cdots, k + 1)$ are pairwise independent.

(5) The c_i $(i = 1, 2, \cdots, k + 1)$ are different linear functions of the parameters.

LEMMA 3. *If an analysis of variance exists then the* $\mathbf{B}_i$: (a) *Commute with* V; (b) *are idempotent*; (c) *are disjoint.*

PROOF: $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}$ being independent and distributed as chi-square implies $\mathbf{B}_i\mathbf{V}\mathbf{B}_i\mathbf{V} = c_i\mathbf{B}_i\mathbf{V}$ and $\mathbf{B}_i\mathbf{V}/c_i \cdot \mathbf{B}_h\mathbf{V}/c_h = \mathbf{0}$ $(i \neq h)$ [6]. Since $\mathbf{V}$ is nonsingular this implies $\mathbf{B}_i\mathbf{V}\mathbf{B}_i = c_i\mathbf{B}_i$ and $\mathbf{B}_i\mathbf{V}\mathbf{B}_h = \mathbf{0}$ $(i \neq h)$. Now $\sum_{h=0}^{k+1}\mathbf{B}_i\mathbf{V}\mathbf{B}_h = \sum_{h\neq i} \mathbf{B}_i\mathbf{V}\mathbf{B}_h + \mathbf{B}_i\mathbf{V}\mathbf{B}_i = c_i\mathbf{B}_i$ but $\sum_{h=0}^{k+1} \mathbf{B}_i\mathbf{V}\mathbf{B}_h = \mathbf{B}_i\mathbf{V}\sum_{h=0}^{k+1} \mathbf{B}_h = \mathbf{B}_i\mathbf{V}$ hence $\mathbf{B}_i\mathbf{V} = c_i\mathbf{B}_i$. Likewise summing over i instead of h we obtain $\mathbf{V}\mathbf{B}_h = c_h\mathbf{B}_h$. Together these results imply that the $\mathbf{B}_i$ commute with $\mathbf{V}$: $\mathbf{V}\mathbf{B}_i = \mathbf{B}_i\mathbf{V}\mathbf{B}_i = \mathbf{V}\mathbf{B}_i\mathbf{B}_i$. Hence $\mathbf{B}_i = \mathbf{B}_i\mathbf{B}_i$ and $\mathbf{B}_i$ is idempotent.

Since the $\mathbf{B}_i$ commutes with $\mathbf{V}$ and $\mathbf{V}$ is nonsingular, we have $\mathbf{B}_i\mathbf{V}\mathbf{B}_h = \mathbf{B}_i\mathbf{B}_h\mathbf{V} = \mathbf{0}$ hence $\mathbf{B}_i\mathbf{B}_h = \mathbf{0}$; $(i \neq h)$. Thus the $\mathbf{B}_i$ are disjoint.

THEOREM 6. *A necessary and sufficient condition for an analysis of variance to exist is* $\mathbf{A}_r$ *and* $\mathbf{A}_j$ $(r, j = 0, \cdots, k + 1)$ *commute and* $\mathbf{W}$ *has* $k + 2$ *distinct characteristic roots.*

PROOF OF THE NECESSITY STATEMENT: $\mathbf{B}_i$ and $\mathbf{V}$ commute as do $\mathbf{B}_i$ and $\mathbf{B}_0$ and since $\mathbf{W}$ can be written in the form $\mathbf{W} = \mathbf{V} + n\mu^2\mathbf{B}_0$, then $\mathbf{B}_i$ commutes with $\mathbf{W}$. We write $\mathbf{B}_i\mathbf{W} = \mathbf{B}_i\sum_{j=0}^{k+1} \sigma_j^2\mathbf{A}_j = (\sum_{j=0}^{k+1} \sigma_j^2\mathbf{A}_j)\mathbf{B}_i = c_i\mathbf{B}_i + \mu^2 n\mathbf{B}_0\mathbf{B}_i$. Equating coefficients of σ_j^2 we have $\mathbf{B}_i\mathbf{A}_j = \mathbf{A}_j\mathbf{B}_i = t_{ji}\mathbf{B}_i$ where t_{ji} are constants and not functions of the parameters. Summing over i we obtain $\mathbf{A}_j = \sum_i t_{ji}\mathbf{B}_i$. $\mathbf{A}_j\mathbf{A}_r = (\sum_i t_{ji}\mathbf{B}_i)(\sum_p t_{rp}\mathbf{B}_p) = \sum_i t_{ji}t_{ri}\mathbf{B}_i = \sum_i t_{ri}t_{ji}\mathbf{B}_i = \mathbf{A}_r\mathbf{A}_j$.

