AD-A090 182 PITTSBURGH UNIV PA INST FOR STATISTICS AND APPLICATIONS F/G 12/1
SIMULTANEOUS PSEUDO CONFIDENCE REGIONS FOR RATIOS OF THE DISCRI--ETC(U)
JUL 80 Y CHIKUSE F49620-79-C-0161
UNCLASSIFIED TR-80-7 AFOSR-TR-80-0939 NL

1 OF 1
AD
A090182

END
DATE
FILMED
11-80
DTIC

AFOSR-TR. 80-0939

LEVEL II

(4)

AD A090182

DTIC ELECTE OCT 6 1980 D

C

Institute for Statistics and Applications

University of Pittsburgh

DDC FILE COPY

Approved for public release; distribution unlimited.

F49620-79-C-0161

80 9 25 125

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 AFOSR-TR-80-0939 19

2. GOVT ACCESSION NO.: AD-A090182

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): 6 Simultaneous Pseudo Confidence Regions for Ratios of the Discriminant Coefficients

5. TYPE OF REPORT & PERIOD COVERED: 9 Interim rept.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): 10 Yasuko Chikuse

17 A5

8. CONTRACT OR GRANT NUMBER(s): 15 F49620-79-C-0161

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
University of Pittsburgh
Department of Mathematics and Statistics
Pittsburgh, PA. 15260

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 16 61102F 2304/A5

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Office of Scientific Research/NM
Bolling AFB, Washington, DC 20332

12. REPORT DATE: 11 Jul[redacted]980

13. NUMBER OF PAGES: 21

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 23 14 TR-80-71

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

pseudo (conditional) confidence regions; multivariate linear hypotheses on coefficients of a linear discriminant function; ratios of discriminant coefficients.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider simultaneous confidence regions for some hypotheses on ratios of the discriminant coefficients of the linear discriminant function when the population means and common covariance matrix are unknown. This problem, involving hypotheses on ratios, yields the so-called 'pseudo' confidence regions valid conditionally in subsets of the parameter space. We obtain the explicit formulae of the regions and give further discussion on the validity of these regions. Illustrations of the pseudo confidence regions are given.

DD FORM 1473, 1 JAN 73 411385 LM

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SIMULTANEOUS PSEUDO CONFIDENCE
REGIONS FOR RATIOS OF THE
DISCRIMINANT COEFFICIENTS

Yasuko Chikuse*

Department of Mathematics and Statistics
University of Pittsburgh

and

Kagawa University, Japan

July 1980

Technical Report No. 80-7

Institute for Statistics and Applications
Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA. 15260

*This work is partially sponsored by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161.

SIMULTANEOUS PSEUDO CONFIDENCE REGIONS FOR RATIOS OF THE DISCRIMINANT COEFFICIENTS

Yasuko Chikuse

Department of Mathematics and Statistics
University of Pittsburgh
and
Kagawa University, Japan

ABSTRACT

We consider simultaneous confidence regions for some hypotheses on ratios of the discriminant coefficients of the linear discriminant function when the population means and common covariance matrix are unknown. This problem, involving hypotheses on ratios, yields the so-called 'pseudo' confidence regions valid conditionally in subsets of the parameter space. We obtain the explicit formulae of the regions and give further discussion on the validity of these regions. Illustrations of the pseudo confidence regions are given.

1. INTRODUCTION

The linear discriminant function based on the vector of measurements $\underset{\sim}{x} = (x_1, x_2, \ldots, x_p)'$ for discriminating two p-variate normal populations with mean vectors μ_1 and μ_2 and common nonsingular covariance matrix Σ is given by $\alpha'\underset{\sim}{x}$ where

$\underset{\sim}{a} = (a_1, a_2, \ldots, a_p)' = \Sigma^{-1}(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$ is the vector of discriminant coefficients. Only the ratios of the discriminant coefficients are uniquely determined in the discriminant analysis. In this paper, we shall investigate simultaneous confidence regions for some hypotheses on the ratios, when the population means and covariance matrix are unknown.

The problem is reduced to using a likelihood ratio F test statistic in the derivation of simultaneous confidence regions. This procedure gives a number of multidimensional regions of different shape, the so-called 'pseudo' confidence regions valid conditionally in subsets of the parameter space. The terminology 'pseudo' confidence region is due to Ogawa (1977). We obtain the explicit formulae of our confidence regions. We focus our discussion on the bounded pseudo confidence regions and consider the conditions for the validity. The meanings of the required conditions are clarified from the practical point of view in connection with problems of testing hypotheses. Illustrations of the pseudo confidence regions are given.

It may be worthwhile to list some literature on the problem of confidence regions for ratios of (components of) normal population means. The univariate case has been considered by Creasy (1954), Fieller (1954), James, Wilkinson and Venables (1974) and Ogawa (1977). Box and Hunter (1954), Roy and Potthoff (1958), Zerbe (1978) and Chikuse (1980) investigated the multivariate case. The tests and confidence regions for hypotheses on ratios of discriminant coefficients were considered by Fisher (1940), Kshirsagar (1963) and Rao (1970a, Chapter 7), (1970b). This paper gives further considerations in this area.

2. PRELIMINARY RESULTS

We consider the multivariate linear hypothesis on the discriminant coefficients

$$B\,\underset{\sim}{a} = \underset{\sim}{0}, \tag{2.1}$$

where B is a $k \times p$ matrix with linearly independent k rows,

$k \geq 1$, $p > k$. The following lemma is fundamental for the subsequent discussion.

