Statistical Research Report
Institute of Mathematics
University of Oslo

No. 1

May 1965

A MIXED MODEL IN THE ANALYSIS OF VARIANCE.

OPTIMAL PROPERTIES.*

By

Emil Spjøtvoll

* Revised March 1967.

0. SUMMARY. A mixed model in the analysis of variance is defined. It contains the general fixed effects model and the general balanced nested classification model as special cases. For the model is given a sufficient condition for transformation to a canonical form of independent random variables. For the parameters of the model are given minimum variance unbiased estimates. There exist UMP unbiased test for certain hypotheses concerning the variances and there exist UMP tests among all invariant and unbiased tests for linear hypotheses concerning the fixed effects. All tests and estimates are expressed by means of the matrices defining the model. In Section 8 it is shown that the estimates of the fixed effects are maximum likelihood estimates, but in general this is not true for the estimates of the variances, and therefore in general the tests are not likelihood ratio tests.

1. NOTATION. The symbols $E(X)$ and $\sum(X)$ are used for the expectation vector and the covariance matrix of a random vector X . Vectors will always be column vectors. For a matrix A the symbols $\mathcal{C}(A)$ and $\mathcal{R}(A)$ denote the spaces spanned by the columns and the rows of A respectively. A' denotes the transpose of A and A^{-1} denotes the inverse of A . $A^{m \times n}$ or A is of type $(m \times n)$ denote that A is a matrix with m rows and n columns. The symbol I is used for the identity matrix.

2. DEFINITION OF THE MODEL. Let X be a random $(n \times 1)$ -vector given by

$$(I) \quad X = A_1 \beta_1 + A_2 \beta_2 + \dots + A_p \beta_p + C_1 V_1 + C_2 V_2 + \dots + C_p V_p$$

where

$$\left. \begin{array}{l} A_i \text{ is a known } (n \times r_i)\text{-matrix} \\ C_i \text{ is a known } (n \times s_i)\text{-matrix} \\ \beta_i \text{ is an unknown } (r_i \times 1)\text{-parameter-vector} \\ V_i \text{ is a random } (s_i \times 1)\text{-vector} \end{array} \right\} i=1, 2, \dots, p,$$

and where the following conditions are imposed on the matrices

$$\left. \begin{array}{ll} (II) & \text{rank } A_i = r_i \\ (III) & \text{rank } C_i = s_i \\ (IV) & \mathcal{C}(A_i) \subset \mathcal{C}(C_i) \end{array} \right\} i = 1, 2, \dots, p,$$

$$\left. \begin{array}{ll} (V) & \mathcal{C}(C_i) \subset \mathcal{C}(C_{i+1}) \\ (VI) & \mathcal{C}(A_{i+1}) \perp \mathcal{C}(C_i) \end{array} \right\} i = 1, 2, \dots, p-1.$$

$V_1, V_2, \dots, V_p$ are furthermore assumed to be independently distributed according to multinormal distributions with expectations $E(V_i) = 0$ and covariance matrices $\sum(V_i) = \sigma_i^2 I$ ($i = 1, 2, \dots, p$) respectively, where $\sigma_1, \sigma_2, \dots, \sigma_p$ are unknown parameters.

A priori for all i the unknown parameter vector β_i may attain any value in an r_i -dimensional Euclidean space.

To avoid a singular multinormal distribution for X

it is assumed that $s_p = n$.

Example. The following example, taken from the author's own practical experience, will be used to illustrate the general theory.

One wants to investigate the variation of the specific weight of wood with the height above sea-level of the growing place, and the mean breadth of the annual rings of the tree. I different growing areas are selected and J trees are taken from each area. The specific weight of the j th tree in the i th area is denoted X_{ij} and one assume the following model

$$X_{ij} = \mu + \beta t_i + U_i + \gamma u_{ij} + V_{ij} \quad i = 1, 2, \dots, I, j = 1, 2, \dots, J,$$

where μ is an unknown parameter, t_i is the elevation of the i th area, and u_{ij} is the mean breadth of the annual rings of the j th tree in the i th area. The parameters β and γ express the increase (or decrease) of specific weight with elevation and mean breadth respectively. Since the areas are taken out of a larger population of growing areas, we let to each area be attached a random effect U_i . Similarly we attach random effects V_{ij} to the trees. We suppose that $U_1, U_2, \dots, U_I$ are independent $N(0, \sigma^2)$, and independent of $V_{11}, V_{12}, \dots, V_{IJ}$ which are independent $N(0, \tau^2)$. We can write

$$X_{ij} = \bar{\mu} + \beta(t_i - \bar{t}) + U_i + \gamma(u_{ij} - \bar{u}) + V_{ij}$$

where $t_{.} = \frac{1}{I} \sum t_i$, $u_{..} = \frac{1}{IJ} \sum u_{ij}$ and $\bar{\mu} = \mu + \beta t_{.} + \gamma u_{..}$

This is a special case of (I) with

$$p = 2, \quad r_1 = 2, \quad s_1 = I, \quad r_2 = 1, \quad s_2 = n = IJ,$$

$$X' = (X_{11}, X_{12}, \dots, X_{IJ}),$$

$$\beta'_1 = (\bar{\mu}, \beta),$$

$$\beta'_2 = (\gamma),$$

$$(\sigma'_1, \sigma'_2) = (\sigma, \tau),$$

$$V'_1 = (U_1, U_2, \dots, U_I),$$

$$V'_2 = (V_{11}, V_{12}, \dots, V_{IJ}),$$

$$A'_1 = \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ \underbrace{t_1 - t_{.} \quad \dots \quad t_1 - t_{.}}_{J \text{ columns}} & t_2 - t_{.} & \dots & t_I - t_{.} \end{bmatrix},$$

J columns

$$A'_2 = (u_{11} - u_{..}, u_{12} - u_{..}, \dots, u_{IJ} - u_{..}),$$

$$C'_1 = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \end{bmatrix},$$

J columns

$$C_2 = I^{n \times n}.$$

The condition (II) of the model is satisfied unless all t_i or all u_{ij} are equal. Conditions (III), (IV) and (V) are easily seen to be satisfied. Condition (VI) implies that we must have $\sum_{j=1}^J (u_{ij} - u_{..}) = 0$ for all i , and that requires $u_{1.} = u_{2.} = \dots = u_{I.} = u_{..}$ where $u_{i.} = \frac{1}{J} \sum_{j=1}^J u_{ij}$. We shall in the following assume this to hold.

3. CANONICAL FORM. In this section we shall prove a theorem giving conditions under which the model can be transformed to a canonical form of independent variables by means of an orthogonal transformation.

THEOREM 1. Suppose that all eigenvalues not equal to 0 of the matrices $C_i C_i'$ are identical, k_i say ($i = 1, 2, \dots, p$). Then there exists an orthogonal matrix P

$$P = \begin{bmatrix} P_{11} \\ P_{12} \\ \vdots \\ P_{p1} \\ P_{p2} \end{bmatrix} \quad \text{where} \quad \begin{array}{ll} P_{i1} & \text{is an } (r_i \times n)\text{-matrix} \\ P_{i2} & \text{is an } ((s_i - s_{i-1} - r_i) \times n)\text{-matrix } (s_0 = 0) \end{array}$$

$i = 1, 2, \dots, p,$

such that the transformation

$$Z = PX$$

gives the following canonical form

$$Z = \begin{bmatrix} Z_{11} \\ Z_{12} \\ \vdots \\ Z_{p1} \\ Z_{p2} \end{bmatrix} \quad \text{where } Z_{ij} = P_{ij}X \quad i = 1, 2, \dots, p, \quad j = 1, 2.$$

Z has a multinormal distribution. All elements of Z are independent, and

$$\left. \begin{aligned} E(Z_{i1}) &= P_{i1}A_i\beta_i \\ E(Z_{i2}) &= 0 \\ \Sigma \begin{pmatrix} Z_{i1} \\ Z_{i2} \end{pmatrix} &= \left(\sum_{\alpha=i}^p k_{\alpha} \sigma_{\alpha}^2 \right) I \end{aligned} \right\} \quad i = 1, 2, \dots, p.$$

PROOF. For any matrix P the transformation $Z = PX$ gives a vector Z which has a multinormal distribution and which is defined by

$$(3.1) \quad Z = PA_1\beta_1 + \dots + PA_p\beta_p + PC_1V_1 + \dots + PV_pC_p.$$

We find

$$(3.2) \quad E(Z) = PA_1\beta_1 + \dots + PA_p\beta_p,$$

$$(3.3) \quad \Sigma(Z) = PC_1C_1'P'\sigma_1^2 + \dots + PC_pC_p'P'\sigma_p^2.$$

We shall now prove that it is possible to construct a matrix P that gives the canonical form of the theorem.

Let P_{11} be a matrix with orthogonal rows such that $\mathcal{R}(P_{11}) = \mathcal{C}(A_1)$. A_1 is of rank r_1 and is an $(n \times r_1)$ -matrix.

It follows that P_{11} is an $(r_1 \times n)$ -matrix. We have

$\mathcal{R}(P_{11}) = \mathcal{C}(A_1) \subset \mathcal{C}(C_1)$. Then there exists an $(s_1 - r_1) \times n$ -matrix P_{12} with orthogonal rows such that $\mathcal{R}(P_{12}) \perp \mathcal{R}(P_{11})$ and the row space of $\begin{bmatrix} P_{11} \\ P_{12} \end{bmatrix}$ is equal to $\mathcal{C}(C_1)$.

Next let P_{21} be an $(r_2 \times n)$ -matrix with orthogonal rows such that $\mathcal{R}(P_{21}) = \mathcal{C}(A_2)$. Since $\mathcal{C}(A_2) \perp \mathcal{C}(C_1) =$

$\mathcal{R}\left(\begin{bmatrix} P_{11} \\ P_{12} \end{bmatrix}\right)$, we have $\mathcal{R}(P_{21}) \perp \mathcal{R}\left(\begin{bmatrix} P_{11} \\ P_{12} \end{bmatrix}\right)$. Hence the matrix

$$\begin{bmatrix} P_{11} \\ P_{12} \\ P_{21} \end{bmatrix}$$

has orthogonal rows, and since $\mathcal{C}(A_2) \subset \mathcal{C}(C_2)$ and $\mathcal{C}(C_1) \subset \mathcal{C}(C_2)$ we have

$$\mathcal{R}\left(\begin{bmatrix} P_{11} \\ P_{12} \\ P_{21} \end{bmatrix}\right) \subset \mathcal{C}(C_2).$$

There exists an $((s_2 - s_1 - r_1) \times n)$ -matrix P_{22} with orthogonal rows such that

$$\mathcal{R}(P_{22}) \perp \mathcal{R}\left(\begin{bmatrix} P_{11} \\ P_{12} \\ P_{21} \end{bmatrix}\right) \quad \text{and} \quad \mathcal{R}\left(\begin{bmatrix} P_{11} \\ P_{12} \\ P_{21} \\ P_{22} \end{bmatrix}\right) = \mathcal{C}(C_2).$$

The matrix

$$\begin{bmatrix} P_{11} \\ P_{12} \\ P_{21} \\ P_{22} \end{bmatrix}$$

has orthogonal rows.

Proceeding in the same way we can prove that it is possible to construct an orthogonal matrix with the following properties

$$P = \begin{bmatrix} P_{11} \\ P_{12} \\ \vdots \\ P_{p1} \\ P_{p2} \end{bmatrix} \quad \text{where} \quad \begin{array}{l} P_{i1} \text{ is an } (r_i \times n)\text{-matrix} \\ P_{i2} \text{ is an } ((s_i - s_{i-1} - r_i) \times n)\text{-matrix} \\ i = 1, 2, \dots, p \end{array}$$

and where

$$\left. \begin{array}{l} R(P_{i1}) = \mathcal{C}(A_i) \\ R \left(\begin{bmatrix} P_{11} \\ P_{12} \\ \vdots \\ P_{i1} \\ P_{i2} \end{bmatrix} \right) = \mathcal{C}(C_i) \end{array} \right\} \quad i = 1, 2, \dots, p.$$

Consider the matrix product $P_{i1}A_\alpha$ for $i \neq \alpha$. For $\alpha < i$ we have $R(P_{i1}) = \mathcal{C}(A_i) \perp \mathcal{C}(C_\alpha) \supset \mathcal{C}(A_\alpha)$. For $\alpha > i$ we have

$\mathcal{R}(P_{i1}) = \mathcal{C}(A_i) \subset \mathcal{C}(C_i) \perp \mathcal{C}(A_\alpha)$. Therefore in both cases we have $\mathcal{R}(P_{i1}) \perp \mathcal{C}(A_\alpha)$, and it follows that $P_{i1}A_\alpha = 0$ when $i \neq \alpha$.

