The Annals of Statistics
1980, Vol. 8, No. 2, 247–263

UNBIASEDNESS OF THE LIKELIHOOD RATIO TESTS FOR EQUALITY OF SEVERAL COVARIANCE MATRICES AND EQUALITY OF SEVERAL MULTIVARIATE NORMAL POPULATIONS[1]

BY MICHAEL D. PERLMAN

University of Washington

Random samples of size N_α are taken from the p-variate normal populations $N_p(\mu^{(\alpha)}, \Sigma^{(\alpha)})$, $1 \leq \alpha \leq k$, with $\mu^{(\alpha)}$ and $\Sigma^{(\alpha)}$ unknown. Bartlett's modification of the likelihood ratio test (LRT) for the hypothesis H_1: $\Sigma^{(1)} = \cdots = \Sigma^{(k)}$ rejects H_1 for large values of $|S|^n/\Pi|S^{(\alpha)}|^{n_\alpha}$, where $S = \Sigma S^{(\alpha)}$, $n_\alpha = N_\alpha - 1$, $n = \Sigma n_i$, and $S^{(\alpha)}$ is the sample covariance matrix from the αth population. The (unmodified) LRT for the hypothesis H_1: $\mu^{(1)} = \cdots = \mu^{(k)}$, $\Sigma^{(1)} = \cdots = \Sigma^{(k)}$ rejects H_2 for large values of $|S + T|^N/\Pi|S^{(\alpha)}|^{N_\alpha}$, where $N = \Sigma N_\alpha$, $T = \Sigma N_\alpha(\bar{X}^{(\alpha)} - \bar{X}^{(+)})(\bar{X}^{(\alpha)} - \bar{X}^{(+)})'$, $\bar{X}^{(\alpha)}$ is the αth sample mean, and $\bar{X}^{(+)}$ is the grand mean. It is proved that each of these tests is unbiased against all alternatives.

1. Introduction. Let $\{X_\beta^{(\alpha)}, 1 \leq \beta \leq N_\alpha\}$ be a random sample of size N_α from the p-dimensional multivariate normal distribution $N_p(\mu^{(\alpha)}, \Sigma^{(\alpha)})$, $1 \leq \alpha \leq k$, with unknown mean vector $\mu^{(\alpha)}$ and unknown covariance matrix $\Sigma^{(\alpha)}$, assumed positive definite. We take $X_\beta^{(\alpha)}$ and $\mu^{(\alpha)}$ to be column vectors of dimension $p \times 1$. The sample mean and sample covariance matrix from the αth sample are

$$\bar{X}^{(\alpha)} = \frac{1}{N_\alpha}\Sigma_{\beta=1}^{N_\alpha} X_\beta^{(\alpha)} \sim N_p\left(\mu^{(\alpha)}, \frac{1}{N_\alpha}\Sigma^{(\alpha)}\right),$$

$$S^{(\alpha)} = \Sigma_{\beta=1}^{N_\alpha}\left(X_\beta^{(\alpha)} - \bar{X}^{(\alpha)}\right)\left(X_\beta^{(\alpha)} - \bar{X}^{(\alpha)}\right)' \sim W_p\left(n_\alpha, \Sigma^{(\alpha)}\right)$$

respectively, where $n_\alpha = N_\alpha - 1$ and $W_p(n, \Sigma)$ denotes the p-dimensional Wishart distribution with n degrees of freedom and expected value $n\Sigma$. We assume that each $n_\alpha \geq p$. We shall show that Bartlett's modification of the likelihood ratio test (LRT) for the hypothesis

$$H_1: \Sigma^{(1)} = \cdots = \Sigma^{(k)}$$

is unbiased, and that the (unmodified) LRT for the hypothesis

$$H_2: \mu^{(1)} = \cdots = \mu^{(k)}, \Sigma^{(1)} = \cdots = \Sigma^{(k)}$$

is also unbiased. These two problems are discussed in sections 10.2 and 10.3 of Anderson (1958).