<u>Lemma 2.1</u>. The hypothesis (2.1) is equivalent to the hypothesis that the coefficients of z_i, $i = 1,2,\ldots,k$, in the discriminant function based on new variables $\underset{\sim}{z} = (z_1, z_2, \ldots, z_p)'$ are zero, where

$$\underset{\sim}{z} = T\underset{\sim}{x}, \quad T = \begin{array}{cc} \begin{array}{cc} \quad k & \qquad\qquad p-k \end{array} & \\ \begin{bmatrix} B_{11}'^{-1} & 0 \\ -(B_{11}^{-1} B_{12})' & I_{p-k} \end{bmatrix} & \begin{array}{l} k \\ p-k \end{array} \end{array}, \quad B = \begin{array}{cc} \begin{array}{cc} k & p-k \end{array} & \\ [B_{11} \quad B_{12}] & k. \end{array} \tag{2.2}$$

Here B_{11} is assumed to be nonsingular without loss of generality and hence T is nonsingular.

The null hypothesis that some of the discriminant coefficients are zero is tested by the well known likelihood ratio F test statistic (see e.g., Kshirsagar (1972, p. 200)). Given samples of sizes n_1 and n_2 from the two normal populations concerned, we apply this result to our problem.

<u>Lemma 2.2</u>. The $100(1-\alpha)$ per cent confidence region for the parameters satisfying the null hypothesis (2.1) is given by

$$\alpha_k - \beta_k (L\underset{\sim}{d})'(LSL')^{-1}(L\underset{\sim}{d}) \leq 0, \tag{2.3}$$

where

$L = [(-B_{11}^{-1} B_{12})' \; I_{p-k}]$,

$\underset{\sim}{d}$; the difference of sample means,

S; the sum of the corrected ss and sp matrices, (2.4)

$$\alpha_k = c^2 \underset{\sim}{d}' S^{-1} \underset{\sim}{d} - k(f-p+1)^{-1} F_{k,f-p+1}(\alpha),$$

$$\beta_k = c^2 [k(f-p+1)^{-1} F_{k,f-p+1}(\alpha) + 1],$$

$$f = n_1 + n_2 - 2, \quad c^2 = n_1 n_2/(n_1 + n_2),$$

and $F_{k,f-p+1}(\alpha)$ denotes the upper 100α percentile of the $F_{k,f-p+1}$ distribution with degrees of freedom k and $f-p+1$.

In the subsequent sections, we shall investigate some hypotheses on the ratios which are expressed in the form (2.1).

3. AN ASSIGNED RATIO OF TWO COEFFICIENTS

For the hypothesis H_1: $a_1/a_2 = \rho$, the L matrix in (2.4) is given by $L_1 = [\underset{\sim}{\Gamma}, I_{p-1}]$, where $\underset{\sim}{\Gamma} = (\rho, 0, \ldots, 0)'$. We obtain, after some algebra and simplification,

$$\begin{aligned}(L_1 \underset{\sim}{d})'(L_1 S L_1')^{-1}(L_1 \underset{\sim}{d}) = |L_1 S L_1'|^{-1} [&|S_{(2\cdot 2)}| \underset{\sim}{d}'_{(2)} S^{-1}_{2\cdot 2} \underset{\sim}{d}_{(2)} \rho^2 \\ &+ 2|S_{(1\cdot 2)}| \underset{\sim}{d}'_{(2)} S^{-1}_{(1\cdot 2)} \underset{\sim}{d}_{(1)} \rho \\ &+ |S_{(1\cdot 1)}| \underset{\sim}{d}'_{(1)} S^{-1}_{(1\cdot 1)} \underset{\sim}{d}_{(1)}],\end{aligned} \tag{3.1}$$

where

$$|L_1 S L_1'| = |S_{(2\cdot 2)}| \rho^2 + 2|S_{(1\cdot 2)}| \rho + |S_{(1\cdot 1)}|,$$

and $S_{(i_1,\ldots,i_k \cdot j_1,\ldots,j_\ell)}$ and $\underset{\sim}{d}_{(i_1,\ldots,i_k)}$ denote respectively the matrix obtained from S by deleting the $i_1,\ldots,i_k$ th rows and the $j_1,\ldots,j_\ell$ th columns and the vector obtained from $\underset{\sim}{d}$ by deleting the $i_1,\ldots,i_k$ th components. Substituting (3.1) into (2.3) gives the confidence region for ρ as

$$g(\rho) = |S_{(2\cdot 2)}|[\alpha_1 - \beta_1\ \underset{\sim}{d}'_{(2)}S^{-1}_{(2\cdot 2)}\underset{\sim}{d}_{(2)}]\rho^2$$

$$+ 2|S_{(1\cdot 2)}|[\alpha_1 - \beta_1\ \underset{\sim}{d}'_2 S^{-1}_{(1\cdot 2)}\underset{\sim}{d}_{(1)}]\rho$$

$$+ |S_{(1\cdot 1)}|[\alpha_1 - \beta_1\ \underset{\sim}{d}'_{(1)}S^{-1}_{(1\cdot 1)}\underset{\sim}{d}_{(1)}] \leq 0. \tag{3.2}$$