If we define c_i^* to equal c_i; $(i \neq 0)$ and c_0^* to equal $c_0 + n\mu^2$, then we can write $\mathbf{B}_i\mathbf{W} = c_i^*\mathbf{B}_i$; $(i = 0, 1, \cdots, k + 1)$. Consider then the equality $\mathbf{P}\mathbf{B}_i'\mathbf{P}'\mathbf{P}\mathbf{W}\mathbf{P}' = c_i^*\mathbf{P}\mathbf{B}_i\mathbf{P}'$ where $\mathbf{P}$ is orthogonal and simultaneously diagonalizes $\mathbf{B}_i$ and $\mathbf{W}$. Letting $\mathbf{D} = \mathbf{P}\mathbf{W}\mathbf{P}'$ and $\mathbf{D}_i = \mathbf{P}\mathbf{B}_i\mathbf{P}'$ we have $\mathbf{D}_i\mathbf{D} = c_i^*\mathbf{D}_i$. $\mathbf{B}_i$ being idempotent implies that the diagonal elements of $\mathbf{D}_i$ are unity or zero. Since the rank of $\mathbf{B}_i$ is p_i unity must appear p_i times in the diagonal elements of $\mathbf{D}_i$. Thus $\mathbf{D}_i\mathbf{D}$ is a diagonal matrix with p_i nonzero diagonal elements all equal to c_i^*. $\sum_{i=0}^{k+1} p_i = n$ [6]. This together with the fact that $p_i \geqq 1$ implies that the c_i^* are the characteristic roots of $\mathbf{W}$. The c_i were assumed to be distinct hence $\mathbf{W}$ has $k + 2$ distinct characteristic roots.

PROOF OF THE SUFFICIENCY STATEMENT: Let $\mathbf{P}$ be orthogonal with first row

$\mathbf{j}'n^{-\frac{1}{2}}$. $\mathbf{Z} = \mathbf{PY}$ is distributed as an n variate normal with a mean vector containing zeros except for the first element equal to $n^{\frac{1}{2}}\mu$. Let $\mathbf{E}_\nu$ be the matrix with unity in the νth diagonal place and zeros elsewhere. Let $\mathbf{D} = \mathbf{PVP}'$ be the diagonal covariance matrix of $\mathbf{Z}$ and let d_ν be the νth diagonal element of $\mathbf{D}$. $\mathbf{Z}'\mathbf{E}_\nu\mathbf{Z}/d_\nu$; $(\nu = 2, \cdots n)$ is distributed as a central chi-square variate with one degree of freedom and $\mathbf{Z}'\mathbf{E}_1\mathbf{Z}/d_1$ is distributed as a noncentral chi-square variate with one degree of freedom and noncentrality parameter $n\mu^2/2d_1$. Since the d_ν are the characteristic roots of $\mathbf{V}$ the characteristic roots of $\mathbf{W}$ are then $d_1 + \mu^2$ and d_ν ; $(\nu = 2, \cdots n)$. Let $d_1 + \mu^2$ and b_i ; $(i = 1, \cdots k + 1)$ be the $k + 2$ distinct characteristic roots of $\mathbf{W}$. Let S_i be the set of ν where $d_\nu = b_i$ and let p_i be the number of roots equal to b_i. Then

$$\sum_{\nu \varepsilon S_i} \mathbf{Z}'\mathbf{E}_\nu\mathbf{Z}/b_i = \mathbf{Z}'\sum_{\nu \varepsilon S_i} \mathbf{E}_\nu\mathbf{Z}/b_i = \mathbf{Z}'\mathbf{F}_i\mathbf{Z}/b_i = \mathbf{Y}'\mathbf{P}'\mathbf{F}_i\mathbf{PY}/b_i \tag{5}$$

where $\mathbf{F}_i$ is defined by the equation. This statistic, since it is the sum of p_i independent chi-square variates, is itself a chi square variate with p_i degrees of freedom. If we let $\mathbf{B}_i = \mathbf{P}'\mathbf{F}_i\mathbf{P}$ condition (2) of the definition is satisfied for $i = 1, \cdots k + 1$ and letting $\mathbf{B}_0 = \mathbf{P}'\mathbf{E}_1\mathbf{P}$ we have condition (2) satisfied for $i = 0$. Since $\mathbf{B}_0 = [n^{-\frac{1}{2}}j, 0]\mathbf{P} = \mathbf{J}/n$ has rank one we have condition (3) satisfied. The b_i were defined to be distinct characteristic roots of $\mathbf{W}$ thus satisfying condition (5). Since $\sum_{i=0}^{k+1} \mathbf{Y}'\mathbf{B}_i\mathbf{Y} = \sum_{i=0}^{k+1} \mathbf{Y}'\mathbf{P}'\mathbf{F}_i\mathbf{PY} = \sum_{i=0}^{k+1} \mathbf{Z}'\mathbf{F}_i\mathbf{Z} = \sum_{\nu=1}^{n} \mathbf{Z}'\mathbf{E}_\nu\mathbf{Z} = \mathbf{Z}'\mathbf{Z} = \mathbf{Y}'\mathbf{Y}$, condition (1) is satisfied. Condition (4) is satisfied by applying Theorem 5, page 684 [6]. Therefore an analysis of variance exists.

