AD-A043 493 ILLINOIS UNIV AT CHICAGO CIRCLE DEPT OF MATHEMATICS F/G 12/1

ON SCHEFFE'S S-METHOD: A REVIEW.(U)

AUG 77 A HEDAYAT AF-AFOSR-3050-76

UNCLASSIFIED AFOSR-TR-77-1147 NL

1 OF 1

AD A043493

END

DATE FILMED

9-77

DDC

AFOSR-TR- 77-1147

(4)

Approved for public release;
distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-77-1147	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON SCHEFFE'S S-METHOD: A REVIEW		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A. Hedayat		8. CONTRACT OR GRANT NUMBER(s) AFOSR 76-3050
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Illinois at Chicago Circle, Department of Mathematics, Box 4348, Chicago, IL 60680		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332		12. REPORT DATE August 1977
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		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)

Scheffe's S-method, simultaneous statistical inference, linear estimation, contrast, connectedness, F test

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

This paper gives in detail the background and derivation of a simultaneous statistical inference in linear models which is commonly known as Scheffe S-method. It is written for a person with a knowledge of linear models and design of experiment. It also gives a list of 46 references related to the subject.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

On Scheffe's S-Method : A Review

by

A.Hedayat

August, 1977

Department of Mathematics

University of Illinois, Chicago

Research sponsored by the Air Force Office of Scientific Research under Grant AFOSR 76-3050

(See 1473)

On Scheffé's S-Method: A Review

A. Hedayat
Department of Mathematics
University of Illinois at Chicago Circle
Chicago, Illinois 60680

Introduction: We shall assume throughout the following model:

$$\Omega: Y = X\beta + e$$

where Y is an $n \times 1$ vector of observations, X is an $n \times p$ known matrix of rank r, β is a $p \times 1$ vector of unknown parameters, and e is distributed $N(0, \sigma^2 I)$.

Definition 1. A linear parametric function $\Psi = c'\beta$, where c is a vector of known constants, is said to be an estimable linear parametric function if there exists an unbiased linear estimator $a'Y$, i.e., such that $E(a'Y) = \Psi$.

The estimability of Ψ solely depends on the design matrix X as the following known tests for estimability of Ψ show:

(i) Ψ is estimable if and only if c' is in the row space of X, i.e., if and only if there exists a vector t such that $c' = t'X$.

(ii) Ψ is estimable if and only if there exists a vector k such that $c' = k'X'X$.

(iii) Ψ is estimable if and only if $c' = c'H$, where $H = (X'X)^{-} X'X$.

It should be noted that in practice it is not a trivial matter to check for estimability due to complicated nature of the design matrix X. This is why the experimenter is well advised to specify his set of linear parametric functions of interest and try to collect his data (i.e., chooses his design matrix X) which guarantees the estimability of his functions of interest.

<u>Definition 2</u>. We say the design X is connected for Ψ if Ψ is estimable under X. Otherwise, X is said to be disconnected for Ψ.

<u>Estimation, Test and Confidence Interval for Ψ</u>.

If Ψ is estimable under X then it's well known that the best linear unbiased estimator of Ψ is given by

$$\hat{\Psi} = c'\hat{\beta} = c'(X'X)^{-} X'Y,$$

with

$$\text{var } \Psi = \sigma^2_{\hat{\Psi}} = c'(X'X)^{-} c\sigma^2.$$

An unbiased estimator of $\sigma^2_{\hat{\Psi}}$ is $c'(X'X)^{-} c\hat{\sigma}^2$ where

$$\hat{\sigma}^2 = \frac{1}{n-r} Y'(I-X(X'X)^{-}X')Y.$$

Therefore

$$\hat{\Psi} \sim N(\Psi, \sigma^2_{\hat{\Psi}}).$$

We know that

$$Q_1 = \frac{1}{\sigma^2} Y'(I-X(X'X)^{-} X')Y = \frac{n-r}{\sigma^2} \hat{\sigma}^2 \sim \chi^2 (n-r).$$

$\hat{\Psi}$ is a linear form in Y and thus $\hat{\Psi}$ and Q_1 are independent if

$$c'(X'X)^- X'[I-X(X'X)^- X'] = c'(X'X)^- X' - c'(X'X)^- X'X (X'X)^- X' = 0.$$

Now since Ψ is estimable thus c' can be expressed as $t'X$. Therefore, by substituting $t'X$ for c' in the above expression we obtain:

$$= t'X(X'X)^- X' - t'X(X'X)^- X'X(X'X)^- X'$$

$$= t'X(X'X)^- X' - t'X(X'X)^- X' = 0.$$

Since $\hat{\Psi} \sim N(\Psi, \sigma^2_{\hat{\Psi}})$, this imples that

$$\frac{\hat{\Psi} - \Psi}{\sigma_{\hat{\Psi}}} = \frac{\hat{\Psi} - \Psi}{\sigma(c'(X'X)^- c)^{1/2}} \sim N(0, 1).$$