Since $\mathcal{R}(P_{i2}) \cap \mathcal{R}(P_{\alpha 1}) = \mathcal{C}(A_\alpha)$ for all i and α , $P_{i2}A_\alpha = 0$ for all i and α . Hence

$$PA_i\beta_i = \begin{bmatrix} P_{11}A_i\beta_i \\ P_{12}A_i\beta_i \\ \vdots \\ P_{i1}A_i\beta_i \\ P_{i2}A_i\beta_i \\ \vdots \\ P_{p1}A_i\beta_i \\ P_{p2}A_i\beta_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ P_{i1}A_i\beta_i \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

Introducing this into equation (3.2) we get

$$E(Z) = E \begin{bmatrix} Z_{11} \\ Z_{12} \\ \vdots \\ Z_{p1} \\ Z_{p2} \end{bmatrix} = \begin{bmatrix} P_{11}A_1\beta_1 \\ 0 \\ \vdots \\ P_{p1}A_p\beta_p \\ 0 \end{bmatrix}$$

where Z_{ij} is defined by $Z_{ij} = P_{ij}X$ ($i = 1, 2, \dots, p$, $j = 1, 2$). This shows that the expected value of Z is of the form given in the theorem.

To find the covariance matrix of Z we examine the matrix product $P_{il}C_\alpha C_\alpha' P_{jm}'$ for different values of i, j, l, m and α .

1) maximum $(i, j) > \alpha$. Suppose $i > \alpha$. Then

$\mathcal{R}(P_{il}) \perp \mathcal{L}(C_\alpha)$, implying $P_{il}C_\alpha = 0$, and thereby $P_{il}C_\alpha C_\alpha' P_{jm}' = 0$ for all j . We get the same result if $j > \alpha$.

2) maximum $(i,j) \leq \alpha$ but $(i,l) \neq (j,m)$. We have $\mathcal{R}(P_{il}) \subset \mathcal{L}(C_\alpha)$ and $\mathcal{R}(P_{jm}) \subset \mathcal{L}(C_\alpha)$. Since the eigenvalues different from 0 for $C_\alpha C_\alpha'$ are identical equal to k_α , the rows of P_{il} and P_{jm} will, according to Lemma 2 in the Appendix, be eigenvectors for $C_\alpha C_\alpha'$ with eigenvalue k_α . The rows of P_{il} and P_{jm} are orthogonal since $(i,l) \neq (j,m)$. We get $P_{il}C_\alpha C_\alpha' P_{jm}' = k_\alpha P_{il}P_{jm}' = 0$.

3) $i = j \leq \alpha$ and $l = m$. The matrix product is now $P_{im}C_\alpha C_\alpha' P_{im}'$. As above it follows that the rows of P_{im} are eigenvectors for $C_\alpha C_\alpha'$ with eigenvalue k_α , and we get $P_{im}C_\alpha C_\alpha' P_{im}' = k_\alpha I$.

We have

$$P C_\alpha C_\alpha' P' = \begin{bmatrix} P_{11}C_\alpha C_\alpha' P_{11}' & P_{11}C_\alpha C_\alpha' P_{12}' & \dots & P_{11}C_\alpha C_\alpha' P_{p2}' \\ P_{12}C_\alpha C_\alpha' P_{11}' & P_{12}C_\alpha C_\alpha' P_{12}' & \dots & P_{12}C_\alpha C_\alpha' P_{p2}' \\ \vdots & & & \vdots \\ P_{p2}C_\alpha C_\alpha' P_{11}' & P_{p2}C_\alpha C_\alpha' P_{12}' & \dots & P_{p2}C_\alpha C_\alpha' P_{p2}' \end{bmatrix}.$$

From what is known about the product $P_{il}C_\alpha C_\alpha' P_{jm}'$ it follows that

$$P C_\alpha C_\alpha' P' = k_\alpha \begin{bmatrix} I_{s_\alpha \times s_\alpha} & 0 \\ 0 & 0_{(n-s_\alpha) \times (n-s_\alpha)} \end{bmatrix}.$$

From equation (3.3) we find the covariance matrix of

Z

$$\Sigma(Z) = \sum_{\alpha=1}^p \sigma_{\alpha}^2 P C_{\alpha} C_{\alpha}' P' = \sum_{\alpha=1}^p k_{\alpha} \sigma_{\alpha}^2 \begin{bmatrix} I_{s_{\alpha} \times s_{\alpha}} & 0 \\ 0 & 0_{(n-s_{\alpha}) \times (n-s_{\alpha})} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 I_{s_1 \times s_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{(s_2-s_1) \times (s_2-s_1)} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \lambda_p I_{(n-s_{p-1}) \times (n-s_{p-1})} \end{bmatrix}$$

where $\lambda_i = k_i \sigma_i^2 + k_{i+1} \sigma_{i+1}^2 + \dots + k_p \sigma_p^2$ ($i = 1, 2, \dots, p$). This matrix is diagonal. All the elements of the vector Z are therefore independent.

From the definition of $\begin{bmatrix} Z_{i1} \\ Z_{i2} \end{bmatrix}$ it follows that the covariance matrix of this subvector of Z is $(k_i \sigma_i^2 + k_{i+1} \sigma_{i+1}^2 + \dots + k_p \sigma_p^2) I_{(s_i - s_{i-1}) \times (s_i - s_{i-1})}$. The theorem is thereby proved.

Example (cont.) We have

$$C_1 C_1' = \begin{bmatrix} 11 & \dots & 100 & \dots & 0 & \dots & 00 & \dots & 0 \\ \vdots & & & & & & & & \vdots \\ 11 & \dots & 100 & \dots & 0 & \dots & 00 & \dots & 0 \\ 00 & \dots & 011 & \dots & 1 & \dots & 00 & \dots & 0 \\ \vdots & & & & & & & & \vdots \\ 00 & \dots & 011 & \dots & 1 & \dots & 00 & \dots & 0 \\ \vdots & & & & & & & & \vdots \\ 00 & \dots & 000 & \dots & 0 & \dots & 11 & \dots & 1 \end{bmatrix}.$$

This matrix has eigenvalues given by J with multiplicity I and 0 with multiplicity $I(J-1)$. The matrix $C_2 C_2' = I^{IJ \times IJ}$ has eigenvalue 1 of multiplicity IJ . Hence the condition of Theorem 1 is satisfied and there exists a reduction to a canonical form with $k_1 = J$ and $k_2 = 1$.

Let P be the matrix of the transformation and Z the transformed vector. A determination of the matrices P_{11} and P_{21} is

$$P_{11} = \begin{bmatrix} \frac{1}{\sqrt{IJ}}, & \frac{1}{\sqrt{IJ}}, & \dots, & \frac{1}{\sqrt{IJ}} \\ \frac{t_1 - t.}{\sqrt{J \sum (t_i - t.)^2}}, & \frac{t_1 - t.}{\sqrt{J \sum (t_i - t.)^2}}, & \dots, & \frac{t_I - t.}{\sqrt{J \sum (t_i - t.)^2}} \end{bmatrix}$$

and

$$P_{21} = \left(\frac{u_{11} - u..}{\sqrt{\sum (u_{ij} - u..)^2}}, \frac{u_{12} - u..}{\sqrt{\sum (u_{ij} - u..)^2}}, \dots, \frac{u_{IJ} - u..}{\sqrt{\sum (u_{ij} - u..)^2}} \right).$$

We get

$$E(Z_{11}) = P_{11} A_1 \beta_1 = \begin{bmatrix} \sqrt{IJ} & 0 \\ 0 & \sqrt{J \sum (t_i - t_{..})^2} \end{bmatrix} \begin{bmatrix} \mu \\ \beta \end{bmatrix} = \begin{bmatrix} \mu \sqrt{IJ} \\ \beta \sqrt{J \sum (t_i - t_{..})^2} \end{bmatrix},$$

$$E(Z_{21}) = P_{21} A_2 \beta_2 = \gamma \sqrt{\sum (u_{ij} - u_{..})^2}.$$

Let Z_{ijk} denote the k th element of Z_{ij} . We have

$$\text{Var}(Z_{1jk}) = J\sigma^2 + \tau^2 \quad \begin{array}{l} j = 1, \quad k = 1, 2, \\ j = 2, \quad k = 1, 2, \dots, I-2, \end{array}$$

$$\text{Var}(Z_{2jk}) = \tau^2 \quad \begin{array}{l} j = 1, \quad k = 1, \\ j = 2, \quad k = 1, 2, \dots, I(J-1)-1, \end{array}$$

and all Z_{ijk} are independent.

4. SUFFICIENT AND COMPLETE STATISTICS.

We make a one-one transformation of the parameters by

$$(4.1) \quad \begin{aligned} \gamma_i &= P_{i1' \cdot i} \beta_i \\ \lambda_i &= \sum_{\alpha=i}^p k_{\alpha} \sigma_{\alpha}^2 \end{aligned} \quad i = 1, 2, \dots, p.$$

From Lemma 3(II) in the Appendix with P_{i1} as Q and A_i as A it follows that $P_{i1} A_i$ is nonsingular for all i . Hence the original parameters are given by

$$\beta_i = (P_{i1} A_i)^{-1} \gamma_i \quad i = 1, 2, \dots, p,$$

$$(4.2) \quad \sigma_i^2 = \frac{1}{k_i} (\lambda_i - \lambda_{i+1}) \quad i = 1, 2, \dots, p-1,$$

$$\sigma_p^2 = \frac{1}{k_p} \lambda_p.$$

To simplify the notation we introduce

$$(4.3) \quad \begin{aligned} t_i &= s_i - s_{i-1} - r_i \\ t_1 &= s_1 - r_1. \end{aligned} \quad i = 2, \dots, p,$$

The density function of Z is given by

$$(4.4) \quad (2\pi)^{-\frac{1}{2}n} \left(\prod_{i=1}^p \lambda_i^{-\frac{1}{2}(t_i + r_i)} \right) \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{(z_{i1} - \gamma_i)'(z_{i1} - \gamma_i) + z_{i2}' z_{i2}}{\lambda_i}\right).$$

THEOREM 2. The statistics $Z_{11}, Z_{21}, \dots, Z_{p1}$, $Z_{11}' Z_{11} + Z_{12}' Z_{12}, Z_{21}' Z_{21} + Z_{22}' Z_{22}, \dots, Z_{p1}' Z_{p1} + Z_{p2}' Z_{p2}$ are sufficient for the family of distributions of Z . The family of distributions of the sufficient statistics is complete.

PROOF. The density of Z can be written

$$(4.5) \quad C \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{z_{i1}' z_{i1} + z_{i2}' z_{i2}}{\lambda_i} + \sum_{i=1}^p \sum_{j=1}^{r_i} \frac{\gamma_{ij} z_{i1j}}{\lambda_i}\right)$$

where C is a constant that depends upon the parameters, and

where

$$\left. \begin{aligned} Y_i' &= (Y_{i1}, Y_{i2}, \dots, Y_{ir_i}) \\ Z_{i1}' &= (Z_{i11}, Z_{i12}, \dots, Z_{i1r_i}) \end{aligned} \right\} \quad i = 1, 2, \dots, p.$$

Introducing $\theta_{ij} = \frac{Y_{ij}}{\lambda_i}$ ($j = 1, 2, \dots, r_i$, $i = 1, 2, \dots, p$), we get

$$(4.6) \quad C \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{Z_{i1}Z_{i1} + Z_{i2}Z_{i2}}{\lambda_i} + \sum_{i=1}^p \sum_{j=1}^{r_i} \theta_{ij} Z_{i1j}\right).$$

The conditions on the parameters are $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$, $-\infty < \theta_{ij} < \infty$ ($j = 1, 2, \dots, r_i$, $i = 1, 2, \dots, p$). The total number of parameters is $p + \sum_{i=1}^p r_i$. For varying values of the parameters the densities (4.6) form an exponential family of densities with $Z_{11}, Z_{21}, \dots, Z_{p1}, Z_{11}'Z_{11} + Z_{12}'Z_{12}, Z_{21}'Z_{21} + Z_{22}'Z_{22}, \dots, Z_{p1}'Z_{p1} + Z_{p2}'Z_{p2}$ as sufficient statistics. The parameter space contains a $(p + \sum_{i=1}^p r_i)$ -dimensional rectangle. From Theorem 1, p.132 of [2] it follows that the family of distributions of the sufficient statistics is complete.

5. ESTIMATION.

In this section the problem of estimating the parameters in the model is treated. First the estimates are expressed by P and Z , and the result of Section 4 is used to show that the estimates are minimum variance unbiased.

Finally the estimates are expressed by X and the matrices defining the model.

