1953

Some fundamental results in analysis of variance theory

Arthur Edwin Davis
The University of Montana

SOME FUNDAMENTAL RESULTS IN
ANALYSIS OF VARIANCE THEORY

by

ARTHUR EDWIN DAVIS

B. A. Montana State University, 1952

Presented in partial fulfillment of the requirements for the degree of

Master of Arts

MONTANA STATE UNIVERSITY

1953

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

Aug 17 1953

Date

ACKNOWLEDGMENT

The author is especially indebted to Professor George Marsaglia for the suggestion of the subject of this thesis and for his generous assistance and constant guidance throughout the entire work.

A. E. D.

TABLE OF CONTENTS

INTRODUCTION

The analysis of variance is a statistical technique introduced by R. A. Fisher in connection with certain experimental designs. The domain of applicability of this technique is very wide and has already been successfully applied in many branches of experimental work. However, no text books give a rigorous development of this theory and the results which have been established in this connection are not well represented in the literature. The purpose of this paper is to develop in a rigorous manner the general theory of linear hypotheses. The development is the author's and the specific results are based upon unpublished research of Professor Marsaglia.

CHAPTER I

PRELIMINARY RESULTS

We begin by stating some theorems which are used in the sequel but whose proofs will be omitted. The results are well known.

By $r(A)$ we mean the rank of the matrix A.

<u>Theorem</u> 1.1: $r(AA') = r(A)$

<u>Theorem</u> 1.2: $r(MN) \leq r(M)$

<u>Theorem</u> 1.3: <u>If</u> $ABA = A$ <u>then</u> $r(A) = r(AB)$

For by theorem 1.2

$$r(AB) \leq r(A)$$

and

$$r(A) = r(ABA) \leq r(AB)$$

hence

$$r(A) = r(AB)$$

The following theorem due to Allen T. Craig (1943) and Harold Hotelling (1944) is of great importance in this development. The proof, due to its length, will not be given here but can be found in the literature.

<u>Theorem</u> 1.4: <u>If A and B are the symmetric matrices of two homogeneous quadratic forms in n variates which are normally and independently distributed with zero means and unit variances, a necessary and sufficient condition for the independence in proba-</u>

bility of these two forms is that $AB = 0$.

Definition: The random variables $x_1, x_2, \ldots, x_n$ have a multivariate normal distribution if their probability density function has the form

$$f(x_1, x_2, \ldots, x_n) = \frac{|A|^{1/2}}{(2\pi)^{n/2}} e^{-1/2 \sum_{i,j} a_{ij}(x_i - \mu_i)(x_j - \mu_j)}$$

where $A = (a_{ij})$ is a positive definite nxn symmetric matrix.

Theorem 1.5: Let $x_1, x_2, \ldots, x_n$ be jointly normally distributed, $\eta = (y_1, y_2, \ldots, y_n)$ and $\xi = (x_1, x_2, \ldots, x_n)$. If $\eta = \xi M$ then the y's in η are jointly normally distributed.

Theorem 1.6: Let $x_1, x_2, \ldots, x_n$ be random variables with means $\mu_1, \mu_2, \ldots, \mu_n$ and covariance matrix S. If $\eta = (y_1, y_2, \ldots, y_n)$, $\xi = (x_1, x_2, \ldots, x_n)$ and $\eta = \xi M$, then $E(\eta) = E(\xi)M$ and $\mathrm{cov}(\eta) = M'SM$.

Definition: The random variable z is said to have the chi-square distribution with n degrees of freedom (χ^2) if it has the moment generating function

$$m_z(t) = (1 - 2t)^{-n/2} .$$

Theorem 1.7: Every nonzero vector space V over the real number field has an orthonormal basis. Moreover, every set of mutually orthogonal, normal vectors of V may be extended to an orthonormal basis of V.