Thus

$$\frac{(\hat{\Psi}-\Psi)/\sigma(c'(X'X)^- c)^{1/2}}{[Q_1/(n-r)]^{1/2}} = \frac{\hat{\Psi} - \Psi}{\hat{\sigma}\,[c'(X'X)^- c]^{1/2}}$$

$$= \frac{\Psi - \Psi}{\hat{\sigma}_{\hat{\Psi}}} \sim t(n-r).$$

or equivalently

$$\left(\frac{\hat{\Psi} - \Psi}{\hat{\sigma}_{\hat{\Psi}}}\right)^2 \sim F(1, n-1).$$

This statistic can be used for testing hypothesis of the form Ho: $\Psi = m$. This statistic can also be used for constructing confidence intervals for Ψ.

$$P\left[\frac{(\hat{\Psi}-\Psi)^2}{\hat{\sigma}^2_{\hat{\Psi}}} \le F_\alpha(1,n-1)\right] = 1 - \alpha$$

or

$$P\left[\hat{\Psi} - \hat{\sigma}_{\hat{\Psi}}(F_\alpha(1,n-1))^{1/2} \le \Psi \le \hat{\Psi} + \hat{\sigma}_{\hat{\Psi}}(F_\alpha(1,n-1))^{1/2}\right] = 1 - \alpha.$$

Suppose we have a set of linear paramteric functions $\Psi_1, \Psi_2, \ldots, \Psi_t$ and we wish to construct $1 - \alpha$ simultaneous confidence intervals for these t linear parametric functions. The above confidence intervals give $1 - \alpha$ confidence intervals for individual Ψ's. Scheffe's S-method answers this problem.

But first we need the following definitions.

Definition 3. Two linear parametric functions Ψ_1 and Ψ_2 are said to be algebraically independent if their corresponding coefficient vectors c_1 and c_2 are independent.

Definition 4. By a q-dimensional subspace L of linear estimable functions under the design matrix X we mean the subspace generated by the coefficient vectors of q independent lieanr estimable functions under X. We say $\Psi = c'\beta \in L$ if $c' \in L$.

The S-method of multiple comparison is based on

<u>Theorem 1</u>. Under Ω the probability is $1-\alpha$ that all estimable functions Ψ in a given q-dimensional space L simultaneously satisfy

$$(1) \qquad \hat{\Psi} - S\hat{\sigma}_{\hat{\Psi}} \leq \Psi \leq \hat{\Psi} + S\hat{\sigma}_{\hat{\Psi}} ,$$

where $S = [q\ F_{\alpha}(q,n-r)]^{1/2}$.

One can rewrite (1) in the following form

$$P\{|\hat{\Psi}-\Psi| \leq \hat{\sigma}_{\hat{\Psi}}\ [q\ F_{\alpha}(q,\ n-1)]^{1/2} \text{ for all } \Psi \in L\} = 1 - \alpha,$$

or

$$P\{\frac{(\hat{\Psi}-\Psi)^2/c'(X'X)^{-}c}{\hat{\sigma}^2} \leq F_{\alpha}(q,n-1) \text{ for all } \Psi \in L\} = 1 - \alpha.$$

Since $\hat{\Psi} = c'\hat{\beta}$, $\Psi = c'\beta$ therefore it suffices to prove that the maximum value of $(c'\hat{\beta} - c'\beta)/c'(X'X)^{-}c$ for all nonzero $c \in L$ is distributed as $\sigma^2 \chi^2(q)$; and this maximum is independent of $(n-1)\hat{\sigma}^2$ which is distributed as $\sigma^2 \chi^2(n-r)$. To prove this we need the following lemmas.

<u>Lemma 1</u>. Let A be a symmetric matrix of order n. The maximum value of $z'Az/z'z$ over all nonzero $z \in E_n$ is ℓ, the largest eigenvalue of A, and this maximum is attained when z is any eigenvector of A corresponding to the root ℓ.

Proof. First we show that the following two problems are equivalent.

(i) $\underset{z \neq 0}{\text{maximize}} \; \frac{z'Az}{z'z}$, (ii) $\underset{z'z=1}{\text{maximize}} \; z'Az$.

Let

$$\max_{z \neq 0} \frac{z'Az}{z'z} = m_1 \quad \text{and} \quad \max_{z'z=1} z'Az = m_2$$

and suppose m_1 is attained for $z = z_1$ and m_2 is attained for $z = z_2$, i.e.,

$$\frac{z_1'Az_1}{z_1'z_1} = m_1 \quad \text{and} \quad z_2'Az_2 = m_2.$$

Let

$$\bar{z}_1 = \frac{1}{\sqrt{z_1'z_1}} z_1, \text{ then } \bar{z}_1' \bar{z}_1 = 1,$$

thus
$$\frac{\bar{z}_1'A\bar{z}_1}{\bar{z}_1'\bar{z}_1} = \frac{z_1'Az_1}{z_1'z_1} = m_1.$$

Therefore, $m_1 \leq m_2$. Also since $z_2'z_2 = 1$, $z_2 \neq 0$ and

$$\frac{z_2'Az_2}{z_2'z_2} = \frac{m_2}{1} = m_2$$

thus $m_2 \leq m_1$. Hence $m_1 = m_2$.