Received May 1978; revised December 1978.

[1]This research was partially supported by National Science Foundation Grant No. MCS76-81435 and U.S. Department of Energy Contract No. EY-76-S-02-2751.*000 at the University of Chicago, and was carried out while the author was a visiting professor in the Department of Statistics, Stanford University.

AMS 1970 *subject classifications.* Primary 62H10, 62H15; secondary 62J05, 62J10.

Key words and phrases. Unbiasedness, likelihood ratio test, Bartlett's modified LRT, equality of covariance matrices, equality of normal populations, monotonicity of power, Schur-convex function, majorization.

The (unmodified) LRT for H_1 rejects this hypothesis for large values of $|S|^N/\Pi|S^{(\alpha)}|^{N_\alpha}$, where $S = \Sigma S^{(\alpha)}$, $N = \Sigma N_\alpha$. When the N_α are not all equal, it is well-known that this test is biased; cf. Brown (1939), Ramachandran (1958), Sugiura and Nagao (1969), and Das Gupta (1969). Bartlett's modified LRT rejects H_1 if

$$\lambda_1 \equiv \lambda_1(S^{(1)}, \cdots, S^{(k)}) \equiv \frac{|S|^n}{\Pi_\alpha |S^{(\alpha)}|^{n_\alpha}} > c_1, \tag{1.1}$$

where $|S|$ denotes the determinant of S, $n = \Sigma n_\alpha$ and c_1 is determined by the significance level. Pitman (1939) established the unbiasedness of this test in the univariate case ($p = 1$) by an elegant invariance argument. This case has also been discussed by Brown (1939), Cohen and Strawderman (1971), and Carter and Srivastava (1977). For the multivariate case ($p \geqslant 2$), Sugiura and Nagao (1968) applied Pitman's argument to prove unbiasedness of the modified LRT in the case of two populations ($k = 2$); also see Das Gupta and Giri (1973) and Srivastava, Khatri, and Carter (1978).

For the multivariate case with $k \geqslant 3$, unbiasedness of the modified LRT for H_1 has remained an open question (cf. Sugiura and Nagao (1968), page 1689; Giri (1973), page 58). Giri (1973) attempted to apply the Pitman-Sugiura-Nagao method to prove unbiasedness against those alternatives such that $\Sigma^{(1)}, \cdots, \Sigma^{(k)}$ are diagonal matrices (or, more generally, simultaneously diagonalizable by the same nonsingular linear transformation) but his proof is in error—see Remark 2.5. The Pitman-Sugiura-Nagao-Giri method does apply, however, when each $\Sigma^{(\alpha)}$ is of the form $\sigma_\alpha I_p$, where σ_α is a scalar and I_p denotes the $p \times p$ identity matrix, as we show in Section 2.

Theorem 2.1, the main result of the present paper, states that the modified LRT for H_1 is unbiased against *all* alternatives. To prove this, we first present two results on the monotonicity of the power function, Lemmas 2.2 and 2.3, which reduce consideration to the case where each $\Sigma^{(\alpha)}$ is of the form $\sigma_\alpha I_p$. This approach yields additional information about the power function, e.g., monotonicity and Schur convexity, even for the case $k = 2$ where unbiasedness is already known—cf. Proposition 2.4. Also, in Lemma 2.6 a monotonicity result of Carter and Srivastava (1977) is extended to the multivariate case.

In Section 3 we present a similar demonstration of the unbiasedness of the (unmodified) LRT for H_2. To our knowledge, this problem has not been treated before, even for the case $k = 2$.

Some of the methods in this paper are applied to related problems in Perlman (1980).