CHAPTER II

THE GENERAL LINEAR HYPOTHESIS

Let $x_1, x_2, \ldots, x_n$ be independent, normally distributed random variables with $E(\xi) = \mu = (m_1, m_2, \ldots, m_n)$, where $\xi = (x_1, x_2, \ldots, x_n)$, and $\operatorname{cov}(\xi) = \sigma^2 I$.

The theory under the heading "linear hypotheses," is concerned with linear assumptions and hypotheses on $m_1, m_2, \ldots, m_n$. The restrictions generally take two forms

(1) Assumptions of the form

$$A\mu' = 0$$

where A is an rxn matrix of rank r.

(2) Hypotheses of the form

$$R\mu' = 0$$

where R is an sxn matrix of rank s and $r + s \leq n$.

The most general method for testing such hypotheses is the likelihood ratio test, that is

let

$$L = (2\pi\sigma^2)^{-n/2}\, e^{-1/2\sigma^2 \, (\xi - \mu)(\xi - \mu)'}$$

and

$$L_a = \text{maximum of } L \text{ under assumptions}$$

$$L_r = \text{maximum of } L \text{ under restrictions (hypothesis)}$$

where L is considered, for a fixed ξ, a function of $\mu = (m_1, m_2, \ldots, m_n)$ and σ^2.

The quantity

$$\lambda = \frac{L_r}{L_a}$$

is the likelihood ratio test and the critical region, for testing the hypothesis, is $\lambda \leq \lambda_0$. λ_0 is chosen to make $P(\lambda \leq \lambda_0) = .05$ if the hypothesis is true. Now the distribution of λ is unknown, in general. We shall give a test which is a monotonic function of λ and thus equivalent to it.

Usually the matrices A, R and $\binom{A}{R}$ are not orthogonal but we shall develop orthogonal matrices which have the same properties as A, R and $\binom{A}{R}$ in order to obtain the desired results of this section.

<u>Theorem</u> 2.1: <u>Let</u> $\alpha_1, \alpha_2, \ldots, \alpha_s$ <u>be linearly independent</u> 1xn <u>vectors. Then there exists orthonormal</u> 1xn <u>vectors</u> $\beta_1, \beta_2, \ldots, \beta_s$ <u>such that for any</u> j, $1 \leq j \leq s$,

$$\{\mu : \alpha_i \mu' = 0, \; i=1, 2, \ldots, j\} = \{\mu : \beta_i \mu' = 0, \; i=1, 2, \ldots, j\}.$$

Define the β's inductively as follows:

(1) $\beta_1 = \dfrac{\alpha_1}{\sqrt{\alpha_1 \alpha_1'}}$

$$(2)\quad \gamma_i = \alpha_i - \alpha_i(\beta_1'\beta_1 + \beta_2'\beta_2 + \cdot\cdot\cdot + \beta_{i-1}'\beta_{i-1})$$

$$(3)\quad \beta_i = \frac{\gamma_i}{\sqrt{\gamma_i\gamma_i'}}$$

Now it is easily verified that

$$\beta_i\beta_i' = 1$$

so that the β's are normal. Certainly $\beta_2\beta_1' = 0$, since

$$\gamma_2\beta_1' = 0.$$

We shall now assume that $\beta_j\,\beta_1' = 0$, for all $j < n$, $j \neq 1$, then we must show that $\beta_n\,\beta_1' = 0$.

Obviously

$$\gamma_n\,\beta_1' = 0$$

hence $\beta_n\,\beta_1' = 0$ and it follows that $\beta_i\,\beta_1' = 0$ for $i = 1, 2, \ldots, s$.

Next assume that $\beta_i\,\beta_j' = 0$ for $j = 1, 2, \ldots, s$ and all $i < m$ where $i \neq j$. Now we must show that $\beta_m\,\beta_j' = 0$ where $j \neq m$, which is clearly true since

$$\gamma_m\,\beta_j' = 0.$$

Therefore we have established the fact that the β's are orthonormal.