Note: Since $\max_{z \neq 0} z'Az/z'z$ is equivalent to $\max_{z'z=1} z'Az$ and $z'Az$ is a continuous function of z and $\{z : z \in E_n, z'z = 1\}$ is a compact set, it follows that $\max_{z \neq 0} z'Az/z'z$ exists.

We shall now present two methods of finishing the proof of

Lemma 1.

Method 1 (Lagrange multiplier method).

We want to maximize $z'Az$ subject to the constraint $z'z = 1$.

Let

$$f(z,\lambda) = z'Az - \lambda(z'z-1)$$

where λ is a Lagrange multiplier. Since $\frac{\partial f}{\partial z} = 2\,Az - 2\,\lambda z = 0$. this imples that $Az = \lambda z$ so λ must be an eigenvalue of A. On the other hand, if λ is an eigenvalue of A, then

$$z'Az = z'(Az) = z'(\lambda z) = \lambda z'z,$$

so that

$$\max_{z'z=1} z'Az = \max_{z'z=1} \lambda z'z = \max \lambda$$

where λ is an eigenvalue of A. Therefore

$$\max_{z\neq 0} \frac{z'Az}{z'z} = \max_{z'z=1} z'Az = \max \lambda = \ell.$$

Now suppose v is any eigenvector corresponding to ℓ, then

$$\frac{v'Av}{v'v} = \frac{v'(Av)}{v'v} = \frac{v'\ell v}{v'v} = \ell\,\frac{v'v}{v'v} = \ell,$$

so that the maximum value is attained for any eigenvector associated with ℓ.

Method 2. Since A is a real symmetric matrix of order n, there exists an orthogonal matrix P such that $P'AP = \Lambda$ where Λ is a diagonal matrix with the (real) eigenvalues of A on the diagonal. Let P_j denote the j-th column of P, i.e., $P = (P_1 \vdots P_2 \vdots \ldots \vdots P_n)$, then P_j is an eigenvalue of A corresponding to λ_j satisfying

$$P_j' P_k = \begin{cases} 0 & \text{when } j \neq k \\ 1 & \text{when } j = k \end{cases} .$$

It follows that

$$A = P\Lambda P' = \lambda_1 P_1 P_1' + \lambda_2 P_2 P_2' + \ldots + \lambda_n P_n P_n'$$

and

$$I = PP' = P_1 P_1' + P_2 P_2' + \ldots + P_n P_n' .$$

The set $\{P_1, P_2, \ldots, P_n\}$ is an orthonormal basis for E_n. Let $z \in E_n$, then $z = Pw$, where $w' = (w_1, w_2, \ldots, w_n)$ and w_i's are the coordinates of the vector z with respect to the basis $\{P_1, P_2, \ldots, P_n\}$. Therefore

$$\frac{z'Az}{z'z} = \frac{(Pw)'P\Lambda P'(Pw)}{(Pw)'(Pw)} = \frac{w'P'P\Lambda P'Pw}{w'P'Pw}$$

$$= \frac{w'\Lambda w}{w'w} = \frac{\lambda_1 w_1^2 + \lambda_2 w_2^2 + \ldots + \lambda_n w_n^2}{w_1^2 + w_2^2 + \ldots + w_n^2} .$$

So maximizing $z'Az/z'z$ for $z \neq 0$, $z \in E_n$ is equivalent to maximizing

$$\frac{\lambda_1 w_1^2 + \lambda_2 w_2^2 + \ldots + \lambda_n w_n^2}{w_1^2 + w_2^2 + \ldots + w_n^2} \quad \text{for } w \in E_n$$

and $w \neq 0$ (since $z \neq 0$).

Suppose $\max\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \ell$, then

$$\lambda_1 w_1^2 + \lambda_2 w_2^2 + \ldots + \lambda_n w_n^2 \leq \ell \sum_{i=1}^{n} w_i^2 ,$$

so

$$\frac{\lambda_1 w_1^2 + \lambda_2 w_2^2 + \ldots + \lambda_n w_n^2}{w_1^2 + w_2^2 + \ldots + w_n^2} \leq \frac{\ell \sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i^2} = \ell .$$

If $\ell = \lambda_j$ then let $w_i = 0$ if $i \neq j$ and $w_j \neq 0$. Then

$$\frac{\lambda_1 w_1^2 + \lambda_2 w_2^2 + \ldots + \lambda_n w_n^2}{w_1^2 + w_2^2 + \ldots + w_n^2} = \frac{\lambda_j w_j^2}{w_j^2} = \lambda_j = \ell$$

so ℓ is attainable and the maximum is attained for $w = \{0,\ldots,0,w_j,0,\ldots,0\}$ or equivalently for $z = Pw = P_j w_j$, $w_j \neq 0$, i.e., for any eigenvalue of A corresponding to ℓ the largest eigenvalue of A.