2. Unbiasedness and monotonicity properties of the modified LRT for H_1.

The power function of the test (1.1) is given by

$$\pi_1 \equiv \pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) = P_{\Sigma^{(1)}, \ldots, \Sigma^{(k)}}[\lambda_1 > c_1]. \tag{2.1}$$

Both the LRT statistic λ_1 and the power function π_1 are invariant under nonsingular linear transformations A, i.e.,

$$\lambda_1(S^{(1)}, \cdots, S^{(k)}) = \lambda_1(AS^{(1)}A', \cdots, AS^{(k)}A'), \tag{2.2}$$

$$\pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) = \pi_1(A\Sigma^{(1)}A', \cdots, A\Sigma^{(k)}A'). \tag{2.3}$$

THEOREM 2.1. *The modified* LRT (1.1) *for* H_1 *is unbiased,* i.e.,

$$\pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) \geqslant \pi_1(I_p, \cdots, I_p) \tag{2.4}$$

for all $\Sigma^{(1)}, \cdots, \Sigma^{(k)}$, *where* I_p *denotes the* $p \times p$ *identity matrix.*

The first step of the proof of Theorem 2.1 is to show that

$$\pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) \geqslant \pi_1(D^{(1)}, \cdots, D^{(k)}), \tag{2.5}$$

where $D^{(\alpha)}$ is an appropriate diagonal matrix depending on $\Sigma^{(\alpha)}$ such that $|D^{(\alpha)}| = |\Sigma^{(\alpha)}|$ (see (2.26)). The second step is to show that

$$\pi_1(D^{(1)}, \cdots, D^{(k)}) \geqslant \pi_1(\sigma_1 I_p, \cdots, \sigma_k I_p), \tag{2.6}$$

where $\sigma_\alpha = |\Sigma^{(\alpha)}|^{1/p}$. The third and final step is to show that

$$\pi_1(\sigma_1 I_p, \cdots, \sigma_k I_p) \geqslant \pi_1(I_p, \cdots, I_p). \tag{2.7}$$

For Lemma 2.2 partition any $p \times p$ symmetric positive definite matrix in the following manner:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad \Sigma_{11} : 1 \times 1, \quad \Sigma_{22} : (p-1) \times (p-1), \tag{2.8}$$

and define

$$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} : 1 \times 1. \tag{2.9}$$

Recall that if Σ represents the population covariance matrix of a normally distributed random vector $(Z_1, \cdots, Z_p)$, i.e., $\Sigma = \text{Cov}(Z_1, \cdots, Z_p)$, then

$$\begin{aligned} \Sigma_{11\cdot 2} &= \text{Var}(Z_1 | Z_2, \cdots, Z_p) \\ \Sigma_{22} &= \text{Cov}(Z_2, \cdots, Z_p). \end{aligned} \tag{2.10}$$

Following Lemma 2.2 we will also need to consider the quantities

$$\begin{aligned} \Sigma_{ii\cdot j} &\equiv \text{Var}(Z_i | Z_j, \cdots, Z_p) : 1 \times 1 \\ \Sigma_{jj} &\equiv \text{Cov}(Z_j, \cdots, Z_p) : (p-j+1) \times (p-j+1) \end{aligned} \tag{2.11}$$

where $1 \leqslant i < j \leqslant p$. The formula for $\Sigma_{ii\cdot j}$ in terms of the elements of Σ is analogous to (2.9).

For any fixed Σ and $0 \leqslant t \leqslant 1$ define

$$\Sigma_t = \begin{pmatrix} \Sigma_{11\cdot 2} + t^2\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & t\Sigma_{12} \\ t\Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{2.12}$$

Note that

$$\Sigma_1 = \Sigma, \quad \Sigma_0 = \begin{pmatrix} \Sigma_{11\cdot 2} & 0 \\ 0 & \Sigma_{22} \end{pmatrix},$$

and

$$(\Sigma_t)_{11\cdot 2} = \Sigma_{11\cdot 2}, \tag{2.13}$$

so Σ_t is positive definite with $|\Sigma_t| = |\Sigma|$.

