*The Annals of Statistics*
1977, Vol. 5, No. 5, 922–931

# SOME VARIATIONAL RESULTS AND THEIR APPLICATIONS IN MULTIPLE INFERENCE

BY D. R. JENSEN[1] AND L. S. MAYER[2]

*Virginia Polytechnic Institute and Princeton University*

Let $(\mathbf{M}, \mathbf{T}, \mathbf{S})$ be random matrices such that $\mathbf{M}$ and $\mathbf{S}$ are Hermitian positive definite almost everywhere. Let $\mathbf{M}_{(t)} = [m_{ij}; 1 \leqq i, j \leqq t]$, $\mathbf{S}_{(t)} = [s_{ij}; 1 \leqq i, j \leqq t]$ and $\mathbf{T}_{(r,s)} = [t_{ij}; 1 \leqq i \leqq r, 1 \leqq j \leqq s]$, and define $Q(r, s) = P[G((\mathbf{M}_{(r)})^{-\frac{1}{2}}\mathbf{T}_{(r,s)}(\mathbf{S}_{(s)})^{-\frac{1}{2}}) \leqq c]$ for some $G$ belonging to the class $\mathscr{G}$ of monotone unitarily invariant functions. The main result is that, for any $c$ and $G \in \mathscr{G}$, $Q(r, s)$ is a decreasing function of $r$ and $s$. Applications yield simultaneous confidence bounds for a variety of multivariate and multiparameter problems.

**1. Introduction.** Invariance considerations in multivariate and multiparameter testing problems often suggest that test functions should depend only on the singular values of certain random matrices. This study develops a natural connection between a Sturmian type separation of singular values and stochastic ordering for members of the class of monotone unitarily invariant test functions, including functions commonly used in multivariate analysis such as the largest characteristic root, the trace, and the ratio of determinants.

The principal results are developed in Section 2 as findings in the study of random matrices. Applications of the results are given in Section 3; each is concerned with the simultaneous inferences associated with a given test procedure and, in the spirit of Scheffé's bounds, each supports the unlimited use of significance tests at a type 1 error rate not exceeding $\alpha$. In contrast to the emphasis found in much of the literature on multiple inference, our results and examples underscore the fact that the methods are strictly variational in character and therefore apply regardless of the underlying distributions.

**2. Some variational results.** Designate by $\mathbb{R}^m$ the $m$-dimensional real Euclidean space, and by $\mathscr{F}_{m\times n}$ the linear space of matrices of order $(m \times n)$ defined over the complex numbers, such that $m \leqq n$. Special arrays are the identity matrix $\mathbf{I}_n$ of order $n$, the unit vector $\mathbf{e}_i$ having unity in the $i$th location and zeros elsewhere, and the matrix $\mathbf{E}_{ij} \in \mathscr{F}_{m\times n}$ whose only nonzero entry is unity in the $(i, j)$ position. Each $\mathbf{A} \in \mathscr{F}_{m\times n}$ admits the *singular decomposition* $\mathbf{A} = \mathbf{U\Lambda V}$ such that $\mathbf{U}(m \times m)$ and $\mathbf{V}(n \times n)$ are unitary and $\mathbf{\Lambda}$ has the form $\mathbf{\Lambda} = [\mathbf{D}_\alpha, \mathbf{0}]$, where $\mathbf{D}_\alpha$ is the diagonal matrix $\mathbf{D}_\alpha = \mathrm{diag}(\alpha_1, \cdots, \alpha_m)$ containing the *singular values*

Received January 1976; revised January 1977.

[1] Supported in part by Contract N00014-72-A-0136-0003 from the U.S. Office of Naval Research.

[2] Supported in part by Grant 26692 from the National Institute of Mental Health.

*AMS* 1970 *subject classifications.* Primary 62E10, 62H05; Secondary 62H15.

*Key words and phrases.* Monotone unitarily invariant functions, separation of singular values, stochastic ordering, simultaneous confidence bounds, symmetric gauge functions.

of $\mathbf{A}$, i.e., the nonnegative square roots of the eigenvalues of $\mathbf{AA}^*$, $\mathbf{A}^*$ being the conjugate transpose of $\mathbf{A}$. We define a partial ordering (due to Charles Loewner) on the space of Hermitian matrices by letting $\mathbf{A} \leqq \mathbf{B}$ whenever $\mathbf{B} - \mathbf{A}$ is the matrix of a positive semidefinite Hermitian form.

Inequalities for the eigenvalues of reduced Hermitian forms are provided by the Sturmian separation theorem (cf. Bellman (1970), page 117). Similar results hold for singular values. Beginning with a matrix $\mathbf{A} = [a_{ij}] \in \mathscr{F}_{m\times n}$ and a Hermitian matrix $\mathbf{H}(t \times t)$, we define the matrices $\mathbf{A}_{(r,s)} = [a_{ij}; 1 \leqq i \leqq r, 1 \leqq j \leqq s]$ and $H_{(r)} = [h_{ij}; 1 \leqq i, j \leqq r]$; we designate by $\{S_1(\mathbf{A}) \geqq \cdots \geqq S_m(\mathbf{A})\}$ the ordered singular values of $\mathbf{A}$; and we denote by $\{C_1(\mathbf{H}) \geqq \cdots \geqq C_t(\mathbf{H})\}$ the ordered characteristic values of $\mathbf{H}$. Using these matrices we give two separation theorems for singular values; the second assumes a central role in our applications.

THEOREM 1. *Let $S_k(\mathbf{A})$ be the* $k$th *largest singular value of* $\mathbf{A}$. *Then*

$$S_{k+1}(\mathbf{A}_{(r+1,s)}) \leqq S_k(\mathbf{A}_{(r,s)}) \leqq S_k(\mathbf{A}_{(r+1,s)}) \leqq S_k(\mathbf{A}_{(r+1,s+1)}) .$$

PROOF. The first two inequalities follow upon (i) fixing $s$, (ii) representing $\mathbf{A}_{(r+1,s+1)}\mathbf{A}^*_{(r+1,s+1)}$ in block-partitioned form with $\mathbf{A}_{(r,s)}\mathbf{A}^*_{(r,s)}$ in the upper left corner, (iii) applying the Sturmian separation theorem, and (iv) taking square roots to get singular values. The final inequality follows similarly upon fixing $r+1$ and applying the foregoing steps to $\mathbf{A}^*_{(r+1,s+1)}\mathbf{A}_{(r+1,s+1)}$.