It is obvious from (2) and (3) that if

$$\beta_1\mu' = \beta_2\mu' = \cdot\cdot\cdot = \beta_j\mu' = 0$$

then

$$\alpha_1\mu' = \alpha_2\mu' = \cdots = \alpha_j\mu' = 0.$$

Now suppose that

$$\alpha_1\mu' = \alpha_2\mu' = \cdots = \alpha_j\mu' = 0$$

but

$$\gamma_i\mu' = 0$$

hence

$$\beta_1\mu' = \beta_2\mu' = \cdots = \beta_j\mu' = 0$$

and therefore

$$\{\mu:\ \alpha_i\mu' = 0,\ i=1,2,\ldots,j\} = \{\mu:\ \beta_i\mu' = 0,\ i=1,2,\ldots,j\}\ .$$

__Corollary__ 2.11: __If__ A __and__ B __are__ axn __and__ bxn __matrices such that__ $r\begin{pmatrix}A\\B\end{pmatrix} = a + b$, __then there exists matrices__ C_{axn} __and__ D_{bxn} __such that__

(1) $$CC' = I_a,\ DD' = I_b$$

$$CD' = O_{ab}$$

(2) $$\{\mu:\ A\mu' = 0\} = \{\mu:\ C\mu' = 0\}$$

$$\{\mu:\begin{pmatrix}A\\B\end{pmatrix}\mu' = 0\} = \{\mu:\begin{pmatrix}C\\D\end{pmatrix}\mu' = 0\}$$

To prove this we apply theorem 2.1 to the rows of $\begin{pmatrix}A\\B\end{pmatrix}$. This gives us a matrix $\begin{pmatrix}C\\D\end{pmatrix}$ such that $\begin{pmatrix}C\\D\end{pmatrix}(C'D') = I_{a+b}$, yielding the relations (1). Also by theorem 2.1

$$\{\mu : A\mu' = 0\} = \{\mu : C\mu' = 0\}$$

$$\{\mu : \binom{A}{B}\mu' = 0\} = \{\mu : \binom{C}{D}\mu' = 0\} .$$

Let B be an sxn matrix of rank s such that $BB' = I_{ss}$ and V be an s dimensional nonzero vector space. By applying theorem 1.7 to the rows of B we obtain an orthonormal basis of V. From this we can clearly see that there exists an (n-s)xn matrix N such that $\binom{B}{N}$ is orthogonal.

Theorem 2.2: Let $Q = \xi\xi' - 2\xi\mu' + \mu\mu'$, and $S = \{\mu : B\mu' = 0\}$ where B is an sxn matrix of rank s such that $BB' = I_{ss}$. Also let N be an (n-s)xn matrix such that $\binom{B}{N}$ is orthogonal. Then the minimum value of Q over S is $\xi B'B\xi'$ and this value is uniquely determined when $\mu = \xi N'N$.

We note that such a matrix N exists by theorem 1.7 and also

$$I_{nn} = B'B + N'N = \begin{pmatrix} BB' & BN' \\ NB' & NN' \end{pmatrix}$$

Now

$$Q = \xi\xi' - 2\xi\mu' + \mu\mu'$$

$$= \xi(B'B + N'N)\xi' - 2\xi(B'B + N'N)\mu' + \mu(B'B + N'N)\mu'$$

and if $\mu \in S$

$$Q = \xi B'B\xi' + \xi N'N\xi' - 2\xi N'N\mu' + \mu N'N\mu'$$

$$= \xi B'B\xi' + (\xi N' - \mu N')(\xi N' - \mu N')'$$

Clearly min (Q) = $\xi B'B\xi'$ since $\xi N' - \mu N' = 0$ when $\mu = \xi N'N$, and $\xi N'N \in S$.

