SOME OPTIMUM CONFIDENCE BOUNDS FOR ROOTS OF DETERMINANTAL EQUATIONS[1]

By T. W. Anderson

Columbia University

0. Summary. A problem considered in this paper is to obtain confidence intervals for all characteristic roots of a population covariance matrix $\mathbf{\Sigma}$ in the form $[\mathrm{ch}_m(\mathbf{S})/u, \mathrm{ch}_M(\mathbf{S})/l]$, where $\mathrm{ch}_m(\mathbf{S})$ and $\mathrm{ch}_M(\mathbf{S})$ are the minimum and maximum characteristic roots, respectively, of a sample covariance matrix $\mathbf{S}$ from a multivariate normal population and u and l are constants. Intervals of this form having probability at least $1 - \epsilon$ can be obtained by basing u and l on certain χ^2-distributions. Among all intervals in a certain class such intervals are shortest.

Another problem treated is to obtain confidence intervals for all characteristic roots $\mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ in the form $[\mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U, \mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L]$, where $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ and $\mathbf{S}_1$ and $\mathbf{S}_2$ are population and sample covariance matrices of two multivariate normal populations, respectively, and U and L are constants, determined from F-distributions to give confidence at least $1 - \epsilon$. Such choices of the constants yield shortest intervals within a certain class.

Comparison is made with other methods of finding such intervals. Various uses of the intervals are suggested, such as simultaneous intervals for variances and correlation coefficients. Some other confidence intervals for related problems are considered.

1. Introduction. The multivariate normal distribution, $N(\boldsymbol{\mu}, \mathbf{\Sigma})$, is characterized by $\boldsymbol{\mu}$, the vector of means of the random variables with this distribution, and $\mathbf{\Sigma}$, the matrix of variances and covariances of these random variables. In this paper we consider confidence bounds for the characteristic roots of $\mathbf{\Sigma}$, namely, the roots $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_p$ of

$$|\mathbf{\Sigma} - \lambda\mathbf{I}| = 0. \tag{1.1}$$

The sufficient set of statistics based on a sample $\mathbf{x}_1, \cdots, \mathbf{x}_N$ from $N(\boldsymbol{\mu}, \mathbf{\Sigma})$ is the sample mean $\bar{\mathbf{x}}$ and sample covariance matrix $\mathbf{S}$ defined by

$$N\bar{\mathbf{x}} = \textstyle\sum_{\alpha=1}^N \mathbf{x}_\alpha, \qquad n\mathbf{S} = \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})', \tag{1.2}$$

and $n = N - 1$. Confidence bounds on all characteristic roots of $\mathbf{\Sigma}$, denoted by $\mathrm{ch}(\mathbf{\Sigma})$, are given by

$$\mathrm{ch}_m(\mathbf{S})/u \leqq \mathrm{ch}(\mathbf{\Sigma}) \leqq \mathrm{ch}_M(\mathbf{S})/l, \tag{1.3}$$

Received 22 November 1963; revised 19 November 1964.

[1] This research was sponsored by the Office of Naval Research under Contract Number Nonr-266(33), Project NR 042-034, and Contract Number Nonr-4195(00). Reproduction in whole or in part is permitted for any purpose of the United States Government.

where $\mathrm{ch}_m(\mathbf{S})$ and $\mathrm{ch}_M(\mathbf{S})$ denote the minimum and maximum characteristic roots of $\mathbf{S}$ and u and l are constants such that

$$(1.4) \qquad \Pr\{nl \leqq \chi_n^2\} \Pr\{\chi^2_{n-p+1} \leqq nu\} = 1 - \epsilon,$$

where χ_m^2 denotes a random variable having a χ^2-distribution with m degrees of freedom; the probability is at least $1 - \epsilon$ that (1.3) holds for any positive definite $\mathbf{\Sigma}$. These confidence bounds for $\mathrm{ch}(\mathbf{\Sigma})$ are optimal within the class of bounds

$$(1.5) \qquad f[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})] \leqq \mathrm{ch}(\mathbf{\Sigma}) \leqq g[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})],$$

where $\mathrm{ch}_i(\mathbf{S})$ is the ith characteristic root of $\mathbf{S}$, $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ are homogeneous of degree 1, and are monotonically nonincreasing in each argument for fixed values of the other arguments $(x_1 \geqq x_2 \geqq \cdots \geqq x_p \geqq 0)$. If (1.5) holds with probability at least $1 - \epsilon$, then a pair of numbers u and l can be found to satisfy (1.4) and

$$(1.6) \quad f[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})] \leqq \mathrm{ch}_m(\mathbf{S})/u, \ \mathrm{ch}_M(\mathbf{S})/l \leqq g[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})].$$

The homogeneity condition means that the confidence bounds for $\mathrm{ch}(\mathbf{\Sigma})$ are multiplied by c^2 if the observed vectors $\mathbf{x}_1, \cdots, \mathbf{x}_N$ are multiplied by c (which is a kind of scale invariance); the monotonicity conditions imply that an increase in the size of $\mathbf{S}$ results in an increase in the limits for $\mathbf{\Sigma}$ (which is a kind of consistency).

Confidence bounds of this type given earlier by Roy (1954) and his colleagues have involved the distributions of the characteristic roots of sample covariance matrices when $\mathbf{\Sigma} = \mathbf{I}$; these distributions are not extensively tabulated. A more detailed comparison between the bounds derived here and those given earlier will be made in Section 4.

We also consider confidence bounds for roots of determinantal equations involving two covariance matrices. The functions of the parameters of two normal distributions, $N(\boldsymbol{\mu}^{(1)}, \mathbf{\Sigma}_1)$ and $N(\boldsymbol{\mu}^{(2)}, \mathbf{\Sigma}_2)$, treated here are the roots $\theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_p$ of

$$(1.7) \qquad |\mathbf{\Sigma}_1 - \theta \mathbf{\Sigma}_2| = 0.$$

Multiplication of (1.7) by $|\mathbf{\Sigma}_2^{-1}|$ shows that the roots of (1.7) are the characteristic roots of $\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1$ and of $\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}$, denoted by $\mathrm{ch}(\mathbf{\Sigma}_2^{-1}\mathbf{\Sigma}_1) = \mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$. We call these roots the characteristic roots of $\mathbf{\Sigma}_1$ in the metric of $\mathbf{\Sigma}_2$.

The sufficient set of statistics based on samples of $N_1 = n_1 + 1$ and $N_2 = n_2 + 1$ from $N(\boldsymbol{\mu}^{(1)}, \mathbf{\Sigma}_1)$ and $N(\boldsymbol{\mu}^{(2)}, \mathbf{\Sigma}_2)$, respectively, are the mean and covariance matrix $\bar{\mathbf{x}}^{(1)}$ and $\mathbf{S}_1$, of the first sample and the mean and covariance matrix, $\bar{\mathbf{x}}^{(2)}$ and $\mathbf{S}_2$, of the second sample. Confidence bounds on $\mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ with confidence at least $1 - \epsilon$ are given by

$$(1.8) \qquad \mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U \leqq \mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) \leqq \mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L,$$

where L and U satisfy

$$(1.9)\quad \Pr\{(n_2 - p + 1)L/n_2 \leqq F_{n_1, n_2-p+1}\} \cdot \Pr\{F_{n_1-p+1, n_2} \leqq n_1 U/(n_1 - p + 1)\} = 1 - \epsilon,$$

where $F_{n,m}$ denotes a random variable with an F-distribution with n and m degrees of freedom. The bounds given earlier by Roy and Gnanadesikan (1957) involved the distribution of $\mathrm{ch}(\mathbf{S}_1\mathbf{S}_2^{-1})$ when $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. The bounds given here are optimal within the class of bounds

$$(1.10)\quad f[\mathrm{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})] \leqq \mathrm{ch}(\mathbf{\Sigma}) \leqq g[\mathrm{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})],$$

where $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ are homogeneous of degree 1 and monotonically nondecreasing in each argument. If (1.10) holds with probability at least $1 - \epsilon$, then a pair of numbers U and L can be found to satisfy (1.9) and

$$(1.11)\quad f[\mathrm{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})] \leqq \mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U, \quad \mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L \leqq g[\mathrm{ch}(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})].$$

Confidence bounds for some other determinantal roots are also given. The methods used here yield some monotonicity properties of certain test procedures.

An important property of the multivariate normal distribution is that a linear transformation $\mathbf{AX} + \mathbf{b}$ of a random vector $\mathbf{X}$ with distribution $N(\boldsymbol{\mu}, \mathbf{\Sigma})$ has a multivariate normal distribution $N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A\Sigma A}')$. One reason for interest in characteristic roots of a covariance matrix $\mathbf{\Sigma}$ is that the roots are the invariants of rotations and translations; that is, under transformations $\mathbf{X} \to \mathbf{PX} + \mathbf{b}$, where $\mathbf{P}$ is orthogonal. In fact, the roots of $\mathbf{P\Sigma P}'$ are the roots of $\mathbf{\Sigma}$, for

$$(1.12)\qquad 0 = |\mathbf{P\Sigma P}' - \lambda \mathbf{I}| = |\mathbf{P}| \cdot |\mathbf{\Sigma} - \lambda\mathbf{I}| \cdot |\mathbf{P}| .$$

Given any covariance matrix $\mathbf{\Sigma}$ there exists an orthogonal matrix $\mathbf{P}$ such that

$$(1.13)\qquad \mathbf{P\Sigma P}' = \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix},$$

where $\lambda_1 \geqq \cdots \geqq \lambda_p$ are the characteristic roots of $\mathbf{\Sigma}$; thus any function of $\mathbf{\Sigma}$ that is invariant under orthogonal transformations is a function of $\lambda_1, \cdots, \lambda_p$. (See Theorem 2 of Appendix A of Anderson (1958), for example.)

The hypothesis that the components of $\mathbf{X}$ are independent and all have variance 1 is the hypothesis that $\mathbf{\Sigma} = \mathbf{I}$ or equivalently that $\lambda_i = 1, i = 1, \cdots, p$. Since this hypothesis is invariant with respect to orthogonal transformations, it is natural to study tests which are also invariant under orthogonal transformations. The power of such a test depends on an alternative $\mathbf{\Sigma}$ through the characteristic roots of this alternative. It seems reasonable to consider the distance of $\mathbf{\Sigma}$ from the null hypothesis in terms of its characteristic roots.