THEOREM 2. *Let $\mathbf{H}(n \times n)$ be a positive definite Hermitian matrix and let $\mathbf{H}^{\frac{1}{2}}$ be its Hermitian square root. Then for any matrix $\mathbf{T} \in \mathscr{F}_{m\times n}$,*

$$S_k(\mathbf{T}_{(m,s)}(\mathbf{H}_{(s)})^{-\frac{1}{2}}) \leqq S_k(\mathbf{T}_{(m,s+1)}(\mathbf{H}_{(s+1)})^{-\frac{1}{2}}) .$$

PROOF. Let $\mathbf{Q}_t = \mathbf{L}_t(\mathbf{L}_t'\mathbf{HL}_t)^{-1}\mathbf{L}_t'$ where $\mathbf{L}_t' = [\mathbf{I}_t, \mathbf{0}]$ is of order $(t \times n)$. Because $\mathbf{T}_{(m,t)} = \mathbf{TL}_t$ and $\mathbf{H}_{(t)} = \mathbf{L}_t'\mathbf{HL}_t$, a direct computation yields

$$\mathbf{T}_{(m,t)}(\mathbf{H}_{(t)})^{-1}\mathbf{T}'_{(m,t)} = \mathbf{TQ}_t\mathbf{T}'$$

and thus it suffices to compare the matrices $\mathbf{Q}_s$ and $\mathbf{Q}_{s+1}$. But the blocks of $\mathbf{Q}_{s+1}$ consisting entirely of zeros are also zero elements of $\mathbf{Q}_s$, and thus we need consider only the upper left submatrix, namely $\mathbf{I}_{s+1}(\mathbf{H}_{(s+1)})^{-1}\mathbf{I}_{s+1}$, of $\mathbf{Q}_{s+1}$. Now for any matrix $\mathbf{U}^* = [\mathbf{U}_1^*, \mathbf{U}_2^*]$ and a positive definite Hermitian matrix $\mathbf{A} = [\mathbf{A}_{ij}; 1 \leqq i, j \leqq 2]$ partitioned conformably, we infer (compare Anderson (1958), pages 28, 42) that

$$\mathbf{U}^*\mathbf{A}^{-1}\mathbf{U} = \mathbf{U}_1^*\mathbf{A}_{11}^{-1}\mathbf{U}_1 + (\mathbf{U}_2 - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{U}_1)^*\mathbf{A}_{22.1}^{-1}(\mathbf{U}_2 - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{U}_1)$$

where $\mathbf{A}_{22.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$. It follows that $\mathbf{U}^*\mathbf{A}^{-1}\mathbf{U} \geqq \mathbf{U}_1^*\mathbf{A}_{11}^{-1}\mathbf{U}_1$ because $\mathbf{A}_{22.1}$ is positive definite. In particular, letting $\mathbf{A} = \mathbf{H}_{(s+1)}$ and $\mathbf{U} = \mathbf{I}_{s+1}$, we have shown that $(\mathbf{H}_{(s+1)})^{-1} \geqq \operatorname{diag}((\mathbf{H}_{(s)})^{-1}, \mathbf{0})$ and hence that $\mathbf{Q}_s \leqq \mathbf{Q}_{s+1}$, which implies $\mathbf{TQ}_s\mathbf{T}' \leqq \mathbf{TQ}_{s+1}\mathbf{T}'$, which, by Theorem 3, page 117 of Bellman (1970), implies $C_k(\mathbf{TQ}_s\mathbf{T}') \leqq C_k(\mathbf{TQ}_{s+1}\mathbf{T}')$. The theorem now follows from successive use of the fact that $S_k(\mathbf{T}_{(m,t)}(\mathbf{H}_{(t)})^{-\frac{1}{2}}) = [C_k(\mathbf{TQ}_t\mathbf{T}')]^{\frac{1}{2}}$, for $t = s$ and $t = s+1$.

We are concerned with stochastic ordering for a class of monotone invariant functions on $\mathscr{F}_{m\times n}$. A function $G: \mathscr{F}_{m\times n} \to \mathbb{R}^1$ is said to be *unitarily invariant* if, for each unitary pair $\mathbf{U}(m \times m)$ and $\mathbf{V}(n \times n)$, $G(\mathbf{UAV}) = G(\mathbf{A})$. We call such a function *monotone* if, as a function of the invariants, it is monotonic in each argument. As the maximal invariants under $\mathbf{A} \to \mathbf{UAV}$ are the singular values of $\mathbf{A}$, it suffices to generate the required functions on $\mathscr{F}_{m\times n}$ by means of a suitably endowed class of functions on $\mathbb{R}^m$. Details of this construction follow.

DEFINITION 1. Let $\Phi$ be the class of functions mapping $\mathbb{R}^m \to \mathbb{R}^1$ such that, if $\phi$ is in $\Phi$, then

(i) $\phi(x_1, \cdots, x_m) \geqq 0$;

(ii) $\phi(\varepsilon_1 x_{i_1}, \cdots, \varepsilon_m x_{i_m}) = \phi(x_1, \cdots, x_m)$, where $\varepsilon_i = \pm 1$, $1 \leqq i \leqq m$, and $(i_1, i_2, \cdots, i_m)$ is any permutation of $(1, 2, \cdots, m)$;

(iii) If $|x_i| \leqq |y_i|$, $1 \leqq i \leqq m$, then $\phi(x_1, \cdots, x_m) \leqq \phi(y_1, \cdots, y_m)$ and, if $0 \leqq u < v$, then $\phi(u, 0, \cdots, 0) < \phi(v, 0, \cdots, 0)$;

(iv) $\phi(x, 0, \cdots, 0) = |x|$.

Further let $\Phi_0$ be the subclass of functions in $\Phi$ which have the additional properties

(v) $\phi(cx_1, \cdots, cx_m) = |c|\phi(x_1, \cdots, x_m)$;

(vi) $\phi(x_1 + y_1, \cdots, x_m + y_m) \leqq \phi(x_1, \cdots, x_m) + \phi(y_1, \cdots, y_m)$.