LEMMA 2.2. *For every fixed set of population covariance matrices* $\Sigma^{(1)}, \cdots, \Sigma^{(k)}$, *the power function*

$$\pi_1(\Sigma_t^{(1)}, \cdots, \Sigma_t^{(k)}) \tag{2.14}$$

is increasing in t, $0 \leqslant t \leqslant 1$; *it is strictly increasing unless*

$$\Sigma_{12}^{(1)}(\Sigma_{22}^{(1)})^{-1} = \cdots = \Sigma_{12}^{(k)}(\Sigma_{22}^{(k)})^{-1}. \tag{2.15}$$

In particular,

$$\pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) \geqslant \pi_1\left(\begin{pmatrix} \Sigma_{11\cdot 2}^{(1)} & 0 \\ 0 & \Sigma_{22}^{(1)} \end{pmatrix}, \cdots, \begin{pmatrix} \Sigma_{11\cdot 2}^{(k)} & 0 \\ 0 & \Sigma_{22}^{(k)} \end{pmatrix}\right). \tag{2.16}$$

PROOF. The proof is based on the decomposition

$$\lambda_1 = \frac{|S_{22}|^n (S_{11\cdot 2})^n}{\Pi_\alpha |S_{22}^{(\alpha)}|^{n_\alpha} (S_{11\cdot 2}^{(\alpha)})^{n_\alpha}}. \tag{2.17}$$

Since $S = \Sigma S^{(\alpha)}$, the scalar random variable $S_{11\cdot 2}$ can be written as

$$S_{11\cdot 2} = (\Sigma_\alpha S_{11\cdot 2}^{(\alpha)}) + \Delta, \tag{2.18}$$

where

$$\begin{aligned} \Delta = {} & \Sigma_\alpha\left[S_{12}^{(\alpha)} - (\Sigma_\beta S_{12}^{(\beta)})(\Sigma_\beta S_{22}^{(\beta)})^{-1} S_{22}^{(\alpha)}\right] \\ & \times (S_{22}^{(\alpha)})^{-1}\left[S_{12}^{(\alpha)} - (\Sigma_\beta S_{12}^{(\beta)})(\Sigma_\beta S_{22}^{(\beta)})^{-1} S_{22}^{(\alpha)}\right]' \\ \equiv {} & YPY' \end{aligned} \tag{2.19}$$

with

$$\begin{aligned} Y &= (Y_1, \cdots, Y_k) && : 1 \times k(p-1) \\ Y_\alpha &= S_{12}^{(\alpha)}(S_{22}^{(\alpha)})^{-\frac{1}{2}} && : 1 \times (p-1) \\ P &= I_{k(p-1)} - W'(WW')^{-1}W && : k(p-1) \times k(p-1) \\ W &= (W_1, \cdots, W_k) && : (p-1) \times k(p-1) \\ W_\alpha &= (S_{22}^{(\alpha)})^{\frac{1}{2}} && : (p-1) \times (p-1). \end{aligned}$$

(Throughout this paper, if A is a symmetric positive semidefinite matrix then $A^{\frac{1}{2}}$ denotes the symmetric positive semidefinite square root of A.) The (random) matrix

P is the orthogonal projection in $R^{k(p-1)}$ onto the $(k-1)(p-1)$-dimensional subspace orthogonal to the row space of W.

We shall show first that for all fixed values of $S_{22}^{(1)}, \cdots, S_{22}^{(k)}$ and every constant c, the conditional probability

$$P_{\Sigma_t^{(1)},\ldots,\Sigma_t^{(k)}}[\Delta > c \mid S_{22}^{(\alpha)}, 1 \leq \alpha \leq k] \tag{2.20}$$

is increasing in t, and if $c > 0$ it is strictly increasing unless (2.15) holds. When the αth population covariance matrix assumes the value $\Sigma_t^{(\alpha)}$, then

$$Y_\alpha \mid S_{22}^{(\alpha)} \sim N_{p-1}\left[t\Sigma_{12}^{(\alpha)}(\Sigma_{22}^{(\alpha)})^{-1}(S_{22}^{(\alpha)})^{\frac{1}{2}}, \Sigma_{11\cdot 2}^{(\alpha)} I_{p-1}\right], \tag{2.21}$$