Now suppose $Q(\mu^*) = \xi B'B\xi'$ and $\mu^* \in S$ then we have $B\mu^{*\prime} = 0$ and $\mu^* N' = \xi N'$, hence

$$\begin{pmatrix} B \\ N \end{pmatrix} \mu^{*\prime} = \begin{pmatrix} 0 \\ N\xi' \end{pmatrix}$$

so that $\mu^{*\prime} = (B' \; N') \begin{pmatrix} 0 \\ N\xi' \end{pmatrix} = N'N\xi'$.

Therefore $\mu = \mu^*$ and min (Q) over S is uniquely determined when $\mu = \xi N'N$.

<u>Lemma</u> 1: <u>Let</u> $Y = \{y: y > 0\}$, $X = \{x: x \geq c \geq 0\}$. <u>If</u> $f(x,y) = (ky)^{-n/2} e^{-1/2 \frac{x}{y}}$, $k > 0$ <u>and</u> $\max_{\substack{x \in X \\ y \in Y}} f(x,y) = f(x_o,y_o)$ <u>then</u>

$$(x_o, y_o) = (c, c/n).$$

Clearly $x_o = c$ since for any (x', y'),

$$f(c,y') \geq f(x',y').$$

Thus we need only to maximize $f(c,y)$ or max $\ln(f(c,y))$ with respect to y. But

$$\ln f(c,y) = -\frac{n}{2}\ln k - \frac{n}{2}\ln y - \frac{1}{2}\frac{c}{y}$$

and

$$\frac{\partial \ln f(c,y)}{\partial y} = -\frac{n}{2y} + \frac{c}{2y^2} = 0$$

then

$$ny - c = 0$$

and

$$y = \frac{c}{n}$$

therefore

$$(x_o, y_o) = (c, c/n).$$

Theorem 2.3: Let $x_1, x_2, \ldots, x_n$ be normally, independently distributed random variables with $cov(\xi) = \sigma^2 I$, $E(\xi) = \mu$ where $\xi = (x_1, x_2, \ldots, x_n)$. Also let $Q_a = \min_{\{\mu: A\mu' = 0\}} Q(\mu)$, $Q_r = \min_{\{\mu: R\mu' = 0\}} Q(\mu)$ and suppose A is a sxn matrix of rank s such that $\mu \in S$, $S = \{\mu: A\mu' = 0\}$. Then the likelihood ratio for testing,

$$H: R\mu' = 0$$

where $r(R) = r$ and $r\begin{pmatrix} A \\ R \end{pmatrix} = s + r$

is

$$\lambda = \left(\frac{Q_r}{Q_a}\right)^{-\frac{n}{2}}.$$

<u>The likelihood ratio test is equivalent to the F test where</u>

$$F = \frac{s(Q_r - Q_a)}{rQ_a}$$

<u>has the</u> F <u>distribution with</u> r, s <u>degrees of freedom.</u> <u>Furthermore,</u>
$Q_a = \xi A'A \xi'$ <u>and</u> $Q_r = \xi R'R \xi'$.

Let $\qquad Q = (\xi - \mu)(\xi - \mu)'$

then the likelihood function is

$$L = (2\pi\sigma^2)^{-n/2} e^{-1/2\sigma^2 Q} .$$

Now the likelihood ratio for testing our hypothesis is

$$\lambda = \frac{L_r}{L_a} .$$

Now by letting B = A in theorem 2.2

$$Q_a = \xi A'A \xi'$$

and by lemma 1

$$L_a = (2\pi\frac{Q_a}{n})^{-n/2} e^{-1/2 \frac{Q_a}{Q_a/n}} .$$

In theorem 2.2 by letting B = R we find that

$$Q_r = \xi R'R \xi'.$$

Also by application of lemma 1 we find that

$$L_r = \left(2\pi \frac{Q_r}{n}\right)^{-n/2} e^{-1/2 \frac{Q_r}{Q_r/n}}$$

so that

$$\lambda = \frac{\left(2\pi \frac{Q_r}{n}\right)^{-n/2} e^{-1/2 \frac{Q_r}{Q_r/n}}}{\left(2\pi \frac{Q_a}{n}\right)^{-n/2} e^{-1/2 \frac{Q_a}{Q_a/n}}}$$

$$= \left(\frac{Q_r}{Q_a}\right)^{-n/2} .$$

Since

$$F = \frac{s(Q_r - Q_a)}{rQ_a}$$

is a monotonic function of λ, it yields a test equivalent to the likelihood ratio test.