Another reason for being interested in the characteristic roots of $\mathbf{\Sigma}$ is that the smallest and largest roots give bounds on quadratic forms in $\mathbf{\Sigma}$. That is, from

$$\text{(1.14)} \qquad \text{ch}_m(\mathbf{\Sigma}) = \min_{\mathbf{a}} (\mathbf{a}'\mathbf{\Sigma}\mathbf{a}/\mathbf{a}'\mathbf{a}), \qquad \max_{\mathbf{a}} (\mathbf{a}'\mathbf{\Sigma}\mathbf{a}/\mathbf{a}'\mathbf{a}) = \text{ch}_M(\mathbf{\Sigma})$$

we derive

$$\text{(1.15)} \qquad \mathbf{a}'\mathbf{a}\,\text{ch}_m(\mathbf{\Sigma}) \leqq \mathbf{a}'\mathbf{\Sigma}\mathbf{a} \leqq \mathbf{a}'\mathbf{a}\,\text{ch}_M(\mathbf{\Sigma})$$

for all $\mathbf{a}$. Thus confidence bounds on $\text{ch}_m(\mathbf{\Sigma})$ and $\text{ch}_M(\mathbf{\Sigma})$ imply bounds on all normalized quadratic forms $\mathbf{a}'\mathbf{\Sigma}\mathbf{a}$. These include the variances, since $\mathbf{a}'\mathbf{a} = 1$ and $\mathbf{a}'\mathbf{\Sigma}\mathbf{a} = \sigma_{ii}$ if all the components of $\mathbf{a}$ are 0 except the ith, which is 1. Intervals for σ_{ii} and σ_{ij} [or $\rho_{ij} = \sigma_{ij}/(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}}$] hold simultaneously. Confidence bounds for $\mathbf{a}'\mathbf{\Sigma}\mathbf{a}$ based on $\mathbf{a}'\mathbf{S}\mathbf{a}$ have been given by Roy and Gnanadesikan; these are discussed in Sections 4 and 5.

The first principal component is defined as the linear combination $\mathbf{a}'\mathbf{X}$, normalized by $\mathbf{a}'\mathbf{a} = 1$, that has maximum variance [Chapter 11 of Anderson (1958), for example]; this maximum variance is λ_1. Other principal components are defined as normalized linear combinations maximizing variances subject to being uncorrelated with other principal components. Their respective variances are $\lambda_2, \cdots, \lambda_p$.

The characteristic roots of $\mathbf{\Sigma}_1$ in the metric of $\mathbf{\Sigma}_2$ are invariants of the distributions $N(\boldsymbol{\mu}^{(1)}, \mathbf{\Sigma}_1)$ and $N(\boldsymbol{\mu}^{(2)}, \mathbf{\Sigma}_2)$ of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, respectively, under transformations $\mathbf{X}^{(1)} \to \mathbf{A}\mathbf{X}^{(1)} + \mathbf{b}_1$ and $\mathbf{X}^{(2)} \to \mathbf{A}\mathbf{X}^{(2)} + \mathbf{b}_2$ ($\mathbf{A}$ nonsingular) since

$$\text{(1.16)} \qquad 0 = |\mathbf{A}\mathbf{\Sigma}_1\mathbf{A}' - \theta\mathbf{A}\mathbf{\Sigma}_2\mathbf{A}'| = |\mathbf{A}| \cdot |\mathbf{\Sigma}_1 - \theta\mathbf{\Sigma}_2| \cdot |\mathbf{A}'|.$$

Any invariant is a function of the set of roots $\theta_1 \geqq \cdots \geqq \theta_p$ because there exists a matrix $\mathbf{A}$ such that

$$\text{(1.17)} \qquad \mathbf{A}\mathbf{\Sigma}_2\mathbf{A}' = \mathbf{I},$$

$$\text{(1.18)} \qquad \mathbf{A}\mathbf{\Sigma}_1\mathbf{A}' = \mathbf{\Theta} = \begin{pmatrix} \theta_1 & 0 & \cdots & 0 \\ 0 & \theta_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \theta_p \end{pmatrix}.$$

The hypothesis that $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ can be stated as the hypothesis that $\theta_1 = \cdots = \theta_p = 1$. The hypothesis is invariant with respect to linear transformations. Tests which are invariant with respect to linear transformations have power functions depending only on $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$.

From the fact that

$$\text{(1.19)} \qquad \text{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) = \min_{\mathbf{a}} (\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}/\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}), \qquad \max_{\mathbf{a}} (\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}/\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}) = \text{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$$

we derive

$$\text{(1.20)} \qquad \mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}\,\text{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) \leqq \mathbf{a}'\mathbf{\Sigma}_1\mathbf{a} \leqq \mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}\,\text{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$$

for all $\mathbf{a}$. This implies that confidence bounds on $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ give bounds on quadratic forms $\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}$ in terms of $\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}$.

To avoid trivialities throughout the paper it is assumed that $p \geqq 2$. One-sided bounds are special cases when l, L, $1/u$, or $1/u$ is set equal to 0.

This study originated from discussions between S. N. Roy and R. Gnanadesikan and the author.

2. Optimum confidence bounds for the characteristic roots of a covariance matrix. The probability of a pair of inequalities (1.5),

$$k(\lambda_1, \cdots, \lambda_p) = \Pr\{f[\text{ch}_1(\mathbf{S}), \cdots, \text{ch}_p(\mathbf{S})] \leqq \text{ch}_m(\mathbf{\Sigma}), \\ \text{ch}_M(\mathbf{\Sigma}) \leqq g[\text{ch}_1(\mathbf{S}), \cdots, \text{ch}_p(\mathbf{S})]\}, \tag{2.1}$$

is a function of $\lambda_1, \cdots, \lambda_p$ $(\lambda_1 \geqq \cdots \geqq \lambda_p > 0)$, the characteristic roots of $\mathbf{\Sigma}$, because ch($\mathbf{S}$) and ch($\mathbf{\Sigma}$) are invariant with respect to orthogonal transformations. For the inequalities to constitute confidence bounds for ch($\mathbf{\Sigma}$) with confidence coefficient $1 - \epsilon$ we require

$$k(\lambda_1, \cdots, \lambda_p) \geqq 1 - \epsilon \tag{2.2}$$

for all positive definite $\mathbf{\Sigma}$ $(\lambda_1 \geqq \cdots \geqq \lambda_p > 0)$.

We shall show that for given $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ homogeneous of degree 1 and monotonically nondecreasing in each argument

$$\inf_{\lambda_1 \geqq \cdots \geqq \lambda_p > 0} k(\lambda_1, \lambda_2, \cdots, \lambda_{p-1}, \lambda_p) \\ = \lim_{\lambda_1 \to \infty, \lambda_p \to 0} k(\lambda_1, \lambda_2, \cdots, \lambda_{p-1}, \lambda_p) \tag{2.3}$$

independent of $\lambda_2, \cdots, \lambda_{p-1}$, and we shall evaluate this quantity.

Since ch($\mathbf{S}$) and ch($\mathbf{\Sigma}$) are invariant with respect to orthogonal transformations we can make the orthogonal transformation to carry $\mathbf{\Sigma}$ to the diagonal matrix $\mathbf{\Lambda}$. The distribution of $n\mathbf{S}$ is the Wishart distribution, $W(\mathbf{\Lambda}, n)$. Some monotonicity properties of the probability (2.1) are given in the following theorem:

THEOREM 2.1. *If* $\lambda_1 \leqq \lambda_1^+$, *then*

$$k(\lambda_1^+, \lambda_2, \cdots, \lambda_p) \leqq k(\lambda_1, \lambda_2, \cdots, \lambda_p); \tag{2.4}$$

if $\lambda_p^- \leqq \lambda_p$, *then*

$$k(\lambda_1, \cdots, \lambda_{p-1}, \lambda_p^-) \leqq k(\lambda_1, \cdots, \lambda_{p-1}, \lambda_p). \tag{2.5}$$

PROOF. We use two lemmas. Lemma 2.1 is given in Anderson and Das Gupta (1963), for example, and Lemma 2.2 is a special case of Section 7.3.3 of Anderson (1958).

LEMMA 2.1. *If* $\mathbf{A}$ *is a positive definite matrix and* $\mathbf{D}$ *is a diagonal matrix with each diagonal element at least equal to* 1, *then the ith ordered characteristic root of* $\mathbf{DAD}$ *is at least equal to the ith ordered characteristic root of* $\mathbf{A}$. *If at least one diagonal element of* $\mathbf{D}$ *is greater than* 1, *then at least one characteristic root of* $\mathbf{DAD}$ *is greater than the corresponding characteristic root of* $\mathbf{A}$.

LEMMA 2.2. *If* $n\mathbf{S}$ *is distributed according to* $W(\mathbf{\Lambda}, n)$, *where* $\mathbf{\Lambda}$ *is a diagonal matrix, then* $n\mathbf{S}$ *is distributed as* $n\mathbf{\Delta}\mathbf{S}^*\mathbf{\Delta}$, *where* $\mathbf{S}^*$ *is distributed according to*

$W(\mathbf{I}, n)$ and $\mathbf{\Delta}$ is a diagonal matrix each of whose diagonal elements is the positive square root of the corresponding diagonal element of $\mathbf{\Lambda}$.

To economize in typesetting we shall write $\mathbf{\Delta}/\delta$ to mean $(1/\delta)\mathbf{\Delta}$. We have

$$\begin{aligned} k(\lambda_1, \cdots, \lambda_p) &= \Pr\{f[\mathrm{ch}_1(\mathbf{\Delta S^* \Delta}), \cdots, \mathrm{ch}_p(\mathbf{\Delta S^* \Delta})] \leqq \delta_p^2, \\ &\qquad \delta_1^2 \leqq g[\mathrm{ch}_1(\mathbf{\Delta S^* \Delta}), \cdots, \mathrm{ch}_p(\mathbf{\Delta S^* \Delta})]\} \\ &= \Pr\{f[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}]\} \end{aligned} \tag{2.6}$$

by the homogeneity of the functions. From Lemma 2.1 we deduce that for any $\mathbf{S}^*$ and any i ($i = 1, \cdots, p$)

$$\mathrm{ch}_i[(\mathbf{\Delta}^-/\delta_p^-)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_p^-)] \geqq \mathrm{ch}_i[(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)], \tag{2.7}$$

$$\mathrm{ch}_i[(\mathbf{\Delta}^-/\delta_1)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_1)] \leqq \mathrm{ch}_i[(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)], \tag{2.8}$$

$$\mathrm{ch}_i[(\mathbf{\Delta}^+/\delta_p)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_p)] \geqq \mathrm{ch}_i[(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)], \tag{2.9}$$

$$\mathrm{ch}_i[(\mathbf{\Delta}^+/\delta_1^+)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_1^+)] \leqq \mathrm{ch}_i[(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)], \tag{2.10}$$

where $\delta_p^- \leqq \delta_p$, $\delta_1^+ \geqq \delta_1$, $\mathbf{\Delta}^-$ has δ_p^- as its pth diagonal element, and $\mathbf{\Delta}^+$ has δ_1^+ as its first diagonal element. Thus

$$\begin{aligned} &\{\mathbf{S}^* \mid f[\mathrm{ch}_1\{(\mathbf{\Delta}^-/\delta_p^-)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_p^-)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}^-/\delta_p^-)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_p^-)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}^-/\delta_1)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}^-/\delta_1)\mathbf{S}^*(\mathbf{\Delta}^-/\delta_1)\}]\} \\ &\subset \{\mathbf{S}^* \mid f[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}]\}, \end{aligned} \tag{2.11}$$

$$\begin{aligned} &\{\mathbf{S}^* \mid f[\mathrm{ch}_1\{(\mathbf{\Delta}^+/\delta_p)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}^+/\delta_p)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_p)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}^+/\delta_1^+)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_1^+)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}^+/\delta_1^+)\mathbf{S}^*(\mathbf{\Delta}^+/\delta_1^+)\}]\} \\ &\subset \{\mathbf{S}^* \mid f[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}]\}. \end{aligned} \tag{2.12}$$

These imply Theorem 2.1.