THEOREM 3. The estimates

$$(5.1) \quad \begin{aligned} \hat{\gamma}_i &= Z_{i1} \\ \hat{\lambda}_i &= \frac{1}{t_i} Z'_{i2} Z_{i2} \end{aligned} \quad i = 1, 2, \dots, p,$$

where $\hat{\gamma}_i$ and $\hat{\lambda}_i$ are estimates for γ_i and λ_i respectively ($i = 1, 2, \dots, p$), are minimum variance unbiased estimates.

PROOF. Z_{i1} is $N(\gamma_i, \lambda_i I)$ and Z_{i2j} is $N(0, \lambda_i)$ ($i = 1, 2, \dots, p, j = 1, 2, \dots, t_i$). It follows that

$$(5.2) \quad \left. \begin{aligned} E(\hat{\gamma}_i) &= E(Z_{i1}) = \gamma_i \\ E(\hat{\lambda}_i) &= E\left(\frac{1}{t_i} Z'_{i2} Z_{i2}\right) = E\left(\frac{1}{t_i} \sum_{j=1}^{t_i} Z_{i2j}^2\right) = \lambda_i \end{aligned} \right\} i=1, 2, \dots, p.$$

Hence the estimates are unbiased. Since

$$\hat{\lambda}_i = \frac{1}{t_i} ((Z'_{i1} Z_{i1} + Z'_{i2} Z_{i2}) - Z'_{i1} Z_{i1}) \quad i = 1, 2, \dots, p,$$

the estimates depend only upon sufficient and complete statistics from which it follows that they are minimum variance estimates (see [3], Theorem 5.1).

THEOREM 4. The estimates

$$\begin{aligned}
 \hat{\beta}_i &= (P_{i1}A_i)^{-1}Z_{i1} & i = 1, 2, \dots, p, \\
 (5.3) \quad \hat{\sigma}_i^2 &= \frac{1}{k_i} \left(\frac{Z_{i2}'Z_{i2}}{t_i} - \frac{Z_{i+1,2}'Z_{i+1,2}}{t_{i+1}} \right) & i = 1, 2, \dots, p-1 \\
 \hat{\sigma}_p^2 &= \frac{1}{k_p} \frac{1}{t_p} Z_{p2}'Z_{p2}
 \end{aligned}$$

are minimum variance unbiased estimates for $\beta_1, \beta_2, \dots, \beta_p$ and $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$ respectively.

PROOF. From (5.1)

$$\begin{aligned}
 \hat{\beta}_i &= (P_{i1}A_i)^{-1}\hat{\gamma}_i & i = 1, 2, \dots, p, \\
 (5.4) \quad \hat{\sigma}_i^2 &= \frac{1}{k_i} (\hat{\lambda}_i - \hat{\lambda}_{i+1}) & i = 1, 2, \dots, p-1, \\
 \hat{\sigma}_p^2 &= \frac{1}{k_p} \hat{\lambda}_p.
 \end{aligned}$$

From (4.2) and Theorem 3 we have

$$\begin{aligned}
 E(\hat{\beta}_i) &= (P_{i1}A_i)^{-1}\gamma_i = \beta_i & i = 1, 2, \dots, p, \\
 E(\hat{\sigma}_i^2) &= \frac{1}{k_i} (\lambda_i - \lambda_{i+1}) = \sigma_i^2 & i = 1, 2, \dots, p-1, \\
 E(\hat{\sigma}_p^2) &= \frac{1}{k_p} \lambda_p = \sigma_p^2.
 \end{aligned}$$

Thus the estimates are unbiased. Since $\hat{\gamma}_i$ and $\hat{\lambda}_i$ depend only upon sufficient and complete statistics, it follows from (5.4) that the same is true for $\hat{\beta}_i$ and $\hat{\sigma}_i^2$

($i = 1, 2, \dots, p$). As in Theorem 3 it follows that the estimates must be minimum variance unbiased.

THEOREM 5. $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ are independent. Each have a multinormal distribution with $E(\hat{\beta}_i) = \beta_i$ and $\sum (\hat{\beta}_i) = \lambda_i (A_i' A_i)^{-1} = (\sum_{\alpha=1}^p k_{\alpha} \sigma_{\alpha}^2) (A_i' A_i)^{-1}$.

The estimates $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_p^2$ are independent of $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$. For $|i-j| > 1$ $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$ are independent. We have

$$\begin{aligned} E(\hat{\sigma}_i^2) &= \sigma_i^2 & i = 1, 2, \dots, p, \\ (5.5) \quad \text{Var}(\hat{\sigma}_i^2) &= \frac{2}{k_i^2} \left(\frac{\lambda_i^2}{t_i} + \frac{\lambda_{i+1}^2}{t_{i+1}} \right) & i = 1, 2, \dots, p-1, \\ \text{Var}(\hat{\sigma}_p^2) &= \frac{2}{k_p^2} \frac{\lambda_p^2}{t_p}, \\ \text{Cov}(\hat{\sigma}_i^2, \hat{\sigma}_{i+1}^2) &= - \frac{2}{k_i k_{i+1}} \frac{\lambda_{i+1}^2}{t_{i+1}} & i = 1, 2, \dots, p-1. \end{aligned}$$

PROOF. That the $\hat{\beta}_i$'s are independent follows from the fact that they depend upon independent subvectors of Z . Unbiasedness of $\hat{\beta}_i$ is proved in Theorem 4. For the covariance matrix we have

$$\begin{aligned} \sum (\hat{\beta}_i) &= (P_{i1} A_i)^{-1} E(Z_{i1} Z_{i1}') ((P_{i1} A_i)^{-1})' \\ &= \lambda_i (P_{i1} A_i)^{-1} ((P_{i1} A_i)^{-1})' = \lambda_i (A_i' A_i)^{-1} = \left(\sum_{\alpha=1}^p k_{\alpha} \sigma_{\alpha}^2 \right) (A_i' A_i)^{-1}. \end{aligned}$$

The second last equality is obtained by using Lemma 3 (IV) in

the Appendix with P_{i1} as Q and A_i as A . From the lemma

$$(P_{i1}A_i)^{-1}P_{i1} = (A_i'A_i)^{-1}A_i'.$$

Post-multiplication with P_{i1}' gives

$$(P_{i1}A_i)^{-1} = (A_i'A_i)^{-1}A_i'P_{i1}',$$

hence

$$(P_{i1}A_i)^{-1}((P_{i1}A_i)^{-1})' = (A_i'A_i)^{-1}.$$

By equation (5.1) $\frac{t_i}{\hat{\lambda}_i} \hat{\lambda}_i = \chi_{t_i}^2$ ($i = 1, 2, \dots, p$) where $\chi_{t_i}^2$ denotes a chi-square distributed random variable with t_i degrees of freedom. Since the $\hat{\lambda}_i$'s depend upon independent subvectors of Z , they are independent. Thus the same is true for the $\chi_{t_i}^2$'s. From (5.4) we have

$$\begin{aligned} \hat{\sigma}_i^2 &= \frac{1}{k_i} \left(\frac{\lambda_i}{t_i} \chi_{t_i}^2 - \frac{\lambda_{i+1}}{t_{i+1}} \chi_{t_{i+1}}^2 \right) \quad i = 1, 2, \dots, p-1, \\ (5.6) \quad \hat{\sigma}_p^2 &= \frac{1}{k_p} \frac{\lambda_p}{t_p} \chi_{t_p}^2, \end{aligned}$$

whereby the probability distribution of $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_p^2$ is completely characterized. It is seen that $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$ are independent when $|i-j| > 1$. Furthermore

$$\begin{aligned}
 \text{Var}(\hat{\sigma}_i^2) &= \frac{1}{k_i^2} \left(\frac{\lambda_i^2}{t_i^2} \text{Var}(\chi_{t_i}^2) + \frac{\lambda_{i+1}^2}{t_{i+1}^2} \text{Var}(\chi_{t_{i+1}}^2) \right) \quad i = 1, 2, \dots, p-1, \\
 (5.7) \\
 \text{Var}(\hat{\sigma}_p^2) &= \frac{1}{k_p^2} \frac{\lambda_p^2}{t_p^2} \text{Var}(\chi_{t_p}^2), \\
 \text{Cov}(\hat{\sigma}_i^2, \hat{\sigma}_{i+1}^2) &= -\frac{1}{k_i} \frac{1}{k_{i+1}} \frac{\lambda_{i+1}^2}{t_{i+1}^2} \text{Var}(\chi_{t_{i+1}}^2) \quad i = 1, 2, \dots, p-1.
 \end{aligned}$$

$\text{Var}(\chi_{t_i}^2) = 2t_i$ inserted in (5.7) gives the result (5.5). Unbiasedness of $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_p^2$ is proved in Theorem 4.

From (5.6) it is seen that $\hat{\sigma}_i^2 < 0$ with positive probability. This corresponds to the situation $\hat{\lambda}_i < \hat{\lambda}_{i+1}$ contrary to the restriction $\lambda_i \geq \lambda_{i+1}$ on the parameter space. As for the estimates $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$ they can, as stated in the proof, be given in terms of independent chi-square variables by $\hat{\lambda}_i = \frac{\lambda_i}{t_i} \chi_{t_i}^2$. Contrary to the probability distribution of $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_p^2$, the distribution of the $\hat{\lambda}_i$'s thus have a simple form.

THEOREM 6. The estimates can be expressed by X and by the matrices defining the model as

$$\begin{aligned}
 \hat{\beta}_i &= (A_i' A_i)^{-1} A_i' X \quad i = 1, 2, \dots, p, \\
 \hat{\sigma}_i^2 &= \frac{1}{k_i} X' \left(\left(\frac{1}{t_i} + \frac{1}{t_{i+1}} \right) C_i (C_i' C_i)^{-1} C_i' - \frac{1}{t_{i+1}} C_{i+1} (C_{i+1}' C_{i+1})^{-1} C_{i+1}' - \right. \\
 &\quad \left. \frac{1}{t_i} C_{i-1} (C_{i-1}' C_{i-1})^{-1} C_{i-1}' - \frac{1}{t_i} A_i (A_i' A_i)^{-1} A_i' + \frac{1}{t_{i+1}} A_{i+1} (A_{i+1}' A_{i+1})^{-1} A_{i+1}' \right) X \\
 &\quad A_{i+1}' X \quad i = 1, 2, \dots, p-1,
 \end{aligned}$$

where the member with C_{i-1} is defined to be 0 when $i=1$.

$$\hat{\sigma}_p^2 = \frac{1}{k_p t_p} X' (C_p (C_p (C_p' C_p)^{-1} C_p' - C_{p-1} (C_{p-1}' C_{p-1})^{-1} C_{p-1}' - A_p (A_p' A_p)^{-1} A_p') X.$$

PROOF. From Lemma 3 (IV) in Appendix with P_{i1} as Q and A_i as A we get

$$\hat{\beta}_i = (P_{i1}' A_i)^{-1} Z_{i1} = (P_{i1}' A_i)^{-1} P_{i1}' X_{i1} = (A_i' A_i)^{-1} A_i' X$$

$i = 1, 2, \dots, p.$

From (III) of the same lemma we get the identity

$$(5.8) \quad P_{i1}' P_{i1} = A_i (A_i' A_i)^{-1} A_i' \quad i = 1, 2, \dots, p.$$

Let Q_i be defined by

$$(5.9) \quad Q_i' = (P_{i1}', P_{i2}', \dots, P_{i1}', P_{i2}')$$

$i = 1, 2, \dots, p.$

Using Lemma 3 (III) now with Q_i as Q and C_i as A we have

$$(5.10) \quad Q_i' Q_i = C_i (C_i' C_i)^{-1} C_i' \quad i = 1, 2, \dots, p.$$

From the definition of Q_i

$$(5.11) \quad Q_i' Q_i = \sum_{\alpha=1}^i (P_{\alpha 1}' P_{\alpha 1} + P_{\alpha 2}' P_{\alpha 2}) \quad i = 1, 2, \dots, p,$$

combining (5.10) and (5.11)

$$(5.12) \quad \begin{aligned} Q_1' Q_i - Q_{i-1}' Q_{i-1} &= P_{i1}' P_{i1} + P_{i2}' P_{i2} = C_i (C_i' C_i)^{-1} C_i' \\ &\quad - C_{i-1} (C_{i-1}' C_{i-1})^{-1} C_{i-1}' \quad i = 2, 3, \dots, p, \end{aligned}$$

$$Q_1' Q_1 = C_1 (C_1' C_1)^{-1} C_1'.$$

Inserting (5.8) in (5.12) we find

$$(5.13) \quad \begin{aligned} P_{12}' P_{i2} &= C_i (C_i' C_i)^{-1} C_i' - C_{i-1} (C_{i-1}' C_{i-1})^{-1} C_{i-1}' \\ &\quad - A_i (A_i' A_i)^{-1} A_i' \quad i = 2, 3, \dots, p, \end{aligned}$$

$$P_{12}' P_{12} = C_1 (C_1' C_1)^{-1} C_1' - A_1 (A_1' A_1)^{-1} A_1'.$$

We have

$$(5.14) \quad \hat{\lambda}_i = \frac{1}{t_i} Z_{i2}' Z_{i2} = \frac{1}{t_i} X' P_{i2}' P_{i2} X \quad i = 1, 2, \dots, p.$$

Thus, by using (5.13), we can express the $\hat{\lambda}_i$'s by X and known matrices. Introducing this in (5.4) we get the result in the theorem.