Put

$$C_1 = |S_{2\cdot 2}|[\alpha_1 - \beta_1\ \underset{\sim}{d}'_{(2)}S^{-1}_{(2\cdot 2)}\underset{\sim}{d}_{(2)}], \text{ and}$$

$$(3.3)$$

$$D_1 = |S_{(1\cdot 2)}|^2[\alpha_1 - \beta_1\ \underset{\sim}{d}'_{(2)}S^{-1}_{(1\cdot 2)}\underset{\sim}{d}_{(1)}]^2$$

$$- \prod_{i=1}^{2} |S_{(i\cdot i)}|[\alpha_1 - \beta_1\ \underset{\sim}{d}_{(i)}S^{-1}_{(i\cdot i)}\underset{\sim}{d}_{(i)}],$$

(the discriminant of the quadratic form $g(\rho)$) and let r_1 and r_2 denote the real roots of $g(\rho) = 0$ $(r_1 \leq r_2)$, if they exist, then the confidence region may be given by

(i) $r_1 \leq \rho \leq r_2$, if Condition $I^{(1)}$: $C_1 > 0$ and Condition $II^{(1)}$: $D_1 > 0$,

(ii) $\rho \leq r_1$ or $\rho \geq r_2$, if $C_1 \leq 0$ and $D_1 > 0$, (3.4)

or

(iii) $-\infty < \rho < \infty$, if $C_1 \leq 0$ and $D_1 \leq 0$.

We may be most interested in the case (i), i.e., when the confidence region is bounded. Taking into account a similar procedure for $1/\rho$ shows that only the case (iii) may be an undesired situation, in which case the value of ρ can not be definitely defined and the whole real line represents the confidence region. The implication of the Conditions $I^{(1)}$ and $II^{(1)}$ are clarified in the following:

Lemma 3.1. D_1 in (3.3) is simplified as

$$D_1 = -|S||S_{(1,2\cdot1,2)}|(\alpha_1 - \beta_1\, d'S^{-1}d)[\alpha_1 - \beta_1\, d'_{(1,2)}S^{-1}_{(1,2\cdot1,2)}d_{(1,2)}], \tag{3.5}$$

for which we must show the following equalities:

$$|S_{(1\cdot2)}|^2 - \prod_{i=1}^{2} |S_{(i\cdot i)}| = -|S||S_{(1,2\cdot1,2)}|, \tag{3.6}$$

$$\prod_{i=1}^{2} |S_{(i\cdot i)}| \sum_{j=1}^{2} d'_{(j)}S^{-1}_{(j\cdot j)}d_{(j)} - 2|S_{(1\cdot2)}|^2\, d'_{(2)}S^{-1}_{(1\cdot2)}d_{(1)}$$

$$= |S||S_{(1,2\cdot1,2)}|[d'S^{-1}d + d_{(1,2)}S^{-1}_{(1,2\cdot1,2)}d_{(1,2)}], \tag{3.7}$$

and

$$|S_{(1\cdot2)}|^2[d'_{(2)}S_{(1\cdot2)}d_{(1)}]^2 - \prod_{i=1}^{2} |S_{(i\cdot i)}|d'_{(i)}S^{-1}_{(i\cdot i)}d_{(i)}$$

$$= -|S||S_{(1,2\cdot1,2)}|d'S^{-1}d\; d'_{(1,2)}S^{-1}_{(1,2\cdot1,2)}d_{(1,2)}. \tag{3.8}$$

Proof. Those equalities are proved by some lengthy but straightforward algebra, and hence the proof is omitted here.

It is shown from (3.5) and the fact $\alpha_k - \beta_k\, d'S^{-1}d < 0$ that the Condition $II^{(1)}$ is equivalent to $\alpha_1 - \beta_1\, d'_{(1,2)} S^{-1}_{(1,2\cdot1,2)}d_{(1,2)} > 0$. It is well known that the Mahalanobis' distance is an increasing function of the number of measurements involved. Hence the Condition $I^{(1)}$: $\alpha_1 - \beta_1\, d'_{(2)}S^{-1}_{(2\cdot2)}d_{(2)} > 0$

implies the Condition $II^{(1)}$.

We summarize in

<u>Theorem 3.1</u>. The $100(1-\alpha)$ per cent confidence region for $\rho = a_1/a_2$ is given by (3.2). Especially this defines a bounded interval on the real line, if the Condition $I^{(1)}$: $\alpha_1 - \beta_1 \underset{\sim}{d}'_{(2)} S^{-1}_{(2\cdot 2)} \underset{\sim}{d}_{(2)} > 0$ is satisfied. The Condition $II^{(1)}$: $\alpha_1 - \beta_1 \underset{\sim}{d}'_{(1,2)} S^{-1}_{(1,2\cdot 1,2)} \underset{\sim}{d}_{(1,2)} > 0$ is the condition for the confidence region not to be the whole real line. These conditions define acceptance regions for the tests of hypotheses $a_2 \neq 0$, and $a_1 \neq 0$ or $a_2 \neq 0$ respectively. These hypotheses are the conditions for the existence of the bounded and definite values of the ratio a_1/a_2. The actual (conditional) confidence coefficients for these pseudo confidence regions are given by the probabilities of (3.2) conditional on those conditions. However, the probability of the conditions approaches one under 'reasonable circumstances' in most practical situations, and hence the bounded confidence region occurs with the confidence coefficient approximately equal to $1-\alpha$ with probability one in 'reasonable' practical situations.

The problem of pseudo confidence regions which are confidence regions for some parametric functions valid conditionally in subsets of the parameter space is rigorously discussed by Chikuse (1980) especially in connection with a multivariate linear functional relationship.