CHAPTER III

MULTIPLE LINEAR REGRESSION

Let $x_1, x_2, \ldots, x_n$ be independently and normally distributed random variables, let $\xi = (x_1, x_2, \ldots, x_n)$ and suppose $\alpha = (a_1, a_2, \ldots, a_k)$ is a vector of parameters and Z a kxn constant matrix of rank k such that $E(\xi) = \alpha Z$ and $\mathrm{cov}(\xi) = \sigma^2 I$.

We wish to find the maximum likelihood estimates of $a_1, a_2, \ldots, a_k$ and σ^2, where the likelihood function is

$$L = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-1/2\sigma^2 (x_i - E(x_i))^2}$$

In order to do this it is desirable to find the minimum of the quadratic form $Q = \sum_{i=1}^{n} (x_i - E(x_i))^2$ as in the following theorem.

<u>Theorem</u> 3.1: <u>Let</u> $\xi = (x_1, x_2, \ldots, x_n)$ <u>and</u> $E(\xi) = (\mu_1, \mu_2, \ldots, \mu_n) = \mu$ <u>and</u> $Q = \sum_{i=1}^{n} (x_i - \mu_i)^2$. <u>Also suppose</u> $\mu = \alpha Z$, $\alpha = (a_1, a_2, \ldots, a_n)$ <u>where</u> Z <u>is</u> kxn, <u>then we have the following results</u>

(1) $$Q = \xi\xi' - 2\xi Z'\alpha' + \alpha ZZ'\alpha'$$

$$(2) \qquad \frac{\partial Q}{\partial a_j} = 2(\alpha ZZ' - \xi Z')e_j'$$

where e_j' is the jth column of the nxn identity matrix.

Now

$$Q = \sum_{i=1}^{n}(x_i - \mu_i)^2$$

$$= \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i \mu_i + \sum_{i=1}^{n}\mu_i^2$$

$$= \xi\xi' - 2\xi\mu' + \mu\mu'$$

$$= \xi\xi' - 2\xi Z'\alpha' + \alpha ZZ'\alpha'$$

and

$$\frac{\partial Q}{\partial a_j} = 0 - 2\xi Z'e_j' + e_j ZZ'\alpha' + \alpha ZZ'e_j'$$

$$= -2\xi Z'e_j' + 2\alpha ZZ'e_j'$$

since $e_j ZZ'\alpha'$ is symmetric of dimension 1x1.

Lemma 2: If $B = Z'(ZZ')^{-1}$, where Z is a kxn matrix of rank k, then BZ is idempotent and symmetric of rank k, ZB = I, and $ZB_i' = e_i'$ where B_i' and e_i' are the ith columns of B and I respectively.

Now

$$(BZ)(BZ)' = (Z'(ZZ')^{-1}Z)(Z'(ZZ')^{-1}Z)$$

$$= Z'(ZZ')^{-1}ZZ'(ZZ')^{-1}Z$$

$$= Z'(ZZ')^{-1}Z$$

$$= BZ$$

also

$$ZB = ZZ'(ZZ')^{-1}$$

$$= I.$$

Now since $ZB = I$, then $ZBZ = Z$; therefore, $r(ZBZ) = r(Z) = k$, and from theorem 1.3

$$r(Z) = r(BZ) = k.$$

Clearly

$$ZB_i' = e_i'$$

since

$$ZB = I.$$

<u>Theorem 3.2: Let</u> $x_1, x_2, \ldots, x_n$ <u>be independent normally distributed random variables with</u> $E(x_i) = \mu_i$. <u>Let</u> $\xi = (x_1, x_2, \ldots, x_n)$, $\alpha = (a_1, a_2, \ldots, a_k)$ <u>and</u> Z <u>be a</u> kxn <u>matrix of rank</u> k <u>and assume</u>

$$E(\xi) = \alpha Z, \ \text{cov}(\xi) = \sigma^2 I.$$

Let the maximum likelihood estimates of α and σ^2 be $\hat{\alpha}$ and $\hat{\sigma}^2$ respectively. Then