Example (cont.). We shall derive the minimum variance estimates of the parameters in the example. We have

$$A_1' A_1 = \begin{bmatrix} n & 0 \\ 0 & J \sum (t_i - t.)^2 \end{bmatrix},$$

$$C_1' C_1 = J \quad I^{1 \times 1},$$

$$A_2' A_2 = \sum (u_{ij} - u_{..})^2,$$

$$C_2' C_2 = I^{n \times n}.$$

The inverse matrices are easily computed. Moreover we need

$$X' A_1 = (\sum X_{ij}, J \sum X_{i.} (t_i - t_{.})),$$

$$X' A_2 = \sum X_{ij} (u_{ij} - u_{..}),$$

$$X' C_1 = (JX_{1.}, JX_{2.}, \dots, jX_{I.}),$$

$$X' C_2 = (X_{11}, X_{12}, \dots, X_{IJ}),$$

where $X_{i.} = \frac{1}{J} \sum_{j=1}^J X_{ij}$ $i = 1, 2, \dots, I$. We define
 $X_{..} = \frac{1}{IJ} \sum X_{ij}.$

From Theorem 6 we have

$$\hat{\beta}_1 = \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = (A_1' A_1)^{-1} A_1' X = \begin{bmatrix} X_{..} \\ \frac{\sum X_{i.} (t_i - t_{.})}{\sum (t_i - t_{.})^2} \end{bmatrix}$$

and

$$\hat{\beta}_2 = \hat{\gamma} = \frac{\sum X_{ij} (u_{ij} - u_{..})}{\sum (u_{ij} - u_{..})^2}.$$

From Theorem 5 we know that these estimates are unbiased, and moreover

$$\sum \left(\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) = (J\sigma^2 + \tau^2) \begin{bmatrix} \frac{1}{IJ} & 0 \\ 0 & \frac{1}{J \sum (t_i - t.)^2} \end{bmatrix}$$

$$\text{Var}(\hat{\gamma}) = \frac{\tau^2}{\sum (u_{ij} - u_{..})^2}.$$

Regarding the estimates of the variances it is most illustrative to find $\hat{\lambda}_1$ and $\hat{\lambda}_2$ given by (5.13) and (5.14). We get

$$(I-2)\hat{\lambda}_1 = J(\sum (x_{i.} - x_{..})^2 - \frac{(\sum x_{i.}(t_i - t.))^2}{\sum (t_i - t.)^2})$$

and

$$(IJ-I-1)\hat{\lambda}_2 = \sum (x_{ij} - x_{i.})^2 - \frac{(\sum x_{ij}(u_{ij} - u_{..}))^2}{\sum (u_{ij} - u_{..})^2}.$$

In the example $t_1 = I-2$, $t_2 = IJ-I-1$, $k_1 = J$ and $k_2 = 1$. The estimates of σ^2 and τ^2 are

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{J}(\hat{\lambda}_1 - \hat{\lambda}_2), \\ \hat{\tau}^2 &= \hat{\lambda}_2. \end{aligned}$$

They are unbiased and

$$\text{Var}(\hat{\sigma}^2) = \frac{2}{J^2} \left(\frac{(J\sigma^2 + \tau^2)^2}{I-2} + \frac{\tau^4}{IJ-I-1} \right),$$

$$\text{Var}(\hat{\tau}^2) = 2 \frac{\tau^4}{IJ-I-1},$$

$$\text{Cov}(\hat{\sigma}^2, \hat{\tau}^2) = -\frac{2}{J} \frac{\tau^4}{IJ-I-1}.$$

6. TESTING HYPOTHESES CONCERNING RANDOM EFFECTS.

There are two kind of hypotheses to be tested in this model, namely, hypotheses concerning the variances or the random effects and hypotheses concerning the parameters representing the fixed effects. In this section we shall derive uniformly most powerful (UMP) unbiased tests for some hypotheses concerning the variances.

We use the notation F_{ν_1, ν_2} and $f_{\alpha; \nu_1, \nu_2}$ for the cumulative F-distribution with ν_1 and ν_2 degrees of freedom and the α -fractile of the same distribution, respectively.

Let $\Delta_1, \Delta_2, \dots, \Delta_{p-1}$ be defined by

$$(6.1) \quad \Delta_i = \frac{k_i \sigma_i^2}{\sum_{\alpha=i+1}^p k_\alpha \sigma_\alpha^2} = \frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1}} \quad i = 1, 2, \dots, p-1,$$

and define

$$(6.2) \quad F_i = \frac{Z'_{i2} Z_{i2}}{Z'_{i+1,2} Z_{i+1,2}} \frac{t_{i+1}}{t_i} \quad i = 1, 2, \dots, p-1.$$

We shall prove the following theorem.

THEOREM 7. For $i = 1, 2, \dots, p-1$ we have:

$\frac{F_i}{1 + \Delta_i}$ have an F-distribution with t_i and t_{i+1} degrees of freedom.

UMP unbiased α -tests of the hypotheses

$$\begin{aligned} H_{i1} : \Delta_i &\leq \Delta_i^0 \quad \text{against} \quad \Delta_i > \Delta_i^0, \\ (6.3) \quad H_{i2} : \Delta_i &\geq \Delta_i^0 \quad \text{against} \quad \Delta_i < \Delta_i^0, \\ H_{i3} : \Delta_i &= \Delta_i^0 \quad \text{against} \quad \Delta_i \neq \Delta_i^0, \end{aligned}$$

where Δ_i^0 is a fixed value of Δ_i , are given by the rejection regions, respectively

$$(6.4) \quad \frac{F_i}{1+\Delta_i^0} > f_{1-\alpha; t_i, t_{i+1}},$$

$$(6.5) \quad \frac{F_i}{1+\Delta_i^0} < f_{\alpha; t_i, t_{i+1}},$$

$$(6.6) \quad \frac{F_i}{1+\Delta_i^0} < f_1 \quad \text{and} \quad \frac{F_i}{1+\Delta_i^0} > f_2.$$

Here f_1 and f_2 are the solutions of the equations

$$(6.7) \quad F_{t_i, t_{i+1}}(f_2) - F_{t_i, t_{i+1}}(f_1) = 1 - \alpha,$$

$$(6.8) \quad F_{t_{i+2}, t_{i+1}}\left(\frac{t_i}{t_{i+2}} f_2\right) - F_{t_{i+2}, t_{i+1}}\left(\frac{t_i}{t_{i+2}} f_1\right) = 1 - \alpha.$$

PROOF. By (4.6) the probability density function of Z is

$$(6.9) \quad C \exp\left(-\frac{1}{2} \sum_{\alpha=1}^p \frac{z'_{\alpha 1} z_{\alpha 1} + z'_{\alpha 2} z_{\alpha 2}}{\lambda_{\alpha}} + \sum_{i=1}^p \sum_{j=1}^{r_i} \theta_{ij} z_{i1j}\right).$$

From (6.1)

$$(6.10) \quad \frac{1}{\lambda_i} = \frac{1}{\lambda_{i+1}(1+\Delta_i)},$$

and (6.9) can be written as

$$\begin{aligned}
 & C \exp\left(-\frac{1}{2} \frac{1}{\lambda_{i+1}} \left(\frac{1}{1+\Delta_i} - \frac{1}{1+\Delta_i^0}\right) (z'_{i1} z_{i1} + z'_{i2} z_{i2})\right) \\
 (6.11) \quad & - \frac{1}{2} \frac{1}{\lambda_{i+1}} (z'_{i+1,1} z_{i+1,1} + z'_{i+1,2} z_{i+1,2} + \frac{1}{1+\Delta_i^0} (z'_{i1} z_{i1} + \\
 & z'_{i2} z_{i2})) - \frac{1}{2} \sum_{\alpha \neq i, i+1} \frac{z'_{\alpha 1} z_{\alpha 1} + z'_{\alpha 2} z_{\alpha 2}}{\lambda_{\alpha}} + \sum_{i=1}^p \sum_{j=1}^{r_i} \theta_{ij} z_{i1j}).
 \end{aligned}$$

This constitutes an exponential family of distributions of the kind treated in Chapter 4.4 of [2]. The parameter space is given by the set of all points $\{\eta_i, -\frac{1}{2} \frac{1}{\lambda_1}, -\frac{1}{2} \frac{1}{\lambda_2}, \dots, -\frac{1}{2} \frac{1}{\lambda_{i-1}}, -\frac{1}{2} \frac{1}{\lambda_{i+1}}, \dots, -\frac{1}{2} \frac{1}{\lambda_p}, \theta_{11}, \theta_{12}, \dots, \theta_{pr_p} |$
 $-\infty < \eta_i < \infty, 0 < \lambda_{\alpha} < \infty \quad \alpha \neq i, -\infty < \theta_{ij} < \infty$
 $i = 1, 2, \dots, p, j = 1, 2, \dots, r_i\}$,

where

$$(6.12) \quad \eta_i = -\frac{1}{2} \frac{1}{\lambda_{i+1}} \left(\frac{1}{1+\Delta_i} - \frac{1}{1+\Delta_i^0}\right).$$

We have presupposed $\Delta_i^0 > 0$. If $\Delta_i^0 = 0$ then η_i varies in the set $\{\eta_i | \eta_i \geq 0\}$. Regarding the hypotheses H_{i2} and H_{i3} we shall assume $\Delta_i^0 \neq 0$.

The total number of parameters is $p + \sum_{i=1}^p r_i$. The parameter space is convex and has dimension $p + \sum_{i=1}^p r_i$. For fixed value of η_i the other parameters constitute a convex set in a $(p + \sum_{i=1}^p r_i - 1)$ -dimensional space. The theory leading to Theorem 3, p. 136 of [2] can therefore be applied. We can find UMP unbiased tests of the hypotheses

$$\begin{aligned}
 (6.13) \quad & \eta_i \leq 0 \quad \text{against} \quad \eta_i > 0, \\
 & \eta_i \geq 0 \quad \text{against} \quad \eta_i < 0, \\
 & \eta_i = 0 \quad \text{against} \quad \eta_i \neq 0,
 \end{aligned}$$

and by the definition of η_i this is equivalent to (6.3).

By analogy to the theory of [2] we introduce the random variable U and the random vector T defined by

$$\begin{aligned}
 (6.14) \quad & U = Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2}, \\
 & T = (Z'_{i+1,1}Z_{i+1,1} + Z'_{i+1,2}Z_{i+1,2} + \frac{1}{1+\Delta_i^0}(Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2}), \\
 & Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2}, \dots, Z'_{i-1,1}Z_{i-1,1} + Z'_{i-1,2}Z_{i-1,2}, \\
 & Z'_{i+2,1}Z_{i+2,1} + Z'_{i+2,2}Z_{i+2,2}, \dots, Z'_{p1}Z_{p1} + Z'_{p2}Z_{p2}, \\
 & Z'_{11}, Z'_{21}, \dots, Z'_{p1}).
 \end{aligned}$$

We define V as a function of U and T as

$$\begin{aligned}
 (6.15a) \quad & \frac{1}{1+\Delta_i^0}(Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2} - Z'_{i1}Z_{i1}) \\
 V = & \frac{\frac{1}{1+\Delta_i^0}(Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2} - Z'_{i1}Z_{i1})}{Z'_{i+1,1}Z_{i+1,1} + Z'_{i+1,2}Z_{i+1,2} + \frac{1}{1+\Delta_i^0}(Z'_{i1}Z_{i1} + Z'_{i2}Z_{i2}) - Z'_{i+1,1}Z_{i+1,1} - \frac{1}{1+\Delta_i^0}Z'_{i1}Z_{i1}} \\
 (6.15b) \quad & = \frac{\frac{1}{1+\Delta_i^0} Z'_{i2}Z_{i2}}{Z'_{i+1,2}Z_{i+1,2} + \frac{1}{1+\Delta_i^0} Z'_{i2}Z_{i2}}
 \end{aligned}$$

$$(6.15c) \quad = \frac{1}{\frac{1}{\lambda_{i+1}} Z'_{i+1,2} Z_{i+1,2} + 1} \frac{1}{\lambda_{i+1} (1+\Delta_i^0)} Z'_{i2} Z_{i2}$$

$$(6.15d) \quad = \frac{1}{(1+\Delta_i^0) \frac{t_{i+1}}{t_i} \frac{1}{F_i} + 1}.$$

When $\Delta_i = \Delta_i^0$ we have $\frac{1}{\lambda_{i+1} (1+\Delta_i^0)} = \frac{1}{\lambda_i}$. Then

$\frac{1}{\lambda_{i+1} (1+\Delta_i^0)} Z'_{i2} Z_{i2}$ and $\frac{1}{\lambda_{i+1}} Z'_{i+1,2} Z_{i+1,2}$ are independent and chi-square distributed with t_i and t_{i+1} degrees of freedom respectively. This implies that

$$\frac{1}{\lambda_{i+1} (1+\Delta_i^0)} Z'_{i2} Z_{i2} + \frac{1}{\lambda_{i+1}} Z'_{i+1,2} Z_{i+1,2} \text{ and } \frac{Z'_{i2} Z_{i2}}{Z'_{i+1,2} Z_{i+1,2}} \frac{1}{1+\Delta_i^0}$$

are independent for $\Delta_i = \Delta_i^0$.

