Ingram Olkin* and Michael Væth**

MAXIMUM LIKELIHOOD ESTIMATION IN A TWO - WAY ANALYSIS OF VARIANCE WITH CORRELATED ERRORS IN ONE CLASSIFICATION

Preprint 1980 No. 1

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

January 1980

Abstract. A two - way analysis of variance model with correlated errors in one classification is discussed. It is assumed that the p measurements in each row have a general covariance matrix Σ. The maximum likelihood estimates of the row and column parameters as well as of the covariance matrix are obtained and their asymptotic distribution are discussed. Finally, the likelihood ratio test statistics for the various hypotheses in a two - way lay out are given.

Key words and phrases: Two - way analysis of variance, correlated errors, multivariate analysis, maximum likelihood theory.

1. Introduction

The origin of this study is an analysis of variance two - way classification model

$$(1.1) \qquad z_{ij} = \alpha_i + \beta_j + \mu + \varepsilon_{ij}, \qquad i=1,\dots,k; \qquad j=1,\dots,p,$$

where now $\varepsilon_i \equiv (\varepsilon_{i1},\dots,\varepsilon_{ip})$, $i=1,\dots,k$ are independently distributed with a multivariate normal distribution with zero means and common covariance matrix Σ. If Σ is a general covariance matrix then not all the variances and covariances are estimable. This difficulty does not arise when replications in each cell are available. It may also not arise if Σ is a patterned matrix. For example, if $\sigma_{ii} = \sigma^2$, and $\sigma_{ij} = \sigma^2\rho$, $i \neq j$, or if Σ is the covariance matrix derived from a first - order autocorrelation model, then Σ is estimable. The latter model is considered by Box (1954), and by Andersen, Jensen and Schou (1978) where it arises from a study with p time points and k subjects, but where each subject has a different level. In both of these papers the behavior of the usual two - way analysis of variance test - statistics are studied and approximations to their distributions are given.

The present analysis gives the maximum likelihood solution assuming that there are N replications in each cell and that the p measurements have a general covariance matrix.

In Section 2 we represent the model in a canonical form, and in Section 3 find maximum likelihood estimates of the parameters; some commentary on the sampling distribution of the maximum likelihood estimates is also included. The asymptotic distribution and Fisher information are discussed in Section 3.1, and in Section 4 likelihood ratio test statistics for the hypotheses usually considered in a two-way analysis of variance are given. The data that generated this study is analyzed in Section 5.

2. A Canonical Form

Since there are $N > 1$ observations per cell we first reduce the model to a consideration of the sufficient statistics $\bar{Z}$, C, where $\bar{Z}$ is the $k \times p$ matrix of cell means, and C is the $p \times p$ pooled cross-product matrix. The rows of $\bar{Z}$ are independently distributed, each having a multivariate normal (MVN) distribution with common covariance matrix Σ / N and with means

$$(2.1) \qquad E\,\bar{Z} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} (1,\ldots,1) + \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} (\beta_1,\ldots,\beta_p) + \mu \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix}$$

$$\equiv \alpha' e_p + e_k' \beta + \mu e_k' e_p ,$$

where $e_m = (1,\ldots,1)$ is an m dimensional vector. The matrix C has a Wishart distribution, $W(\Sigma,p,n)$, where $n = k(N-1)$.

We first transform $(\bar{Z},C)$ to (Z^*,C^*)

$$(2.2) \qquad Z^* = \sqrt{N}\,\Gamma\,\bar{Z}\,\Delta, \qquad C^* = \Delta'\,C\,\Delta ,$$

where $\Gamma : k \times k$ is an orthogonal matrix with first row $e_k / \sqrt{k}$ and $\Delta : p \times p$ is an orthogonal matrix with first column $e_p / \sqrt{p}$. Then

$$(2.3) \qquad EZ^* = \sqrt{N}(\Gamma\alpha')\sqrt{p}\;(1\;0\;\ldots\;0) + \sqrt{N}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\sqrt{k}\beta\Delta + \sqrt{N}\mu\sqrt{kp}\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

$$\equiv \alpha^{*\prime}\,(1\ 0\ \ldots\ 0) + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \beta^* + \mu^* \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_1^* + \beta_1^* + \mu^* & \beta_2^* & \ldots & \beta_p^* \\ \alpha_2^* & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ \alpha_k^* & 0 & \ldots & 0 \end{pmatrix},$$

where

$$\alpha^* = \sqrt{pN}\,\alpha\,\Gamma' , \qquad \beta^* = \sqrt{kN}\,\beta\,\Delta , \qquad \mu^* = \mu\sqrt{kpN} .$$