THEOREM 2.2.

$$\begin{aligned} &\lim_{\lambda_1 \to \infty, \lambda_p \to 0} k(\lambda_1, \lambda_2, \cdots, \lambda_{p-1}, \lambda_p) \\ &\qquad = \Pr\{n \leqq \chi_n^2 g(1, 0, \cdots, 0)\} \Pr\{\chi_{n-p+1}^2 f(\infty, \cdots, \infty, 1) \leqq n\}. \end{aligned} \tag{2.13}$$

PROOF. For any fixed $\mathbf{S}^*$,

$$\lim_{\delta_1 \to \infty} (\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1) = \begin{pmatrix} s_{11}^* & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{2.14}$$

since

$$(2.15)\qquad \lim_{\delta_1\to\infty}(\mathbf{\Delta}/\delta_1) = \lim_{\delta_1\to\infty}\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & \delta_2/\delta_1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \delta_p/\delta_1\end{pmatrix} = \begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix};$$

and therefore

$$(2.16)\qquad \lim_{\delta_1\to\infty} \mathrm{ch}_1[(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)] = s_{11}^*,$$

$$(2.17)\qquad \lim_{\delta_1\to\infty} \mathrm{ch}_i[(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)] = 0, \qquad i = 2, \cdots, p.$$

Similarly

$$(2.18)\qquad \lim_{\delta_p\to 0} \mathrm{ch}_1[(\delta_p\mathbf{\Delta}^{-1})(\mathbf{S}^*)^{-1}(\delta_p\mathbf{\Delta}^{-1})] = s^{*pp},$$

$$(2.19)\qquad \lim_{\delta_p\to 0} \mathrm{ch}_i[(\delta_p\mathbf{\Delta}^{-1})(\mathbf{S}^*)^{-1}(\delta_p\mathbf{\Delta}^{-1})] = 0, \qquad i = 2, \cdots, p,$$

where s^{*pp} is the pth diagonal element of $(\mathbf{S}^*)^{-1}$, since

$$(2.20)\qquad \lim_{\delta_p\to 0}(\delta_p\mathbf{\Delta}^{-1}) = \lim_{\delta_p\to 0}\begin{pmatrix}\delta_p/\delta_1 & \cdots & 0 & 0\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & \delta_p/\delta_{p-1} & 0\\ 0 & \cdots & 0 & 1\end{pmatrix} = \begin{pmatrix}0 & \cdots & 0 & 0\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & 1\end{pmatrix}.$$

Then

$$(2.21)\qquad \begin{aligned}\lim_{\delta_p\to 0} \mathrm{ch}_p[(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)] &= 1/\lim_{\delta_p\to 0} \mathrm{ch}_1[(\mathbf{\Delta}/\delta_p)^{-1}(\mathbf{S}^*)^{-1}(\mathbf{\Delta}/\delta_p)^{-1}]\\ &= 1/s^{*pp}.\end{aligned}$$

Thus

$$(2.22)\qquad \begin{aligned}\lim_{\delta_1\to\infty,\delta_p\to 0}\{\mathbf{S}^* \mid f[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_p)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}] &\leqq 1,\\ 1 \leqq g[\mathrm{ch}_1\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{\Delta}/\delta_1)\mathbf{S}^*(\mathbf{\Delta}/\delta_p)\}]\}&\\ = \{\mathbf{S}^* \mid f(\infty, \cdots, \infty, 1/s^{*pp}) \leqq 1 \leqq g[s_{11}^*, 0, \cdots, 0]\}&\\ = \{\mathbf{S}^* \mid f(\infty, \cdots, \infty, 1)/s^{*pp} \leqq 1 \leqq s_{11}^*\, g(1, 0, \cdots, 0)\}.&\end{aligned}$$

The theorem follows because s_{11}^* is the sample variance of the first component of the random vector and

$$(2.23)\qquad 1/s^{*pp} = s_{pp}^* - (s_{1p}^*, \cdots, s_{p-1,p}^*)\begin{pmatrix}s_{11}^* & \cdots & s_{1,p-1}^*\\ \vdots & & \vdots\\ s_{p-1,1}^* & \cdots & s_{p-1,p-1}^*\end{pmatrix}^{-1}\begin{pmatrix}s_{p1}^*\\ \vdots\\ s_{p,p-1}^*\end{pmatrix}$$

is independently distributed as $1/n$ times a χ^2-variable with $n - p + 1$ degrees of freedom [Theorem 4.3.2 of Anderson (1958)]. [We note that unless $f(\infty, \cdots, \infty, 1)$ is finite, $k(\lambda_1, \cdots, \lambda_{p-1}, 0) = 0$.]

THEOREM 2.3. *If $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ are homogeneous functions* *of degree* 1 *and monotonically nondecreasing in each argument for the other arguments fixed* $(x_1 \geqq x_2 \geqq \cdots \geqq x_p \geqq 0)$, *then the interval*

$$(2.24) \quad f[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})] \leqq \mathrm{ch}(\mathbf{\Sigma}) \leqq g[\mathrm{ch}_1(\mathbf{S}), \cdots, \mathrm{ch}_p(\mathbf{S})]$$

is a confidence interval for $\mathrm{ch}(\mathbf{\Sigma})$ *of confidence* $1 - \epsilon$ *if and only if*

$$(2.25) \quad \Pr\{n \leqq \chi_n^2 g(1, 0, \cdots, 0)\} \Pr\{\chi^2_{n-p+1} f(\infty, \cdots, \infty, 1) \leqq n\} \geqq 1 - \epsilon.$$

This is the main result, the proof of which was indicated at the beginning of the section; it shows how confidence limits may be obtained from tables of the χ^2-distribution. In particular x_p/u and x_1/l are functions satisfying the conditions of Theorem 2.3. If l and u satisfy (1.4) then (1.3) is a confidence interval of confidence $1 - \epsilon$.

THEOREM 2.4. *If* $f(x_1, \cdots, x_p)$ *and* $g(x_1 \cdots, x_p)$ *are homogeneous functions of degree* 1 *and monotonically nondecreasing in each argument for the other arguments fixed* $(x_1 \geqq x_2 \geqq \cdots \geqq x_p \geqq 0)$, *such that* (2.24) *holds with probability at least equal to* $1 - \epsilon$, *then there exist numbers* l *and* u *satisfying* (1.4) *such that* (1.3) *holds with probability* $1 - \epsilon$ *and the interval* (1.3) *is contained within the interval* (2.24).

PROOF. Let $1/u = f(\infty, \cdots, \infty, 1)$. Then by Theorem 2.3 $1/l \leqq g(1, 0, \cdots, 0)$. Thus

$$(2.26) \quad x_p/u = x_p f(\infty, \cdots, \infty, 1) = f(\infty, \cdots, \infty, x_p) \geqq f(x_1, \cdots, x_{p-1}, x_p),$$

$$(2.27) \quad x_1/l \leqq x_1 g(1, 0, \cdots, 0) = g(x_1, 0, \cdots, 0) \leqq g(x_1, x_2, \cdots, x_p).$$

The interval $(x_p/u, x_1/l)$ is contained by the interval (2.24).

3. Optimum confidence bounds for the characteristic roots of one covariance matrix in the metric of another. Since $\mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ and $\mathrm{ch}(\mathbf{S}_1\mathbf{S}_2^{-1})$ are invariant with respect to linear transformations,

$$(3.1) \quad h(\theta_1, \cdots, \theta_p) = \Pr\{f[\mathrm{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})] \leqq \mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}),$$
$$\mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) \leqq g[\mathrm{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \mathrm{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})]\}$$

is a function of $\theta_1 = \mathrm{ch}_1(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}), \cdots, \theta_p = \mathrm{ch}_p(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$. For the inequalities (1.10) to constitute confidence limits for all $\mathrm{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ with confidence coefficient $1 - \epsilon$ we require

$$(3.2) \quad h(\theta_1, \cdots, \theta_p) \geqq 1 - \epsilon$$

for all $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ (that is, all $\theta_1 \geqq \cdots \geqq \theta_p > 0$). We shall show that for any given $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ homogeneous of degree 1 and monotonically nondecreasing in each argument

$$(3.3) \quad \inf_{\theta_1 \geqq \cdots \geqq \theta_p > 0} h(\theta_1, \theta_2, \cdots, \theta_{p-1}, \theta_p)$$
$$= \lim_{\theta_1 \to \infty, \theta_p \to 0} h(\theta_1, \theta_2, \cdots, \theta_{p-1}, \theta_p)$$

independent of $\theta_2, \cdots, \theta_{p-1}$, and we shall evaluate this probability in terms of

F-distributions. After the introductory theorem the derivations are similar to those of Section 2. As noted in Section 1, we can take $\mathbf{\Sigma}_1 = \mathbf{\Theta}$ (composed of $\theta_1 \geqq \cdots \geqq \theta_p > 0$ as diagonal elements) and $\mathbf{\Sigma}_2 = \mathbf{I}$.

THEOREM 3.1. *The distribution of* $\mathrm{ch}(\mathbf{S}_1\mathbf{S}_2^{-1})$ *is the same as the distribution of* $\mathrm{ch}(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1)$, *where* $\mathbf{D}_1$ *is a diagonal matrix with* $[\mathrm{ch}(\mathbf{S}_1)]^{\frac{1}{2}}$ *as diagonal elements, if* $n_2\mathbf{S}_2$ *is distributed according to* $W(\mathbf{I}, n_2)$.

PROOF. Let

$$\mathbf{S}_1 = \mathbf{Q}\mathbf{D}_1^2\mathbf{Q}', \tag{3.4}$$

where $\mathbf{Q}$ is orthogonal. Then

$$\begin{aligned} \mathrm{ch}_i(\mathbf{S}_1\mathbf{S}_2^{-1}) &= \mathrm{ch}_i(\mathbf{Q}\mathbf{D}_1^2\mathbf{Q}'\mathbf{S}_2^{-1}) \\ &= \mathrm{ch}_i(\mathbf{D}_1\mathbf{Q}'\mathbf{S}_2^{-1}\mathbf{Q}\mathbf{D}_1) \\ &= \mathrm{ch}_i[\mathbf{D}_1(\mathbf{Q}'\mathbf{S}_2\mathbf{Q})^{-1}\mathbf{D}_1]. \end{aligned} \tag{3.5}$$

For any orthogonal $\mathbf{Q}$ the matrix $n_2\mathbf{Q}'\mathbf{S}_2\mathbf{Q}$ has the distribution $W(\mathbf{I}, n_2)$. Hence, $\mathbf{D}_1(\mathbf{Q}'\mathbf{S}_2\mathbf{Q})^{-1}\mathbf{D}_1$ has the distribution of $\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1$, and the theorem follows.