using (2.14). Therefore,

$$Y \mid S_{22}^{(1)}, \cdots, S_{22}^{(k)} \sim N_{k(p-1)}(t\Phi, \Lambda), \tag{2.22}$$

where

$$\begin{aligned} \Phi &= \left(\Sigma_{12}^{(\alpha)}(\Sigma_{12}^{(1)})^{-1}(S_{22}^{(1)})^{\frac{1}{2}}, \cdots, \Sigma_{12}^{(k)}(\Sigma_{22}^{(k)})^{-1}(S_{22}^{(k)})^{\frac{1}{2}}\right) &&: 1 \times k(p-1) \\ \Lambda &= \operatorname{diag}\left(\Sigma_{11\cdot 2}^{(1)} I_{p-1}, \cdots, \Sigma_{11\cdot 2}^{(k)} I_{p-1}\right) &&: k(p-1) \times k(p-1). \end{aligned} \tag{2.23}$$

Since P depends only on $S_{22}^{(1)}, \cdots, S_{22}^{(k)}$, when these variables are held fixed the event $\{\Delta \leq c\} \equiv \{YPY' \leq c\}$ determines a circular cylinder in Y-space $\equiv R^{k(p-1)}$, the cylinder whose base is the sphere of radius $c^{\frac{1}{2}}$ in the $(k-1)(p-1)$-dimensional subspace orthogonal to the row space of W. Since this cylinder is convex and symmetric about the origin in $R^{k(p-1)}$, the theorem of Anderson (1955) implies that the conditional probability of the event $\{YPY' \leq c\}$ is decreasing in t, and if $c > 0$ it is in fact strictly decreasing unless the conditional mean vector Φ of Y is parallel to the axis of the cylinder, i.e., $\Phi P = 0$, which is equivalent to (2.15).

To complete the proof, note that (2.14) can be written as

$$E_{\Sigma_t^{(1)},\ldots,\Sigma_t^{(k)}} P_{\Sigma_t^{(1)},\ldots,\Sigma_t^{(k)}}[\Delta > c^* \mid S_{11\cdot 2}^{(\alpha)}, S_{22}^{(\alpha)}; 1 \leq \alpha \leq k], \tag{2.24}$$

where, from (2.17) and (2.18),

$$c^* = \left[\frac{\Pi_\alpha |S_{22}^{(\alpha)}|^{n_\alpha}(S_{11\cdot 2}^{(\alpha)})^{n_\alpha}}{|S_{22}|^n} \cdot c_1\right]^{1/n} - (\Sigma_\alpha S_{11\cdot 2}^{(\alpha)}). \tag{2.25}$$

Since $(\Delta, S_{22}^{(1)}, \cdots, S_{22}^{(k)})$ is independent of $(S_{11\cdot 2}^{(1)}, \cdots, S_{11\cdot 2}^{(k)})$, the conditional distribution of Δ given $\{S_{11\cdot 2}^{(\alpha)}, S_{22}^{(\alpha)}; 1 \leq \alpha \leq k\}$ is the same as that given $\{S_{22}^{(\alpha)}, 1 \leq \alpha \leq k\}$. Therefore, by the result of the preceding paragraph, the conditional probability in (2.24) is increasing in t. Furthermore, by (2.12) and (2.13), the marginal distribution of $\{S_{11\cdot 2}^{(\alpha)}, S_{22}^{(\alpha)}; 1 \leq \alpha \leq k\}$ does not depend on t, so (2.24) itself is increasing in t, hence so is (2.14). Finally, since $c^* > 0$ with positive probability, (2.14) is strictly increasing in t unless (2.15) holds. This completes the proof of Lemma 2.2.