The rows of Z^* are independently distributed, each having a MVN distribution with $E\,Z^*$ given by (2.3) and with common covariance matrix $\Sigma^* = \Delta'\,\Sigma\,\Delta$. Furthermore, C^* has a Wishart distribution $W(\Sigma^*,p,n)$.

Remark. In the event that $N = 1$, then C^* does not exist, and we see from (2.3) that σ_{11}^* is not estimable.

We can use (2.3) as the starting point for any further analysis. However, before doing this, we extend the model (2.1) and carry out the analysis on the more general version.

2.1 An Extension

In (2.1) we can consider the vectors e_p and e_k as design matrices. But these can be made more general by replacing (2.1) by

(2.4) $$E\,\overline{Z} = A\,X_1 + X_2\,B ,$$

where $A : k \times m$, $X_1 : m \times p$, $X_2 : k \times q$, $B : q \times p$, X_1 is of rank $m \leqq p$, X_2 is of rank $q \leqq k$, and $k \geqq p$. Write X_1 and X_2 as

$$(2.5) \qquad X_1 = T\,(I_m 0)\,\Delta', \qquad X_2' = U\,(I_q 0)\,\Gamma',$$

where $T : m \times m$ and $U : q \times q$ are nonsingular square roots of $X_1 X_1'$ and $X_2' X_2$, respectively; $\Delta : p \times p$ and $\Gamma : k \times k$ are orthogonal.

Make the transformation

$$(2.6) \qquad Z^* = \sqrt{N}\,\Gamma\,\bar{Z}\,\Delta\,, \qquad C^* = \Delta'\,C\,\Delta\,,$$

and write

$$(2.7) \qquad A^* = A\,T\,, \qquad B^* = U'\,B\,.$$

Then

$$(2.8) \qquad E\,Z^* = A^*\,(I_m 0) + \begin{pmatrix} I_q \\ 0 \end{pmatrix} B^* = \begin{pmatrix} A_1^* & 0 \\ A_2^* & 0 \end{pmatrix} + \begin{pmatrix} B_1^* & B_2^* \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} A_1^* + B_1^* & B_2^* \\ A_2^* & 0 \end{pmatrix} \equiv \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & 0 \end{pmatrix},$$

where

$$A_1^* : q \times m, \quad A_2^* : k-q \times m, \quad B_1^* : q \times m, \quad B_2^* : q \times p-m\,,$$

$$\Theta_{11} = A_1^* + B_1^*, \quad \Theta_{12} = B_2^*, \quad \Theta_{21} = A_2^*\,.$$

Thus our starting point is (Z^*,C^*) with the rows of Z^* independent, each having a multivariate normal distribution with a common covariance matrix Σ and means

$$E\, Z^* = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & 0 \end{pmatrix}.$$

The matrix C*, which is independent of Z*, has a Wishart distribution $W(\Sigma,p,n)$. For simplicity of notation we omit the asterisk on C* and write

$$Z^* \equiv \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix},$$

where $Y_1 : q \times p$, $Y_2 : k-q \times p$;

$$Y_{11} : q \times m, \quad Y_{12} : q \times p-m, \quad Y_{21} : k-q \times m, \quad Y_{22} : k-q \times p-m;$$

$$E\,Y_1 = (\Theta_{11},\Theta_{12}) \equiv \Theta_1, \qquad E\,Y_2 = (\Theta_{21},0) = \Theta_2 .$$