We might note that nothing need be assumed about the distribution of $\mathbf{S}_1$ except that $\mathbf{S}_1$ is positive semi-definite. (If we replaced $\mathbf{D}_1^2$ by a diagonal matrix with possibly negative diagonal elements, we would only need to assume $\mathbf{S}_1$ symmetric to show the distribution of $\mathrm{ch}(\mathbf{S}_1\mathbf{S}_2^{-1})$ is the same as the distribution of $\mathrm{ch}(\mathbf{D}_1^2\mathbf{S}_2^{-1})$; we use $\mathbf{D}_1^2$ here to simplify notation later.) The only property of $\mathbf{S}_2$ that is used is that its distribution is invariant with respect to transformation to $\mathbf{Q}'\mathbf{S}_2\mathbf{Q}$.

THEOREM 3.2. *If* $\theta_1 \leqq \theta_1^+$, *then*

$$h(\theta_1^+, \theta_2, \cdots, \theta_p) \leqq h(\theta_1, \theta_2, \cdots, \theta_p); \tag{3.6}$$

if $\theta_p^- \leqq \theta_p$, *then*

$$h(\theta_1, \cdots, \theta_{p-1}, \theta_p^-) \leqq h(\theta_1, \cdots, \theta_{p-1}, \theta_p). \tag{3.7}$$

PROOF. Theorem 3.1 implies

$$\begin{aligned} h(\theta_1, \cdots, \theta_p) &= \Pr\{f[\mathrm{ch}_1(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1), \cdots, \mathrm{ch}_p(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1)] \leqq \delta_p^2, \\ &\qquad \delta_1^2 \leqq g[\mathrm{ch}_1(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1), \cdots, \mathrm{ch}_p(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1)]\}, \\ &= \Pr\{f[\mathrm{ch}_1\{(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)\}, \cdots, \mathrm{ch}_p\{(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)\}] \leqq 1, \\ &\qquad 1 \leqq g[\mathrm{ch}_1\{(\mathbf{D}_1/\delta_1)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_1)\}, \cdots, \mathrm{ch}_p\{(\mathbf{D}_1/\delta_1)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_1)\}]\}, \end{aligned} \tag{3.8}$$

where $\mathbf{D}_1$ consists of $[\mathrm{ch}(\mathbf{S}_1)]^{\frac{1}{2}}$, $\mathbf{S}_1 = \mathbf{\Delta}\mathbf{S}_1^*\mathbf{\Delta}$, $n_1\mathbf{S}_1^*$ has distribution $W(\mathbf{I}, n_1)$, and $\mathbf{\Delta}$ is a diagonal matrix with diagonal elements $\delta_i = \theta_i^{\frac{1}{2}}$. We note that $\mathbf{D}_1/\delta_p$ consists of $\{\mathrm{ch}[(\mathbf{\Delta}/\delta_p)\mathbf{S}_1^*(\mathbf{\Delta}/\delta_p)]\}^{\frac{1}{2}}$ and $\mathbf{D}_1/\delta_1$ consists of $\{\mathrm{ch}[(\mathbf{\Delta}/\delta_1)\mathbf{S}_1^*(\mathbf{\Delta}/\delta_1)]\}^{\frac{1}{2}}$. Let $\delta_1^+ = (\theta_1^+)^{\frac{1}{2}} \geqq \delta_1$, $\delta_p^- = (\theta^-)^{\frac{1}{2}} \leqq \delta_p$, $\mathbf{\Delta}^+$, be $\mathbf{\Delta}$ with δ_1 replaced by δ_1^+, and $\mathbf{\Delta}^-$ be $\mathbf{\Delta}$ with δ_p replaced by δ_p^-. Let $\mathbf{D}_1^+$ and $\mathbf{D}_1^-$ be derived from $\mathbf{S}_1^*$ by replacing $\mathbf{\Delta}$ by $\mathbf{\Delta}^+$ and $\mathbf{\Delta}^-$, respectively. Then the inequalities (2.7) to (2.10) hold with

S^* replaced by S_1^*. Then the ith diagonal element of D_1/δ_p is not greater than the ith diagonal element of D_1^-/δ_p^- [by (2.7)] or of D_1^+/δ_p [by (2.9)], and the ith diagonal element of D_1/δ_1 is not less than the ith diagonal element of D_1^-/δ_1 [by (2.8)] or of D_1^+/δ_1^+ [by (2.10)]. Application again of Lemma 2.1 yields for any S_1^* and S_2,

$$\text{ch}_i[(D_1^-/\delta_p^-)S_2^{-1}(D_1^-/\delta_p^-)] \geqq \text{ch}_i[(D_1/\delta_p)S_2^{-1}(D_1/\delta_p)], \tag{3.9}$$

$$\text{ch}_i[(D_1^-/\delta_1)S_2^{-1}(D_1^-/\delta_1)] \leqq \text{ch}_i[(D_1/\delta_1)S_2^{-1}(D_1/\delta_1)], \tag{3.10}$$

$$\text{ch}_i[(D_1^+/\delta_p)S_2^{-1}(D_1^+/\delta_p)] \geqq \text{ch}_i[(D_1/\delta_p)S_2^{-1}(D_1/\delta_p)], \tag{3.11}$$

$$\text{ch}_i[(D_1^+/\delta_1)S_2^{-1}(D_1^+/\delta_1)] \leqq \text{ch}_i[(D_1/\delta_1)S_2^{-1}(D_1/\delta_1)]. \tag{3.12}$$

Thus for any S_1^* we have the inclusions (2.11) and (2.12) with Δ, Δ^- and Δ^+ replaced by D_1, D_1^- and D_1^+, respectively and S^* replaced by S_2^{-1}. These imply Theorem 3.2.

THEOREM 3.3.

$$\begin{aligned} &\lim_{\theta_1\to\infty,\theta_p\to 0} h(\theta_1, \cdots, \theta_p) \\ &\quad = \Pr\{(n_2 - p + 1)/n_2 \leqq F_{n_1,n_2-p+1}g(1, 0, \cdots, 0)\} \\ &\qquad \cdot \Pr\{F_{n_1-p+1,n_2}f(\infty, \cdots, \infty, 1) \leqq n_1/(n_1 - p + 1)\}. \end{aligned} \tag{3.13}$$

PROOF. For any S_1^* it follows from (2.14) that

$$\lim_{\delta_1\to\infty}(D_1/\delta_1) = \begin{pmatrix} (s_{11}^{(1)*})^{\frac{1}{2}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{3.14}$$

where $s_{11}^{(1)*}$ is the upper left hand element of S_1^*. For any fixed S_1^* and S_2,

$$\lim_{\delta_1\to\infty}(D_1/\delta_1)S_2^{-1}(D_1/\delta_1) = \begin{pmatrix} s_{11}^{(1)*}s_{(2)}^{11} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{3.15}$$

and

$$\lim_{\delta_1\to\infty} \text{ch}_1[(D_1/\delta_1)S_2^{-1}(D_1/\delta_1)] = s_{11}^{(1)*}s_{(2)}^{11}, \tag{3.16}$$

$$\lim_{\delta_1\to\infty} \text{ch}_i[(D_1/\delta_1)S_2^{-1}(D_1/\delta_1)] = 0, \qquad i = 2, \cdots, p. \tag{3.17}$$

Similarly, for any S_1^* and S_2

$$\lim_{\delta_p\to 0} \delta_p D_1^{-1} = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & (s_{(1)}^{*pp})^{\frac{1}{2}} \end{pmatrix}, \tag{3.18}$$

$$(3.19)\qquad \lim_{\delta_p\to 0}(\delta_p \mathbf{D}_1^{-1})\mathbf{S}_2(\delta_p \mathbf{D}^{-1}) = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & s_{(1)}^{*pp}s_{pp}^{(2)} \end{pmatrix},$$

$$(3.20)\qquad \lim_{\delta_p\to 0} \text{ch}_p[(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)] = 1/\lim_{\delta_p\to 0} \text{ch}_M[(\delta_p\mathbf{D}_1^{-1})\mathbf{S}_2(\delta_p\mathbf{D}^{-1})] = 1/(s_{(1)}^{*pp}s_{pp}^{(2)}),$$

$$(3.21)\qquad \lim_{\delta_p\to\infty} \text{ch}_i[(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)] = 1/\lim_{\delta_p\to\infty} \text{ch}_{p+1-i}[(\delta_p\mathbf{D}_1^{-1})\mathbf{S}_2(\delta_p\mathbf{D}_1^{-1})] = \infty.$$

Thus for any $\mathbf{S}_1^*$

$$(3.22)\qquad \begin{aligned} \lim_{\delta_1\to\infty,\delta_p\to 0}\{\mathbf{S}_2 \mid f[\text{ch}_1\{(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)\}, \cdots, \text{ch}_p\{(\mathbf{D}_1/\delta_p)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_p)\}] \leqq 1, \\ 1 \leqq g[\text{ch}_1\{(\mathbf{D}_1/\delta_1)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_1)\}, \cdots, \text{ch}_p\{(\mathbf{D}_1/\delta_1)\mathbf{S}_2^{-1}(\mathbf{D}_1/\delta_1)\}]\} \\ = \{\mathbf{S}_2 \mid f[\infty, \cdots, \infty, 1/(s_{(1)}^{*pp}s_{pp}^{(2)})] \leqq 1 \leqq g[s_{11}^{(1)*}s_{(2)}^{11}, 0, \cdots, 0]\} \\ = \{\mathbf{S}_2 \mid f(\infty, \cdots, \infty, 1)/(s_{(1)}^{*pp}s_{pp}^{(2)}) \leqq 1 \leqq s_{11}^{(1)*}s_{(2)}^{11}g(1, 0, \cdots, 0)\}. \end{aligned}$$

The probability of the set on the right hand side of (3.22) relative to the distribution of $\mathbf{S}_1^*$ and $\mathbf{S}_2$ is obtained on the basis that $n_1s_{11}^{(1)*}$, $n_2/s_{(2)}^{11}$, $n_1/s_{(1)}^{*pp}$, and $n_2s_{pp}^{(2)}$ are independently distributed according to χ^2-distributions with n_1, $n_2 - p + 1$, $n_1 - p + 1$, and n_2 degrees of freedom, respectively. The probability is

$$(3.23)\qquad \begin{aligned} \Pr\{\mathbf{S}_1^*, \mathbf{S}_2 \mid 1 \leqq s_{11}^{(1)*}s_{(2)}^{11}g(1, 0, \cdots, 0), f(\infty, \cdots, \infty, 1)/(s_{(1)}^{*pp}s_{pp}^{(2)}) \leqq 1\} \\ = \Pr\{n_1/n_2 \leqq [n_1s_{11}^{(1)*}/(n_2/s_{(2)}^{11})]g(1, 0, \cdots, 0)\} \\ \cdot\Pr\{[(n_1/s_{(1)}^{*pp})/n_2s_{pp}^{(2)}]f(\infty, \cdots, \infty, 1) \leqq n_1/n_2\}, \end{aligned}$$

which is (3.13).