REMARK 1. For any $\psi(\cdot)$ having properties (i)—(iii) and any even function $\xi: \mathbb{R}^1 \to \mathbb{R}^1$ which is strictly increasing on $\mathbb{R}_+{}^1 = [0, \infty)$ such that $\xi(0) = 0$, the exterior composition $(\psi \wedge \xi)(x_1, \cdots, x_m) \equiv \psi(\xi(x_1), \cdots, \xi(x_m))$ also has properties (i)—(iii).

REMARK 2. Although not completely descriptive, we call $\Phi$ the class of *symmetric monotone functions*. Each $\phi \in \Phi$ essentially is determined by properties (i)—(iii); property (iv) avoids the need to consider monotonic functions of monotone functions in our applications. To support these assertions consider any $\psi: \mathbb{R}^m \to \mathbb{R}^1$ having properties (i)—(iii). To each such function we define an *associate function* $\psi_1: \mathbb{R}_+{}^1 \to \mathbb{R}_+{}^1$ by $\psi_1(|x|) = \psi(x, 0, \cdots, 0)$, which is strictly increasing by property (iii) and thus is invertible. Upon replacing $\psi$ by $\phi = \psi_1{}^{-1} \circ \psi$, we find that property (iv) holds and thus $\phi \in \Phi$, i.e., $\phi(x, 0, \cdots, 0) = (\psi_1{}^{-1} \circ \psi)(x, 0, \cdots, 0) = (\psi_1{}^{-1} \circ \psi_1)(|x|) = |x|$. The exterior composition mentioned earlier can be given property (iv) in a similar manner by letting $\phi = \xi_1{}^{-1} \circ \psi_1{}^{-1} \circ (\psi \wedge \xi)$, in which case $\phi(x, 0, \cdots, 0) = (\xi_1{}^{-1} \circ \psi_1{}^{-1} \circ (\psi_1 \wedge \xi_1))(|x|) = |x|$.

Turning to functions on $\mathscr{F}_{m\times n}$, let $\sigma: \mathscr{F}_{m\times n} \to \mathbb{R}^m$ be the map which associates with $\mathbf{A} \in \mathscr{F}_{m\times n}$ its singular values $[S_1(\mathbf{A}), \cdots, S_m(\mathbf{A})]$.

DEFINITION 2. Let $\mathscr{G}$ be the class of monotone unitarily invariant functions $G: \mathscr{F}_{m\times n} \to \mathbb{R}^1$ generated by compositions of the type $G = \phi \circ \sigma$ such that $\phi \in \Phi$, i.e., $\mathscr{G} = \{G \mid G = \phi \circ \sigma, \phi \in \Phi\}$. Let $\mathscr{G}_0$ be the subclass of functions in $\mathscr{G}$ of the form $G = \phi \circ \sigma$ such that $\phi \in \Phi_0$, i.e., $\mathscr{G}_0 = \{G \mid G = \phi \circ \sigma, \phi \in \Phi_0\}$.

REMARK 3. In addition to its monotonicity and unitary invariance, $\mathscr{G}$ consists of functions standardized such that $G(u\mathbf{E}_{ij}) = |u|$ for each $G \in \mathscr{G}$. In particular, if $\mathbf{A}$ has unit rank and the nonvanishing singular value $S_1(\mathbf{A}) = \alpha$, then $G(\mathbf{A}) = \alpha$; this is a consequence of the definition of $G$ and property (iv) of Definition 1.

REMARK 4. Apart from standardization, the class $\Phi_0$ consists of the symmetric gauge functions on $\mathbb{R}^m$, and $\mathscr{G}_0$ is the class of all standardized unitarily invariant norms on $\mathscr{F}_{m\times n}$; von Neumann (1937) showed that these classes generate each other. For further comments see Schatten (1970).

REMARK 5. Examples of functions in $\Phi_0$ are the $l_p$ norms $\phi_{(p)}(x_1, \cdots, x_m) = (\sum_{i=1}^m |x_i|^p)^{1/p}$, $1 \leqq p < \infty$, and the functions $\phi_{[s]}(x_1, \cdots, x_m) = \sum_{i=1}^s x_{(i)}$, $1 \leqq s \leqq m$, where $\{x_{(1)} \geqq x_{(2)} \geqq \cdots \geqq x_{(m)}\}$ are the ordered values of $\{|x_1|, |x_2|, \cdots, |x_m|\}$. These functions include the $l_\infty$ norm $\phi_{[1]}(x_1, \cdots, x_m) = \max\{|x_1|, \cdots, |x_m|\}$, which generates the largest-root statistic of S. N. Roy in multivariate analysis, and the Euclidean norm $\phi_{(2)}(x_1, \cdots, x_m) = (x_1^2 + \cdots + x_m^2)^{\frac{1}{2}}$, which generates the Lawley–Hotelling trace statistic.

REMARK 6. Examples of functions in $\Phi$ but not $\Phi_0$ are $\phi(x_1, \cdots, x_m) = [\prod_{i=1}^m (1 + x_i^2) - 1]^{\frac{1}{2}}$ and

$$\phi(x_1, \cdots, x_m) = \{\textstyle\sum_{i=1}^m x_i^2/(1 + x_i^2)/[1 - \sum_{i=1}^m x_i^2/(1 + x_i^2)]\}^{\frac{1}{2}} .$$

The first of these generates the likelihood ratio statistic and the second another trace statistic used in the analysis of multivariate linear models.

Against this background we now investigate the stochastic ordering of functions $G \in \mathscr{G}$ which follows as a consequence of the almost sure ordering of the singular values of their random arguments. More precisely, let $\mathscr{F} = \mathscr{F}_{m\times m} \times \mathscr{F}_{m\times n} \times \mathscr{F}_{n\times n}$; suppose $(\mathbf{M}, \mathbf{T}, \mathbf{S}) \in \mathscr{F}$; and let $(\mathscr{F}, \mathscr{B}(\mathscr{F}), P)$ be a probability space such that $\mathbf{M}$ and $\mathbf{S}$ are Hermitian positive definite a.e.