4. ASSIGNED RATIOS OF SEVERAL COEFFICIENTS; CASE I

For the hypothesis H_2: $a_1/a_{k+1} = \rho_1$, $a_2/a_{k+1} = \rho_2, \ldots,$ $a_k/a_{k+1} = \rho_k$ $(k \geq 2)$, the L matrix in (2.4) is given by $L_2 = [\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_k, I_{p-k}]$, where $\underset{\sim}{\Gamma}_i = (\rho_i, 0, \ldots, 0)'$, $i = 1, \ldots, k$. Some algebra similar to, but more complicated than that for the hypothesis H_1 yields the following results,

$$(L_2\underset{\sim}{d})'(L_2SL_2')^{-1}(L_2\underset{\sim}{d}) = |L_2SL_2'|^{-1}[\sum_{i=1}^{k} B_{ii}\rho_i^2$$

$$+ 2\sum_{i,j=1,i<j}^{k} B_{ij}\rho_i\rho_j + 2\sum_{i=1}^{k} B_i\rho_i + B_0], \quad (4.1)$$

where

$$|L_2SL_2'| = \sum_{i=1}^{k} A_{ii}\rho_i^2 + 2\sum_{i,j=1,i<j}^{k} A_{ij}\rho_i\rho_j + 2\sum_{i=1}^{k} A_i\rho_i + A_0,$$

and, putting $J = \{1,2,\ldots,k\}$,

$$A_{ii} = |S_{(J-\{i\},k+1\cdot J-\{i\},k+1)}|,$$

$$B_{ii} = A_{ii}\underset{\sim}{d}'_{(J-\{i\},k+1)}S^{-1}_{(J-\{i\},k+1\cdot J-\{i\},k+1)}\underset{\sim}{d}_{(J-\{i\},k+1)},\quad i=1,\ldots,k,$$

$$A_{ij} = |S_{(J-\{j\},k+1\cdot J-\{i\},k+1)}|,$$

$$B_{ij} = A_{ij}\underset{\sim}{d}_{(J-\{i\},k+1)}S^{-1}_{(J-\{j\},k+1\cdot J-\{i\},k+1)}\underset{\sim}{d}_{(J-\{j\},k+1)'} \quad (4.2)$$
$$i < j,\quad i,j = 1,\ldots,k.$$

$$A_i = |S_{(J\cdot J-\{i\},k+1)}|,$$

$$B_i = A_i\underset{\sim}{d}'_{(J-\{i\},k+1)}S^{-1}_{(J\cdot J-\{i\},k+1)}\underset{\sim}{d}_{(J)},\quad i = 1,\ldots,k,$$

$$A_0 = |S_{(J\cdot J)}|,$$

$$B_0 = A_0\underset{\sim}{d}'_{(J)}S^{-1}_{(J\cdot J)}\underset{\sim}{d}_{(J)}.$$

Substituting (4.1) into (2.3) yields the confidence region for $\underset{\sim}{\rho} = (\rho_1,\ldots,\rho_k)'$ as

$$(\underset{\sim}{\rho} + C_k^{-1}\underset{\sim}{c}_k)' C_k (\underset{\sim}{\rho} + C_k^{-1}\underset{\sim}{c}_k) - (\underset{\sim}{c}_k' C_k^{-1} \underset{\sim}{c}_k - c_0) \le 0, \qquad (4.3)$$

where $C_k = (c_{ij})$, with $c_{ij} = c_{ji}$, $\underset{\sim}{c}_k = (c_1,\ldots,c_k)'$,

$$\begin{aligned} c_{ii} &= \alpha_k A_{ii} - \beta_k B_{ii}, \quad i = 1,\ldots,k, \\ c_{ij} &= \alpha_k A_{ij} - \beta_k B_{ij}, \quad i < j, \quad i,j = 1,\ldots,k, \\ c_i &= \alpha_k A_i - \beta_k B_i, \quad i = 1,\ldots,k, \\ c_0 &= \alpha_k A_0 - \beta_k B_0, \end{aligned} \qquad (4.4)$$

with the A's and B's given in (4.2).

The inequality (4.3) gives the interior of a hyperellipsoid in R^k as the only bounded region, if the following two conditions are satisfied:

$$\begin{aligned} &\text{Condition } I^{(k)}: \; C_k > 0 \text{ positive definite and,} \\ &\text{Condition } II^{(k)}: \; D_k = \underset{\sim}{c}_k' C_k^{-1} \underset{\sim}{c}_k - c_0 > 0. \end{aligned} \qquad (4.5)$$

(the discriminant of the quadratic form in (4.3))

Lemma 4.2.

$$|C_k| = |S_{(k+1\cdot k+1)}||S_{(J,k+1\cdot J,k+1)}|^{k-1}[\alpha_k - \beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)}]$$

$$[\alpha_k - \beta_k \underset{\sim}{d}'_{(J,k+1)} S^{-1}_{(J,k+1\cdot J,k+1)} \underset{\sim}{d}_{(J,k+1)}]^{k-1},$$

and (4.6)

$$D_k = -|S||S_{(J,k+1\cdot J,k+1)}|(\alpha_k - \beta_k \underset{\sim}{d}' S^{-1} \underset{\sim}{d})$$

$$[\alpha_k - \beta_k \underset{\sim}{d}'_{(J,k+1)} S^{-1}_{(J,k+1\cdot J,k+1)} \underset{\sim}{d}_{(J,k+1)}]$$

$$/|S_{(k+1\cdot k+1)}|[\alpha_k-\beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)}].$$

It follows that

(i) the Condition $I^{(k)}$ is equivalent to

$$\alpha_k-\beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)} > 0,$$

(ii) $D_k \underset{(\leq)}{>} 0$ is equivalent to

$$[\alpha_k-\beta_k \underset{\sim}{d}'_{(J,k+1\cdot J,k+1)} S^{-1}_{(J,k+1\cdot J,k+1)} \underset{\sim}{d}_{(J,k+1)}]$$

$$/[\alpha_k-\beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)}] \underset{(\leq)}{>} 0,$$

and therefore

(iii) the Condition $I^{(k)}$ implies the Condition $II^{(k)}$.