Then the joint density of Y_1, Y_2 and C are given by

$$(2.9) \qquad p(Y_1,Y_2,C) = c(p,k,N)\,|\Sigma|^{-\frac{kN}{2}}\, e^{-\frac{1}{2}\mathrm{tr}\,\Sigma^{-1}(Y_1-\Theta_1)'(Y_1-\Theta_1)}$$

$$\times\, e^{-\frac{1}{2}\mathrm{tr}\,\Sigma^{-1}(Y_2-(\Theta_{21},0))'(Y_2-(\Theta_{21},0))}$$

$$\times\, e^{-\frac{1}{2}\mathrm{tr}\,\Sigma^{-1}C} .$$

3. Maximum Likelihood Estimation (MLE)

From (2.9) it is immediate that $\hat{\Theta}_1 = Y_1$. To obtain the MLE of Θ_{21}, let $\Lambda = \Sigma^{-1}$ and note that

$$\text{(3.1)} \quad \operatorname{tr}(Y_2 - (\Theta_{21},0))\Lambda(Y_2 - (\Theta_{21},0))'$$

$$= \operatorname{tr} Y_2\Lambda Y_2' + \operatorname{tr} (\Theta_{21},0)\Lambda(\Theta_{21},0)' - 2\operatorname{tr} (\Theta_{21},0)\Lambda Y_2'$$

$$= \operatorname{tr} Y_2\Lambda Y_2' + \operatorname{tr} \Theta_{21}\Lambda_{11}\Theta_{21}' - 2\operatorname{tr} \Theta_{21}(\Lambda_{11}\Lambda_{12})Y_2'$$

$$= \operatorname{tr} Y_2\Lambda Y_2' + \operatorname{tr} (\Theta_{21} - Y_2(I,\Lambda_{11}^{-1}\Lambda_{12})')\Lambda_{11}(\Theta_{21} - Y_2(I,\Lambda_{11}^{-1}\Lambda_{12})')'$$

$$- \operatorname{tr} Y_2 \begin{pmatrix} I \\ \Lambda_{21}\Lambda_{11}^{-1} \end{pmatrix} \Lambda_{11}(I,\Lambda_{11}^{-1}\Lambda_{12})Y_2',$$

where Λ is partitioned to be conformable. Consequently,

$$\text{(3.2)} \qquad \hat{\Theta}_{21} = Y_2 \begin{pmatrix} I \\ \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1} \end{pmatrix},$$

and from (3.1) we obtain

$$\text{(3.3)} \qquad \operatorname{tr} Y_2'Y_2 \left[\begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} - \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12} \end{pmatrix} \right]$$

$$= \operatorname{tr} Y_2'Y_2 \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} = \operatorname{tr} Y_{22}' Y_{22} \Sigma_{22}^{-1} .$$

From (2.9) and (3.3) we need to determine

$$\text{(3.4)} \qquad \operatorname*{Max}_{\Sigma} |\Sigma|^{-\frac{kN}{2}} e^{-\frac{1}{2}\operatorname{tr} Y_{22}' Y_{22} \Sigma_{22}^{-1} - \frac{1}{2}\operatorname{tr} C \Sigma^{-1}} .$$

This maximization is carried out by Gleser and Olkin (1970, Lemma 4.1). The maximum is equal to

$$\text{(3.5)}\qquad \operatorname*{Max}_{\Omega}\ p(Y_1, Y_2, C) = \frac{c(p,k,N)\,(kN)^{\frac{pkN}{2}}\,|C_{22}|^{\frac{kN}{2}}\,e^{-\frac{1}{2}pkN}}{|C_{22} + Y_{22}'Y_{22}|^{\frac{kN}{2}}\,|C|^{\frac{kN}{2}}},$$

where $\Omega = \{\Theta_{11}, \Theta_{12}, \Theta_{21}, \Sigma\}$. The maximum of Σ is achieved by

$$\text{(3.6)}\qquad \hat{\Sigma} = \{C + \begin{pmatrix} C_{12}C_{22}^{-1} \\ I \end{pmatrix} (Y_{22}'Y_{22})\,(C_{22}^{-1}C_{21}, I)\} / kN .$$

Note that $E\,\hat{\Sigma} = \Sigma\,[(kN - q) / kN]$.