The following corollaries follow from Theorem 6.

COROLLARY 1. *The terms c_i which appear in the analysis of variance are the distinct characteristic roots of the covariance matrix* $\mathbf{V}$.

COROLLARY 2. *The quadratic forms $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}/c_i$ are central chi-square variates with p_i degrees of freedom.* $(i \geqq 1)$

COROLLARY 3. *The $\mathbf{A}_j$ are linear combinations of the $\mathbf{B}_i$.*

COROLLARY 4. *The $\mathbf{B}_i$ are linear combinations of the $\mathbf{A}_j$.*

7. Best Quadratic Unbiased Estimators. In this section quadratic estimates of variance components are considered. Hsu [1] under certain conditions has shown that the best (minimum variance) quadratic unbiased estimate of σ_e^2 is given by the analysis of variance method of estimating σ_e^2. Graybill [9] has shown for the general balanced nested classification in the Model II situation that the method in [10] gives best unbiased estimates. In this paper we state conditions under which the best quadratic unbiased estimates of variance components can be obtained from the analysis of variance.

THEOREM 7. *If under assumption* (i) *the following analysis of variance exists for a vector* Y *of observations: Sum of Squares* $= \mathbf{Y}'\mathbf{B}_i\mathbf{Y}$; $E(\mathbf{Y}'\mathbf{B}_i\mathbf{Y}) = \alpha_i^2$; $i = 0, 1, \cdots, k + 1$; *then $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}$ is the uniformly best quadratic unbiased estimate of α_i^2 under assumptions* (ii) *and* (iv).

This theorem states that if an analysis of variance exists under the assumption of normality for the random variables, then uniformly best quadratic unbiased estimates of the parameters exist when less stringent assumptions (ii) and (iv) are imposed in place of the assumption of normality.

PROOF: Let the general quadratic estimate of α_i^2 be $\hat{\alpha}_i^2$, let the symmetric matrix $\mathbf{C}_i$ be defined by the equation $\hat{\alpha}_i^2 = \mathbf{Y}'\mathbf{B}_i\mathbf{Y} + \mathbf{Y}'\mathbf{C}_i\mathbf{Y}$, and let the elements of $\mathbf{C}_i$ be constants. We wish to restrict this general quadratic estimate to the class of all unbiased estimates and then obtain the estimate with variance less than that of any other unbiased estimate of α_i^2.

Unbiasedness implies that $E(\mathbf{Y}'\mathbf{C}_i\mathbf{Y}) = 0$ and "best" implies $E[\hat{\alpha}_i^2]^2 = E[\mathbf{Y}'\mathbf{B}_i\mathbf{Y}]^2 + 2E[\mathbf{Y}'\mathbf{B}_i\mathbf{Y}]\,[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}] + E[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}]^2$ is a minimum. By straightforward evaluation of the expected values involved it can be shown that if $E(\mathbf{Y}'\mathbf{B}_i\mathbf{Y}) = 0$, then $E[\mathbf{Y}'\mathbf{B}_i\mathbf{Y}]\,[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}] = 0$. We shall not present the numerous details of the proof of this statement but using this fact we can write $E[\hat{\alpha}_i^2]^2 = E[\mathbf{Y}'\mathbf{B}_i\mathbf{Y}]^2 + E[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}]^2$. $E[\hat{\alpha}_i^2]^2$ then takes on its minimum value when $E[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}]^2 = 0$. $E[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}]^2$ and $E[\mathbf{Y}'\mathbf{C}_i\mathbf{Y}]$ both equal to zero implies $\mathbf{C}_i = 0$. Hence the best quadratic unbiased estimate of α_i^2 is $\mathbf{Y}'\mathbf{B}_i\mathbf{Y}$.