THEOREM 3. *For each $G \in \mathscr{G}$ and each fixed $c > 0$ define*

$$Q(r, s) = P[G((\mathbf{M}_{(r)})^{-\frac{1}{2}}\mathbf{T}_{(r,s)}(\mathbf{S}_{(s)})^{-\frac{1}{2}}) \leqq c] ; \qquad 1 \leqq r \leqq m, 1 \leqq s \leqq n$$

*such that* $\mathbf{M}$ *and* $\mathbf{S}$ *are positive definite* a.e. *Then $Q(r, s)$ is a decreasing function of $r$ and $s$.*

PROOF. Designate by $S_k(r, s)$ the $k$th largest singular value of $[(\mathbf{M}_{(r)})^{-\frac{1}{2}}\mathbf{T}_{(r,s)}(\mathbf{S}_{(s)})^{-\frac{1}{2}}]$. Theorem 2 applied twice assures that, for almost all $(\mathbf{M}, \mathbf{T}, \mathbf{S}) \in \mathscr{F}$,

$$S_k(r, s) \leqq \{S_k(r + 1, s), S_k(r, s + 1)\} \leqq S_k(r + 1, s + 1)$$

where each of the central expressions is understood to satisfy both inequalities. Because $G$ is a monotonic increasing function of the singular values of its argument, it follows that

$$\{G((\mathbf{M}_{(r)})^{-\frac{1}{2}}\mathbf{T}_{(r,s)}(\mathbf{S}_{(s)})^{-\frac{1}{2}}) \leqq c\} \Rightarrow \{G((\mathbf{M}_{(r')})^{-\frac{1}{2}}\mathbf{T}_{(r',s')}(S_{(s')})^{-\frac{1}{2}}) \leqq c\}$$

for all $r' \leqq r$ and $s' \leqq s$. The fact that measures are monotone with respect to set inclusion establishes the result.

A version of Theorem 3 directed specifically toward our applications is the following.

COROLLARY 1. *Corresponding to each* $G \in \mathscr{G}$ *let* $c_\alpha > 0$ *be a constant such that* $P[G(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}}) \leqq c_\alpha] = 1 - \alpha$. *Then*

$$P[\sup_{(\mathbf{H},\mathbf{C}) \in \mathscr{F}_0} G((\mathbf{H}^*\mathbf{M}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}^*\mathbf{T}\mathbf{C}(\mathbf{C}^*\mathbf{S}\mathbf{C})^{-\frac{1}{2}}) \leqq c_\alpha] = 1 - \alpha$$

*where* $\mathscr{F}_0 = \{\mathscr{F}_{m\times h} \times \mathscr{F}_{n\times c};\ 1 \leqq h \leqq m,\ 1 \leqq c \leqq n\}$ *such that* $\mathbf{H}$ *and* $\mathbf{C}$ *are of rank* $h$ *and* $c$, *respectively.*

PROOF. It suffices to show that

$$\sup_{\mathscr{F}_0} G((\mathbf{H}^*\mathbf{M}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}^*\mathbf{T}\mathbf{C}(\mathbf{C}^*\mathbf{S}\mathbf{C})^{-\frac{1}{2}}) = G(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}}) .$$

If necessary let $\mathbf{E} = [\mathbf{H}, \mathbf{F}]$ and $\mathbf{B} = [\mathbf{C}, \mathbf{D}]$ be the full-rank completions of $\mathbf{H}$ and $\mathbf{C}$ of orders $(m \times m)$ and $(n \times n)$, respectively. It follows that

$$(\mathbf{H}^*\mathbf{M}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}^*\mathbf{T}\mathbf{C}(\mathbf{C}^*\mathbf{S}\mathbf{C})^{-\frac{1}{2}} = ((\mathbf{E}^*\mathbf{M}\mathbf{E})_{(h)})^{-\frac{1}{2}}(\mathbf{E}^*\mathbf{T}\mathbf{B})_{(h,c)}((\mathbf{B}^*\mathbf{S}\mathbf{B})_{(c)})^{-\frac{1}{2}} .$$

Moreover, the increasing character of the singular values of the foregoing expressions as functions of $h$ and $c$, together with the monotonicity of $G$, now assure that

$$\sup_{\mathscr{F}_0} G((\mathbf{H}^*\mathbf{M}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}^*\mathbf{T}\mathbf{C}(\mathbf{C}^*\mathbf{S}\mathbf{C})^{-\frac{1}{2}}) = G((\mathbf{E}^*\mathbf{M}\mathbf{E})^{-\frac{1}{2}}\mathbf{E}^*\mathbf{T}\mathbf{B}(\mathbf{B}^*\mathbf{S}\mathbf{B})^{-\frac{1}{2}}) .$$

But the expression on the right is precisely $G(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}})$, and the proof is complete.

**3. Some applications.** This study was motivated largely by interest in the simultaneous inferences generated by certain statistical procedures, including those which we now investigate. In all cases the matrices in question are real.

3.1. *Multivariate linear models.* Suppose $\mathbf{Y} = \mathbf{X}\mathbf{\Theta} + \mathbf{E}$ such that (i) the rows of $\mathbf{Y}(n \times m)$ are mutually independent Gaussian vectors all having the same definite dispersion matrix $\mathbf{\Sigma}(m \times m)$, (ii) $E(\mathbf{Y}) = \mathbf{X}\mathbf{\Theta}$, where (iii) $\mathbf{X}(n \times r)$ is a known matrix of rank $r < n$ and (iv) $\mathbf{\Theta}(r \times m)$ is a matrix of unknown parameters. The problem of testing $H\colon \mathbf{\Theta} = \mathbf{\Theta}_0$ can be reduced by invariance considerations (cf. Lehmann (1959), Section 7.10), the maximal invariant statistics being essentially the singular values of $\mathbf{M}^{-\frac{1}{2}}(\hat{\mathbf{\Theta}} - \mathbf{\Theta}_0)\mathbf{S}^{-\frac{1}{2}}$, and the maximal parametric invariants being the singular values of $\mathbf{M}^{-\frac{1}{2}}(\mathbf{\Theta} - \mathbf{\Theta}_0)\mathbf{\Sigma}^{-\frac{1}{2}}$. Here $\mathbf{\Theta}_0$ is fixed, $\hat{\mathbf{\Theta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, $\mathbf{S} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{\Theta}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{\Theta}})/(n - r)$, and $\mathbf{M} = (\mathbf{X}'\mathbf{X})^{-1}$.