Using (3.6) in (3.2) the MLE of Θ_{21} becomes

$$\text{(3.7)}\qquad \hat{\Theta}_{21} = Y_2 \begin{pmatrix} I \\ -\,C_{22}^{-1}C_{21} \end{pmatrix}.$$

Notice that $\hat{\Theta}_1$ and $(\hat{\Theta}_{21}, \hat{\Sigma})$ are independent and that $\hat{\Theta}_{21}$ is an unbiased estimate of Θ_{21}. Furthermore one can show that

$$\operatorname{Var}(\hat{\Theta}_{21}) = (I_{k-q} \otimes \Lambda_{11}^{-1})\,(1 + \frac{p-m}{(N-1)\,k - p - 1}),$$

where $\otimes$ is the Kronecker product.

The sampling distribution of $\hat{\Theta}_{21}$ is rather complicated. If we let

$$V = (\hat{\Theta}_{21} - \Theta_{21})\,\Lambda_{11}^{\frac{1}{2}},$$

then after considerable calculation, when $m \leq p - m$, the density of V is proportional to

$$\int_{0<G<I_m} e^{-\frac{1}{2}\operatorname{tr} G V' V}\,|G|^{\frac{a}{2}}\,|I - G|^{\frac{b}{2}}\,dG,$$

where $a = Nk - q - p + 3m + 1$ and $b = p - 2m - 1$. The details of this derivation are omitted.

3.1 The Information Matrix

The asymptotic distribution of the maximum likelihood estimates follows from standard large sample theory once the Fisher information matrix is derived.

From the joint density (2.9) of Y_1, Y_2 and C the first partial derivatives of the log likelihood are

$$\frac{\partial \log L}{\partial \Theta_1} = (Y_1 - \Theta_1)\,\Lambda\,, \qquad \frac{\partial \log L}{\partial \Theta_{21}} = (Y_{21} - \Theta_{21})\,\Lambda_{11} - Y_{22}\,\Lambda_{21}\,,$$

$$\frac{\partial \log L}{\partial \Lambda} = \frac{kN}{2}\Sigma - \frac{1}{2}(Y_1-\Theta_1)'(Y_1-\Theta_1) - \frac{1}{2}(Y_{21}-\Theta_{21},Y_{22})'(Y_{21}-\Theta_{21},Y_{22}) - \frac{1}{2}C\,.$$

The Fisher information, $I(\Theta,\Lambda)$, can now be obtained as the matrix of covariances of the first partial derivatives, or as the negative of the expected value of the second (cross) partial derivatives. In either case,

$$I(\Theta,\Lambda) = \mathrm{diag}\,((I_q \otimes \Lambda),(I_{k-q} \otimes \Lambda_{11}),\frac{Nk}{2}(\Sigma \otimes \Sigma))\,,$$

from which

$$[I(\Theta,\Lambda)]^{-1} = \mathrm{diag}\,((I_q \otimes \Sigma),(I_{k-q} \otimes \Lambda_{11}^{-1}),\frac{2}{Nk}(\Lambda \otimes \Lambda))\,.$$

The transformation introduced in (2.6) implies that the factor N^{-1} only occurs in the bottom right block of $[I(\Theta,\Lambda)]^{-1}$.

4. Tests of Hypotheses

In the context of the analysis of variance model that motivated this study, the hypotheses of "no row effects" and "no column effects" are equivalent to $H_R : \Theta_{21} = 0$ and $H_C : \Theta_{12} = 0$, respectively. The alternative hypotheses of interest normally are $A_R : \Theta_{21} \neq 0$ and $A_C : \Theta_{12} \neq 0$. The hypothesis "no column and no row effects" is $H_{RC} = H_R \cap H_C$ and could be tested if H_R is not rejected or if H_C is not rejected, that is, we test H_{RC} versus the alternative H_R or the alternative H_C.

For each model we obtain the maximum of the likelihood, from which the likelihood ratio statistics can be constructed. The first order approximation to the asymptotic distribution is given. A more accurate approximation to the null distribution of the likelihood ratio statistic can be obtained using the method of Box (1949) (see also Anderson (1958) Section 8.6.1).