8. An Example. An examination of the matrices for various experimental designs reveals that most of the commonly used designs with equal numbers in the subclasses possess the conditions of Theorem 7. Consider the randomized block design with interaction having b blocks of t treatments. The treatments and blocks can be labeled in such a way that in the model $\mathbf{Y} = \mu\mathbf{j} + \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\tau + \mathbf{e}$; $\mathbf{X}_1(bt \times b) = \text{Diag.}\,[\mathbf{j}_t, \mathbf{j}_t, \cdots, \mathbf{j}_t]$; $\mathbf{X}_2'(bt \times b) = [\mathbf{I}_t, \mathbf{I}_t, \cdots, \mathbf{I}_t]$; $\mathbf{A}_1(bt \times bt) = \text{Diag.}\,[\mathbf{J}_t, \mathbf{J}_t, \cdots, \mathbf{J}_t]$; and $\mathbf{A}_2(bt \times bt) = [\mathbf{X}_2, \mathbf{X}_2, \cdots, \mathbf{X}_2]$. Matrix multiplication will verify that $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{J}$ commute. It then follows that $\mathbf{W} = \mu^2\mathbf{J} + \sigma_1^2\mathbf{A}_1 + \sigma_2^2\mathbf{A}_2 + \sigma_3^2\mathbf{I}$, where $\sigma_1^2\mathbf{I} = E(\boldsymbol{\beta}\boldsymbol{\beta}')$, $\sigma_2^2\mathbf{I} = E(\boldsymbol{\tau}\boldsymbol{\tau}')$ and $\sigma_3^2\mathbf{I} = E(\mathbf{ee}')$. The characteristic roots of $\mathbf{W}$ can be shown to be σ_3^2, $t\sigma_1^2 + \sigma_3^2$, $b\sigma_2^2 + \sigma_3^2$, and $tb\mu^2 + t\sigma_1^2 + b\sigma_2^2 + \sigma_3^2$. Since k in this model is 2 the number $k + 2 = 4$ agrees with the number of distinct characteristic roots. Thus if $\boldsymbol{\beta}$, $\boldsymbol{\tau}$, and $\mathbf{e}$ have distributions satisfying assumption (ii), then minimum variance quadratic estimates of σ_1^2, σ_2^2, and σ_3^2 can be obtained by the analysis of variance technique.

9. Estimable Functions. In this section we shall define estimable functions for our model and give a necessary and sufficient condition for the σ_i^2 to be estimable.

DEFINITION 2. *The parameter σ_s^2 is said to be estimable if a quadratic form $\mathbf{Y}'\mathbf{B}_s^*\mathbf{Y}$ exists such that $E[\mathbf{Y}'\mathbf{B}_s^*\mathbf{Y}] = \sigma_s^2$.*

THEOREM 8. *A necessary and sufficient condition that the σ_s^2 are estimable is that the $\mathbf{A}_s$ are linearly independent.*

PROOF. If the σ_s^2 are estimable there exists matrices $\mathbf{B}_s^*$; $(s = 1, \cdots k + 1)$ such that $E[\sum \mathbf{X}_i\boldsymbol{\beta}_i]'\mathbf{B}_s[\sum \mathbf{X}_i\boldsymbol{\beta}_i] = \sigma_s^2$. It then follows that $\sum \sigma_i^2 \operatorname{tr} \mathbf{X}_i'\mathbf{B}_s^*\mathbf{X}_i = \sum \sigma_i^2 \operatorname{tr} \mathbf{A}_i\mathbf{B}_s^* = \sigma_s^2$. If the coefficients of σ_p^2 are equated we obtain $\operatorname{tr} \mathbf{A}_i\mathbf{B}_s^* = 0$, $(i \neq s)$ and $\operatorname{tr} \mathbf{A}_s\mathbf{B}_s^* = 1$. Now let $c_0^*, c_1^*, \cdots c_{k+1}^*$ be any set of constants such that $\sum_{i=0}^{k+1} c_i^*\mathbf{A}_i = \mathbf{0}$, then $\sum_{i=0}^{k+1} c_i^* \operatorname{tr} \mathbf{A}_i\mathbf{B}_s = c_s^*$, hence $\operatorname{tr} \mathbf{B}_s^*\left(\sum_{i=0}^{k+1} c_i^*\mathbf{A}_i\right) = c_s^*$ which implies $c_s^* = 0$. But if $c_s^* = 0$; $(s = 1, \cdots k + 1)$, then since $\sum_{i=0}^{k+1} c_i^*\mathbf{A}_i = \mathbf{O}$ we also have $c_0^* = 0$, which implies the $\mathbf{A}_i$ are linearly independent.