THEOREM 3.4. *If $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ are homogeneous functions of degree* 1 *and monotonically nondecreasing in each argument for the other arguments fixed* ($x_1 \geqq x_2 \geqq \cdots \geqq x_p \geqq 0$), *then the interval*

$$(3.24)\qquad f[\text{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \text{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})] \leqq \text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) \leqq g[\text{ch}_1(\mathbf{S}_1\mathbf{S}_2^{-1}), \cdots, \text{ch}_p(\mathbf{S}_1\mathbf{S}_2^{-1})]$$

is a confidence interval for $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ *of confidence* $1 - \epsilon$ *if and only if*

$$(3.25)\qquad \Pr\{(n_2 - p + 1)/n_2 \leqq F_{n_1,n_2-p+1}g(1, 0, \cdots, 0)\} \cdot\Pr\{F_{n_1-p+1,n_2}f(\infty, \cdots, \infty, 1) \leqq n_1/(n_1 - p + 1)\} \geqq 1 - \epsilon.$$

THEOREM 3.5. *If $f(x_1, \cdots, x_p)$ and $g(x_1, \cdots, x_p)$ are homogeneous functions of degree* 1 *and monotonically nondecreasing in each argument for the other arguments fixed* ($x_1 \geqq x_2 \geqq \cdots \geqq x_p \geqq 0$) *such that* (3.24) *holds with probability at least equal to* $1 - \epsilon$, *then there exist numbers L and U satisfying* (1.9) *such that*

(1.8) *holds with probability at least* $1 - \epsilon$ *and the interval* (1.8) *is contained within the interval* (3.24).

4. Discussion of bounds. Other bounds for $\text{ch}(\mathbf{\Sigma})$ based on χ^2-distributions can be derived fairly easily, but they are not as tight as those given in Section 2. Let $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_M$ be vectors such that

$$\boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\boldsymbol{\alpha}_m = \text{ch}_m(\mathbf{\Sigma}), \tag{4.1}$$

$$\boldsymbol{\alpha}_M'\mathbf{\Sigma}\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\boldsymbol{\alpha}_M = \text{ch}_M(\mathbf{\Sigma}) \tag{4.2}$$

and $\boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_M = 0 = \boldsymbol{\alpha}_m'\boldsymbol{\alpha}_M$. Let l' and u' be numbers such that

$$1 - \epsilon = \Pr\{nl' \leqq \chi_n^2\} \Pr\{\chi_n^2 \leqq nu'\}. \tag{4.3}$$

Since $\boldsymbol{\alpha}_m'\mathbf{X}$ and $\boldsymbol{\alpha}_M\mathbf{X}$ are uncorrelated and have variances $\boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_M'\mathbf{\Sigma}\boldsymbol{\alpha}_M$, respectively, $n\boldsymbol{\alpha}_m'\mathbf{S}\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_m$ and $n\boldsymbol{\alpha}_M'\mathbf{S}\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\mathbf{\Sigma}\boldsymbol{\alpha}_M$ are independently distributed according to χ^2-distributions with n degrees of freedom. Hence,

$$\begin{aligned} 1 - \epsilon &= \Pr\{l' \leqq \boldsymbol{\alpha}_M'\mathbf{S}\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\mathbf{\Sigma}\boldsymbol{\alpha}_M,\ \boldsymbol{\alpha}_m'\mathbf{S}\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_m \leqq u'\} \\ &= \Pr\{(1/u')\boldsymbol{\alpha}_m'\mathbf{S}\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\boldsymbol{\alpha}_m \leqq \boldsymbol{\alpha}_m'\mathbf{\Sigma}\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\boldsymbol{\alpha}_m, \\ &\qquad \boldsymbol{\alpha}_M'\mathbf{\Sigma}\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\boldsymbol{\alpha}_M \leqq (1/l')\boldsymbol{\alpha}_M'\mathbf{S}\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\boldsymbol{\alpha}_M\} \\ &\leqq \Pr\{(1/u')\min_{\mathbf{a}}(\mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a}) \leqq \text{ch}_m(\mathbf{\Sigma}), \\ &\qquad \text{ch}_M(\mathbf{\Sigma}) \leqq (1/l')\max_{\mathbf{a}}(\mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a})\} \end{aligned} \tag{4.4}$$

by (1.14). Thus confidence bounds of the form of (1.3) with confidence $1 - \epsilon$ can be obtained by choosing $u = u'$ and $l = l'$ to satisfy (4.3). However, since

$$\Pr\{\chi_n^2 \leqq nu\} < \Pr\{\chi_{n-p+1}^2 \leqq nu\} \tag{4.5}$$

for $p \geqq 2$, the bounds given in Section 2 [based on (1.4)] are tighter than those derived here.

The bounds that Roy (1954) has given are of the form (1.3) and are obtained from

$$\begin{aligned} 1 - \epsilon &= \Pr\{l^* \leqq \text{ch}_m(\mathbf{S}^*),\ \text{ch}_M(\mathbf{S}^*) \leqq u^*\} \\ &= \Pr\{l^* \leqq \min_{\mathbf{a}}(\mathbf{a}'\mathbf{S}^*\mathbf{a}/\mathbf{a}'\mathbf{a}),\ \max_{\mathbf{b}}(\mathbf{b}'\mathbf{S}^*\mathbf{b}/\mathbf{b}'\mathbf{b}) \leqq u^*\} \\ &= \Pr\{l^* \leqq \mathbf{a}'\mathbf{S}^*\mathbf{a}/\mathbf{a}'\mathbf{a} \leqq u^*, \text{ all } \mathbf{a}\} \\ &= \Pr\{l^* \leqq \mathbf{b}'\mathbf{S}\mathbf{b}/\mathbf{b}'\mathbf{\Sigma}\mathbf{b} \leqq u^*, \text{ all } \mathbf{b}\} \\ &= \Pr\{\mathbf{b}'\mathbf{S}\mathbf{b}/\mathbf{b}'\mathbf{b}u^* \leqq \mathbf{b}\mathbf{\Sigma}\mathbf{b}/\mathbf{b}'\mathbf{b} \leqq \mathbf{b}'\mathbf{S}\mathbf{b}/\mathbf{b}'\mathbf{b}l^*, \text{ all } \mathbf{b}\} \\ &= \Pr\{\text{ch}_m(\mathbf{S})/u^* \leqq \text{ch}(\mathbf{\Sigma}) \leqq \text{ch}_M(\mathbf{S})/l^*\}, \end{aligned} \tag{4.6}$$

where $\mathbf{\Sigma} = \mathbf{A}'\mathbf{A}$, $\mathbf{S} = \mathbf{A}'\mathbf{S}^*\mathbf{A}$, and $\mathbf{a} = \mathbf{A}\mathbf{b}$ for suitable $\mathbf{A}$. Comparison of (4.6) and (4.4) shows that either $l^* \leqq l'$, $u' \leqq u^*$ or both; hence bounds for $\text{ch}(\mathbf{\Sigma})$ based on (4.4) are tighter than those based on (4.6). Then the preceding paragraph implies that the bounds given in Section 2 are tighter than those of Roy.

TABLE 1

Constants for one-sided confidence limits for $\text{ch}(\mathbf{\Sigma})$

$p = 2, \quad 1 - \epsilon = .975$

n	$u = \infty$		$l = 0$		
	$l = l'$	l^*	u	u'	u^*
3	.0719	.00844	2.46	3.12	4.14
4	.121	.0304	2.34	2.79	3.61
5	.166	.0590	2.23	2.57	3.27
10	.325	.195	1.90	2.05	2.48
20	.479	.358	1.64	1.71	1.98
40	.611	.510	1.45	1.48	1.66
100	.742	.669	1.28	1.30	1.40

Table 1 records some numerical values of l, l', and l^* for $u = \infty$ (upper confidence bounds only) and of u, u', and u^* for $l = 0$ (lower confidence bounds only) when $1 - \alpha = .975$. Together u and l (or u' and l') yield two-sided confidence bounds at confidence level $(.975)^2 = .950625$. The values of u^* and l^* are taken from Thompson (1962). He also gives the value of u^* corresponding to l^* in Table 1 for two-sided confidence intervals with 95% confidence; these values are close to u^* in Table 1. (One reason that both upper and lower limits are developed in Sections 2 and 3 is that the statistical independence involved must be proved.) It will be observed that the optimal constants are considerably nearer 1 than the constants based on the distribution of roots. (It may be noted that as $n_2 \to \infty$, $U \to u$, $U' \to u'$ and $U^* \to u^*$.)

Confidence bounds for $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ based on F-distributions can be obtained more easily than in Section 3, but the result is not as good. Let $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_M$ be vectors such that

$$\boldsymbol{\alpha}_m'\mathbf{\Sigma}_1\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\mathbf{\Sigma}_2\boldsymbol{\alpha}_m = \text{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}), \tag{4.7}$$

$$\boldsymbol{\alpha}_M'\mathbf{\Sigma}_1\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\mathbf{\Sigma}_2\boldsymbol{\alpha}_M = \text{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}). \tag{4.8}$$

[See Chapter 12 of Anderson (1958), for example.] Then $\boldsymbol{\alpha}_m'\mathbf{X}$ and $\boldsymbol{\alpha}_M'\mathbf{X}$ are independent when $\mathbf{X}$ is distributed according to either $N(\boldsymbol{\mu}^{(1)}, \mathbf{\Sigma}_1)$ or $N(\boldsymbol{\mu}^{(2)}, \mathbf{\Sigma}_2)$ because $\boldsymbol{\alpha}_m'\mathbf{\Sigma}_1\boldsymbol{\alpha}_M = 0$ and $\boldsymbol{\alpha}_m'\mathbf{\Sigma}_2\boldsymbol{\alpha}_M = 0$. (These conditions must be satisfied except when $\mathbf{\Sigma}_1$ is proportional to $\mathbf{\Sigma}_2$ and then these conditions can be imposed.) It follows that

$$F_m = \frac{\boldsymbol{\alpha}_m'\mathbf{S}_1\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\mathbf{\Sigma}_1\boldsymbol{\alpha}_m}\Big/\frac{\boldsymbol{\alpha}_m'\mathbf{S}_2\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\mathbf{\Sigma}_2\boldsymbol{\alpha}_m}, \qquad F_M = \frac{\boldsymbol{\alpha}_M'\mathbf{S}_1\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\mathbf{\Sigma}_1\boldsymbol{\alpha}_M}\Big/\frac{\boldsymbol{\alpha}_M'\mathbf{S}_2\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\mathbf{\Sigma}_2\boldsymbol{\alpha}_M} \tag{4.9}$$

are independently distributed, each according to an F-distribution with n_1 and n_2 degrees of freedom. Let L' and U' be two numbers such that