We summarize in

<u>Theorem 4.1</u>. The $100(1-\alpha)$ per cent confidence region for $\underset{\sim}{\rho} = (\rho_1,\rho_2,\ldots,\rho_k)'$ is given by (4.3). This especially gives the bounded region defined by the interior of a hyperellipsoid in R^k, if the Condition $I^{(k)}$: $\alpha_k-\beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)} > 0$ is satisfied. The condition $I^{(k)}$ defines an acceptance region for the test of hypothesis $a_{k+1} \neq 0$. The whole space R^k represents the confidence region, if $\alpha_k-\beta_k \underset{\sim}{d}'_{(j)} S^{-1}_{(j\cdot j)} \underset{\sim}{d}_{(j)} < 0$, $j = 1,2,\ldots,k+1$, which defines an acceptance region for the test of hypothesis $a_1 = a_2 = \ldots = a_{k+1} = 0$. The last part of Theorem 3.1 immediately generalizes to the present case.

We notice here that the significance level of the above test for the hypothesis $a_{k+1} = 0$, the critical region of which is given by the Condition $I^{(k)}$: $\alpha_k-\beta_k \underset{\sim}{d}'_{(k+1)} S^{-1}_{(k+1\cdot k+1)} \underset{\sim}{d}_{(k+1)} > 0$, is much larger than α.

5. ASSIGNED RATIOS OF SEVERAL COEFFICIENTS: CASE II

We consider the hypothesis H_3: $a_1/a_{k+1} = \rho_1$, $a_2/a_{k+2} = \rho_2,\ldots,a_k/a_{2k} = \rho_k$ $(k \geq 2)$. The L matrix in (2.4) is given by $L_3 = [\underset{\sim}{\Delta}_1,\underset{\sim}{\Delta}_2,\ldots,\underset{\sim}{\Delta}_k,I_{p-k}]$, where $\underset{\sim}{\Delta}_i = (0,\ldots,0,\rho_i,0,\ldots,0)'$, the vector of p-k components zero except the ith component which is ρ_i. Here for simplicity we consider the case $k = 2$, i.e., the hypothesis $H_3^{(2)}$: $a_1/a_3 = \rho_1$, $a_2/a_4 = \rho_2$.

Theorem 5.1. The $100(1-\alpha)$ per cent confidence region for (ρ_1,ρ_2) is given by

$$g_{21}(\rho_2|\rho_1) = (C_{11\cdot 22}\rho_1^2+2C_{1\cdot 22}\rho_1+C_{22})\rho_2^2 + 2(C_{11\cdot 2}\rho_1^2+C_{1\cdot 2}\rho_1+C_2)\rho_2$$

$$+ (C_{11}\rho_1^2+2C_1\rho_1+C_0) \leq 0, \tag{5.1}$$

or, equivalently, $g_{12}(\rho_1|\rho_2) \leq 0$, where $g_{12}(\rho_1|\rho_2)$ is obtained from $g_{21}(\rho_2|\rho_1)$ being rewritten as a quadratic form in terms of ρ_1, where

$$\begin{aligned}
C_{11\cdot 22} &= |S_{(3,4\cdot 3,4)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(3,4)}S^{-1}_{(3,4\cdot 3,4)}\underset{\sim}{d}_{(3,4)}],\\
C_{11\cdot 2} &= |S_{(2,3\cdot 3,4)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(3,4)}S^{-1}_{(2,3\cdot 3,4)}\underset{\sim}{d}_{(2,3)}],\\
C_{1\cdot 22} &= |S_{(1,4\cdot 3,4)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(3,4)}S^{-1}_{(1,4\cdot 3,4)}\underset{\sim}{d}_{(1,4)}],\\
C_{1\cdot 2} &= |S_{(1,2\cdot 3,4)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(3,4)}S^{-1}_{(1,2\cdot 3,4)}\underset{\sim}{d}_{(1,2)}],\\
C_{11} &= |S_{(2,3\cdot 2,3)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(2,3)}S^{-1}_{(2,3\cdot 2,3)}\underset{\sim}{d}_{(2,3)}], \qquad (5.2)\\
C_{22} &= |S_{(1,4\cdot 1,4)}|[\alpha_2 - \beta_2\underset{\sim}{d}'_{(1,4)}S^{-1}_{(1,4\cdot 1,4)}\underset{\sim}{d}_{(1,4)}],
\end{aligned}$$

$$C_1 = |S_{(1,2\cdot 2,3)}|[\alpha_2 - \beta_2 \underset{\sim}{d}'_{(2,3)} S^{-1}_{(1,2\cdot 2,3)} \underset{\sim}{d}_{(1,2)}],$$

$$C_2 = |S_{(1,2\cdot 1,4)}|[\alpha_2 - \beta_2 \underset{\sim}{d}'_{(1,4)} S^{-1}_{(1,2\cdot 1,4)} \underset{\sim}{d}_{(1,2)}],$$

$$C_0 = |S_{(1,2\cdot 1,2)}|[\alpha_2 - \beta_2 \underset{\sim}{d}'_{(1,2)} S^{-1}_{(1,2\cdot 1,2)} \underset{\sim}{d}_{(1,2)}].$$