For each test function $G \in \mathscr{G}$ choose $c_\alpha$ such that

$$P[G(\mathbf{M}^{-\frac{1}{2}}(\hat{\mathbf{\Theta}} - \mathbf{\Theta})\mathbf{S}^{-\frac{1}{2}}) \leqq c_\alpha] = 1 - \alpha . \tag{3.1}$$

An application of Corollary 1 immediately gives

$$P[\sup_{\mathscr{F}_0} G((\mathbf{H}'\mathbf{M}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}'(\hat{\mathbf{\Theta}} - \mathbf{\Theta})\mathbf{C}(\mathbf{C}'\mathbf{S}\mathbf{C})^{-\frac{1}{2}}) \leqq c_\alpha] = 1 - \alpha \tag{3.2}$$

where $(\mathbf{H}, \mathbf{C}) \in \mathscr{F}_0 = \{\mathscr{F}_{r\times h} \times \mathscr{F}_{m\times c};\ 1 \leqq h \leqq r,\ 1 \leqq c \leqq m\}$; this expression in turn generates simultaneous confidence sets as preimages in the parameter space $\mathscr{F}_{r\times m}$. Equivalently, the same expression generates simultaneous tests for all hypotheses

$$H\colon \mathbf{H}'(\Theta - \Theta_0)\mathbf{C} = \mathbf{0}\,, \qquad (\mathbf{H}, \mathbf{C}) \in \mathscr{F}_0 \tag{3.3}$$

at a type 1 probability error rate not exceeding $\alpha$, using any test function $G \in \mathscr{G}$.

Confidence sets generated by (3.2) are especially useful in some particular cases. When the argument of $G$ has unit rank there is but one nonvanishing singular value, in which case the standardization $G(u\mathbf{E}_{ij}) = |u|$ enables us to invert the corresponding probability statement explicitly; the end results are ellipsoidal confidence sets obtained as follows.

Specializing first to the case $\mathbf{H} = \mathbf{a} \in \mathbb{R}^r$, and then to the case $\mathbf{C} = \mathbf{b} \in \mathbb{R}^m$, we have

$$\mathbf{a}'(\hat{\Theta} - \Theta)\mathbf{C}(\mathbf{C}'\mathbf{S}\mathbf{C})^{-1}\mathbf{C}'(\hat{\Theta} - \Theta)'\mathbf{a} \leqq c_\alpha^2\mathbf{a}'\mathbf{M}\mathbf{a} \tag{3.4}$$

and

$$\mathbf{b}'(\hat{\Theta} - \Theta)'\mathbf{H}(\mathbf{H}'\mathbf{M}\mathbf{H})^{-1}\mathbf{H}'(\hat{\Theta} - \Theta)\mathbf{b} \leqq c_\alpha^2\mathbf{b}'\mathbf{S}\mathbf{b} \tag{3.5}$$

simultaneously for all $(\mathbf{a}, \mathbf{C}) \in \mathscr{F}_0$ and $(\mathbf{H}, \mathbf{b}) \in \mathscr{F}_0$, respectively, all such statements holding with confidence coefficient at least $1 - \alpha$. Now writing the columns of $\Theta$ as $\Theta = [\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_m]$ and its rows as $\Theta' = [\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_r]$, choosing $\mathbf{C} = \mathbf{I}_m$, and letting $\mathbf{a}$ successively take the values $\{\mathbf{e}_1, \cdots, \mathbf{e}_r\}$, we obtain from (3.4) the confidence ellipsoids of Hotelling's type

$$(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i)'\mathbf{S}^{-1}(\hat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i) \leqq c_\alpha^2 m_{ii}\,, \qquad 1 \leqq i \leqq r\,. \tag{3.6}$$

A similar development from (3.5) yields the confidence ellipsoids

$$(\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta}_j)'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta}_j) \leqq c_\alpha^2 s_{jj}\,, \qquad 1 \leqq j \leqq m \tag{3.7}$$

simultaneously for $\boldsymbol{\theta}_j(r \times 1)$, $1 \leqq j \leqq m$.

Further letting $\mathbf{H} = \mathbf{a} \in \mathbb{R}^r$ and $\mathbf{C} = \mathbf{b} \in \mathbb{R}^m$, we get confidence limits of the type of Roy and Bose (1953)

$$\mathbf{a}'\Theta\mathbf{b} \in \{\mathbf{a}'\hat{\Theta}\mathbf{b} \pm c_\alpha(\mathbf{a}'\mathbf{M}\mathbf{a}\mathbf{b}'\mathbf{S}\mathbf{b})^{\frac{1}{2}}\} \tag{3.8}$$

simultaneously for all $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^r \times \mathbb{R}^m$ with coefficient at least $1 - \alpha$, for each $G \in \mathscr{G}$.

REMARK 7. If inferences are required only for a finite collection of statements such as (3.5) and (3.6), it sometimes is possible to evaluate the actual joint confidence coefficient in terms of the joint distribution of appropriate statistics. In this setting Khatri (1967) provided tighter bounds than $1 - \alpha$ for confidence sets of type (3.5) and, assuming special structure for $\boldsymbol{\Sigma}$, for sets of type (3.6) as well.

As all invertible confidence sets given here have the coefficient $1 - \alpha$ for each $G$, it is pertinent to ask whether tightest bounds can be achieved through choice of $G \in \mathscr{G}$. Results are known for the bilinear functions $\mathbf{a}'\Theta\mathbf{b}$; Mudholkar (1966) showed that the Roy–Bose bounds at level $1 - \alpha$ are tightest, in the class

$\mathscr{G}_0$, when $G$ is generated by $\phi_{[1]}(x_1, \cdots, x_m) = \max\{|x_1|, \cdots, |x_m|\}$, and Gabriel (1968) reached identical conclusions in a larger class of test functions depending only on the eigenvalues of certain random matrices. These findings are now extended to any region $R_G(\Theta)$ of the types (3.4) through (3.8).