V depends only upon $\frac{Z'_{i2} Z_{i2}}{Z'_{i+1,2} Z_{i+1,2}}$. It follows that

V is independent of $\frac{1}{1+\Delta_i^0} Z'_{i2} Z_{i2} + Z'_{i+1,2} Z_{i+1,2}$. From (6.14)

it is now seen that V is independent of T for $\Delta_i = \Delta_i^0$. Furthermore V is linear and increasing in U for fixed value of T . According to Theorem 1, p. 161 of [2] the UMP unbiased rejection regions for the hypotheses H_{i1}, H_{i2} and H_{i3} are given by

$$(6.16) \quad \begin{aligned} V &> C_1, \\ V &< C_2, \\ V &< C_3 \text{ and } V > C_4, \end{aligned}$$

respectively, where C_1, C_2, C_3 and C_4 are constants. From (6.15d) it is seen that this is equivalent to the rejection regions

$$(6.17) \quad \begin{aligned} \frac{F_i}{1+\Delta_i^0} &> C_1', \\ \frac{F_i}{1+\Delta_i^0} &< C_2', \\ \frac{F_i}{1+\Delta_i^0} &< C_3' \quad \text{and} \quad \frac{F_i}{1+\Delta_i^0} > C_4', \end{aligned}$$

where C_1', C_2', C_3' and C_4' are new constants.

From the independence of the chi-square distributed variables $\frac{1}{\lambda_i} Z_{i2}' Z_{i2}$ and $\frac{1}{\lambda_{i+1}} Z_{i+1,2}' Z_{i+1,2}$ it follows that

$$(6.18) \quad \frac{F_i}{1+\Delta_i} = \frac{\frac{1}{t_i} \frac{Z_{i2}' Z_{i2}}{\lambda_i}}{\frac{1}{t_{i+1}} \frac{Z_{i+1,2}' Z_{i+1,2}}{\lambda_{i+1}}}$$

have an F-distribution with t_i and t_{i+1} degrees of freedom. Particularly when $\Delta_i = \Delta_i^0$ it is seen that $\frac{F_i}{1+\Delta_i^0}$ is F-distributed with t_i and t_{i+1} degrees of freedom.

Let the critical functions corresponding to (6.17) be denoted $\varphi_{i1}, \varphi_{i2}$ and φ_{i3} respectively. They depend only upon F_i , and the distribution of F_i depend only upon Δ_i . The constants C_1', C_2', C_3' and C_4' are determined by the condition of unbiasedness i.e. $E_{\Delta_i^0}(\varphi_{i1}(F_i)) = \alpha$ giving (6.4), $E_{\Delta_i^0}(\varphi_{i2}(F_i)) = \alpha$ giving (6.5), $E_{\Delta_i^0}(\varphi_{i3}(F_i)) = \alpha$

and the derivative of the power function equal to 0 at Δ_i^0 giving (6.6), (6.7) and (6.8). This completes the proof.

The power function $E_{\Delta_i} \varphi_{i1}(F_i)$ is given by

$$\begin{aligned}
 \beta_{i1}(\Delta_i) &= P_{\Delta_i} \left(\frac{F_i}{1+\Delta_i^0} > f_{1-\alpha; t_i, t_{i+1}} \right) \\
 (6.19) \quad &= 1 - P_{\Delta_i} \left(\frac{F_i}{1+\Delta_i} \leq \frac{1+\Delta_i^0}{1+\Delta_i} f_{1-\alpha; t_i, t_{i+1}} \right) \\
 &= 1 - F_{t_i, t_{i+1}} \left(\frac{1+\Delta_i^0}{1+\Delta_i} f_{1-\alpha; t_i, t_{i+1}} \right).
 \end{aligned}$$

It is seen that the power depends only upon Δ_i . In the same way the power functions $E_{\Delta_i} \varphi_{i2}(F_i)$ and $E_{\Delta_i} \varphi_{i3}(F_i)$ can be written down, and these too depend only upon Δ_i .

From the definition of F_i we have

$$(6.20) \quad F_i = \frac{\hat{\lambda}_i}{\hat{\lambda}_{i+1}} = \frac{k_i \hat{\sigma}_i^2}{\sum_{\alpha=i+1}^p k_{\alpha} \hat{\sigma}_{\alpha}^2} + 1.$$

This gives the rejection regions an intuitive interpretation.

To get F_i expressed by X and the matrices defining the model, we can use equations (5.13) and (5.14) for the $\hat{\lambda}_i$'s.

7. TESTING HYPOTHESES CONCERNING FIXED EFFECTS.

In this section we shall treat the problem of testing certain hypotheses concerning the β_i 's.

Suppose we shall derive a test of the hypothesis

$$H : M_i \beta_i = 0$$

where M_i is a $(q_i \times r_i)$ -matrix where $\text{rank } M_i = q_i \leq r_i$.

It will be proved that among all tests that are both invariant (under a certain group of transformations) and unbiased there exists a UMP test of H .

First we express the hypothesis by means of γ_i . When constructing the canonical form for the model we can choose P_{i1} so that

$$(7.1) \quad P_{i1} = \begin{bmatrix} q_i \times n \\ P_{i11} \\ (r_i - q_i) \times n \\ P_{i12} \end{bmatrix} \quad \text{where } \mathcal{R}(P_{i11} A_i) = \mathcal{R}(M_i).$$

That this is possible follows from Lemma 4 in Appendix with A_i as A , M_i as B and P_{i11} as Q . We define $Z_{i1j} = P_{i1j} X (j = 1, 2)$, and γ_{i1} and γ_{i2} to be the corresponding expectation vectors i.e.

$$(7.2) \quad \gamma_i = \begin{bmatrix} \gamma_{i1} \\ \gamma_{i2} \end{bmatrix} = \begin{bmatrix} P_{i11} \\ P_{i12} \end{bmatrix} A_i \beta_i = \begin{bmatrix} P_{i11} A_i \beta_i \\ P_{i12} A_i \beta_i \end{bmatrix}$$

The notation here differs from the notation in Sections 4 and 5, since here γ_{i1} and γ_{i2} refer to two subvectors of γ_i while in the preceding sections γ_{i1} and γ_{i2} were the two first elements of γ_i . The hypothesis H is that $\beta_i \perp \mathcal{R}(M_i) = \mathcal{R}(P_{i11} A_i)$ which by (7.2) is equivalent to

$$H' : \gamma_{i1} = 0.$$

We define the following group G of transformations of Z

$$(7.3) \quad \begin{aligned} Y_{\alpha 1} &= Z_{\alpha 1} + d_{\alpha} \quad \alpha = 1, 2, \dots, i-1, i+1, \dots, p, \\ Y_{i12} &= Z_{i12} + e_i, \\ Y_{i11} &= Q_i Z_{i11}, \\ Y_{\alpha 2} &= R_{\alpha} Z_{\alpha 2} \quad \alpha = 1, 2, \dots, p, \end{aligned}$$

where $d_{\alpha}^{r_{\alpha} \times 1}$ ($\alpha \neq i$) and $e_i^{(r_i - q_i) \times 1}$ are vectors, and $Q_i^{q_i \times q_i}$ and $R_{\alpha}^{t_{\alpha} \times t_{\alpha}}$ ($\alpha = 1, 2, \dots, p$) are orthogonal matrices. The problem of testing the hypothesis H' is invariant under G . It is easily seen that a maximal invariant is given by

$$(7.4) \quad S' = (Z'_{12}Z_{12}, Z'_{22}Z_{22}, \dots, Z'_{p2}Z_{p2}, Z'_{i11}Z_{i11})$$

We introduce

$$(7.5) \quad \begin{aligned} S_{\alpha} &= Z'_{\alpha 2}Z_{\alpha 2} \quad \alpha = 1, 2, \dots, p, \\ S_{p+1} &= Z'_{i11}Z_{i11}. \end{aligned}$$

For the probability distribution of $S_1, S_2, \dots, S_{p+1}$ we have that $\frac{S_{\alpha}}{\lambda_{\alpha}}$ is chi-square distributed with t_{α} degrees of freedom ($\alpha = 1, 2, \dots, p$), and $\frac{S_{p+1}}{\lambda_i}$ has a noncentral chi-square distribution with q_i degrees of freedom and noncentrality parameter $\delta = \sqrt{\frac{1}{\lambda_i} \gamma'_{i1} \gamma_{i1}}$. In particular if H' holds $\frac{S_{p+1}}{\lambda_i}$ has a central chi-square distribution with q_i degrees of freedom. The probability density function of $S_1, S_2, \dots, S_{p+1}$ is given by

$$f(s; \lambda, \delta) = \left(\prod_{\alpha=1}^p \frac{1}{2^{\frac{1}{2}t_{\alpha}} \Gamma(\frac{1}{2}t_{\alpha})} s_{\alpha}^{\frac{1}{2}t_{\alpha}-1} \left(\frac{1}{\lambda_{\alpha}} \right)^{\frac{1}{2}t_{\alpha}} \exp\left(-\frac{1}{2} \frac{s_{\alpha}}{\lambda_{\alpha}}\right) \right) \frac{1}{2^{\frac{1}{2}q_i} \Gamma(\frac{1}{2}q_i)} \left(\frac{1}{\lambda_i} \right)^{\frac{1}{2}q_i}$$

(7.6)

$$s_{p+1}^{\frac{1}{2}q_i-1} \exp\left(-\frac{1}{2}\left(\delta^2 + \frac{s_{p+1}}{\lambda_i}\right)\right) \sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} s_{p+1}^{\beta} \Gamma(\beta + \frac{1}{2})}{\lambda_i^{\beta} (2\beta)! \Gamma(\frac{1}{2}q_i + \beta)}.$$

Under H' we have

$$f(s; \lambda, 0) = \frac{1}{2^{\frac{1}{2}q_i} \Gamma(\frac{1}{2}q_i)} \left(\frac{1}{\lambda_i} \right)^{\frac{1}{2}q_i} \left(\prod_{\alpha=1}^p \frac{1}{2^{\frac{1}{2}t_{\alpha}} \Gamma(\frac{1}{2}t_{\alpha})} \left(\frac{1}{\lambda_{\alpha}} \right)^{\frac{1}{2}t_{\alpha}} s_{\alpha}^{\frac{1}{2}t_{\alpha}-1} \right) \exp\left(-\frac{1}{2}\right)$$

(7.7)

$$\left(\frac{1}{\lambda_i} (s_i + s_{p+1}) + \sum_{\alpha \neq i} \frac{s_{\alpha}}{\lambda_{\alpha}} \right).$$

This shows that under H' we have an exponential family of distribution with $s_i + s_{p+1}, s_1, s_2, \dots, s_{i-1}, s_{i+1}, \dots, s_p$ as sufficient and complete statistics. According to Theorem 2, p. 134 of [2] it is no restriction, when considering similar tests (unbiasedness implies similarity here), to derive the tests from the conditional distribution of S given

$$R' = (s_i + s_{p+1}, s_1, s_2, \dots, s_{i-1}, s_{i+1}, \dots, s_p).$$

Let $V = s_i + s_{p+1}$. The density of R is given by

(7.8)

$$g(r; \lambda, \delta) = \left(\prod_{\alpha \neq i} \frac{1}{2^{\frac{1}{2}t_{\alpha}} \Gamma(\frac{1}{2}t_{\alpha})} \left(\frac{1}{\lambda_{\alpha}} \right)^{\frac{1}{2}t_{\alpha}} s_{\alpha}^{\frac{1}{2}t_{\alpha}-1} \right)$$

$$\exp\left(-\frac{1}{2} \frac{s_{\alpha}}{\lambda_{\alpha}}\right) \frac{1}{2^{\frac{1}{2}(q_i+t_i)} \Gamma(\frac{1}{2}(q_i+t_i))} \left(\frac{1}{\lambda_i} \right)^{\frac{1}{2}(q_i+t_i)-1} \exp\left(-\frac{1}{2}\left(\delta^2 + \frac{v}{\lambda_i}\right)\right) \sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} v^{\beta} \Gamma(\beta + \frac{1}{2})}{\lambda_i^{\beta} (2\beta)! \Gamma(\frac{1}{2}(q_i+t_i) + \beta)}.$$

$$\frac{\delta^{2\beta} v^\beta \Gamma(\beta + \frac{1}{2})}{\lambda_i^\beta (2\beta)! \Gamma(\frac{1}{2}(q_i + t_i) + \beta)}.$$