(5.1), in general, defines asymmetric regions on the (ρ_1,ρ_2) plane. The boundedness of the region (5.1) is equivalent to the condition that there is no solution of (5.1) at the infinite points, namely,

$$C_{11\cdot 22}\rho_1^2 + 2C_{1\cdot 22}\rho_1 + C_{22} > 0 \quad \text{for any} \quad \rho_1 \tag{5.3}$$

and

$$C_{11\cdot 22}\rho_2^2 + 2C_{11\cdot 2}\rho_2 + C_{11} > 0 \quad \text{for any} \quad \rho_2, \tag{5.4}$$

which are respectively equivalent to

$$\alpha_2 - \beta_2 \underset{\sim}{d}'_{(3)} S^{-1}_{(3\cdot 3)} \underset{\sim}{d}_{(3)} > 0 \tag{5.5}$$

and

$$\alpha_2 - \beta_2 \underset{\sim}{d}'_{(4)} S^{-1}_{(4\cdot 4)} \underset{\sim}{d}_{(4)} > 0. \tag{5.6}$$

<u>Theorem 5.2</u>. The confidence regions (5.1) are bounded when the conditions (5.5) and (5.6) are both satisfied. These conditions define an intersection of rejection regions against the tests of hypotheses $a_3 = 0$ and $a_4 = 0$. The last part of Theorem 3.1 immediately applies to the present case.

6. ILLUSTRATIONS

We shall illustrate the confidence regions considered in the previous sections, using the well known data of Fisher (1938). The data consists of 50 observations each with four characters, sepal length (x_1), sepal width (x_2), petal length (x_3) and petal width (x_4) in two species of plants Iris versicolor and Iris setosa. We utilize the tables of observed mean values and covariances shown in Rao (1970a, Chapter 7, Section 7b.3) in our computation. The sample coefficients of the discriminant function are given by

$$\begin{aligned} \hat{a}_1 &= -3.0692, \quad \hat{a}_2 = -18.0006, \\ \hat{a}_3 &= 21.7641, \quad \hat{a}_4 = 30.7549. \end{aligned} \tag{6.1}$$

Fig. 1, shows the 97.5% and 99.5% confidence regions, given by (4.3) with $k = 2$, for $\rho_1 = a_1/a_3$ and $\rho_2 = a_2/a_3$. The point of the sample ratios $\hat{\rho}_1 = \hat{a}_1/\hat{a}_3$ and $\hat{\rho}_2 = \hat{a}_2/\hat{a}_3$ with the $\hat{a}_i$, $i = 1,2,3$, given by (6.1) is plotted for reference. It is noted that the values of $\alpha_2 - \beta_2 \underset{\sim}{d}'_{(3)} S^{-1}_{(3 \cdot 3)} \underset{\sim}{d}_{(3)}$ are positive for both cases. Fig. 2 shows the 97.5% and 99.5% confidence regions, given by (5.1), for $\rho_1 = a_1/a_3$ and $\rho_2 = a_2/a_4$. The point of the sample ratios $\hat{\rho}_1 = \hat{a}_1/\hat{a}_3$ and $\hat{\rho}_2 = \hat{a}_2/\hat{a}_4$ with the $\hat{a}_i$, $i = 1,\ldots,4$, given by (6.1) is plotted. Here, the values of $\alpha_2 - \beta_2 \underset{\sim}{d}'_{(4)} S^{-1}_{(4 \cdot 4)} \underset{\sim}{d}_{(4)}$ are positive for the case of 97.5% and negative for the case of 99.5%, and the values of $\alpha_2 - \beta_2 \underset{\sim}{d}'_{(3)} S^{-1}_{(3 \cdot 3)} \underset{\sim}{d}_{(3)}$ are positive for both cases.

ACKNOWLEDGEMENT

This work is partially sponsored by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161.

BIBLIOGRAPHY

Box, G.E.P. and Hunter, J.S. (1954). A confidence region for the solution of a set of simultaneous equations with an application to experimental design. Biometrika 41, 190-199.

Chikuse, Y. (1980). Pseudo confidence regions for the solution of a multivariate linear functional relationship. Submitted for publication.

Creasy, M.A. (1954). Limits for the ratio of means. J. Roy. Statist. Soc. B16, 186-194.

Fieller, E.C. (1954). Some problems in interval estimation. J. Roy. Statist. Soc. B16, 175-185.

Fisher, R.A. (1938). The statistical utilization of multiple measurements. Ann. Eugenics 8, 376-386.

Fisher, R.A. (1940). The precision of discriminant functions. Ann. Eugenics 10, 422-429.

James, A.T., Wilkinson, G.N. and Venables, W.N. (1974). Interval estimates for a ratio of means. Sankhya A36, 177-183.

Kshirsagar, A.M. (1963). Confidence intervals for discriminant function coefficients. J. Indian Statist. Assoc. 1, 1-7.

Kshirsagar, A.M. (1973). Multivariate Analysis. New York: Marcel Dekker.

Ogawa, J. (1977). On the 'confidence bounds' of the ratio of the means of a bivariate normal distribution. Mimeograph series No. 19. Science Laboratory, Soka University, Tokyo.

Rao, C.R. (1970a). Advanced Statistical Methods in Biometric Research, reprinted by arrangement. Darien, Conn.: Hafner.