$$
\begin{aligned}
1-\epsilon &= \Pr\{L' \leqq F_M\}\Pr\{F_m \leqq U'\} \\
&= \Pr\{L' \leqq F_M, F_m \leqq U'\} \\
&= \Pr\left\{L' \leqq \frac{\boldsymbol{\alpha}_M'\mathbf{S}_1\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\boldsymbol{\Sigma}_1\boldsymbol{\alpha}_M}\Big/\frac{\boldsymbol{\alpha}_M'\mathbf{S}_2\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\boldsymbol{\Sigma}_2\boldsymbol{\alpha}_M}, \frac{\boldsymbol{\alpha}_m'\mathbf{S}_1\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\boldsymbol{\Sigma}_1\boldsymbol{\alpha}_m}\Big/\frac{\boldsymbol{\alpha}_m'\mathbf{S}_2\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\boldsymbol{\Sigma}_2\boldsymbol{\alpha}_m} \leqq U'\right\} \qquad (4.10) \\
&= \Pr\left\{\frac{\boldsymbol{\alpha}_M'\boldsymbol{\Sigma}_1\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\boldsymbol{\Sigma}_2\boldsymbol{\alpha}_M} \leqq \frac{1}{L'}\cdot\frac{\boldsymbol{\alpha}_M'\mathbf{S}_1\boldsymbol{\alpha}_M}{\boldsymbol{\alpha}_M'\mathbf{S}_2\boldsymbol{\alpha}_M}, \frac{1}{U'}\cdot\frac{\boldsymbol{\alpha}_m'\mathbf{S}_1\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\mathbf{S}_2\boldsymbol{\alpha}_m} \leqq \frac{\boldsymbol{\alpha}_m'\boldsymbol{\Sigma}_1\boldsymbol{\alpha}_m}{\boldsymbol{\alpha}_m'\boldsymbol{\Sigma}_2\boldsymbol{\alpha}_m}\right\} \\
&\leqq \Pr\{\mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U' \leqq \mathrm{ch}_m(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}), \\
&\qquad\qquad \mathrm{ch}_M(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1}) \leqq \mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L'\}.
\end{aligned}
$$

Thus confidence bounds on $\mathrm{ch}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})$ of the form (1.8) can be obtained by choosing L' and U' to satisfy (4.10). The bounds are not as tight as those given in Section 3 because those F-distributions have n_1 and $n_2 - p + 1$ and $n_1 - p + 1$ and n_2 degrees of freedom instead of n_1 and n_2 degrees of freedom as in (4.10).

These bounds are of the same form as those of Roy and Gnanadesikan [also (14.10.9) and (14.10.12) of Roy (1957)]. Their bounds, however, are based on the distribution of the largest and smallest roots of $\mathbf{S}_1\mathbf{S}_2^{-1}$ when $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$; that is, $L = L^*$ and $U = U^*$ are chosen so

$$
\begin{aligned}
1-\epsilon &= \Pr\{L^* \leqq \mathrm{ch}_m(\mathbf{S}_1^*\mathbf{S}_2^{*-1}), \mathrm{ch}_M(\mathbf{S}_1^*\mathbf{S}_2^{*-1}) \leqq U^*\} \qquad (4.11) \\
&= \Pr\{L^* \leqq \min_{\mathbf{a}}(\mathbf{a}'\mathbf{S}_1^*\mathbf{a}/\mathbf{a}'\mathbf{S}_2^*\mathbf{a}), \max_{\mathbf{b}}(\mathbf{b}'\mathbf{S}_1^*\mathbf{b}/\mathbf{b}'\mathbf{S}_2^*\mathbf{b}) \leqq U^*\},
\end{aligned}
$$

where $\mathbf{S}_1^*$ and $\mathbf{S}_2^*$ represent $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively, when $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \mathbf{I}$. Since (4.10) can be written

$$
1-\epsilon = \Pr\{L' \leqq \mathbf{a}'\mathbf{S}_1^*\mathbf{a}/\mathbf{a}'\mathbf{S}_2^*\mathbf{a}, \mathbf{b}'\mathbf{S}_1^*\mathbf{b}/\mathbf{b}'\mathbf{S}_2^*\mathbf{b} \leqq U'\} \qquad (4.12)
$$

for any $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a}'\mathbf{b} = 0$, we see that $L^* \leqq L'$, $U' \leqq U^*$, or both. These facts and the discussion of the preceding paragraph imply that the bounds of Section 3 are tighter than the bounds with L^* and U^*.

Tables 2 and 3 give some examples of the numerical values of the constants.

TABLE 2

Constants for one-sided confidence limits for $\mathrm{ch}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1})$

$p = 2, \quad 1 - \epsilon = .95$

n_2	$n_1 = 3$			$n_1 = 5$		
	U	U'	U^*	U	U'	U^*
13	2.54	3.41	5.63	2.54	3.03	4.85
23	2.28	3.03	4.58	2.24	2.64	3.84
120	2.05	2.68		1.96	2.29	
2003			3.59			2.90
∞	2.00	2.60		1.90	2.21	

TABLE 3

Constants for one-sided confidence limits for $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$

$p = 6, \quad 1 - \epsilon = .95$

n_2	$n_1 = 7$			$n_1 = 11$		
	U	U'	U^*	U	U'	U^*
17	1.026	2.61	11.42	1.47	2.41	10.21
27	.958	2.37	7.33	1.34	2.17	6.33
120	.878	2.09		1.19	1.87	
2007			4.06			3.29
∞	.856	2.01		1.14	1.79	

The values of U^* are from Pillai (1960). (The degrees of freedom n_2 are limited; lower significance points L^* are not given, but they could be derived by interchanging $\mathbf{S}_1$ and $\mathbf{S}_2$ and L^* and U^*.) It will be noted that as p increases the advantage of the optimum bounds increases.

S. N. Roy has informed the author while this paper was in preparation that he had given the bounds based on (4.10) in unpublished lecture notes of 1952 and had verified that the upper bound alone or the lower bound alone based on the F-distribution was better than the corresponding one-sided bound based on the distribution of the maximum or minimum root of $\mathbf{S}_1^*\mathbf{S}_2^{*-1}$.

In the situation considered here—as is usually the case—tests of hypotheses can be obtained from confidence regions; a null hypothesis is rejected if the hypothesized parameter values are not contained in the confidence region. Here the hypothesis $\mathbf{\Sigma} = \mathbf{I}$ is rejected if $\text{ch}_m(\mathbf{\Sigma}) = \text{ch}_M(\mathbf{\Sigma}) = 1$ is not in the confidence bounds, and the hypothesis $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ is rejected if $\text{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) = \text{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) = 1$ is not in the confidence bounds. In the latter case, for instance, the hypothesis is rejected if

$$\text{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1}) > U \tag{4.13}$$

or if

$$\text{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1}) < L; \tag{4.14}$$

that is, the hypothesis is rejected if either all roots are large or if all roots are small. The test does not seem to have good power against alternatives in which only $\text{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ is large or in which only $\text{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ is small; yet these are the parameter values which were crucial in establishing the confidence coefficients.

It is curious that the rejection region implied by the confidence interval consists of inequalities similar to those of the acceptance region of the test associated with the original probability statement. For instance, a test at significance level ϵ based on (4.11) has the acceptance region

$$L^* \leqq \text{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1}), \qquad \text{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1}) \leqq U^*. \tag{4.15}$$

5. Some uses of confidence bounds. The confidence bounds for $\text{ch}(\mathbf{\Sigma})$ can be written

$$(5.1)\qquad \mathrm{ch}_m(\mathbf{S})/u \leqq \mathbf{a}'\mathbf{\Sigma}\mathbf{a}/\mathbf{a}'\mathbf{a} \leqq \mathrm{ch}_M(\mathbf{S})/l$$

for all vectors **a** or in the form of (1.15). When the ith component of **a** is 1 and the other components are 0, (5.1) is

$$(5.2)\qquad \mathrm{ch}_m(\mathbf{S})/u \leqq \sigma_{ii} \leqq \mathrm{ch}_M(\mathbf{S})/l.$$

Setting the ith component of **a** equal to 1, the jth component equal to 1 and -1 alternately, and the other components equal to 0 gives

$$(5.3)\qquad 2\mathrm{ch}_m(\mathbf{S})/u \leqq \sigma_{ii} + \sigma_{jj} + 2\sigma_{ij} \leqq 2\mathrm{ch}_M(\mathbf{S})/l, \qquad i \neq j,$$

$$(5.4)\qquad 2\mathrm{ch}_m(\mathbf{S})/u \leqq \sigma_{ii} + \sigma_{jj} - 2\sigma_{ij} \leqq 2\mathrm{ch}_M(\mathbf{S})/l, \qquad i \neq j.$$

From these last two pairs of inequalities we deduce

$$(5.5)\qquad -\tfrac{1}{2}[\mathrm{ch}_M(\mathbf{S})/l - \mathrm{ch}_m(\mathbf{S})/u] \leqq \sigma_{ij} \leqq \tfrac{1}{2}[\mathrm{ch}_M(\mathbf{S})/l - \mathrm{ch}_m(\mathbf{S})/u], \qquad i \neq j.$$

Thus we arrive at simultaneous confidence bounds for all elements of the covariance matrix $\mathbf{\Sigma}$.

A feature of the approach of this paper is that the bounds for $\mathbf{a}'\mathbf{\Sigma}\mathbf{a}/\mathbf{a}'\mathbf{a}$ do not depend on **a** and only depend on the observed sample covariance matrix **S** through the two functions $\mathrm{ch}_m(\mathbf{S})$ and $\mathrm{ch}_M(\mathbf{S})$. A consequence is that the bounds for σ_{ii} are the same for all i and the bounds for σ_{ij}, $i \neq j$, are the same for all pairs i and j. While these facts imply computational advantages, they also imply the inferential disadvantages of not reflecting closely the characteristics of the sample covariance matrix.