The conditional density of S given R is then

$$h(s; \lambda, \delta | r) = \frac{f(s; \lambda, \delta)}{g(r; \lambda, \delta)} = \frac{1}{\Gamma(\frac{1}{2}t_i)} \frac{s_{p+1}^{\frac{1}{2}q_i-1} s_i^{\frac{1}{2}t_i-1}}{v^{\frac{1}{2}(q_i+t_i)-1}}$$

(7.9)

$$\exp(-\frac{1}{2} \frac{s_{p+1} + s_i}{\lambda_i} + \frac{1}{2} \frac{v}{\lambda_i}) \cdot \frac{\sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} s_{p+1}^\beta}{\lambda_i^\beta} \frac{\Gamma(\beta + \frac{1}{2})}{(2\beta)! \Gamma(\frac{1}{2}q_i + \beta)}}{\sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} v^\beta}{\lambda_i^\beta} \frac{\Gamma(\beta + \frac{1}{2})}{(2\beta)! \Gamma(\frac{1}{2}(q_i + t_i) + \beta)}}$$

This is for $s_i + s_{p+1} = v$. The exponent of e vanishes, and substituting $s_i = v - s_{p+1}$ we get for $0 \leq s_{p+1} \leq v$

$$(7.10) \quad h(s; \lambda, \delta | r) =$$

$$C(v; \lambda, \delta) s_{p+1}^{\frac{1}{2}q_i-1} (v - s_{p+1})^{\frac{1}{2}t_i-1} \frac{\sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} s_{p+1}^\beta}{\lambda_i^\beta} \frac{\Gamma(\beta + \frac{1}{2})}{(2\beta)! \Gamma(\frac{1}{2}q_i + \beta)}}$$

where

$$C(v; \lambda, \delta) =$$

$$\frac{1}{\Gamma(\frac{1}{2}t_i) v^{\frac{1}{2}(q_i+t_i)-1}} \frac{\sum_{\beta=0}^{\infty} \frac{\delta^{2\beta} v^\beta}{\lambda_i^\beta} \frac{\Gamma(\beta + \frac{1}{2})}{(2\beta)! \Gamma(\frac{1}{2}(q_i + t_i) + \beta)}}$$

The hypothesis H' is equivalent to the hypothesis $\delta^2 = 0$ against $\delta^2 > 0$. According to Neyman-Pearson's fundamental lemma, the most powerful test against an alternative $\delta_1^2 > 0$ is given by the rejection region

$$(7.11) \quad \frac{h(s; \lambda, \delta_1 | r)}{h(s; \lambda, 0 | r)} > \text{constant},$$

or

$$(7.12) \quad \frac{C(v; \lambda, \delta_1)}{C(v; \lambda, 0)} \sum_{\beta=0}^{\infty} \frac{\delta_1^{2\beta} s_{p+1}^{\beta}}{\lambda_i^{\beta}} \frac{\Gamma(\beta + \frac{1}{2})}{(2\beta)! \Gamma(\frac{1}{2}q_i + \beta)} > \text{constant}.$$

The left-hand side is an increasing function of s_{p+1} . Thus the rejection region is given by

$$(7.13) \quad s_{p+1} > \text{constant}.$$

This is for given value of R . The constant may be a function of R . It must be determined from the condition

$$(7.14) \quad \alpha = P_{\delta=0}(s_{p+1} > C(R) | R)$$

for all values of R where $C(R)$ is the constant.

We get

$$\begin{aligned} \alpha &= P_{\delta=0} \left(\frac{s_{p+1}}{s_i + s_{p+1}} > \frac{C(R)}{s_i + s_{p+1}} \mid R \right) \\ (7.15) \quad &= P_{\delta=0} \left(\frac{\frac{s_{p+1}}{s_i}}{1 + \frac{s_{p+1}}{s_i}} > C_1(R) \mid R \right) \\ &= P_{\delta=0} \left(\frac{s_{p+1}}{s_i} > C_2(R) \mid R \right) \end{aligned}$$

where C_1 and C_2 are new constant. When $\delta = 0$, $\frac{S_{p+1}}{\lambda_i}$ and $\frac{S_i}{\lambda_i}$ are independent and chi-square distributed with q_i and t_i degrees of freedom respectively. Then $\frac{S_{p+1}}{S_i}$ and $S_i + S_{p+1}$ are independent, and from the definition of R it follows that $\frac{S_{p+1}}{S_i}$ is independent of R . This proves that the function $C_2(R)$ must be a constant C^* not depending on R . Thus we have

$$\begin{aligned} \alpha &= P_{\delta=0} \left(\frac{S_{p+1}}{S_i} > C^* \right) \\ (7.16) \quad &= P_{\delta=0} \left(\frac{\frac{1}{q_i} \frac{S_{p+1}}{\lambda_i}}{\frac{1}{t_i} \frac{S_i}{\lambda_i}} > \frac{t_i}{q_i} C^* \right). \end{aligned}$$

The left-hand side within the parenthesis has an F -distribution with q_i and t_i degrees of freedom for $\delta = 0$. It follows that $\frac{t_i}{q_i} C^* = f_{1-\alpha; q_i, t_i}$.

We can now formulate the following theorem.

THEOREM 8. Given the hypothesis

$$H : M_{i \setminus i} = 0 \quad \text{where rank } M_i^{q_i \times r_i} = q_i \leq r_i.$$

Then we have:

The transformation to the canonical form can be performed such that the hypothesis is equivalent to

$$H' : \chi_{i1} = 0 \quad \text{where } \chi_{i1} = E(Z_{i11}^{q_i \times 1})$$

and where Z_{i11} is a sub-vector of Z_{i1} .

The hypothesis problem H' is invariant under a

group G of transformations defined by (7.3). A UMP α -test for H' among all tests that are both invariant and unbiased is given by the rejection region

$$(7.17) \quad \frac{t_i}{q_i} \frac{Z'_{i11} Z_{i11}}{Z'_{i2} Z_{i2}} > f_{1-\alpha; q_i, t_i},$$

where the statistic of the left-hand side has an F -distribution with q_i and t_i degrees of freedom and non-centrality parameter $\sqrt{\frac{1}{\lambda_i} \gamma'_{i1} \gamma_{i1}}$.

When $q_i = 1$, the test (7.17) is UMP unbiased.

PROOF. From what is earlier said the theorem is proved with the exception of the last statement.

When $q_i = 1$, the density function of Z can be written

$$(7.18) \quad C(\theta, \lambda) \exp(\theta_{i1} z_{i11} - \frac{1}{2} \sum_{\alpha=1}^p \frac{z'_{\alpha 1} z_{\alpha 1} + z'_{\alpha 2} z_{\alpha 2}}{\lambda_{\alpha}} + \sum_{\alpha=1}^p \sum_{\beta=1}^{r_{\alpha}} \theta_{\alpha\beta} z_{\alpha 1\beta})$$

($\alpha, \beta \neq (i, 1)$)

where $\theta_{i1} = \frac{\gamma_{i1}}{\lambda_i}$. The hypothesis $\gamma_{i1} = 0$ is equivalent to the hypothesis $\theta_{i1} = 0$.

It is seen that (7.18) constitutes an exponential family of distributions, and it follows that we can use the same theory as under the proof of Theorem 7. We define

$$(7.19a) \quad U = \frac{Z_{i11}^2}{(Z'_{i1} Z_{i1} + Z'_{i2} Z_{i2}) - Z'_{i12} Z_{i12}} = \frac{Z_{i11}^2}{Z'_{i2} Z_{i2} + Z_{i11}^2}$$

$$(7.19b) \quad = \frac{\frac{Z_{i11}^2}{Z_{i2}'Z_{i2}}}{1 + \frac{Z_{i11}^2}{Z_{i2}'Z_{i2}}}$$

When $\Theta_{i1} = 0$, $\frac{Z_{i11}^2}{\lambda_i}$ and $\frac{1}{\lambda_i} Z_{i2}'Z_{i2}$ are independent and chi-square distributed with 1 and t_i degrees of freedom, respectively. Then $\frac{Z_{i11}^2}{Z_{i2}'Z_{i2}}$ and $Z_{i2}'Z_{i2} + Z_{i11}^2$ are independent. From (7.19b) it now follows that U and $Z_{i2}'Z_{i2} + Z_{i11}^2$ are independent. In general U and $Z_{i12}'Z_{i12}$ are independent. For $\Theta_{i1} = 0$ it then follows that U is independent of $Z_{i2}'Z_{i2} + Z_{i11}^2 + Z_{i12}'Z_{i12} = Z_{i2}'Z_{i2} + Z_{i1}'Z_{i1}$. Thus for $\Theta_{i1} = 0$, U is independent of all sufficient and complete statistics under the hypothesis. The same is true for $T = \frac{Z_{i11}}{|Z_{i11}|} \sqrt{U}$, the sign of Z_{i11} being independent of Z_{i11}^2 . From (7.19a) it is seen that T is linear and increasing as a function of Z_{i11} for fixed values of the sufficient and complete statistics. We can apply Theorem 1, p. 161 of [2], and get the rejection region $T > C_1$ and $T < C_2$ where C_1 and C_2 are constants determined from the condition of unbiasedness. Under the hypothesis the density of

$$(7.20) \quad T = \frac{\frac{Z_{i11}}{\sqrt{Z_{i2}'Z_{i2}}}}{\sqrt{1 + \frac{Z_{i11}^2}{Z_{i2}'Z_{i2}}}}$$

is symmetric with respect to the origin. Then unbiasedness implies $C_2 = -C_1$, such that the rejection region is given by $|T| > \text{constant}$ or equivalently $T^2 > \text{constant}$. But T^2 is an increasing function of $\frac{Z_{i11}^2}{Z_{i2}'Z_{i2}}$. Thus we shall reject

the hypothesis when

$$(7.21) \quad \frac{Z_{i11}^2}{Z_{i2}'Z_{i2}} > \text{constant}.$$

Under the hypothesis $\frac{t_i Z_{i11}^2}{Z_{i2}'Z_{i2}}$ have an F-distribution with 1 and t_i degrees of freedom. From the condition of unbiasedness we get the rejection region to be

$$(7.22) \quad \frac{t_i Z_{i11}^2}{Z_{i2}'Z_{i2}} > f_{1-\alpha; 1, t_i}.$$

It is seen that it is a special case of the general test (7.17). The theorem is proved.

The power function depends on the parameters only through $\delta = \sqrt{\frac{1}{\lambda_i} Z_{i11}' \delta_{i11}}$, and is given by

$$\begin{aligned} \beta(\delta) &= P \left(\frac{t_i}{q_i} \frac{Z_{i11}' Z_{i11}}{Z_{i2}' Z_{i2}} > f_{1-\alpha; q_i, t_i} \right) \\ &= 1 - F_{\delta; q_i, t_i}(f_{1-\alpha; q_i, t_i}), \end{aligned}$$

where $F_{\delta; q_i, t_i}$ is the cumulative F-distribution with q_i and t_i degrees of freedom and noncentrality parameter δ .