Lemma 2. Let A be a symmetric matrix of order n and let B be any positive symmetric matrix of order n. The maximum value of $z'Az/z'Bz$ over all nonzero $z \in E_n$ is ℓ, the largest eigenvalue of $B^{-1}A$, and this maximum is attained for any eigenvector of $B^{-1}A$ corresponding to the root ℓ.

Proof. First let us solve the following problem. Maximize $z'Az$ subject to $z'Bz = 1$. Using the method of Lagrange's multiplier let

$$f(z,\lambda) = z'Az - \lambda(z'Bz-1)$$

where, λ is a Lagrange multiplier. A necessary condition is

$$\partial f/\partial z = 2Az - 2\lambda Bz = 0$$

so

$$Az = \lambda Bz = B\lambda z \Rightarrow B^{-1}Az = \lambda z.$$

So that λ is an eigenvalue of B^{-1} and z is the corresponding eigenvector. Therefore,

$$\max_{z'Bz=1} z'Az = \max_{z'Bz=1} z'(Az) = \max_{z'Bz=1} z'(\lambda Bz)$$

$$= \max_{z'Bz=1} \lambda z'Bz = \max \lambda = \ell .$$

Now we shall show that the following two problems are equivalent:

(i) $\max_{\substack{z\in E_n \\ z\neq 0}} \frac{z'Az}{z'Bz}$, (ii) $\max_{z'Bz=1} z'Az$.

Proof of this is very much like the counterpart proof given in the beginning of Lemma 1.

Let

$$\max_{z\neq 0} \frac{z'Az}{z'Bz} = m_1' \quad \text{and} \quad \max_{z'Bz=1} z'Az = m_2'$$

and suppose m_1' is attained for $z = v_1$ and m_2' is attained for $z = v_2$, i.e.,

$$\frac{v_1'Av_1}{v_1'Bv_1} = m_1' \quad \text{and} \quad v_2'Av_2 = m_2' \quad \text{under} \quad v_2'Bv_2 = 1.$$

Let

$$\bar{v}_1 = \frac{v_1}{(v_1'Bv_1)^{1/2}}$$

then

$$\bar{v}_1'B\bar{v}_1 = \frac{v_1'Bv_1}{v_1'Bv_1} = 1,$$

also

$$\frac{\bar{v}_1'A\bar{v}_1}{\bar{v}_1'Bv_1} = \frac{v_1'Av_1}{v_1'Bv_1} = \lambda.$$

Therefore, $m_1 \leq m_2$. On the other hand, since $v_2'Bv_2 = 1$ and since B is positive definite $v_2 \neq 0$,

$$\frac{v_2'Av_2}{v_2'Bv_2} = \frac{v_2'Av_2}{1} = m_2 ,$$

thus $m_2 \leq m_1$. Hence $m_1 = m_2$.

We shall shortly give a generalization of the preceeding results. Let L be a vector subspace of E_n of dimension q. Let the columns of $C = [c_1, c_2, \ldots, c_q]$ be a basis for L.

<u>Lemma 3</u>. The maximum value of $z'Az/z'z$ over all nonzero $z \in L$ is ℓ, the largest eigenvalue of CC^+A, and is attained when z is any eigenvector of CC^+A corresponding to the root ℓ, where C^+ is the Moore-Penrose generalized inverse of C.

<u>Proof</u>. C is an n x q matrix of rank q thus $(C'C)^{-1}$ exists and one can check that

$$C^+ = (CC')^{-1} C',$$

is the Moore-Penrose generalized inverse of C. Also note that matrices CD and DC have the same eigenvalues. Now let $v \in E_q$, then $z = Cv \in L$, hence

$$\max_{z \in L,\ z \neq 0} \frac{z'Az}{z'z} \geq \max_{v \in E_q,\ v \neq 0} \frac{(Cv)'A(Cv)}{(Cv)'(Cv)}$$

$$= \max_{v \in E_q,\ v \neq 0} \frac{v'(C'AC)v}{v'(C'C)v} .$$

Since $C'C$ is positive definite we can use Lemma 2 and conclude that

$$\max_{v \in E_q, v \neq 0} \frac{v'(C'AC)v}{v'(C'C)v}$$

$=$ largest eigenvalue of $(C'C)^{-1} C'AC$.

$=$ largest eigenvalue of $C^{+}AC$

$=$ largest eigenvalue of $C(C^{+}A)$ (see the remark about CD and DC)

$= \ell$.