On the other hand the approach by Roy and Gnanadesikan and other colleagues of Roy leads to bounds that depend on **a** and the sample quantities in more detail. From (4.6) we see that

$$(5.6)\qquad \mathbf{a}'\mathbf{S}\mathbf{a}/u^* \leqq \mathbf{a}'\mathbf{\Sigma}\mathbf{a} \leqq \mathbf{a}'\mathbf{S}\mathbf{a}/l^*$$

for all **a** with (exact) confidence $1 - \epsilon$. From (5.6) we deduce the following bounds which hold simultaneously:

$$(5.7)\qquad s_{ii}/u^* \leqq \sigma_{ii} \leqq s_{ii}/l^*,$$

$$(5.8)\qquad (a_i^2 s_{ii} + a_j^2 s_{jj} + 2a_i a_j s_{ij})/u^* \leqq a_i^2\sigma_{ii} + a_j^2\sigma_{jj} + 2a_i a_j \sigma_{ij} \leqq (a_i^2 s_{ii} + a_j^2 s_{jj} + 2a_i a_j s_{ij})/l^*, \quad i \neq j,$$

$$(5.9)\qquad (a_i^2 s_{ii} + a_j^2 s_{jj} - 2a_i a_j s_{ij})/u^* \leqq a_i^2\sigma_{ii} + a_j^2\sigma_{jj} - 2a_i a_j \sigma_{ij} \leqq (a_i^2 s_{ii} + a_j^2 s_{jj} - 2a_i a_j s_{ij})/l^*, \quad i \neq j,$$

$$(5.10)\qquad \tfrac{1}{2}(l^{*-1} + u^{*-1})s_{ij} - (l^{*-1} - u^{*-1})(a_i^2 s_{ii} + a_j^2 s_{jj})/4a_i a_j \leqq \sigma_{ij} \leqq \tfrac{1}{2}(l^{*-1} + u^{*-1})s_{ij} + (l^{*-1} + u^{*-1})(a_i^2 s_{ii} + a_j^2 s_{jj})/4a_i a_j, \quad a_i a_j > 0,\ i \neq j.$$

The length of the interval (5.10) is minimized by taking $a_i = (s_{jj})^{\frac{1}{2}}$ and $a_j = (s_{ii})^{\frac{1}{2}}$; the resulting interval is

$$(5.11)\quad \tfrac{1}{2}(l^{*-1} + u^{*-1})s_{ij} - (l^{*-1} - u^{*-1})(s_{ii}s_{jj})^{\frac{1}{2}}/2 \leqq \sigma_{ij} \leqq \tfrac{1}{2}(l^{*-1} + u^{*-1})s_{ij} + (l^{*-1} - u^{*-1})(s_{ii}s_{jj})^{\frac{1}{2}}/2.$$

If we set $a_i = 1/(\sigma_{ii})^{\frac{1}{2}}$ in (5.8) and (5.9) and apply (5.7) we obtain for $r_{ij} \geqq 0$,

$$(5.12)\quad (l^*/u^*)r_{ij} + l^*/u^* - 1 \leqq \rho_{ij} \leqq r_{ij} + 1 - l^*/u^*,$$

where r_{ij} and ρ_{ij} are the i, jth sample and population correlations, respectively; if $r_{ij} < 0$, the interval is

$$(5.13)\quad r_{ij} + l^*/u^* - 1 \leqq \rho_{ij} \leqq (l^*/u^*)r_{ij} + 1 - l^*/u^*.$$

Thompson (1962) has given (5.7) and (5.11), has shown that the set of $\mathbf{\Sigma}$ satisfying (5.6) is the intersection of two convex cones with vertices at $\mathbf{S}/l^*$ and $\mathbf{S}/u^*$, and has given tables for $p = 2$.

Comparison of (5.1) and (5.6) shows that the lower bound of (5.1) is better for $\mathbf{a}'\mathbf{\Sigma a}$ when

$$(5.14)\quad \mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a} < (u^*/u)\,\mathrm{ch}_m(\mathbf{S})$$

and the upper bound of (5.1) is better when

$$(5.15)\quad (l^*/l)\,\mathrm{ch}_M(\mathbf{S}) < \mathbf{a}'\mathbf{S}\mathbf{a}/\mathbf{a}'\mathbf{a}.$$

If $\mathrm{ch}_m(\mathbf{S}) > ul^*\,\mathrm{ch}_M(\mathbf{S})/(u^*l)$, then (5.1) is better for all $\mathbf{a}$; otherwise there will be some $\mathbf{a}$ for which one or both bounds of (5.6) are better.

Confidence bounds for ch $(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ can be written

$$(5.16)\quad \mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U \leqq \mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}/\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a} \leqq \mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L$$

for all $\mathbf{a}$ or in the form of (1.20). This is formally similar to (5.6) with $\mathbf{\Sigma}$ replaced by $\mathbf{\Sigma}_1$, $\mathbf{S}$ replaced by $\mathbf{\Sigma}_2$, $1/u^*$ replaced by $\mathrm{ch}_m\,(\mathbf{S}_1\mathbf{S}_2^{-1})/U$ and $1/l^*$ replaced by $\mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L$. The inequalities (5.7) to (5.13) can be translated into these terms to obtain inequalities on $\sigma_{ij}^{(1)}$ and $\rho_{ij}^{(1)}$ in terms of $\sigma_{ij}^{(2)}$ and $\rho_{ij}^{(2)}$, respectively. The analogue of (5.7), for example, is

$$(5.17)\quad \sigma_{ii}^{(2)}\,\mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})/U \leqq \sigma_{ii}^{(1)} \leqq \sigma_{ii}^{(2)}\,\mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})/L.$$

Similarly, (4.11) gives bounds of the form (5.16) with U and L replaced by U^* and L^*, respectively. Roy (1958) has shown that (5.16) simultaneously holds when pairs of corresponding rows and columns are deleted from $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$, $\mathbf{S}_1$, and $\mathbf{S}_2$ [as given by Roy and Gnanadesikan (1957)].

Other bounds can be derived from (5.6), which yields the confidence statement

$$(5.18)\quad l^* \leqq \mathrm{ch}(\mathbf{S}\mathbf{\Sigma}^{-1}) \leqq u^*.$$

For example, the inequality (2.12) of Anderson and Das Gupta (1963) can be stated

$$(5.19)\quad \mathrm{ch}_i(\mathbf{S})\,\mathrm{ch}_p(\mathbf{\Sigma}\mathbf{S}^{-1}) \leqq \mathrm{ch}_i(\mathbf{\Sigma}) \leqq \mathrm{ch}_i(\mathbf{S})\,\mathrm{ch}_1(\mathbf{\Sigma}\mathbf{S}^{-1})$$

or equivalently

$$(5.20)\quad \mathrm{ch}_i(\mathbf{S})/\mathrm{ch}_1(\mathbf{S}\mathbf{\Sigma}^{-1}) \leqq \mathrm{ch}_i(\mathbf{\Sigma}) \leqq \mathrm{ch}_i(\mathbf{S})/\mathrm{ch}_p(\mathbf{S}\mathbf{\Sigma}^{-1}).$$

Then (5.18) implies the confidence statement (pointed out to me by Das Gupta)

$$\text{(5.21)} \qquad \mathrm{ch}_i(\mathbf{S})/u^* \leqq \mathrm{ch}_i(\mathbf{\Sigma}) \leqq \mathrm{ch}_i(\mathbf{S})/l^*.$$

6. Confidence bounds for some ordered roots in a two-sample problem. We now consider a lower bound for $\mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ and an upper bound for $\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$. Here let L and U be two numbers such that

$$\text{(6.1)} \quad 1 - \epsilon = \Pr\{L \leqq F_{n_1,n_2} \leqq U\} = \Pr\left\{L \leqq \frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}} \Big/ \frac{\mathbf{a}'\mathbf{S}_2\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}} \leqq U\right\},$$

where $\mathbf{a}$ is an arbitrary vector. Then

$$\begin{aligned}1 - \epsilon &= \Pr\left\{\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}} \leqq \frac{1}{L}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}},\ \frac{1}{U}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}} \leqq \frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}}\right\} \\ \text{(6.2)} \qquad &\leqq \Pr\left\{\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) = \min_{\mathbf{a}}\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}} \leqq \frac{1}{L}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}},\right. \\ &\qquad \left.\frac{1}{U}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}} \leqq \max_{\mathbf{a}}\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_2\mathbf{a}} = \mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})\right\}.\end{aligned}$$

Thus simultaneous upper bound on $\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ and lower bound on $\mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ of confidence at least $1 - \epsilon$ are

$$\text{(6.3)} \qquad \mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}) \leqq \frac{1}{L}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}},\ \frac{1}{U}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{S}_2\mathbf{a}} \leqq \mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}).$$

Single bounds can be obtained by setting $L = 0$ or $U = \infty$.

It appears that the bound for $\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ is a good one if $\mathbf{a}$ is close to $\boldsymbol{\alpha}_m$ and that the bound for $\mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ is good if $\mathbf{a}$ is close to $\boldsymbol{\alpha}_M$. More precisely, if $L = 0$, the left hand inequality of (6.3) is vacuous, and if $\mathbf{a} = \boldsymbol{\alpha}_M$, the probability of the right hand inequality is exactly $1 - \epsilon$; correspondingly if $U = \infty$, the right hand inequality is vacuous, and if $\mathbf{a} = \boldsymbol{\alpha}_m$, the probability of the left hand inequality is exactly $1 - \epsilon$. The arbitrariness of the vector $\mathbf{a}$ could be eliminated by replacing $\mathbf{a}'\mathbf{S}_1\mathbf{a}/\mathbf{a}'\mathbf{S}_2\mathbf{a}$ by $\mathrm{ch}_M(\mathbf{S}_1\mathbf{S}_2^{-1})$ in the bound for $\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$ and by $\mathrm{ch}_m(\mathbf{S}_1\mathbf{S}_2^{-1})$ in the bound for $\mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1})$, but these would give the worst possible bounds (that is, correspond to the worst selection of $\mathbf{a}$). Another way of eliminating the arbitrariness is by a randomized choice of $\mathbf{a}$.

Another result of this approach follows from

$$\begin{aligned}1 - \epsilon &= \Pr\left\{L \leqq \frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}} \Big/ \frac{\mathbf{b}'\mathbf{S}_2\mathbf{b}}{\mathbf{b}'\mathbf{\Sigma}_2\mathbf{b}} \leqq U\right\} \\ \text{(6.4)} \qquad &= \Pr\left\{\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{b}'\mathbf{\Sigma}_2\mathbf{b}} \leqq \frac{1}{L}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{b}'\mathbf{S}_2\mathbf{b}},\ \frac{1}{U}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{b}'\mathbf{S}_2\mathbf{b}} \leqq \frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{b}'\mathbf{\Sigma}_2\mathbf{b}}\right\} \\ &\leqq \Pr\left\{\frac{\mathrm{ch}_m(\mathbf{\Sigma}_1)}{\mathrm{ch}_M(\mathbf{\Sigma}_2)} = \min_{\mathbf{a}}\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{a}} \Big/ \max_{\mathbf{b}}\frac{\mathbf{b}'\mathbf{\Sigma}_2\mathbf{b}}{\mathbf{b}'\mathbf{b}} \leqq \frac{1}{L}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{b}'\mathbf{S}_2\mathbf{b}}\cdot\frac{\mathbf{b}'\mathbf{b}}{\mathbf{a}'\mathbf{a}},\right. \\ &\qquad \left.\frac{1}{U}\cdot\frac{\mathbf{a}'\mathbf{S}_1\mathbf{a}}{\mathbf{b}'\mathbf{S}_2\mathbf{b}}\cdot\frac{\mathbf{b}'\mathbf{b}}{\mathbf{a}'\mathbf{a}} \leqq \max_{\mathbf{a}}\frac{\mathbf{a}'\mathbf{\Sigma}_1\mathbf{a}}{\mathbf{a}'\mathbf{a}} \Big/ \min_{\mathbf{b}}\frac{\mathbf{b}'\mathbf{\Sigma}_2\mathbf{b}}{\mathbf{b}'\mathbf{b}} = \frac{\mathrm{ch}_M(\mathbf{\Sigma}_1)}{\mathrm{ch}_m(\mathbf{\Sigma}_2)}\right\}.\end{aligned}$$

Some similar inequalities were given in Appendix C of Anderson (1963).