Rao, C.R. (1970b). Inference on discriminant function coefficients. In Essays in Probability and Statistics, Ed. R.C. Bose, I.M. Chakravarti, P.C. Mahalanobis, C.R. Rao and K.J.C. Smith: Univ. of North Carolina Press, 587-602.

Roy, S.N. and Potthoff, R.F. (1958). Confidence bounds on vector analogues on the "ratio of means" and the "ratio of variances" for two correlated normal variates and some associated tests. Ann. Math. Statist. 29, 829-841.

Zerbe, G.O. (1978). On Fieller's Theorem and the general linear model. Amer. Statist. 32, 103-105.

FIG. 1. Confidence regions for $\rho_1 = a_1/a_3$ and $\rho_2 = a_2/a_3$.

(1a) 97.5% confidence region

(1b) 99.5% confidence region

FIG. 2. Confidence regions for $\rho_1 = a_1/a_3$ and $\rho_2 = a_2/a_4$.

(2a) 97.5% confidence region

(2b) 99.5% confidence region

Fig. 1a

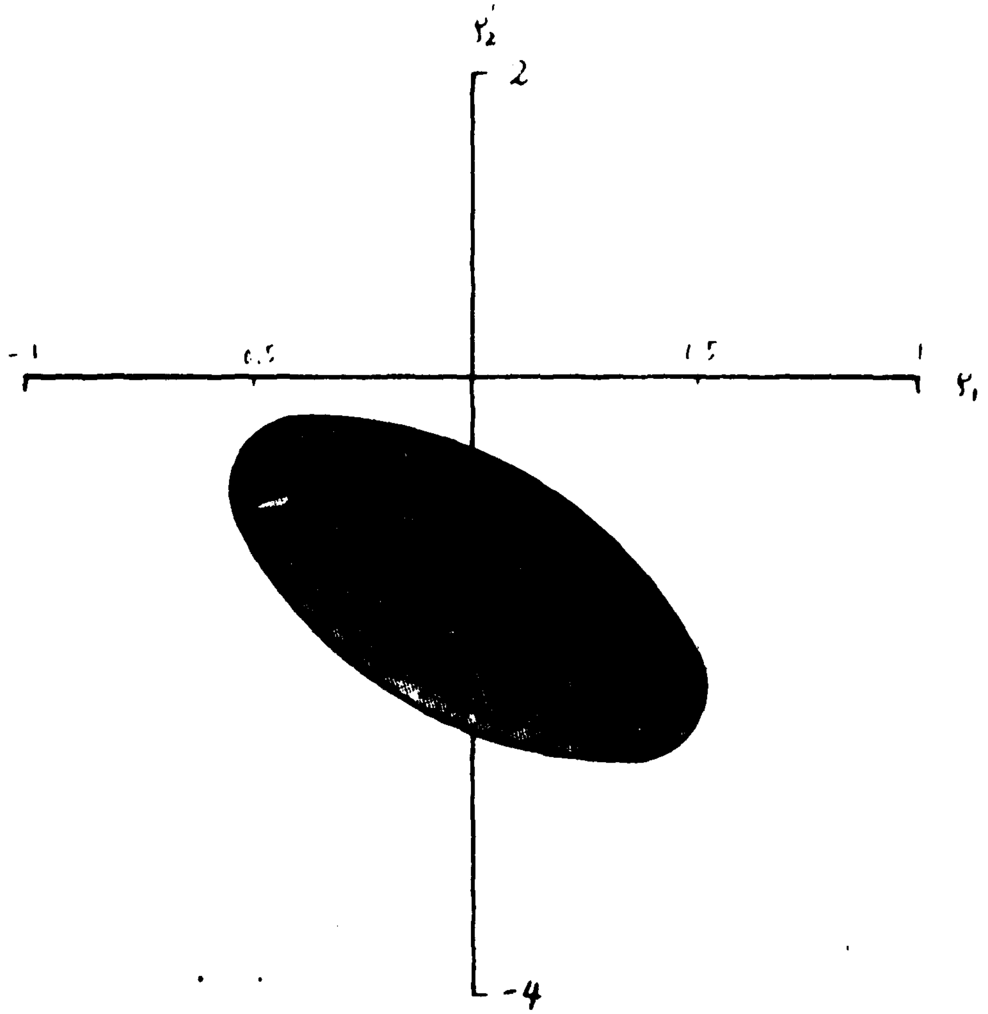

Fig. 1b

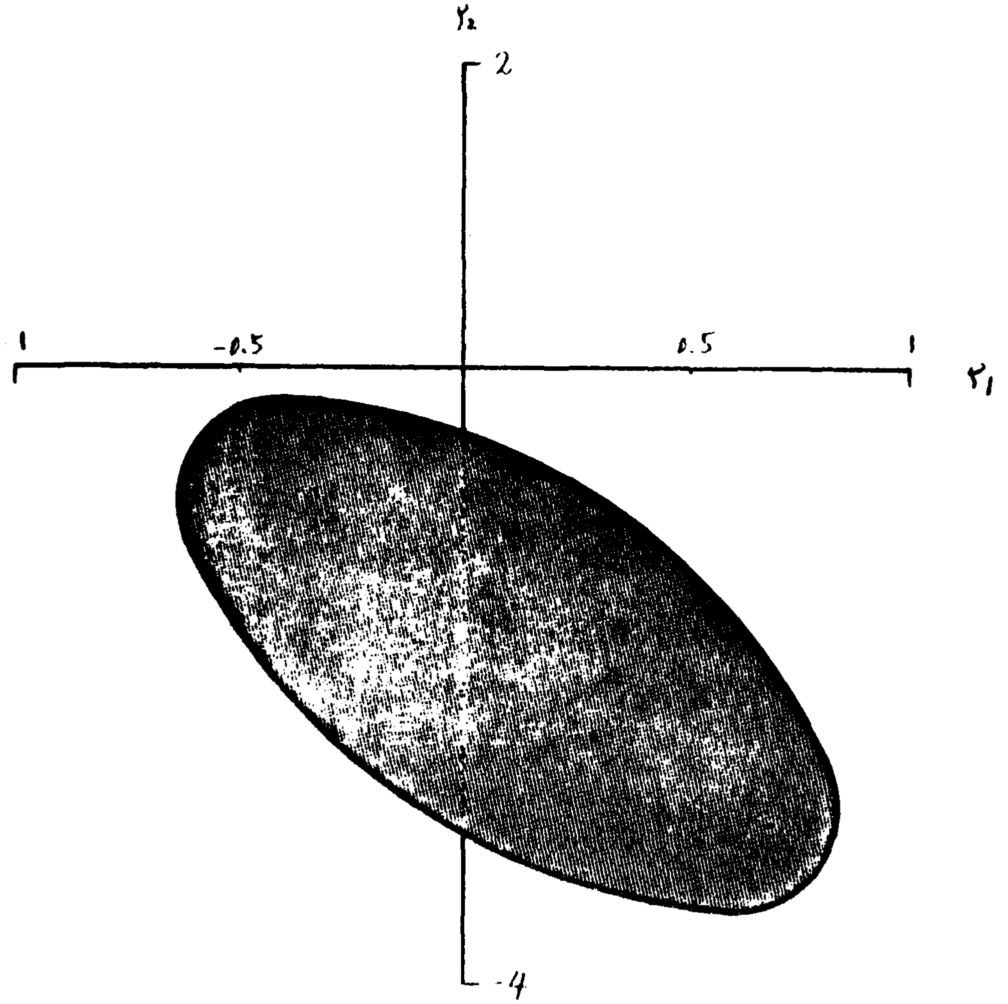

Fig. 2a

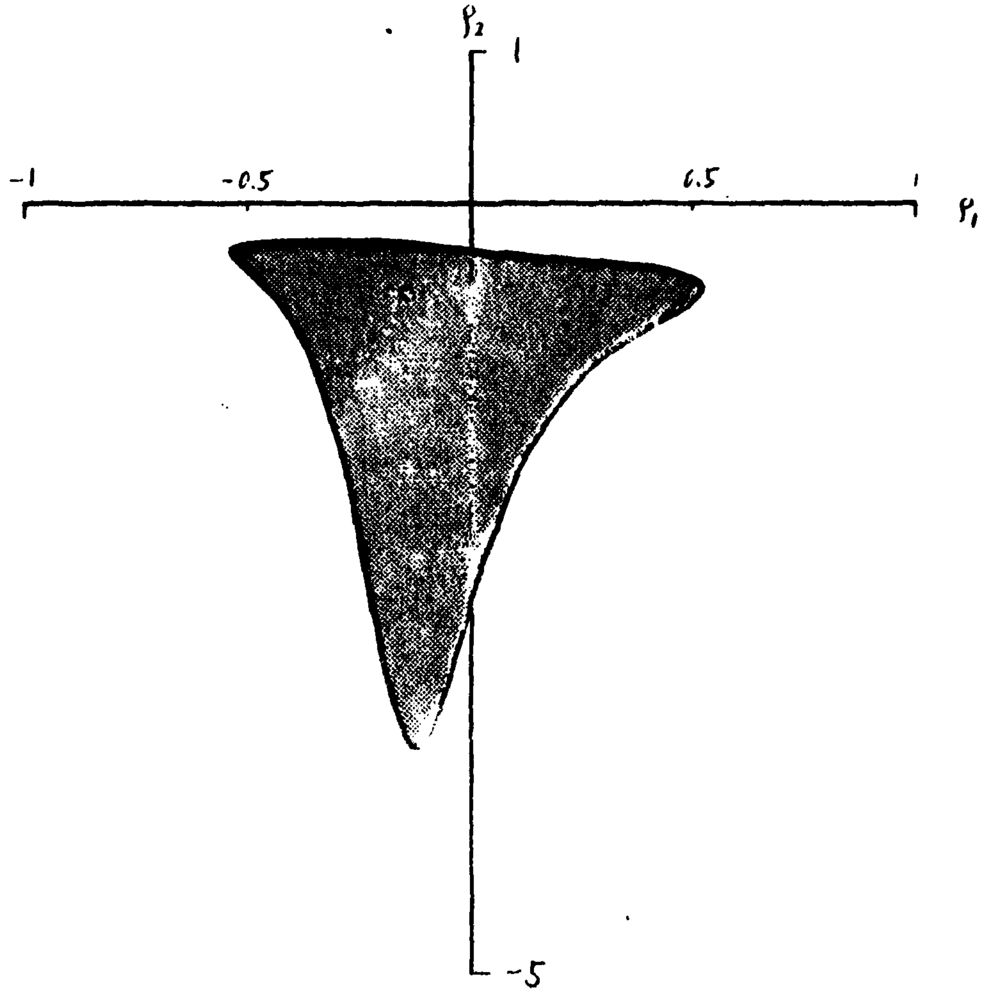

Fig. 2b