To obtain the inequality in the other direction, let $z \in L$. This implies that there exists a $v \in E_q$ such that $z = Cv$. Thus

$$\max_{z \in L,\ z \neq 0} \frac{z'Az}{z'z} \leq \max_{v \in E_q,\ v \neq 0} \frac{(Cv)'A(Cv)}{(Cv)'(Cv)}$$

$=$ largest eigenvalue of $CC^{+}A = \ell$.

Now let z be any eigenvalue of $CC^{+}A$ corresponding to the root ℓ. Then $CC^{+}Az = \ell z$ which imples $z'CC^{+}Az = \ell z'z$ which in turn implies $z'Az/z'z = \ell$ because $z \in L$ implies that $z'CC^{+} = z'$.

Corollary 1. Let H be any $q \times n$ matrix of rank q. Then the maximum value of $z'Az/z'z$ over all nonzero z in E_n satisfying $Hz = 0$ is ℓ, the largest eigenvalue of $(I-H^+H)A$, and is obtained when z is any eigenvector of $(I-H^+H)A$ corresponding to the root ℓ. Here H^+ denotes the Moore-Penrose generalized inverse of H.

Proof. $Hz = 0$ if and only if z belongs to the column space of $I-H^+H$. This is seen as follows:

If $Hz = 0$, then z is in the column space of $I-H^+H$, i.e., there exists a w such that $(I-H^+H)w = z$. Set $w = z$ then $(I-H^+H)z = z - H^+Hz = z - H^+(Hz) = z - H^+(0) = z$. On the other hand, if z is in the column space of $I-H^+H$ then $Hz = H(I-H^+H)w = Hw - HH^+Hw = Hw - Hw = 0$. Now the proof follows from Lemma 3.

Lemma 4. Let B be a positive definite matrix of order n. Then the maximum value of $z'Az/z'Bz$ over all nonzero z in L is ℓ, the largest eigenvalue of $C(C'BC)^{-1}C'A$, and is attained when z is any eigenvector of $C(C'BC)^{+}C'A$ corresponding to the root ℓ.

Proof. An argument similar to the one used in the proof of Lemma 3 gives us

$$\max_{z \in L,\, z \neq 0} \frac{z'Az}{z'Bz} = \max_{v \in E_q,\, v \neq 0} \frac{(Cv)'A(Cv)}{(Cv)'B(Cv)} .$$

By an earlier result one gets

$$\max_{v \in E_q,\ v \neq 0} \frac{v'C'ACv}{v'C'BCv} = \text{largest eivenvalue of } [(C'BC)^+ C'AC]$$

$$= \text{largest eigenvalue of } [C(C'BC)^+ C'A] = \ell$$

(recall the argument about CD and DC).

Now let z be any eigenvector of $C(C'BC)^+ C'A$ corresponding to the root ℓ. Then

$C(C'BC)^+ C'Az = \ell z$ which implies that

$z'BC(C'BC)^+ C'Az = \ell z = \ell z'Bz$, which implies that

$z'Az/z'Bz = \ell$, since $z \in L$ implies that

$$z'BC(C'BC)^+ C' = z'.$$

This latter claim is seen as follows. Since $z \in L$ then there exists a w such that $Cw = z$, i.e., z is a linear combination of the columns of C which generate L. Then

$$z'BC(C'BC)^+ C' = w'C'BC(C'BC)^+ C' = w'C' = (Cw)' = z'.$$

<u>Lemma 5</u>. The Moore-Penrose of X is given by $(X'X)^+X'$ where A^+ denotes the Moore-Penrose of the matrix A.

<u>Proof</u>. By definition K is the Moore-Penrose generalized inverse of A if $AKA = A$, $KAK = K$, $(AK)' = AK$ and $(KA)' = KA$. Therefore, we shall check these four conditions for X^+. In what follows we use the following well known facts:

F_1: $X(X'X)^-X'$ is symmetric and $X(X'X)^+X' = X(X'X)^-X'$ where $(X'X)^-$ is any generalized inverse of $(X'X)$.

F_2: $X(X'X)^-X'X = X$ and $X'X(X'X)^-X' = X'$.

(i) $XX^+X = X(X'X)^+X'X = X(X'X)^-X'X = X$,

(ii) $X^+XX^+ = (X'X)^+X'X(X'X)^+X' = (X'X)^+X'X(X'X)^-X' = (X'X)^+X'$.

(iii) $(XX^+)' = (X(X'X)^+X')' = (X(X'X)^-X')' = X(X'X)^- = X(X'X)^+X'$
$= XX^+$,

(iv) $(X^+X)' = ((X'X)^+X'X)' = (X'X)^+$ since $(X'X)^+$ is the Moore-Penrose inverse of $X'X$ and thus $(X'X)^+X'X$ is symmetric.

<u>Lemma 6</u>. The Moore-Penrose generalized inverse of X' is $(X^+)'$.