7. Monotonicity properties of some power functions. Let ω be a set of $c_1, \cdots, c_p$ such that when a point $(c_1, \cdots, c_p)$ is in ω so is every point $(c_1^-, \cdots, c_p^-)$ with $c_i^- \leqq c_i$, $i = 1, \cdots, p$. It was shown by Anderson and Das Gupta (1964) that if $c_i = \text{ch}_i(\mathbf{S})$, $i = 1, \cdots, p$, and ω is the acceptance region of a test of the hypothesis that $\mathbf{\Sigma} = \mathbf{I}$, then the power function of that test is monotonically increasing in each characteristic root of $\mathbf{\Sigma}$. It was also shown that if $c_i = \text{ch}_i(\mathbf{S}_1\mathbf{S}_2^{-1})$, $i = 1, \cdots, p$, and ω is the acceptance region of the hypothesis that $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$, then the power function of that test is monotonically increasing in each characteristic root of $\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}$. The methods of the present paper afford the basis of an alternative proof of the second result.

Lemma 2.1 implies that

$$\{\mathbf{S}^* \mid \text{ch}(\mathbf{\Delta}^+\mathbf{S}^*\mathbf{\Delta}^+) \ \varepsilon\ \omega\} \subset \{\mathbf{S}^* \mid \text{ch}(\mathbf{\Delta}\mathbf{S}^*\mathbf{\Delta}) \ \varepsilon\ \omega\} \tag{7.1}$$

for $\mathbf{\Delta}$ and $\mathbf{\Delta}^+$ diagonal with $\delta_i \leqq \delta_i^+$; hence, the probability of the left hand side is less than that of the right hand side. Theorem 1 of Anderson and Das Gupta (1964) and the monotonicity of the power function were deduced in this fashion.

Lemma 2.1 similarly implies that if $\mathbf{D}_1$ has as diagonal elements $[\text{ch}_i(\mathbf{\Delta}\mathbf{S}_1^*\mathbf{\Delta})]^{\frac{1}{2}}$ and $\mathbf{D}_1^+$ has as diagonal elements $[\text{ch}_i(\mathbf{\Delta}^+\mathbf{S}_1^*\mathbf{\Delta}^+)]^{\frac{1}{2}}$, then each diagonal element of $\mathbf{D}_1^+$ is at least as large as the corresponding element of $\mathbf{D}_1$ and hence that

$$\{\mathbf{S}_2 \mid \text{ch}(\mathbf{D}_1^+\mathbf{S}_2^{-1}\mathbf{D}_1^+) \ \varepsilon\ \omega\} \subset \{\mathbf{S}_2 \mid \text{ch}(\mathbf{D}_1\mathbf{S}_2^{-1}\mathbf{D}_1) \ \varepsilon\ \omega\} \tag{7.2}$$

for each $\mathbf{S}_1^*$. From this fact and Theorem 3.1 we can deduce Theorem 2 of Anderson and Das Gupta (1964) and the monotoncity of the power function.

The results and proofs are not stated fully here because they are treated in detail in the paper referred to above and Theorem 2 is proved more directly.

8. Confidence bounds for all roots in a problem of several samples. A problem that Gnanadesikan has studied [Gnanadesikan (1959) and (1960)] is that of confidence bounds simultaneously on $\text{ch}(\mathbf{\Sigma}_1\mathbf{\Sigma}_0^{-1}), \cdots, \text{ch}(\mathbf{\Sigma}_k\mathbf{\Sigma}_0^{-1})$ based on covariance matrices $\mathbf{S}_0, \mathbf{S}_1, \cdots, \mathbf{S}_k$ of samples of sizes $n_0 + 1, n_1 + 1, \cdots, n_k + 1$ from normal distributions with covariance matrices $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \cdots, \mathbf{\Sigma}_k$, respectively. These roots are a complete set of invariants under linear transformations only if there exists a nonsingular matrix $\mathbf{A}$ such that $\mathbf{A\Sigma}_0\mathbf{A}', \cdots, \mathbf{A\Sigma}_k\mathbf{A}'$ are diagonal. We can deal with a slightly more general case in which there exist vectors $\boldsymbol{\alpha}_m$ and $\boldsymbol{\alpha}_M$ such that $\boldsymbol{\alpha}_m'\mathbf{\Sigma}_i\boldsymbol{\alpha}_m/\boldsymbol{\alpha}_m'\mathbf{\Sigma}_0\boldsymbol{\alpha}_m = \text{ch}_m(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1})$, $\boldsymbol{\alpha}_M'\mathbf{\Sigma}_i\boldsymbol{\alpha}_M/\boldsymbol{\alpha}_M'\mathbf{\Sigma}_0\boldsymbol{\alpha}_M = \text{ch}_M(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1})$, $i = 1, \cdots, k$. Let L_i and $U_i (i = 1, \cdots, k)$ be a set of numbers such that

$$1 - \epsilon = \Pr\{L_i \leqq s_i^2/s_0^2, i = 1, \cdots, k\} \Pr\{s_i^2/s_0^2 \leqq U_i, i = 1, \cdots, k\}, \tag{8.1}$$

where $n_is_i^2$ are distributed independently according to χ^2-distributions with n_i degrees of freedom, respectively. From this we deduce the following simultaneous bounds:

$$\text{ch}_m(\mathbf{S}_i\mathbf{S}_0^{-1})/U_i \leqq \text{ch}_m(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}), \quad \text{ch}_M(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}) \leqq \text{ch}_M(\mathbf{S}_i\mathbf{S}_0^{-1})/L_i, \quad i = 1, \cdots, k, \tag{8.2}$$

which hold with confidence at least $1 - \epsilon$. Values of U_i for which the second factor in (8.1) is a specified probability can be determined from Nair's tables [Pearson and Hartley (1958), p. 164] in some cases.

Gnandesikan's bounds are

$$(8.3)\quad (1/U_i^*)\cdot \mathrm{ch}_m(\mathbf{S}_i)/\mathrm{ch}_M(\mathbf{S}_0) \leqq \mathrm{ch}_m(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}),$$

$$\mathrm{ch}_M(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}) \leqq (1/L_i^*)\cdot \mathrm{ch}_M(\mathbf{S}_i)/\mathrm{ch}_m(\mathbf{S}_0), \quad i = 1, \cdots, k,$$

where L_i^* and $U_i^*(i = 1, \cdots, k)$ are a set of numbers such that

$$(8.4)\quad 1 - \epsilon = \Pr\left\{L_i^* \leqq \min_{\mathbf{a}} \frac{\mathbf{a}'\mathbf{S}_i^*\mathbf{a}}{\mathbf{a}'\mathbf{a}} \Big/ \max_{\mathbf{b}} \frac{\mathbf{b}'\mathbf{S}_0^*\mathbf{b}}{\mathbf{b}'\mathbf{b}},\right.$$

$$\left.\max_{\mathbf{a}} \frac{\mathbf{a}'\mathbf{S}_i^*\mathbf{a}}{\mathbf{a}'\mathbf{a}} \Big/ \min_{\mathbf{b}} \frac{\mathbf{b}'\mathbf{S}_0\mathbf{b}}{\mathbf{b}'\mathbf{b}} \leqq U_i^*, i = 1, \cdots, k\right\}.$$

These bounds, which do not require the restrictive conditions used in obtaining (8.2), follow from the confidence bounds

$$(8.5)\quad (1/U_i^*)\cdot \mathrm{ch}_m(\mathbf{S}_i)/\mathrm{ch}_M(\mathbf{S}_0) \leqq \mathrm{ch}_m(\mathbf{\Sigma}_i)/\mathrm{ch}_M(\mathbf{\Sigma}_0),$$

$$\mathrm{ch}_M(\mathbf{\Sigma}_i)/\mathrm{ch}_m(\mathbf{\Sigma}_0) \leqq (1/L_i^*)\cdot \mathrm{ch}_M(\mathbf{S}_i)/\mathrm{ch}_m(\mathbf{S}_0), \quad i = 1, \cdots, k,$$

and the inequalities

$$(8.6)\quad \mathrm{ch}_m(\mathbf{\Sigma}_i)/\mathrm{ch}_M(\mathbf{\Sigma}_0) \leqq \mathrm{ch}_m(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}), \mathrm{ch}_M(\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1}) \leqq \mathrm{ch}_M(\mathbf{\Sigma}_i)/\mathrm{ch}_m(\mathbf{\Sigma}_0).$$

The bounds (8.3), however, can be very ineffective because the inequalities (8.6) can be very ineffective. For example, if

$$(8.7)\qquad \mathbf{\Sigma}_0 = \mathbf{\Sigma}_1 = \begin{pmatrix} \theta & 0 \\ 0 & \varphi \end{pmatrix}$$

with $\theta > \varphi$, then $\mathrm{ch}_m(\mathbf{\Sigma}_1\mathbf{\Sigma}_0^{-1}) = \mathrm{ch}_M(\mathbf{\Sigma}_1\mathbf{\Sigma}_0^{-1}) = 1$, but $\mathrm{ch}_m(\mathbf{\Sigma}_1)/\mathrm{ch}_M(\mathbf{\Sigma}_0) = \varphi/\theta$ and $\mathrm{ch}_M(\mathbf{\Sigma}_1)/\mathrm{ch}_m(\mathbf{\Sigma}_0) = \theta/\varphi$; these ratios can be arbitrarily small and large, respectively. Similar difficulties hold for the test procedures derived from Gnanadesikan's bounds.

REFERENCES

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

ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34** 122–148.

ANDERSON, T. W. and DAS GUPTA, S. (1963). Some inequalities on the characteristic roots of matrices. *Biometrika* **50** 522–523.

ANDERSON, T. W. and DAS GUPTA, S. (1964). A monotonicity property of the power functions of some tests of the equality of two covariance matrices. *Ann. Math. Statist.* **35** 1059–1063.

GNANADESIKAN, R. (1959). Equality of more than two variances and more than two dispersion matrices against certain alternatives. *Ann. Math. Statist.* **30** 177–184.

GNANADESIKAN, R. (1960). Correction to and comment on "Equality of more than two variances and of more than two dispersion matrices against certain alternatives." *Ann. Math. Statist.* **31** 227–228.

PEARSON, E. S. and HARTLEY, H. O. (1958). *Biometrika Tables for Statisticians.* **1** (Second Edition) Cambridge Univ. Press.

PILLAI, K. C. S. (1960). *Statistical Tables for Tests of Multivariate Hypotheses.* The Statistical Center, University of the Philippines, Manila.

ROY, S. N. (1954). Some further results in simultaneous confidence interval estimation. *Ann. Math. Statist.* **25** 752–761.

ROY, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

ROY, S. N. (1958). A note on confidence bounds connected with the hypothesis of equality of two dispersion matrices. Institute of Statistics Mimeograph Series No. 190, University of North Carolina.

ROY, S. N. and GNANADESIKAN, R. (1957). Further contributions to multivariate confidence bounds. *Biometrika* **44** 399–410. (Corrigenda (1961) *Biometrika* **48** 474.)

THOMPSON, W. A., JR. (1962). Estimation of dispersion parameters. *Journ. Res. Nat. Bur. Standards* **66B** 161–164.