THEOREM 4. *Let $G^*$ be the test function generated by $\phi_{[1]}(x_1, \cdots, x_m) = \max\{|x_1|, \cdots, |x_m|\}$ and, for each $G \in \mathscr{G}$, let $R_G(\Theta)$ represent any one region of the types* (3.4) *through* (3.8). *Then the infimum for all regions of the type $R_G(\Theta)$ having probability $1 - \alpha$ is given by $\inf_{\mathscr{G}} R_G(\Theta) \equiv \bigcap_{\mathscr{G}} R_G(\Theta) = R_{G^*}(\Theta)$.*

PROOF. Let $\{\alpha_1 \geqq \cdots \geqq \alpha_m \geqq 0\}$ be the ordered singular values of $\mathbf{A} \in \mathscr{F}_{m\times n}$ and, for each $G \in \mathscr{G}$, write $G(\mathbf{A}) = \phi(\alpha_1, \cdots, \alpha_m)$ for some $\phi \in \Phi$. Clearly $G^*(\mathbf{A}) = \phi_{[1]}(\alpha_1, \cdots, \alpha_m) = \alpha_1$ and, owing to properties (ii) and (iii) of Definition 1, it follows that $\phi(\alpha_1, 0, \cdots, 0) \leqq \phi(\alpha_1, \cdots, \alpha_m)$ for each $\phi \in \Phi$. But property (iv) assures that $\phi(\alpha_1, 0, \cdots, 0) = \alpha_1$. That $G^*$ is stochastically minimal in the class $\mathscr{G}$ is now clear; we have $G^*(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}}) \leqq G(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}})$ for almost all $(\mathbf{M}, \mathbf{T}, \mathbf{S}) \in \mathscr{F}$, so that

$$\sup_{\mathscr{G}} P[G(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}}) \leqq c] = P[G^*(\mathbf{M}^{-\frac{1}{2}}\mathbf{T}\mathbf{S}^{-\frac{1}{2}}) \leqq c]$$

for each $c > 0$. The proof is completed upon noting that each region $R_G(\Theta)$ consists of the boundary and interior of an ellipsoid, these having the same center and orientation for all $G$ and radius depending on $c_\alpha$, a function of $G$.

3.2. *Canonical correlations.* Let $\mathbf{Y}_1(m \times 1)$ and $\mathbf{Y}_2(n \times 1)$ be jointly Gaussian vectors having the nonsingular dispersion matrix $\mathbf{\Sigma} = [\mathbf{\Sigma}_{ij}]$ partitioned conformably, and let $\mathbf{S} = [\mathbf{S}_{ij}]$ be a sample dispersion matrix having $\nu > m + n$ degrees of freedom such that $m \leqq n$. The problem of testing $H$: $\mathbf{\Sigma}_{12} = \mathbf{0}$ is invariant under nonsingular linear transformations acting separately on $\mathbf{Y}_1$ and $\mathbf{Y}_2$; the maximal parametric invariants are the canonical correlations $\{\rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_m \geqq 0\}$, and the maximal invariant statistics are the sample coefficients $\{\hat\rho_1 \geqq \hat\rho_2 \geqq \cdots \geqq \hat\rho_m \geqq 0\}$ (cf. Eaton (1972), page 10.37). These maximal invariants in turn are the singular values $\rho_k = S_k(\mathbf{\Sigma}_{11}^{-\frac{1}{2}}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-\frac{1}{2}})$ and $\hat\rho_k = S_k(\mathbf{S}_{11}^{-\frac{1}{2}}\mathbf{S}_{12}\mathbf{S}_{22}^{-\frac{1}{2}})$.

The simultaneous inferences generated by invariant tests for $H$ are summarized in the following corollary.

COROLLARY 2. *For some $G \in \mathscr{G}$ choose $c_\alpha$ such that $\{G(\mathbf{S}_{11}^{-\frac{1}{2}}\mathbf{S}_{12}\mathbf{S}_{22}^{-\frac{1}{2}}) \leqq c_\alpha\}$ is an $\alpha$-level acceptance region for testing $H$: $\mathbf{\Sigma}_{12} = \mathbf{0}$. If*

$$\{G((\mathbf{H}'\mathbf{S}_{11}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}'\mathbf{S}_{12}\mathbf{C}(\mathbf{C}'\mathbf{S}_{22}\mathbf{C})^{-\frac{1}{2}}) \leqq c_\alpha\}$$

*is used as an acceptance region for testing $H$: $\mathbf{H}'\mathbf{\Sigma}_{12}\mathbf{C} = \mathbf{0}$, then all such invariant tests hold simultaneously at level $\alpha$ for $(\mathbf{H}, \mathbf{C}) \in \mathscr{F}_0$.*

PROOF. Invariant tests for each hypothesis of the type $H$: $\mathbf{H}'\mathbf{\Sigma}_{12}\mathbf{C} = \mathbf{0}$ depend only on the singular values of the matrix $(\mathbf{H}'\mathbf{S}_{11}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}'\mathbf{S}_{12}\mathbf{C}(\mathbf{C}'\mathbf{S}_{22}\mathbf{C})^{-\frac{1}{2}}$. That the type 1 probability error rate for the family of such tests is equal to $\alpha$, follows from Theorem 3 and Corollary 1.

It may be noted that the singular values of $(\mathbf{H}'\mathbf{\Sigma}_{11}\mathbf{H})^{-\frac{1}{2}}\mathbf{H}'\mathbf{\Sigma}_{12}\mathbf{C}(\mathbf{C}'\mathbf{\Sigma}_{22}\mathbf{C})^{-\frac{1}{2}}$, for all $(\mathbf{H}, \mathbf{C}) \in \mathscr{F}_0$, constitute the totality of canonical correlations between linear functions of the elements of $\mathbf{Y}_1$ and $\mathbf{Y}_2$; these include the canonical correlations between all subsets of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ and, in particular, all pair-wise simple correlations between their elements. Corollary 2 supports tests regarding all such parameters, these tests holding simultaneously at level $\alpha$.

REMARK 8. A test function belonging to $\mathscr{G}_0$ is the often used largest sample coefficient $G_{[1]}(\mathbf{S}_{11}^{-\frac{1}{2}}\mathbf{S}_{12}\mathbf{S}_{22}^{-\frac{1}{2}}) = \hat{\rho}_1$. A function belonging to $\mathscr{G}$ but not $\mathscr{G}_0$ is $G(\mathbf{S}_{11}^{-\frac{1}{2}}\mathbf{S}_{12}\mathbf{S}_{22}^{-\frac{1}{2}}) = [1 - \prod_{i=1}^{m}(1 - \hat{\rho}_i^2)]^{\frac{1}{2}}$, which is a monotonic function of the likelihood ratio statistic for testing $H\colon \mathbf{\Sigma}_{12} = \mathbf{0}$.