<u>Proof</u>.

(i) $X'(X^+)'X' = [XX^+X]' = [X]' = X'$,

(ii) $(X^+)'X'(X^+)' = [X^+XX^+]' = [X^+]'$,

(iii) $[X'(X^+)']' = [(X^+X)']' = [X^+X]' = X'(X^+)'$,

(iv) $[(X^+)'X']' = [(XX^+)']' = [XX^+]' = (X^+)'X'$.

<u>Lemma 7</u>. If X^+ is the Moore-Penrose of X, then $XX^+(X^+)'$ $= (X^+)'$.

<u>Proof</u>. From Lemma 5 $X^+ = (X'X)^+X'$. Thus

$$\begin{aligned} XX^+(X^+)' &= X(X'X)^+X' \, [(X'X)^+X']' \\ &= X(X'X)^+ \, [X'X)^+X'X]' \\ &= X(X'X)^+ \, [X'X[(X'X)^+]'] \\ &= X(X'X)^+ \, [X'X[X'X)']^+] \quad \text{by Lemma 6} \end{aligned}$$

$= X(X'X)^+ [X'X(X'X^+]$ by a property of Moore-Penrose

$= X(X'X)^+$ generalized inverse

$= [[(X'X)^+]'X']' = [[(X'X)']^+X']'$

$= [(X'X)^+X']' = (X^+)'$.

Lemma 8. The set of nonzero eigenvalues of DC coincides with the set of nonzero eigenvalues of CD.

Proof. Let Λ_1 and Λ_2 be the set of nonzero eigenvalues of DC and CD respectively.

If $(DC)x = \lambda x \Rightarrow C(DC)x = \lambda Cx$

$\Rightarrow CD(Cx) = \lambda(Cx) = CDy = \lambda y$,

so if λ is an eigenvalue of DC it is an eigenvalue of CD,

$\Rightarrow \Lambda_1 \subset \Lambda_2$. Similarly, $\Lambda_2 \subset \Lambda_1$. Thus $\Lambda_1 = \Lambda_2$.

Proof of Theorem 1.

(1). $$\max_{c \in L} \frac{(c'\hat{\beta} - c'\beta)^2}{c'(X'X)^- c} = \max_{a \in \mathcal{L}[(X')^+ C]} \frac{(a'X\hat{\beta} - a'X\beta)^2}{a'X(X'X)^- X'a} .$$

Reason. $c'\beta$ is estimable $\Rightarrow \exists$ an a such that $c' = a'X$. But $c \in L \Rightarrow c' = \Sigma t_i c_i'$, $c_i' = b_i'X$, c_i's were chosen $\Rightarrow c' = \Sigma t_i b_i' X = [\Sigma t_i b_i']X = a'X$ so a' is a linear combination of b_i's. But from $c_i' = b_i'X \Rightarrow [\Sigma t_i b_i']X = a'X$ so a' is a combination of b_i's. But from $c_i' = b_i'X \Rightarrow X'b_i = c_i$ or $b_i = (X')^+ c_i$.

(2). $\dfrac{(a'X\hat{\beta}-a'X\beta)^2}{a'X(X'X)^{-}X'a} = \dfrac{(a'X(X'X)^{-}X'Y-a'X\beta)^2}{a'X(X'X)^{-}X'a} = \dfrac{(a'X(X'X)^{+}X'Y-a'X\beta)^2}{a'X(X'X)^{+}X'a}$

$= \dfrac{(a'XX^{+}Y-a'X\beta)^2}{a'XX^{+}a}$. Reason. See Lemmas 5 and 6.

(3) $\dfrac{(a'XX^{+}Y-a'X\beta)^2}{a'XX^{+}a} = \dfrac{(a'Y-a'X\beta)^2}{a'a} = \dfrac{a'(Y-X\beta)(Y-X\beta)'a}{a'a}$.

Reason. Since $a \in \mathcal{L}[(X')^{+}C] \Rightarrow a \in$ column space of $(X')^{+}$, i.e., $\exists$ an f such that $a = (X')^{+}f = (X^{+})'f$ $\Rightarrow a' = f'X^{+}$. Thus $a'XX^{+} = f'X^{+}XX^{+} = f'X^{+} = a'$.

(4) From (1) and (3)

$$\max_{c\in L} \frac{(c'\hat{\beta}-c'\beta)^2}{c'(X'X)^{-}c} = \max_{a\in\mathcal{L}[(X')^{+}C]} \frac{a'(Y-X\beta)(Y-X\beta)'a}{a'a}$$

$$= \max_{a\in\mathcal{L}[(X')^{+}C]} \frac{a'Aa}{a'a}, \quad A = (Y-X\beta)(Y-X\beta)'$$

= largest eigenvalue of $[(X')^{+}C][(X')^{+}C]^{+}(Y-X\beta)(Y-X\beta)'$ by Lemma 3.