(1) $\hat{\alpha} = \xi Z'(ZZ')^{-1}$

(2) $n\hat{\sigma}^2 = \xi(I - BZ)\xi'$

(3) $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k$ are each independent of $\hat{\sigma}^2$

(4) $\hat{a}_i$ is normally distributed and $E(\hat{\alpha}) = \alpha$,

$\text{cov}(\hat{a}) = (ZZ')^{-1}\sigma^2$

(5) $\frac{n\hat{\sigma}^2}{\sigma^2}$ is χ^2 distributed with n - k degrees of freedom.

The likelihood function is

$$L = \prod_{i=1}^{n}(2\pi\sigma^2)^{-1/2} e^{-1/2\sigma^2 (x_i - E(x_i))^2} .$$

Let

$$Q = \sum_{i=1}^{n}(x_i - E(x_i))^2$$

then

$$\ln L = c - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}Q$$

and

$$\frac{\partial \ln L}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}Q$$

and by theorem 3.1

$$\frac{\partial \ln L}{\partial a_j} = \frac{1}{\sigma^2}(\alpha ZZ' - \varsigma Z')e_j'$$

hence the equations to be solved are

$$n\sigma^2 = Q$$

$$(\alpha ZZ' - \varsigma Z')e_j' \quad \text{for } j = 1, 2, \ldots, k.$$

Then if

$$\hat{\alpha} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k)$$

hence

$$\alpha ZZ' - \varsigma Z' = 0$$

$$n\hat{\sigma}^2 = Q\big|_{\alpha=\hat{\alpha}}$$

and

$$\alpha = \varsigma Z'(ZZ')^{-1} = \varsigma B \text{ where } B = Z'(ZZ')^{-1}$$

we note that

$$n\hat{\sigma}^2 = Q\big|_{\alpha=\hat{\alpha}}$$

$$= \varsigma\varsigma' - 2\varsigma Z'B'\varsigma' + \varsigma BZZ'B'\varsigma'$$

$$= \varsigma\varsigma' - \varsigma Z'B'\varsigma'$$

$$= \xi(I - Z'B')\xi' .$$

Also it is easily verified that

$$(\xi - \alpha Z)(I - Z'B')(\xi' - Z'\alpha') = \xi(I - Z'B')\xi'$$

and thus

$$n\hat{\sigma}^2 = \xi(I - Z'B')\xi'$$

$$= (\xi - \alpha Z)(I - Z'B')(\xi' - Z'\alpha')$$

Therefore, $\frac{n\hat{\sigma}^2}{\sigma^2}$ is χ^2 distributed with n - k degrees of freedom.

Furthermore

$$a_j = \xi B e_j' = \xi B_j'$$

and

$$a_j^2 = \xi B_j' B_j \xi'$$

since

$$B_j' B_j Z' B' = B_j' e_j B'$$

$$= B_j' B_j$$

it follows from theorem 1.4 that since

$$B_j' B_j (I - Z'B') = 0$$

then a_j is independent of $\hat{\sigma}^2$. Since $\hat{\alpha} = \xi B$, then by theorems 1.5 and 1.6 the $\hat{a}_i$'s in $\hat{\alpha}$ are normally distributed with $E(\hat{\alpha}) = \alpha$ and $\text{cov}(\hat{\alpha}) = \sigma^2 B'B = \sigma^2 (ZZ')^{-1}$.