A general discussion of related tests is provided by Gleser and Olkin (1970) and we rely on some of these results. We note that a reduction of each problem can be accomplished using invariance since most of the problems are left invariant by a group of transformations similar to that discussed in Gleser and Olkin (1970). A detailed study of the general MANOVA problem from a decision-theoretic point of view was made by Kariya (1978).

4.1 No Row Effects Model

From (2.9) with $H_R : \Theta_{21} = 0$, we obtain

$$\hat{\Theta}_1 = Y_1, \quad \hat{\Sigma} = (C + Y_2'Y_2) / kN,$$

and

$$(4.1) \qquad \max_{H_R} p(Y_1, Y_2, C) = \frac{c(p,k,N)\ (kN)^{pkN/2}\ e^{-\frac{1}{2}pkN}}{|C + Y_2'Y_2|^{kN/2}} .$$

The likelihood ratio statistic is obtained as the ratio of (4.1) to (3.5):

$$L_R = \frac{\max_{H_R} p(Y_1, Y_2, C)}{\max_{\Omega} p(Y_1, Y_2, C)} = \frac{|C_{22} + Y_{22}'Y_{22}|^{\frac{kN}{2}}\ |C|^{\frac{kN}{2}}}{|C_{22}|^{\frac{kN}{2}}\ |C + Y_2'Y_2|^{\frac{kN}{2}}}$$

$$= \frac{|I + Y_{22}\, C_{22}^{-1}\, Y_{22}'|^{kN/2}}{|I + Y_2\, C^{-1}\, Y_2'|^{kN/2}} .$$

When H_R holds, $-2 \log L_R$ is approximately distributed as a χ^2 variable with $(k-q)m$ degrees of freedom.

4.2 No Column Effects Model

Write

$$\Phi = \begin{pmatrix} \Theta_{11} \\ \Theta_{21} \end{pmatrix} : k \times m, \qquad \dot{Y} = \begin{pmatrix} Y_{11} \\ Y_{21} \end{pmatrix}, \qquad \ddot{Y} = \begin{pmatrix} Y_{12} \\ Y_{22} \end{pmatrix},$$

then, with $H_C : \Theta_{12} = 0$, (2.9) becomes

$$(4.2) \qquad p(Y_1, Y_2, C) = c(p,k,N)\ |\Sigma|^{-\frac{kN}{2}}\ e^{-\frac{1}{2}\operatorname{tr} \Sigma^{-1} C}$$

$$\times\ e^{-\frac{1}{2}\operatorname{tr} \Sigma^{-1} (\dot{Y} - \Phi, \ddot{Y})' (\dot{Y} - \Phi, \ddot{Y})} .$$

A maximization similar to that in Section 3 yields

$$(4.3)\qquad \underset{H_C}{\operatorname{Max}}\, p(Y_1, Y_2, C) = c(p,k,N)\,(kN)^{\frac{pkN}{2}}\, e^{-\frac{1}{2}pkN} \frac{|C_{22}|^{\frac{kN}{2}}}{|C|^{\frac{kN}{2}}\, |C_{22} + \ddot{Y}'\ddot{Y}|^{\frac{kN}{2}}}.$$

The maximum is achieved by

$$\hat{\Sigma} = \{C + \begin{pmatrix} C_{12}C_{22}^{-1} \\ I \end{pmatrix} (\ddot{Y}'\ddot{Y})\,(C_{22}^{-1}C_{21}, I)\} / kN ,$$

$$\hat{\Phi} = \dot{Y} + \ddot{Y}\hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1} = \dot{Y} - \ddot{Y}C_{22}^{-1}C_{21} .$$

The likelihood ratio statistic is obtained as the ratio of (4.3) to (3.5):

$$L_C = \frac{\underset{H_C}{\operatorname{Max}}\ p(Y_1, Y_2, C)}{\underset{\Omega}{\operatorname{Max}}\ p(Y_1, Y_2, C)} = \frac{|C_{22} + Y_{22}'Y_{22}|^{kN/2}}{|C_{22} + Y_{22}'Y_{22} + Y_{12}'Y_{12}|^{kN/2}}$$

$$= \frac{1}{|I + Y_{12}\,(C_{22} + Y_{22}'Y_{22})^{-1}\, Y_{12}'|^{kN/2}} .$$

When H_C holds, $-2 \log L_C$ is approximately distributed as a χ^2 variable with $q(p-m)$ degrees of freedom.