= largest eigenvalue of $(Y-X\beta)'[(X')^{+}C][(X')^{+}C]^{+}(Y-X\beta)$ by Lemma 8, but this is a scalar,

$= (Y-X\beta)'[(X')^{+}C][(X')^{+}C]^{+}(Y-X\beta) = Q_1$ which is a quadratic in $(Y-X\beta) \sim N(0, \sigma^2 I)$.

The claim is that $Q \sim \sigma^2\chi^2(q)$. This will be the case if we prove that $[(X')^{+}C][(X')^{+}C)^{+}$ is idempotent and its rank is q. The idempotency is obvious since in general $(BB^+)(BB^+) = BB^+BB^+ = BB^+$. We shall now show that rank $[(X')^+C][(X')^+C]^+ = q$. This can be seen as follows:

$$r[(X')^+C][(X')^+C]^+ \leq r[(X')^+C] \leq r[C] = q,$$

on the other hand,

$$r[(X'X)^+C][(X')^+C]^+ \geq r[(X')^+C][(X')^+C]^+[(X')^+C]$$

$$\geq r[(X')^+C] = r[(X^+)'C] = r[[(X'X)^+X']'C] = r[X[(X'X)^+]'C]$$

$$\geq r[X'X[(X'X)^+]'C] = r[X'X[(X'X)^+]'X'K] \quad \text{since } C = X'K$$

$$= r[[X(X'X)^+X'X]'K] = r[[X)X'X)^-X'X)'K] = r[X'X(X'X)^-X'K]$$

$$= r[X'K] = r[C] = q.$$

The proof of Theorem 1 will be complete if we show that Q_1 and Q_2 are independent where,

$$Q_2 = (n-r)\,\hat{\sigma}^2 = Y'(I-X)X'X)^-X')Y$$

$$= (Y-X\beta)'(I-X(X'X)^-X')(Y-X\beta).$$

It is sufficient to prove that

$$[I-X(X'X)^-X'][(X')^+C][(X')^+C]^+ = [I-X(X'X)^-X'][(X^+)'C][(X^+)'C]^+ = 0,$$

by Lemma 6

$$\text{LHS} = [(X^+)'C][(X^+)'C]^+ - X(X'X)^-X'(X^+)'C[(X^+)'C]^+$$

$$= [(X^+)'C][(X^+)'C]^+ - X(X'X)^+X'(X^+)'C[(C^+)'C]^+$$

$$= [(X^+)'C][(X^+)'C]^+ - XX^+(X^+)'C[(X^+)'C]^+$$

$$= [(X^+)'C][(X^+)'C] - (X^+)'C[(X^+)'C] = 0.$$

The relation of the S-method or S-intervals for Ψ in L and the standard F-test of the hypothesis

$$H_0: \Psi_1 = \Psi_2 = \ldots = \Psi_q = 0$$

is stated in

<u>Theorem 2</u>. Under Ω the α-level F-test of H_0 will accept H_0 if and only if for all Ψ in L the intervals (1) in Theorem 1 cover zero.

BIBLIOGRAPHY

(1) Aitchison, J. (1964). Confidence-region tests. J. Roy. Statist. Soc. Ser. B 26 462-476.

(2) _____. (1965). Likelihood-ratio and confidence region tests. J. Roy. Statist. Soc. Ser. B 27 245-250.

(3) Brown, L., "The Conditional Level of the t Test," Annals of Mathematical Statistics, 38 (August 1967), 1068-71.

(4) Buehler, R.J., "Some Validty Criteria for Statistical Inferences," Annals of Mathematical Statistics, 30 (December 1959), 845-63.

(5) _____ and Fedderson, A. P., "Note on a Conditional Property of Student's t," Annals of Mathematical Statistics, 34 (Sepetmber 1963), 1098-100.

(6) Cohen, A., "Improved Confidence Intervals for the Variance of a Normal Distribution," Journal of the American Statistical Association, 67 (June 1972), 382-7.

(7) Cornfield, J., "The Bayesian Outlook and Its Application," Biometrics, 24 (December 1969), 617-42.

(8) Duncan, D.B., (1951). A significance test for differences between ranked treatments in an analysis of variance. Virginia J. Sci. 2 171-189.

(9) _____, (1955). Multiple range and multiple F-tests. Biometrics 13 164-176.

(10) Dunnett, Charles W., (1955). A multiple comparison procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50 1096-1121.

(11) Edwards, W., Lindman, H. and Savage, L.J., "Bayesian Statistical Inference for Psychological Research," Psychological Review, 70 (May 1963), 193-242.

(12) Fisher, R.A., "On a Test of Significance in Pearson's Biometrika Tables (No. 11)," Journal of the Royal Society, Ser. B. 18 (1956), 56-60.

(13) _____, Statistical Methods and Scientific Inference, Second ed., New York: Hafner Publishing Co., 1959.