Now suppose we have, instead of the situation just considered, the case of repeated samples. That is, for each set $Z_{1i}, Z_{2i}, \ldots, Z_{ki}$ we have not only one quantity x_i, but a set of quantities, i.e., $(x_{i1}, x_{i2}, \ldots, x_{in})$. We want the maximum likelihood estimates of $a_1, a_2, \ldots, a_k$ and σ^2 under this additional assumption. The theory of this situation will now be considered.

<u>Theorem</u> 3.3: <u>Let</u> $\xi_i = (x_{i1}, x_{i2}, \ldots, x_{in})$, $i = 1, 2, \ldots, m$, <u>and</u> $\eta = (x_{11}, x_{21}, \ldots, x_{m1}, \ldots, x_{1n}, x_{2n}, \ldots, x_{mn})$ <u>where the x's are independent normally distributed random variables. Let</u> $\bar{\xi} = \frac{1}{m}\sum_{i=1}^{m} \xi_i$, $\alpha = (a_1, a_2, \ldots, a_k)$ <u>and</u> Z <u>be a</u> kxn <u>matrix of rank</u> k. <u>Assume that</u> $E(\xi_i) = \alpha Z$ <u>and</u> $\text{cov}(\xi_i) = \sigma^2 I$. <u>Let the maximum likelihood estimates of</u> α <u>and</u> σ^2 <u>be</u> $\hat{\alpha}$ <u>and</u> $\hat{\sigma}^2$ <u>respectively. Then</u>

(1) $\hat{\alpha} = \bar{\xi} Z'(ZZ')^{-1}$

(2) $nm\hat{\sigma}^2 = \eta(I - mM\, Z'B'M')\eta'$

$$= \sum_{i=1}^{m} \xi_i \xi_i' - m\bar{\xi} Z'B' \bar{\xi}'$$

(3) $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k$ <u>are each independent of</u> $\hat{\sigma}^2$

(4) $\hat{a}_i$ <u>is normally distributed and</u> $E(\hat{\alpha}) = \alpha$,

$$\mathrm{cov}(\hat{\alpha}) = \frac{\sigma^2}{m}(ZZ')^{-1}$$

(5) $\frac{mn\hat{\sigma}^2}{\sigma^2}$ <u>has the</u> χ^2 <u>distribution with</u> $nm - k$ <u>degrees of freedom</u>.

The likelihood function is

$$L = \prod_{i=1, j=1}^{m \quad n} (2\pi\sigma^2)^{-1/2} e^{-1/2\sigma^2 (x_{ij} - E(x_{ij}))^2} .$$

Let

$$Q_i = \sum_{j=1}^{n} (x_{ij} - E(x_{ij}))^2$$

then

$$\ln L = c - \frac{nm}{2} \ln\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{m} Q_i$$

and

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{nm}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{m} Q_i$$

also

$$\frac{\partial \ln L}{\partial a_j} = -\frac{1}{2\sigma^2} \sum_{i=1}^{m} \frac{\partial Q_i}{\partial a_j} .$$

If we let $\varsigma = \varsigma_i$ in theorem 3.1, we have

$$\frac{\partial Q_i}{\partial a_j} = 2(\alpha ZZ' - \varsigma_i Z')e_j' .$$

Hence, we have the equations

$$nm\sigma^2 = \sum_{i=1}^{m} Q_i$$

$$-\frac{1}{\sigma^2}\sum_{i=1}^{m} (\alpha ZZ' - \varsigma_i Z')e_j' = 0 \quad \text{for } j = 1,2,\ldots,k.$$