3.3. *Generalized Friedman tests.* Let $\mathbf{Y}_{ij} = [Y_{ij}^1, \cdots, Y_{ij}^m]'$ be the vector response at the $j$th of $k$ treatments on the $i$th of $N$ replications in a complete two-way classification scheme, and let $\{F_{ij}(\mathbf{z}); 1 \leqq j \leqq k, 1 \leqq i \leqq N\}$ designate their cumulative distribution functions, assumed here to satisfy conditions set forth in Gerig (1969). The problem is to test $\{H\colon F_{i1} \equiv F_{i2} \cdots \equiv F_{ik} \equiv F_i, 1 \leqq i \leqq N\}$ against the translation alternatives $\{A\colon F_{ij}(\mathbf{z}) = F_i(\mathbf{z} - \boldsymbol{\mu}_j), 1 \leqq j \leqq k, 1 \leqq i \leqq N\}$.

Following Gerig (1969), for each $s \in \{1, 2, \cdots, m\}$ and $i \in \{1, 2, \cdots, N\}$ we replace $\{Y_{i1}^s, \cdots, Y_{ik}^s\}$ by their respective ranks $\{R_{i1}^s, \cdots, R_{ik}^s\}$ and define the matrices $\mathbf{T}_N(k \times m)$ and $\mathbf{V}_N(m \times m)$ in terms of their typical $(s, s')$ elements as $\mathbf{T}_N = [(N^{-1}\sum_{i=1}^{N} R_{is}^{s'} - \frac{1}{2}(k+1))]$ and $\mathbf{V}_N = [\sum_{i=1}^{N}(\sum_{j=1}^{k} R_{ij}^{s} R_{ij}^{s'} - k(k+1)^2/4)/N(k-1)]$. The following class of invariant tests for $H$ is based on the joint permutation distribution of $\{R_{ij}^s; 1 \leqq j \leqq k, 1 \leqq s \leqq m, 1 \leqq i \leqq N\}$.

DEFINITION 3. A *generalized Friedman test* is any invariant permutation test for $H$ having acceptance region of the form $\{G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}}) \leqq c_\alpha\}$ for some $G \in \mathscr{G}$.

Our terminology stems from the fact (using $G(\mathbf{E}_{ij}) = 1$) that each such test is equivalent, when $m = 1$, to Friedman's (1937) test. Gerig (1969) studied one such test in detail using essentially the test function $G_{(2)}(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}}) = (N \operatorname{tr} \mathbf{T}_N'\mathbf{T}_N\mathbf{V}_N^{-1})^{\frac{1}{2}}$ belonging to $\mathscr{G}_0$.

Each generalized Friedman test generates simultaneous tests for equality of $k$ marginal distributions as follows at level not exceeding $\alpha$. Let $\{\mathbf{C}_t; t \in \tau\}$ be the collection of $2^m - 1$ matrices of order $(m \times r)$, $1 \leqq r \leqq m$, which operate to delete elements of $\mathbf{Y} \in \mathbb{R}^m$; let $\{F_{ij}^t(\cdot); i \leqq j \leqq k; 1 \leqq i \leqq N, t \in \tau\}$ be the corresponding marginal distribution functions; and consider the family $\{H_t\colon F_{i1}^t \equiv F_{i2}^t \equiv \cdots \equiv F_{ik}^t \equiv F_i^t, 1 \leqq i \leqq N; t \in \tau\}$ of hypotheses. The principal result is the following.

THEOREM 5. *For some $G \in \mathscr{G}$ let $\{G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}}) \leqq c_\alpha\}$ be an $\alpha$-level acceptance region of a permutation test for $H$. Then acceptance regions of the type $\{G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{C}_t(\mathbf{C}_t'\mathbf{V}_N\mathbf{C}_t)^{-\frac{1}{2}}) \leqq c_\alpha\}$ provide tests for all $\{H_t; t \in \tau\}$ simultaneously at a type 1 error rate not exceeding $\alpha$.*

PROOF. The required permutation test for $H$ follows constructively for any

$G \in \mathscr{G}$ along the lines of Gerig (1969). Because such tests depend explicitly on $\{R_{ij}^s; 1 \leqq j \leqq k, 1 \leqq s \leqq m, 1 \leqq i \leqq N\}$, the exact conditional permutation distribution of $G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}})$ can be found for small samples, thereby determining a randomized test for $H$ which is strictly distribution-free under $H$ and which has exact size $\alpha$. For large samples, a normal-theory approximation to the distribution of $G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}})$ stems from the limiting Wishart character of $N\mathbf{T}_N'\mathbf{T}_N\mathbf{V}_N^{-1}$ having parameters $k-1$ and $\mathbf{I}_m$; further details are supplied in Jensen (1974). Accordingly, choose $c_\alpha$ such that $P(G(N^{\frac{1}{2}}\mathbf{T}_N\mathbf{V}_N^{-\frac{1}{2}}) \leqq c_\alpha) = 1-\alpha$; that the type 1 probability error rate for testing $\{H_t; t \in \tau\}$ is at most $\alpha$ follows from Theorem 3 and Corollary 1.

REMARK 9. Letting $\mathbf{C}_t$ successively take the values $\{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$, we get the Friedman (1937) statistics $\{N\mathbf{e}_s'\mathbf{T}_N'\mathbf{T}_N\mathbf{e}_s(\mathbf{e}_s'\mathbf{V}_N\mathbf{e}_s)^{-1}; 1 \leqq s \leqq m\}$ for testing equality of treatment effects in each one-dimensional marginal distribution. The square roots of these functions belong to $\mathscr{G}_0$.

3.4. *Pearson's tests for goodness of fit.* Consider a multidimensional array of frequencies obtained upon cross-classifying each of $N$ independent observations taken from some multivariate distribution. Let $\mathbf{n} = [n_1, \cdots, n_\nu]'$ be the vector of frequencies from such an array arranged in arbitrary order, and let $\boldsymbol{\pi}(\boldsymbol{\theta}) = [\pi_1(\boldsymbol{\theta}), \cdots, \pi_\nu(\boldsymbol{\theta})]'$ be the corresponding cell probabilities, assumed here to be determined completely by some parameter $\boldsymbol{\theta} \in \mathbb{R}^s$. The problem of goodness of fit is to test $H: \boldsymbol{\pi}(\boldsymbol{\theta}) = \boldsymbol{\pi}(\boldsymbol{\theta}_0)$ for fixed $\boldsymbol{\theta}_0$. In addition to $H$, suppose we are concerned with testing goodness of fit for various marginal distributions associated with the original array. Suppose $\{\mathbf{T}_t; t \in \tau\}$ is a collection of matrices of order $(\nu \times \kappa_t)$ having zero and unit elements such that $\mathbf{n}_t = [n_{t1}, \cdots, n_{t\kappa_t}]' = \mathbf{T}_t'\mathbf{n}$ is multinomial having the probabilities $\boldsymbol{\pi}_t(\boldsymbol{\theta}) = \mathbf{T}_t'\boldsymbol{\pi}(\boldsymbol{\theta})$, and consider the family $\{H_t: \boldsymbol{\pi}_t(\boldsymbol{\theta}) = \boldsymbol{\pi}_t(\boldsymbol{\theta}_0); t \in \tau\}$ of hypotheses for these marginal arrays. Simultaneous tests at level $\alpha$ are generated by using the Pearson (1900) test statistic