(14) Freedman, D.A. and Purves, R.A., "Bayes' Method for Bookies," Annals of Mathematical Statistics, 40 (August 1969), 1177-86.

(15) Gabriel, K.R.. (1964). A procedure for testing the homogeneity of all sets of means in analysis of variance. Biometrics 20 459-477.

(16) _____, (1966). Simultaneous test procedures for multiple comparisons on categorical data. J. Amer. Statist. Assoc. 61 1081-1096.

(17) _____, (1968). Simultaneous test procedures in multivariate analysis of variance.Biometrika 55. 489-504.

(18) _____, (1968). On the relation between union intersection and likelihood ratio tests. To appear in S.N. Roy memorial volume (R.C. Bose, ed.).

(19) Johnson, D.E., (1973). A derivati of Scheffe's S-method by maximizing a quadratic form. The American Statistician 27, 27-29.

(20) Keuls, M. (1952). The use of the 'studentized range' in connection with an analysis of variance. Euphytica 1 112-122.

(21) Knight, William (1965). A lemma for multiple inference, Ann. Math. Statist. 36 1873-1874.

(22) Krishnaiah, P.R., (1964). Multiple comparison tests in multivariate case. U.S.A.F. Office of Aerospace Research, ARL- 64-124.

(23) _____, (1965). On the simultaneous Anova and Manova tests. Ann. Inst. Statist. Math. 17 35-53.

(24) Lehmann, E.L., (1957). A theory of some multiple decision problems, I. Ann. Math. Statist. 28 1-25.

(25) _____, (1957). A theory of some multiple decision problems, II. Ann. Math. Statist. 28 547-572.

(26) Miller, R.G., Jr., Simultaneous Statistical Inference, New York: McGraw Hill Book Co., 1966.

(27) Nemenyi, Peter (1963). Distribution-free Multiple Comparisons. (mimeographed).

(28) Newman, D. (1939). The distribution of the range in samples from the normal population, expressed in terms of an independent estimate of standard deviation. Biometrika 31 20-30.

(29) Olshen, R.A., (1973). The conditional level of the F-test. J. Amer. Statist. Assoc. 68, 692-698.

(30) Rao, C.R., Jr., Linear Statistical Inference and Its Applications, New York: John Wiley and Sons, Inc. 1965.

(31) Roy, S.N., (1957). Some Aspects of Multivariate Analysis. Wiley, New York.

(32) _____, (1961). A Survey of some recent results in normal multivariate confidence bounds. Proc. of the 33rd Ses. of the Interna. Statist. Inst., Paris.

(33) _____, and Bose, R.C., (1953). Simultaneous confidence interval estimation. Ann. Math. Statist. 24 513-536.

(34) _____, and Gnanadesikan, R., (1959). Some contributions to Anova in one or more dimensions: I. Ann. Math. Statist. 30 304-317.

(35) _____, and Shrivastava, J.N., (1961). Inference on treatment effects and design of experiments in relation to such inference. Univ. of North Carolina, Inst. of Statist., Mimeo Series, No. 274.

(36) Scheffe, H., (1953). A method for judging all contrasts in the analysis of variance. Biometrika 40 87-104.

(37) _____, "A method for judging all contrasts in the analysis of variance," Biometrika, 40 (June 1953), 87-104; Corrigenda, Biometrkia, 56 (March 1969), 229.

(38) _____, The Analysis of Variance, New York: John Wiley and Sons, Inc., 1959.

(39) _____, "Multiple Testing Versus Multiple Estimation. Improper Confidence Sets. Estimation of Directions and Ratios," Annals of Mathematical Statistics, 41 (February 1970), 1-29.

(40) Sclove, S.L., Morris, C. and Radhakrishnan, R., "Non-Optimality of Preliminary-Test Estimators for the Mean of a Multivariate Normal Distribution," Annals of Mathematical Statistics, 43 (October 1972), 1481-90.

(41) Stein, C., "Approximation of Prior Measures by Probability Measures," (Multilithed) Notes of Lecture 2 of the Wald Lectures, 1961.

(42) _____, "Inadmissibility of the Usual Estimator for the Variance of a Normal Distribution with Unknown Mean," Annals of the Institute of Statistical Mathematics, 16 (1964), 155-60.

(43) _____, "An Approach to the Recovery of Inter-Block Information in Balanced Incomplete Block Designs," in F.N. David, ed., Festschrift for J. Neyman, New York: John Wiley and Sons, Inc., 1966.

(44) Tukey, J.W., (1951). Quick and dirty methods in statistics. Part II. Simple analyses for standard designs. Proc. Fifth Annual Conven. Amer. Soc. Quality Control, 189-197.

(45) _____, (1953). The problem of multiple comparisons. Unpublished manuscript.

(46) Wallace, D.L., "Conditional Confidence Level Properties," Annals of Mathematical Statistics, 38 (December 1959), 864-76 (corrected in [16, p. 8]).