Since $\bar{\varsigma} = \frac{1}{m}\sum_{i=1}^{m}\varsigma_i$, then the solutions to the equations are

$$nm\hat{\sigma}^2 = \sum_{i=1}^{m} Q_i \Big|_{\alpha=\hat{\alpha}}$$

$$\hat{\alpha} = \bar{\varsigma}\, Z'(ZZ')^{-1}$$

Now $\bar{\varsigma} = \eta M$ and M has dimensions nmxn, where

$$M = \frac{1}{m}\begin{pmatrix} m\left\{\begin{matrix}1\\ \cdot\\ \cdot\\ \cdot\\ 1\end{matrix}\right. & \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} & \cdots & \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} \\ \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} & m\left\{\begin{matrix}1\\ \cdot\\ \cdot\\ \cdot\\ 1\end{matrix}\right. & \cdots & \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} \\ \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} & \begin{matrix}0\\ \cdot\\ \cdot\\ \cdot\\ 0\end{matrix} & \cdots & m\left\{\begin{matrix}1\\ \cdot\\ \cdot\\ \cdot\\ 1\end{matrix}\right. \end{pmatrix}$$

then

$$M'M = \frac{1}{m} I_{nn}$$

and

$$\eta I \eta' = \sum_{i=1}^{m} \xi_i \xi_i' \ .$$

According to theorem 3.1

$$Q_i = \xi_i \xi_i' - 2 \xi_i Z' \alpha' + \alpha Z Z' \alpha'$$

so that

$$Q_i|_{\alpha=\hat{\alpha}} = \xi_i \xi_i' - 2 \xi_i Z' B' \bar{\xi}' + \bar{\xi} B Z Z' B' \bar{\xi}'$$

and

$$nm\hat{\sigma}^2 = \sum_{i=1}^{n} Q_i|_{\alpha=\hat{\alpha}}$$

$$= \sum_{i=1}^{m} \xi_i \xi_i' - 2(\sum_{i=1}^{m} \xi_i) Z' B' \bar{\xi}' + m \ \bar{\xi} \ B Z Z' B' \ \bar{\xi}'$$

$$= \eta \eta' - 2m \bar{\xi} Z' B' \ \bar{\xi}' + m \ \bar{\xi} \ Z' B' \ \bar{\xi}'$$

$$= \eta \eta' - m \ \bar{\xi} Z' B' \ \bar{\xi}'$$

$$= \eta \eta' - m \eta M Z' B' M' \eta'$$

$$= \eta (I - m M Z' B' M') \eta'$$

Now

$$(MZ'B'M')(MZ'B'M')' = MZ'B'M'MBZM'$$

$$= MZ'B'M'$$

hence MZ'B'M' is idempotent.

Also

$$r(mMZ'B'M') = k$$

since

$$MZ'BM' = (MZ'B')(MZ'B')'$$

and

$$(MZ'B')M'(MZ'B') = \frac{1}{m} MZ'B' \ .$$

The result follows from theorems 1.1 and 1.3.

Now

$$a_j^2 = \bar{\varsigma}\, B_j'B_j \ \bar{\varsigma}'$$

$$= \eta MB_j'B_jM' \eta'$$

and since

$$(MB_j'B_jM')(I - mMZ'B'M') = 0,$$

therefore $\hat{a}_j$ and $\hat{\sigma}^2$ are independent by theorem 1.4.

Finally, by theorems 1.5 and 1.6, $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n$ are normally distributed with $E(\hat{\alpha}) = \alpha$ and $\mathrm{cov}(\hat{\alpha}) = \frac{\sigma^2}{m}(ZZ')^{-1}$.

Also $\frac{mn\hat{\sigma}^2}{\sigma^2}$ has the χ^2 distribution with mn - k degrees of freedom.

BIBLIOGRAPHY

Craig, Allen T. "A Note on the Independence of Certain Quadratic Forms," *Annals of Mathematical Statistics*, Vol. 14, pp. 195-197, 1943.

Craig, Allen T. "Bilinear Forms in Normally Correlated Variables," *Annals of Mathematical Statistics*, Vol. 18, pp. 565-573, 1947.

Cramér, Harald. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, New Jersey, 1951.

Hotelling, Harold. "Note on a Matrix Theorem of A. T. Craig," *Annals of Mathematical Statistics*, Vol. 15, pp. 427-429, 1944.