Using Lemma 4 of the Appendix with A_i as A , M_i as B and P_{i11} as Q it follows that

$$(7.23) \quad \begin{aligned} Z_{i11}' Z_{i11} &= X' P_{i11}' P_{i11} X \\ &= X' A_i (A_i' A_i)^{-1} M_i' (M_i (A_i' A_i)^{-1} M_i')^{-1} M_i (A_i' A_i)^{-1} A_i' X. \end{aligned}$$

From Section 5 it is known how $\hat{\lambda}_i = \frac{1}{t_i} Z'_{i2} Z_{i2}$ can be expressed by X and the matrices defining the model. Hence we have an expression for the test statistic (7.17) by means of the matrices defining the model and the matrix M_i defining the hypothesis. We also get

$$\begin{aligned} \delta^2 &= \frac{1}{\lambda_i} \gamma'_{i1} \gamma_{i1} = \frac{1}{\lambda_i} \beta_i' A_i' P_{i11} P_{i11} A_i \beta_i \\ (7.24) \quad &= \frac{1}{\lambda_i} \beta_i' M_i' (M_i (A_i' A_i)^{-1} M_i')^{-1} M_i \beta_i. \end{aligned}$$

With a simple modification of the statistic (7.17) we can derive tests for hypotheses of the form

$$(7.25) \quad H^* : M_i \beta_i = a$$

where a is a known vector. Let β_i^0 be a solution of the equation $M_i \beta_i = a$. Then the hypothesis can be given as

$$(7.26) \quad H^* : M_i (\beta_i - \beta_i^0) = 0.$$

We use the same P_{i11} as in the proof of Theorem 8, but instead of Z_{i11} we introduce

$$(7.27) \quad Z_{i11}^* = Z_{i11} - P_{i11} A_i \beta_i^0 = P_{i11} (X - A_i \beta_i^0),$$

and define γ_{i1}^* by

$$(7.28) \quad E(Z_{i11}^*) = \gamma_{i1}^* = P_{i11} A_i (\beta_i - \beta_i^0).$$

We have $\gamma_{i1}^* = 0$ if and only if $\beta_i - \beta_i^0 \perp \mathcal{R}(P_{i11} A_i) = \mathcal{R}(M_i)$ which is true if and only if $M_i (\beta_i - \beta_i^0) = 0$. Hence the hypo-

thesis is equivalent to

$$(7.29) \quad H^*: \gamma_{i1}^* = 0.$$

The problem is thereby reduced to that of Theorem 8. We only have to replace Z_{i11} by Z_{i11}^* in the statistic (7.17). From (7.27) it follows that when expressing the test statistic by means of X , we have to replace X by $X - A_i \beta_i^0$ in the numerator of the statistic. Hence to test the hypothesis $M_i \beta_i = a$ we shall use the statistic

$$(7.30) \quad \frac{1}{q_i \hat{\lambda}_i} (X - A_i \beta_i^0)' A_i (A_i' A_i)^{-1} M_i' (M_i (A_i' A_i)^{-1} M_i')^{-1} M_i (A_i' A_i)^{-1} A_i' (X - A_i \beta_i^0),$$

or

$$(7.31) \quad \frac{1}{q_i \hat{\lambda}_i} (M_i (A_i' A_i)^{-1} A_i' X - M_i \beta_i^0)' (M_i (A_i' A_i)^{-1} M_i')^{-1} (M_i (A_i' A_i)^{-1} A_i' X - M_i \beta_i^0).$$

From Section 5 we have $\hat{\beta}_i = (A_i' A_i)^{-1} A_i' X$ and $\sum (\hat{\beta}_i) = (A_i' A_i)^{-1} \lambda_i$. Hence

$$(7.32) \quad M_i \hat{\beta}_i = M_i (A_i' A_i)^{-1} A_i' X,$$

$$\sum (M_i \hat{\beta}_i) = M_i (A_i' A_i)^{-1} M_i' \lambda_i.$$

Using this and the condition $M_i \beta_i^0 = a$, (7.31) can be written

$$(7.33) \quad \frac{1}{q_i \hat{\lambda}_i} (M_i \hat{\beta}_i - a)' \left(\frac{1}{\lambda_i} \sum (M_i \hat{\beta}_i) \right)^{-1} (M_i \hat{\beta}_i - a),$$

which is the usual test statistic in such cases.

Example (cont.). We have

$$\Delta = \frac{J\sigma^2}{J\sigma^2 + \tau^2}$$

From (6.4) and (6.20) follows that the hypothesis $\Delta_1 \leq \Delta_0$ should be rejected when

$$\frac{IJ-I-1}{I-2} \frac{1}{1+\Delta_0} \frac{J(\sum (x_{i.} - x_{..})^2 - \frac{(\sum x_{i.}(t_i - t_{..}))^2}{\sum (t_i - t_{..})^2})}{\sum (x_{ij} - x_{i.})^2 - \frac{(\sum x_{ij}(u_{ij} - u_{..}))^2}{\sum (u_{ij} - u_{..})^2}} > f_{1-\alpha; I-2, IJ-I-1}.$$

For fixed effects let us first consider the hypothesis

$$H_{\beta} : \beta = 0$$

This corresponds to $M_1 \beta_1 = 0$ where $M_1 = (0, 1)$ and $\beta_1' = (\bar{\mu}, \beta)$. We find

$$M_1 (A_1' A_1)^{-1} A_1' = \frac{1}{J \sum (t_i - t_{..})^2} (t_1 - t_{..}, \dots, t_1 - t_{..}, t_2 - t_{..}, \dots, t_I - t_{..}),$$

and

$$M_1 (A_1' A_1)^{-1} M_1' = \frac{1}{J \sum (t_i - t_{..})^2}.$$

From the theory it follows that we shall reject the hypothesis when

$$(I-2) \frac{(\sum X_{i.}(t_i - t_{..}))^2}{\sum (t_i - t_{..})^2} \frac{J_{\beta}^2}{J_{\beta}^2 + L^2} > f_{1-\alpha; 1, I-2}$$

The noncentrality parameter is given by the square root of $\frac{J_{\beta}^2}{J_{\beta}^2 + L^2} \sum (t_i - t_{..})^2$.

Another hypothesis of interest is

$$H_{\gamma} : \gamma = 0$$

which corresponds to $M_2 \beta_2 = 0$ when $M_2 = (1)$ and $\beta_2 = (\gamma)$. We get the rejection region

$$(IJ-I-1) \frac{(\sum X_{ij}(u_{ij} - u_{..}))^2}{\sum (u_{ij} - u_{..})^2} \frac{J_{\beta}^2}{J_{\beta}^2 + L^2} > f_{1-\alpha; 1, IJ-I-1}.$$

The noncentrality parameter is $\frac{|\gamma|}{L} \sqrt{\sum (u_{ij} - u_{..})^2}$.

Since $q_1 = q_2 = 1$ the test for H_{β} and H_{γ} are UMP unbiased. This fact can be used to find uniformly most accurate unbiased confidence intervals for the parameters β and γ (see [2], pp.176-177).

8. MAXIMUM LIKELIHOOD ESTIMATES AND LIKELIHOOD RATIO TESTS.

It is well known from multivariate normal theory that maximum likelihood estimates and likelihood ratio tests give reasonable and often optimal estimates and tests. In the model treated here, however, the likelihood ratio estimates cannot all be given explicitly. This is because of the conditions $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ on the variances. We encounter the same difficulties as Herbach [1].

The probability density function of Z is given by

$$(8.1) \quad f(z; \gamma, \lambda) = (2\pi)^{-\frac{n}{2}} \left(\prod_{i=1}^p \lambda_i - \frac{t_i + s_i}{2} \right) \exp \left(-\frac{1}{2} \sum_{i=1}^p \frac{(z_{i1} - \gamma_i)'(z_{i1} - \gamma_i) + z_{i2}' z_{i2}}{\lambda_i} \right)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ and with no restrictions on $\gamma_1, \gamma_2, \dots, \gamma_p$.

If one or more of the λ_i 's approach 0 or ∞ , then $f(z; \gamma, \lambda) \rightarrow 0$. The same is true if one or more elements of the γ_i 's approach $-\infty$ or ∞ . Not taking into account the conditions on the λ_i 's, the maximum of $f(z, \gamma, \lambda)$ is obtained by setting equal to 0 the first derivatives of $f(z; \gamma, \lambda)$ with respect to the parameters, and solving the equations so obtained. If the solution is unique, it will give the maximum of $f(z; \gamma, \lambda)$, and if, in addition, it satisfies the conditions on the λ_i 's, the solution gives the likelihood estimates.

It is easily seen that the maximum likelihood

estimates of $\gamma_1, \gamma_2, \dots, \gamma_p$ are given by

$$(8.2) \quad \hat{\gamma}_i = z_{i1} \quad i = 1, 2, \dots, p.$$

These coincide with the minimum variance estimates obtained in Section 5.

We then proceed to maximize

$$(8.3) \quad h(z; \lambda) = (2\pi)^{-\frac{n}{2}} \left(\prod_{i=1}^p \lambda_i^{-\frac{t_i+s_i}{2}} \right) \exp\left(-\frac{1}{2} \sum_{i=1}^p \frac{z'_{i2} z_{i2}}{\lambda_i}\right).$$

To maximize $h(z; \lambda)$ is the same as to maximize

$$(8.4) \quad L = \ln h(z; \lambda) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^p (t_i + r_i) \ln \lambda_i - \frac{1}{2} \sum_{i=1}^p \frac{z'_{i2} z_{i2}}{\lambda_i}.$$

We get

$$(8.5) \quad \begin{aligned} \frac{\partial L}{\partial \lambda_i} &= -\frac{1}{2} \frac{t_i + r_i}{\lambda_i} + \frac{1}{2} \frac{z'_{i2} z_{i2}}{\lambda_i^2} \\ &= -\frac{1}{2} \frac{t_i + r_i}{\lambda_i^2} \left(\lambda_i - \frac{z'_{i2} z_{i2}}{t_i + r_i} \right) \quad i = 1, 2, \dots, p. \end{aligned}$$

We define $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$ as the solutions of $\frac{\partial L}{\partial \lambda_i} = 0$ ($i = 1, 2, \dots, p$). From (8.5),

$$(8.6) \quad \tilde{\lambda}_i = \frac{z'_{i2} z_{i2}}{t_i + r_i} \quad i = 1, 2, \dots, p.$$

If

$$(8.7) \quad \frac{z'_{12} z_{12}}{t_1 + r_1} \geq \frac{z'_{22} z_{22}}{t_2 + r_2} \geq \dots \geq \frac{z'_{p2} z_{p2}}{t_p + r_p},$$

then $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$ satisfy the side conditions, and are maximum likelihood estimates. In such cases the maximum likelihood estimates and the minimum variance estimates are

connected by

$$(8.8) \quad \tilde{\lambda}_i = \frac{t_i}{t_i + r_i} \hat{\lambda}_i \quad i = 1, 2, \dots, p.$$

But with positive probability it can happen that $\tilde{\lambda}_i < \tilde{\lambda}_j$ for the same $i < j$.

In such cases we have to modify the estimates $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$ to get the maximum likelihood estimates. For the latter we use the notation $\hat{\hat{\lambda}}_1, \hat{\hat{\lambda}}_2, \dots, \hat{\hat{\lambda}}_p$. It can be shown that if $\tilde{\lambda}_i < \tilde{\lambda}_{i+1}$ for some value of i , then we must have $\hat{\hat{\lambda}}_i = \hat{\hat{\lambda}}_{i+1}$. The proof is analogous to one of Herbach [1] for a similar situation in the random effects model. The hypothesis

$$H : \lambda_i = \lambda_{i+1} \text{ against } \lambda_i > \lambda_{i+1}$$

is the same as the hypothesis $\Delta_i = 0$. We shall try to derive the likelihood ratio test. First we show that the exponent $-\frac{1}{2} \sum_{i=1}^p \frac{1}{\hat{\lambda}_i} Z_{i2}^2 Z_{i2}$ is equal to $-\frac{n}{2}$. It can be shown that the general expression for the maximum likelihood estimates will be of the form

$$(8.9) \quad \hat{\hat{\lambda}}_{i_1} = \hat{\hat{\lambda}}_{i_2} = \dots = \hat{\hat{\lambda}}_{i_{s_j}} = \frac{Z'_{i_1} Z_{i_1}^2 + Z'_{i_2} Z_{i_2}^2 + \dots + Z'_{i_{s_j}} Z_{i_{s_j}}^2}{t_{i_1} + r_{i_1} + t_{i_2} + r_{i_2} + \dots + t_{i_{s_j}} + r_{i_{s_j}}}$$

$$j = 1, 2, \dots, k,$$

for some k and where $\sum_{j=1}^k s_j = p$. Hence

$$\begin{aligned}
 (8.10) \quad & -\frac{1}{2} \sum_{i=1}^p \frac{Z'_{i2} Z_{i2}}{\hat{\lambda}_i} = -\frac{1}{2} \sum_{j=1}^k \sum_{h=1}^{s_j} \frac{Z'_{ih} Z_{ih}}{\hat{\lambda}_{i_h}} \\
 & = -\frac{1}{2} \sum_{j=1}^k (t_{i_1} + r_{i_1} + t_{i_2} + r_{i_2} + \dots + t_{i_{s_j}} + r_{i_{s_j}}) = -\frac{1}{2} n
 \end{aligned}$$

It follows that the maximum of the probability density function can be written

$$(8.11) \quad (2\pi)^{-\frac{n}{2}} \left(\prod_{i=1}^p \hat{\lambda}_i^{-\frac{p}{2}} - \frac{t_i + r_i}{2} \right) \exp\left(-\frac{1}{2} n\right).$$

Let $\hat{\lambda}_{i\Omega}$ and $\hat{\lambda}_{i\omega}$ ($i = 1, 2, \dots, p$) denote the maximum likelihood estimates under a priori conditions and under the hypothesis respectively. (8.11) is true in both cases. The likelihood ratio is based upon

$$(8.12) \quad R = \prod_{i=1}^p \left(\frac{\hat{\lambda}_{i\Omega}}{\hat{\lambda}_{i\omega}} \right)^{\frac{t_i + r_i}{2}}.$$

But since we have no explicit expressions for the $\hat{\lambda}_i$'s, R cannot be given explicitly in terms of Z .