4.3 No Row and No Column Effects Model

With $\Theta_{21} = 0$ and $\Theta_{12} = 0$, (2.9) becomes

$$(4.4) \quad p(Y_1, Y_2, C) = c(p,k,N)\, |\Sigma|^{-kN/2}\, e^{-\frac{1}{2}\operatorname{tr}\Sigma^{-1}(C + Y_2'Y_2)}$$

$$\times\, e^{-\frac{1}{2}\operatorname{tr}\Sigma^{-1}(Y_{11} - \Theta_{11}, Y_{12})'(Y_{11} - \Theta_{11}, Y_{12})} .$$

By analogy with Section 3, we obtain

$$(4.5) \quad \operatorname*{Max}_{H_{RC}} p(Y_1, Y_2, C)$$

$$= c(p,k,N)\,(kN)^{\frac{pkN}{2}}\, e^{-\frac{1}{2}pkN} \frac{|C_{22} + Y_{22}'Y_{22}|^{\frac{kN}{2}}}{|C + Y_2'Y_2|^{\frac{kN}{2}}\, |C_{22} + Y_{22}'Y_{22} + Y_{12}'Y_{12}|^{\frac{kN}{2}}} ,$$

and is achieved by

$$\hat{\Sigma} = \{Q + \begin{pmatrix} Q_{12} & Q_{22}^{-1} \\ & I \end{pmatrix} (Y_{12}'Y_{12})\,(Q_{22}^{-1}Q_{21}, I)\}/kN ,$$

$$\hat{\Theta}_{11} = Y_{11} + Y_{12}\hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1} = Y_{11} - Y_{12}Q_{22}^{-1}Q_{21} ,$$

where $Q = C + Y_2'Y_2$.

The likelihood ratio statistic for testing H_{RC} versus H_C is given by the ratio of (4.5) to (4.3):

$$L_{RC|C} = \frac{|C_{22} + Y_{22}'Y_{22}|^{\frac{kN}{2}}\, |C|^{\frac{kN}{2}}}{|C_{22}|^{\frac{kN}{2}}\, |C + Y_2'Y_2|^{\frac{kN}{2}}} = \frac{|I + Y_{22}C_{22}^{-1}Y_{22}'|^{\frac{kN}{2}}}{|I + Y_2C^{-1}Y_2'|^{\frac{kN}{2}}} .$$

When H_{RC} holds, $-2 \log L_{RC|C}$ is approximately distributed as a χ^2 variable with $(k-q)m$ degrees of freedom.

To test H_{RC} versus H_R we need the ratio of (4.5) to (4.1):

$$L_{RC|R} = \frac{|C_{22} + Y_{22}'Y_{22}|^{\frac{kN}{2}}}{|C_{22} + Y_2'Y_{22} + Y_{12}'Y_{12}|^{\frac{kN}{2}}} = \frac{1}{|I + Y_{12}(C_{22} + Y_2'Y_2)^{-1}Y_{12}'|^{\frac{kN}{2}}} .$$

When H_{RC} holds, $-2 \log L_{RC|C}$ is approximately distributed as a χ^2 variable with $q(p-m)$ degrees of freedom.

Remark. Note that

$$L_R = L_{RC|C}, \qquad L_C = L_{RC|R}$$

as expected.

5. An Example

To illustrate the proposed analysis we consider some data from an unpublished report by T. Toftegaard Nielsen, N. Schwartz Sørensen and E. Bjørn Jensen. The analysis of these data was originally carried out using the methods described in Andersen, Jensen and Schou (1978). Measurements of the concentration of plasma-citrate on 10 patients given a certain diet was taken every hour from 8 a.m. to 9 p.m. The 10 patients fall naturally in two groups, a "high level" group and a "low level" group, each with 5 patients. For the present analysis the measurements corresponding to 8 a.m., 11 a.m. and 3 p.m. were chosen. The data and computations are given below.