In its present form, Lemma 2.2 states that the power function π_1 decreases monotonically if the covariances between the *first variate* and the remaining $p-1$

variates decrease at the same rate in each of the k populations, with the conditional variances $\Sigma_{11\cdot 2}^{(\alpha)}$ and the marginal covariance matrices $\Sigma_{22}^{(\alpha)}$ held fixed. Clearly, by the invariance property (2.3), the same is true of π_1 if the covariance between the ith *variate* and the remaining $p - 1$ variates decreases at the same rate in each population, with the appropriate conditional variances and marginal covariance matrices held fixed, for *each* $i = 1, \cdots, p$. By applying this fact for $i = 2$ to the right-hand side of (2.16), then for $i = 3$ to the right-hand side of the resulting inequality, and so on, through $i = p - 1$, we obtain (2.5), where $D^{(\alpha)}$ is the $p \times p$ *diagonal* matrix given by

$$D^{(\alpha)} = \text{diag}\big(\Sigma_{11\cdot 2}^{(\alpha)}, \Sigma_{22\cdot 3}^{(\alpha)}, \cdots, \Sigma_{p-1,p-1\cdot p}^{(\alpha)}, \Sigma_{pp}^{(\alpha)}\big). \tag{2.26}$$

(Note that $|D^{(\alpha)}| = |\Sigma^{(\alpha)}|$.)

We remark that the diagonal matrices $D^{(1)}, \cdots, D^{(k)}$ satisfying (2.5) are not uniquely determined, because a preliminary permutation of the p variates (take A to be a permutation matrix in (2.3)) before application of Lemma 2.2 can lead to diagonal matrices other than $D^{(1)}, \cdots, D^{(k)}$. In fact, by taking A to be an appropriate *orthogonal* matrix (see (2.27)) in (2.3) with $\Sigma^{(\alpha)}$ replaced by $D^{(\alpha)}$ therein, and then applying Lemma 2.2, we will arrive at inequality (2.6), the second step in the proof of Theorem 2.1, wherein the matrices $D^{(\alpha)}$ are reduced to diagonal matrices of the form $\sigma_\alpha I_p$. The details now follow.

Let $d_i^{(\alpha)}$ denote the ith diagonal element of $D^{(\alpha)}$. For $0 \leqslant u \leqslant 1$ define the block-diagonal $p \times p$ orthogonal matrix A_u by

$$A_u = \text{diag}\left\{\begin{bmatrix} (1-u)^{\frac{1}{2}} & -u^{\frac{1}{2}} \\ u^{\frac{1}{2}} & (1-u)^{\frac{1}{2}} \end{bmatrix}, I_{p-2}\right\}. \tag{2.27}$$

Hence, by (2.3) and (2.16),

$$\begin{aligned} \pi_1(D^{(1)}, \cdots, D^{(k)}) &= \pi_1\big(A_u D^{(1)} A_u', \cdots, A_u D^{(k)} A_u'\big) \\ &\geqslant \pi_1\big(D_u^{(1)}, \cdots, D_u^{(k)}\big) \end{aligned} \tag{2.28}$$

for all $0 \leqslant u \leqslant 1$, where

$$\begin{aligned} D_u^{(\alpha)} = \text{diag}\big\{ & d_1^{(\alpha)} d_2^{(\alpha)} (u d_1^{(\alpha)} + (1-u) d_2^{(\alpha)})^{-1}, \\ & u d_1^{(\alpha)} + (1-u) d_2^{(\alpha)}, d_3^{(\alpha)}, \cdots, d_p^{(\alpha)} \big\}. \end{aligned} \tag{2.29}$$

Furthermore, note that the inequality (2.28) also holds if, in the definition (2.29) of $D_u^{(\alpha)}$, the pair of components $d_1^{(\alpha)}, d_2^{(\alpha)}$ is replaced by any other pair $d_i^{(\alpha)}, d_j^{(\alpha)}$, $1 \leqslant i \neq j \leqslant p$.