9. APPLICATIONS.

To avoid repetition the defined model will be denoted by $\mathcal{M}$.

a) Setting $p = 1$, $C_1 = I^{n \times n}$, $\sum(V_1) = \sigma^2 I^{n \times n}$, we get

$$X = A_1 \beta_1 + V_1$$

which is the usual fixed effects model. This is a special case of $\mathcal{M}$ provided we use a parametrization giving the matrix A_1 maximal rank. The test statistics for any hypotheses $M_1\beta_1 = a$ is given by (7.31) or (7.33).

b) The balanced nested classification variance components model we may define by

$$X_{i_1 i_2 \dots i_p} = +A_{i_1} + B_{i_1 i_2} + \dots + K_{i_1 i_2 \dots i_p}$$

$$\begin{aligned} i_1 &= 1, 2, \dots, I_1, \\ i_2 &= 1, 2, \dots, I_2, \\ &\vdots \\ i_p &= 1, 2, \dots, I_p, \end{aligned}$$

where A_{i_1} , $B_{i_1 i_2}$, ..., $K_{i_1 i_2 \dots i_p}$ are all independently and normally distributed with expectations 0 and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$, respectively. This is a special case of $\mathcal{M}$ with $V_1' = (A_1, A_2, \dots, A_{I_1})$, $V_2' = (B_{11}, B_{12}, \dots, B_{I_1 I_2})$ etc.

Both these models, however, have been treated in literature by simpler methods. Typical for $\mathcal{M}$ is that both fixed and random effects occur. From the literature we give two examples where optimal tests have been obtained.

c) Lehmann [2] chapter 7.8 treats the model

$$X_{ijk} = \mu_i + B_{ij} + C_{ijk} \quad i=1, 2, \dots, I, \quad j=1, 2, \dots, J, \quad k=1, 2, \dots, K,$$

where B_{ij} and C_{ijk} all are independently and normally distributed with expectations 0 and variances σ_B^2 and σ^2 respectively, and where $\mu_1, \mu_2, \dots, \mu_I$ are unknown parameters. It is easily seen that $\mathcal{M}$ is satisfied and that in particular $\beta_1' = (\mu_1, \mu_2, \dots, \mu_I)$. The hypothesis $\mu_1 = \mu_2 = \dots = \mu_I$ may be expressed by

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & -1 \\ 0 & 1 & 0 & \dots & 0 & 0 & -1 \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & -1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{I-1} \\ \mu_I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

where the matrix to the left is the matrix M specifying the hypothesis.

d) Torgersen [4] considers the model

$$X_{ij} = (\beta_j + W_j + V_{ij}) \quad i = 1, 2, \dots, I, \quad j = 1, 2, \dots, J,$$

where $(\beta_1, \beta_2, \dots, \beta_J)$ is an unknown point in an $s < J$ dimensional subspace of Euclidean J -space, and W_i and V_{ij} are independently and normally distributed with expectations 0 and variances τ^2 and σ^2 respectively. If we express the s -dimensional subspace by means of a matrix, we are back to $\mathcal{M}$.

e) Frequently we have models of the type

$$\begin{aligned} X_{ijk} &= \mu + U_i + \alpha t_{ij} + V_{ijk} & i &= 1, 2, \dots, I, \\ & & j &= 1, 2, \dots, J, \\ X_{ijk} &= \mu + U_i + \alpha_i t_{ij} + V_{ijk} & k &= 1, 2, \dots, K, \end{aligned}$$

with self-evident notations, and the usual normality assumptions. Both are special cases of $\mathcal{M}$ if

$t_{1.} = t_{2.} = \dots = t_{I.}$. In some cases we may define the model as

$$X_{ijk} = \mu + U_i + \alpha_i (t_{ij} - t_{i.}) + V_{ijk}$$

which clearly satisfies $\mathcal{M}$.

f) A model of another type is

$$X_{ijk} = \mu + U_i + \alpha_j + V_{ij} + W_{ijk} \quad \begin{array}{l} i = 1, 2, \dots, I, \\ j = 1, 2, \dots, J, \\ k = 1, 2, \dots, K, \end{array}$$

with the usual normality assumption, and where the unknown parameters $\alpha_1, \alpha_2, \dots, \alpha_J$ satisfy $\sum_{j=1}^J \alpha_j = 0$. When writing this in matrix notation with $\beta_1 = (\mu)$, $\beta_2' = (\alpha_1, \alpha_2, \dots, \alpha_{J-1})$, $V_1' = (U_1, U_2, \dots, U_I)$, $V_2' = (V_{11}, V_{12}, \dots, V_{IJ})$ and $V_3' = (W_{111}, W_{112}, \dots, W_{IJK})$, it is seen that it is a special case of $\mathcal{M}$.

g) Special cases of $\mathcal{M}$ are in general models of the type

$$X_{ijklm\dots} = \mu + U_i + \alpha_{ij} + V_{ijk} + \beta_{ijkl} + W_{ijklm} + \dots \quad \begin{array}{l} i = 1, 2, \dots, I, \\ j = 1, 2, \dots, J, \\ \vdots \\ m = 1, 2, \dots, M, \\ \vdots \end{array}$$

where U, V, W are the random effects, and where the fixed effects satisfy $\sum_{j=1}^J \alpha_{ij} = 0$, $\sum_{l=1}^L \beta_{ijkl} = 0$ etc. The model in f) is of this type with $\alpha_{ij} = \alpha_j$ for all i and j .

APPENDIX

SOME LEMMAS CONCERNING MATRICES.

Four matrix lemmas used in the preceding sections are proved here.

LEMMA 1. Given a matrix $C^{m \times n}$. A vector $b^{n \times 1}$ is an eigenvector of the matrix $C'C$ with eigenvalue 0 if and only if $b \perp \mathcal{R}(C)$.

PROOF. Let $b \perp \mathcal{R}(C)$, then we have $(C'C - 0I)b = C'Cb = 0$. Conversely suppose that b is an eigenvector with eigenvalue 0. Then $0 = (C'C - 0I)b = C'(Cb)$, which gives $Cb \perp \mathcal{R}(C')$. But since also $Cb \in \mathcal{C}(C) = \mathcal{R}(C')$, we must have $Cb = 0$, and hence $b \perp \mathcal{R}(C)$.

LEMMA 2. Let the matrix C be defined as in Lemma 1. Suppose that all eigenvalues of $C'C$ not equal to 0 are identical, λ say. A vector $b^{n \times 1}$ is an eigenvector with eigenvalue λ for $C'C$ if and only if $b \in \mathcal{R}(C)$.

PROOF. If b is an eigenvector with eigenvalue λ , then b must be orthogonal to all eigenvectors with eigenvalue 0. According to Lemma 1 we must have $b \in \mathcal{R}(C)$. Conversely, suppose $b \in \mathcal{R}(C)$. Let $a_1, a_2, \dots, a_p$ be a basis for the space of eigenvectors with eigenvalue λ , and hence a basis for $\mathcal{R}(C)$. It follows that b can be written $b = \alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_p a_p$ for some $\alpha_1, \alpha_2, \dots, \alpha_p$. We have

$(C'C - \lambda I)b = \sum_{i=1}^p \alpha_i (C'C - \lambda I)a_i = 0$ which shows that b is an eigenvector with eigenvalue λ .

LEMMA 3. Given two matrices $A^{n \times m}$ and $Q^{m \times n}$ where $\text{rank } A = \text{rank } Q = m \leq n$, $\mathcal{R}(Q) = \mathcal{L}(A)$ and $QQ' = I$. Then we have:

(I) There exists a nonsingular matrix $B^{m \times m}$ such that

$$Q = BA',$$

(II) QA is nonsingular,

$$(III) \quad Q'Q = A(A'A)^{-1}A',$$

$$(IV) \quad (QA)^{-1}Q = (A'A)^{-1}A'.$$

PROOF. (I) Follows from the fact that the rows of Q are linear combinations of the columns of A . $\text{Rank } B = m$ we get from $m = \text{rank } Q = \text{rank } BA' \leq \text{rank } B \leq m$.

(II) From (I) we get $QA = BA'A$. Both B and $A'A$ are nonsingular with rank m . The product must have rank m , and since QA is an $(m \times m)$ -matrice, it is nonsingular.

(III) We have $I = QQ' = BA'AB'$, hence $B'B = (A'A)^{-1}$. This together with $Q = BA'$ implies $Q'Q = AB'BA' = A(A'A)^{-1}A'$.

$$(IV) \quad (QA)^{-1}Q = (BA'A)^{-1}BA' = (A'A)^{-1}A'.$$

LEMMA 4. Given two matrices $A^{n \times m}$ and $B^{p \times m}$ where $\text{rank } B = p \leq \text{rank } A = m \leq n$.

(I) There exists a matrix $Q^{p \times n}$ with orthogonal rows such that $\mathcal{R}(Q) \subset \mathcal{L}(A)$ and $\mathcal{R}(QA) = \mathcal{R}(B)$.

$$(II) \quad Q'Q = A(A'A)^{-1}B'(B(A'A)^{-1}B')^{-1}B(A'A)^{-1}A'.$$

PROOF. (I) From $\text{rank } B \leq \text{rank } A = m$ we get $\mathcal{R}(B) \subset \mathcal{R}(A)$. Then there exists a matrix $R^{p \times n}$ such that $RA = B$. It is

possible to choose R such that $\mathcal{R}(R) \subset \mathcal{L}(A)$. For, given any matrix R , let r_i be the i th row of R . We write $r_i = r_{i1} + r_{i2}$ where $r_{i1} \in \mathcal{L}(A)$ and $r_{i2} \perp \mathcal{L}(A)$. Correspondingly we may write the whole matrix $R = R_1 + R_2$. Since $\mathcal{R}(R_2) \perp \mathcal{L}(A)$, we get $RA = R_1A + R_2A = R_1A$. We may therefore replace R by R_1 where $\mathcal{R}(R_1) \subset \mathcal{L}(A)$. Rank $R = p$ is proved from the relation $p = \text{rank } B \leq \text{rank } RA \leq \text{rank } R \leq p$.

There exists a nonsingular matrix $S^{p \times p}$ such that the matrix $Q^{p \times n}$ defined by $Q = SR$ has orthogonal rows. It follows that $\mathcal{R}(Q) \subset \mathcal{R}(R)$, and hence $\mathcal{R}(Q) \subset \mathcal{L}(A)$. We have $QA = SRA = SB$. Hence $\mathcal{R}(QA) = \mathcal{R}(SB) = \mathcal{R}(B)$.

(II) Since $\mathcal{R}(Q) \subset \mathcal{L}(A)$, there exists a matrix $D^{p \times m}$ satisfying $Q = DA'$. From $QA = SB$ we now get $DA'A = SB$, or since $A'A$ is nonsingular $D = SB(A'A)^{-1}$. It follows that $Q = SB(A'A)^{-1}A'$. Since Q has orthogonal rows $I = QQ' = SB(A'A)^{-1}B'S'$, which implies $S'S = (B(A'A)^{-1}B')^{-1}$. We get $Q'Q = A(A'A)^{-1}B'S'SB(A'A)^{-1}A' = A(A'A)^{-1}B'(B(A'A)^{-1}B')^{-1}B(A'A)^{-1}A'$.

REFERENCES

- [1] HERBACH, L.H. (1959). Properties of model II - Type analysis of variance tests, A: Optimum nature of the F-test for model II in the balanced case.
Ann.Math.Statist. 30 939-959.

- [2] LEHMANN, E.L. (1959). Testing Statistical Hypotheses.
Wiley, New York.

- [3] LEHMANN, E.L. and SCHEFFÉ, H. (1950). Completeness, similar regions, and unbiased estimation - Part I.
Sankhyā 10 305-340.

- [4] TORGERSEN, E. (1963). On a class of combined "One-way classification (model II)-covariance" models.
Skand.Aktuarietidskr. 46 167-172.