TIME

Groups		8 a.m.	11 a.m.	3 p.m.
	High	125	137	121
		144	173	147
		105	119	125
		151	149	128
		137	139	109
	Low	93	121	107
		116	135	106
		109	83	100
		89	95	83
		116	128	100

Cell Means	132.40	143.40	126.00
	104.60	112.40	99.20

Cross product matrix C	1964.40	1621.00	519.40
		3558.40	1255.60
			1130.80

Model with No Row Effect: Estimation

Means

118.50	127.90	112.60
118.50	127.90	112.60

Covariances

389.65	377.55	238.20
	596.09	333.26
		292.64

Model with No Column Effect: Estimation

Means

124.42	124.42	124.42
98.23	98.23	98.23

Covariances

248.54	282.91	61.32
	636.29	147.39
		114.79

Test for No Row Effects

$$-2 \log L R = 10.92 \qquad \chi_1^2(.01) = 6.63 , \quad \chi_1^2(.05) = 3.84$$

Test for No Column Effects

$$-2 \log L R = 7.22 \qquad \chi_2^2(.01) = 9.21 , \quad \chi_2^2(.05) = 5.99$$

The following Table is based on the usual univariate analysis of variance without regard to the fact that the column measurements are correlated.

ANALYSIS TABLE

Variation	SSD	d.f.	M SSD	F	
R	6106.1	1	6106.1	22	$F_{1,24}(.05) = 4.26$
C	1190.9	2	595.45	2.15	$F_{2,24}(.05) = 3.40$
R × C	24.1	2	12.05	.043	$F_{2,24}(.05) = 3.40$
Within Cells	6653.6	24	277.23		
Total	13974.7	29			

For this particular set of data the univariate and multivariate analysis both yield comparable results, namely, that there is no interaction effect, but a strong row effect and weak column effect.

Acknowledgements. Stimulating discussions with Eva Bjørn Jensen are gratefully acknowledged.

References

Andersen, A.H., Jensen, E.B. and Schou, G. (1978): Two - way analysis of variance with correlated errors. Research report No.33, Department of Theoretical Statistics, University of Aarhus.

Anderson, T.W. (1958): An Introduction to Multivariate Statistical Analysis. John Wiley & Sons. New York.

Box, G.E.P. (1949): A general distibution theory for a class of likelihood criteria. Biometrika 36, 317-346.

Box, G.E.P. (1954): Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two - way classification. Ann. Math. Statist. 25, 484-498.

Gleser, L.J. and Olkin, I. (1970): Linear Models in Multivariate Analysis. Essays in Probability and Statistics. (ed. R.C. Bose, I.M. Chakravarti, P.C. Mahalanobis, C.R. Rao and K.J.C. Smith) p. 267-292. The University of North Carolina Press. Chapel Hill.

Kariya, T. (1978): The general MANOVA problem. Ann. Statist. 6, 200-214.

PREPRINTS 1979

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Edwards, David: Large Sample Tests for Stationarity and Reversibility in Finite Markov Chains.

No. 2 Andersson, Steen A. : Distribution of Maximal Invariants Using Proper Action and Quotient Measures.

No. 3 Johansen, Søren: A Note on the Welch-James Approximation to the Distribution of the Residual Sum of Squares in a Weighted Linear Regression.

No. 4 Björnsson, Ottó: Four Simple Charcterizations of Standard Borel Spaces.

No. 5 Johansen, Søren: Some Comments on Robustness

No. 6 Hald, Anders: T.N. Thiele's Contributions to Statistics.

No. 7 Jacobsen, Martin: Markov Chains: Birth and Death Times with Conditional Independence.

PREPRINTS 1980

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Olkin, Ingram and Væth, Michael: Maximum Likelihood Estimation in a Two-way Analysis of Variance with Correlated Errors in One Classification.