$$X_N^2 = \sum_{i=1}^{\nu} (n_i - N\pi_i(\boldsymbol{\theta}_0))^2/N\pi_i(\boldsymbol{\theta}_0)$$

in conjunction with the marginal statistics

$$X_{Nt}^2 = \sum_{i=1}^{\kappa_t} (n_{ti} - N\pi_{ti}(\boldsymbol{\theta}_0))^2/N\pi_{ti}(\boldsymbol{\theta}_0)\,, \qquad t \in \tau$$

as follows.

COROLLARY 3. *For each $N$ let $\{X_N^2 \leqq c_\alpha\}$ be an $\alpha$-level acceptance region for testing $H$. If $\{X_{Nt}^2 \leqq c_\alpha\}$ is used as an acceptance region for testing $H_t$ for each $t \in \tau$, then the type 1 error rate is no greater than $\alpha$.*

PROOF. Replace $\kappa_t$ by $r$ and let $\mathbf{L}' = [\mathbf{I}_{\nu-1}, \mathbf{0}]$ and $\mathbf{L}_t' = [\mathbf{I}_{r-1}, \mathbf{0}]$ be of orders $(\nu-1) \times \nu$ and $(r-1) \times r$, respectively. Upon writing $\boldsymbol{\Sigma} = \operatorname{diag}(\pi_1(\boldsymbol{\theta}), \cdots, \pi_\nu(\boldsymbol{\theta})) - \boldsymbol{\pi}(\boldsymbol{\theta})\boldsymbol{\pi}'(\boldsymbol{\theta})$ and using a standard construction (cf. Kendall and Stuart (1963), pages 355–356) for removing the singularity of $\boldsymbol{\Sigma}$, we infer that $X_N^2 = \mathbf{z}'\boldsymbol{\Omega}^{-1}\mathbf{z}$, where $\mathbf{z} = N^{-\frac{1}{2}}\mathbf{L}'(\mathbf{n} - N\boldsymbol{\pi}(\boldsymbol{\theta}_0))$ and $\boldsymbol{\Omega} = \mathbf{L}'\boldsymbol{\Sigma}\mathbf{L}$. It follows similarly that

$$X_{Nt}^2 = N^{-1}(\mathbf{n}_t - N\boldsymbol{\pi}_t(\boldsymbol{\theta}_0))'\mathbf{L}_t(\mathbf{L}_t'\boldsymbol{\Sigma}_t\mathbf{L}_t)^{-1}\mathbf{L}_t'(\mathbf{n}_t - N\boldsymbol{\pi}_t(\boldsymbol{\theta}_0))$$

where $\mathbf{\Sigma}_t = \mathbf{T}_t'\mathbf{\Sigma}\mathbf{T}_t$. But $(\mathbf{n}_t - N\boldsymbol{\pi}_t(\boldsymbol{\theta}_0))$ is a linear function of $(\mathbf{n} - N\boldsymbol{\pi}(\boldsymbol{\theta}))$; thus $\mathbf{L}_t'(\mathbf{n}_t - N\boldsymbol{\pi}_t(\boldsymbol{\theta}_0))$ is some linear function of $\mathbf{L}'(\mathbf{n} - N\boldsymbol{\pi}(\boldsymbol{\theta}_0))$ which we can express as $\mathbf{z}_t \equiv N^{-\frac{1}{2}}\mathbf{L}_t'(\mathbf{n}_t - N\boldsymbol{\pi}_t(\boldsymbol{\theta}_0)) = \mathbf{C}_t'\mathbf{z}$ for some $\mathbf{C}_t((\nu - 1) \times (r - 1))$. It follows that

$$X^2_{Nt} = \mathbf{z}'\mathbf{C}_t(\mathbf{C}_t'\mathbf{\Omega}\mathbf{C}_t)^{-1}\mathbf{C}_t'\mathbf{z}\,.$$

An application of Corollary 1 completes the proof.

## REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

BELLMAN, R. (1970). *Introduction to Matrix Analysis*, 2nd ed. McGraw-Hill, New York.

EATON, M. L. (1972). *Multivariate Statistical Analysis.* Institute of Mathematical Statistics, Univ. of Copenhagen.

FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–699.

GABRIEL, K. R. (1968). Simultaneous test procedures in multivariate analysis of variance. *Biometrika* **55** 489–504.

GERIG, T. M. (1969). A multivariate extension of Friedman's $\chi_r^2$-test. *J. Amer. Statist. Assoc.* **64** 1595–1608.

JENSEN, D. R. (1974). On the joint distribution of Friedman's $\chi_r^2$ statistics. *Ann. Statist.* **2** 311–322.

KENDALL, M. G. and STUART, A. (1963). *The Advanced Theory of Statistics* I, 2nd ed. Hafner, New York.

KHATRI, C. G. (1967). On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. *Ann. Math. Statist.* **38** 1853–1867.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

MUDHOLKAR, G. S. (1966). On confidence bounds associated with multivariate analysis of variance and non-independence between two sets of variates. *Ann. Math. Statist.* **37** 1736–1746.

VON NEUMANN, J. (1937). Some matrix inequalities and metrization of matric space. *Tomsk Univ. Rev.* **1** 286–300.

PEARSON, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series* 5, **50** 157–172.

ROY, S. N. and BOSE, R. C. (1953). Simultaneous confidence interval estimation. *Ann. Math. Statist.* **24** 513–536.

SCHATTEN, R. (1970). *Norm Ideals of Completely Continuous Operators.* Springer-Verlag, Berlin.

DEPARTMENT OF STATISTICS
VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
BLACKSBURG, VIRGINIA 24061