It is now convenient to restate the inequality (2.28) in terms of the column vectors

$$\delta^{(\alpha)} \equiv \big(\delta_1^{(\alpha)}, \cdots, \delta_p^{(\alpha)}\big)' \equiv \big(\log d_1^{(\alpha)}, \cdots, \log d_p^{(\alpha)}\big)', \qquad 1 \leqslant \alpha \leqslant k. \tag{2.30}$$

Express the power function in terms of the $\delta^{(\alpha)}$ by defining $\tilde{\pi}_1$ as follows:

$$\tilde{\pi}_1(\delta^{(1)}, \cdots, \delta^{(k)}) = \pi_1(D^{(1)}, \cdots, D^{(k)}). \tag{2.31}$$

We now require a concept from the theory of *majorization* (cf. Hardy, Littlewood, and Polya (1952)): for indices $1 \leqslant i < j \leqslant p$ and for $0 \leqslant \theta \leqslant 1$, the T-transform T_θ^{ij} is defined to be the $p \times p$ doubly stochastic matrix given by

$$T_\theta^{ij} = \theta I_p + (1-\theta)R_{ij}, \tag{2.32}$$

where R_{ij} is the permutation matrix which interchanges the ith and jth coordinates, leaving the others fixed. Inequality (2.28) now can be restated as follows:

$$\tilde{\pi}_1(\delta^{(1)}, \cdots, \delta^{(k)}) \geqslant \tilde{\pi}_1\left(T^{12}_{\theta_1(u)}\delta^{(1)}, \cdots, T^{12}_{\theta_k(u)}\delta^{(k)}\right), \tag{2.33}$$

where

$$\begin{aligned} \theta_\alpha(u) &\equiv \theta(u;\, \delta_1^{(\alpha)}, \delta_2^{(\alpha)}) \\ &= \frac{\delta_1^{(\alpha)} - \log\left[u \exp\{\delta_1^{(\alpha)}\} + (1-u)\exp\{\delta_2^{(\alpha)}\}\right]}{\delta_1^{(\alpha)} - \delta_2^{(\alpha)}} \qquad \text{if } \delta_1^{(\alpha)} \neq \delta_2^{(\alpha)} \\ &= 1 \qquad \text{if } \delta_1^{(\alpha)} = \delta_2^{(\alpha)} \end{aligned} \tag{2.34}$$

(recall that $\delta_i^{(\alpha)} = \log d_i^{(\alpha)}$). Again, inequality (2.33) also holds with any pair of variables i, j substituted for 1, 2.

LEMMA 2.3. *Let $T \equiv T_\theta^{ij}$ be an arbitrary T-transform. Then there exist T-transforms $T^{(\alpha)} \equiv T_{\theta_\alpha}^{ij}$, $2 \leqslant \alpha \leqslant k$, such that*

$$\tilde{\pi}_1(\delta^{(1)}, \cdots, \delta^{(k)}) \geqslant \tilde{\pi}_1(T\delta^{(1)}, T^{(2)}\delta^{(2)}, \cdots, T^{(k)}\delta^{(k)}). \tag{2.35}$$

Here θ_α, and hence $T^{(\alpha)}$, depends on θ, $\delta_i^{(1)}$, $\delta_j^{(1)}$, $\delta_i^{(\alpha)}$, and $\delta_j^{(\alpha)}$. Furthermore, the role of $\delta^{(1)}$ may be taken by any other $\delta^{(\alpha)}$, $2 \leqslant \alpha \leqslant k$.

PROOF. Taking $(i, j) = (1, 2)$ for simplicity, (2.35) follows from (2.33) and the fact that $\theta_1(u)$ strictly decreases from 1 to 0 as u increases from 0 to 1, provided that $\delta_1^{(1)} \neq \delta_2^{(1)}$.