To prove that the σ_i^2 are estimable consider $E(\mathbf{YY}') = \sum_{i=0}^{k+1} \mathbf{A}_i\sigma_i^2$. Define $z_{rq} = y_r y_q$ and let the $[n(n+1)/2] \times 1$ vector $\mathbf{Z}$ be defined by

$$\mathbf{Z}' = (z_{11}, \cdots z_{1p}, z_{22}, \cdots z_{2p}, \cdots z_{pp})'.$$

$\mathbf{Z}$ has as elements the quantities on and above the main diagonal of $\mathbf{YY}'$ ordered in a particular fashion. Now let the rqth element of $\mathbf{A}_i$ be denoted by a_{rq}^i and let the $[n(n+1)/2] \times 1$ vector $\boldsymbol{\alpha}_i = (a_{11}^i, a_{12}^i, \cdots a_{1p}^i, a_{22}^i, \cdots a_{2p}^i, \cdots a_{pp}^i)'$.

The expected value of $\mathbf{Z}$ is $\sum_{i=0}^{k+1} \sigma_i^2\boldsymbol{\alpha}_i$. By hypothesis the $\mathbf{A}_i$ are linearly independent, thus since the elements of $\boldsymbol{\alpha}_i$ are elements in $\mathbf{A}_i$, the $\boldsymbol{\alpha}_i$ are also linearly independent. Denoting the $[n(n+1)/2] \times (k+2)$ matrix $[\boldsymbol{\alpha}_0, \cdots \boldsymbol{\alpha}_{k+1}]$ by $\boldsymbol{\alpha}$ and the vector $(\sigma_0^2, \cdots \sigma_{k+1}^2)'$ by $\boldsymbol{\Sigma}$ we can write $E(\mathbf{Z}) = \boldsymbol{\alpha\Sigma}$. $\boldsymbol{\alpha}$ has column rank $k+2$ and hence has row rank $k+2$. Let $\boldsymbol{\alpha}^*$ be the $(k+2) \times (k+2)$ matrix which consists of $k+2$ linearly independent rows of $\boldsymbol{\alpha}$. Let $\mathbf{Z}^*$ be the corresponding rows of $\mathbf{Z}$, then $E(\mathbf{Z}^*) = \boldsymbol{\alpha}^*\boldsymbol{\Sigma}$. Now $\boldsymbol{\alpha}^*$ has an inverse so that $(\boldsymbol{\alpha}^*)^{-1}E(\mathbf{Z}^*) = \boldsymbol{\Sigma}$. Thus $(\boldsymbol{\alpha}^*)^{-1}\mathbf{Z}^*$ is an unbiased estimate of $\boldsymbol{\Sigma} = [\sigma_i^2]$. This completes the proof.

Of course, if the $\mathbf{A}_i$ are linearly independent, then this implies certain conditions on the $\mathbf{X}_i$, but this will not be discussed here.

REFERENCES

[1] P. L. Hsu, "On the best unbiased quadratic estimate of the variance," *London Univ. Stat. Research Memoirs*, Vol. 2 (1938), pp. 91–104.

[2] Franklin A. Graybill and A. W. Wortham, "A note on uniformly best unbiased estimators for variance components," *J. Amer. Stat. Assn.*, Vol. 51 (1956), pp. 266–268.

[3] S. Lee Crump, "The estimation of variance components in the analysis of variance," *Biometrics Bull.*, Vol. 2 (1946), pp. 7–11.

[4] Churchill Eisenhart, "The assumptions underlying the analysis of variance," *Biometrics*, Vol. 3 (1947), pp. 1–21.

[5] B. O. Koopman, "On distributions admitting a sufficient statistic," *Transactions. Amer. Math. Soc.*, Vol. 39 (1936), pp. 399–409.

[6] Franklin A. Graybill and George Marsaglia, "Idempotent matrices and quadratic forms in the general linear hypothesis," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 678–686.

[7] E. L. Lehmann and Henry Scheffé, "Completeness, similar regions, and unbiased estimation, part I," *Sankhyā*, Vol. 10 (1950), pp. 305–340.

[8] Werner Gautschi, "Some remarks on Herbach's paper, 'Optimum nature of the F-test for model II in the balanced case,'" *Ann. Math. Stat.*, Vol. 30 (1959), pp. 960–963.

[9] Franklin A. Graybill, "On quadratic estimates of variance components," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 367–372.

[10] S. Lee Crump, "The present status of variance components," *Biometrics*, Vol. 7 (1951), pp. 1–16.

[11] Robert M. Thrall and Leonard Tornheim, *Vector Spaces and Matrices*, John Wiley and Sons, New York, 1957.