We digress to note that Lemma 2.3 immediately provides a new monotonicity property in the two-population case. When $k = 2$ the invariance property (2.3) implies that

$$\pi_1(\Sigma^{(1)}, \Sigma^{(2)}) = \pi_1(I_p, \operatorname{diag}(d_1, \cdots, d_p)) = \tilde{\pi}_1(0, \delta), \tag{2.36}$$

where $d_1, \cdots, d_p$ are the characteristic roots of $(\Sigma^{(1)})^{-1}\Sigma^{(2)}$, $\delta = (\delta_1, \cdots, \delta_p)'$, and $\delta_i = \log d_i$. Then Lemma 2.3 implies that

$$\tilde{\pi}_1(0, \delta) \geqslant \tilde{\pi}_1(0, T\delta) \tag{2.37}$$

for every T-transform T. A fundamental result in the theory of majorization states that a vector ν is majorized by a vector δ (i.e., $\nu = Q\delta$ for some doubly stochastic matrix Q) if and only if ν can be obtained from δ by a finite number of T-transforms, i.e., iff $\nu = T_m \cdots T_2T_1\delta$ for some m and $T_1, \cdots, T_m$ (cf. Hardy, Littlewood, and Polya (1952), page 47). Thus (2.37) implies that $\tilde{\pi}_1(0, \delta) \geqslant \tilde{\pi}_1(0, \nu)$ whenever ν is majorized by δ, i.e., $\tilde{\pi}_1(0, \delta)$ is a *Schur-convex function of* δ. This is summarized as follows:

PROPOSITION 2.4. *When $k = 2$, $\pi_1(\Sigma^{(1)}, \Sigma^{(2)})$ is a Schur-convex function of the logarithms of the characteristic roots of* $(\Sigma^{(1)})^{-1}\Sigma^{(2)}$.

We now apply Lemma 2.3 to obtain inequality (2.6) for the general case $k \geq 2$. Put $e = (1, \cdots, 1)' : p \times 1$ and define the scalars $\bar{\delta}^{(\alpha)}$ and σ_α by

$$(2.38)\quad \bar{\delta}^{(\alpha)} \equiv \frac{1}{p} e'\delta^{(\alpha)} = \frac{1}{p}\Sigma_{i=1}^{p}\delta_i^{(\alpha)} = \frac{1}{p}\log|D^{(\alpha)}| = \frac{1}{p}\log|\Sigma^{(\alpha)}| \equiv \log \sigma_\alpha.$$

Since $\bar{\delta}^{(1)}e = (p^{-1}ee')\delta^{(1)}$ and $p^{-1}ee'$ is a doubly stochastic matrix, $\bar{\delta}^{(1)}e$ is majorized by $\delta^{(1)}$. Therefore, Lemma 2.3 and the subsequent discussion imply that

$$(2.39)\quad \tilde{\pi}_1(\delta^{(1)}, \delta^{(2)}, \cdots, \delta^{(k)}) \geq \tilde{\pi}_1(\bar{\delta}^{(1)}e, \gamma^{(2)}, \cdots, \gamma^{(k)}),$$

where each vector $\gamma^{(\alpha)}$ is of the form $\gamma^{(\alpha)} = Q_\alpha\delta^{(\alpha)}$, for some doubly stochastic $p \times p$ matrix Q_α, $2 \leq \alpha \leq k$ (the product of T-transforms is doubly stochastic). Next, since $\bar{\delta}^{(2)} = \bar{\gamma}^{(2)}$ and $Te = e$ for any T-transform T, this procedure can be repeated to obtain

$$(2.40)\quad \tilde{\pi}_1(\bar{\delta}^{(1)}e, \gamma^{(2)}, \gamma^{(3)}, \cdots, \gamma^{(k)}) \geq \tilde{\pi}_1(\bar{\delta}^{(1)}e, \bar{\delta}^{(2)}e, \varphi^{(3)}, \cdots, \varphi^{(k)}),$$

where $\varphi^{(\alpha)}$ is again of the form $\varphi^{(\alpha)} = R_\alpha\delta^{(\alpha)}$ for some doubly stochastic matrix R_α, $3 \leq \alpha \leq k$ (the product of doubly stochastic matrices is doubly stochastic). Continuing this procedure $k - 2$ times more, we obtain

$$(2.41)\quad \tilde{\pi}_1(\delta^{(1)}, \cdots, \delta^{(k)}) \geq \tilde{\pi}_1(\bar{\delta}^{(1)}e, \cdots, \bar{\delta}^{(k)}e).$$

By (2.31) and (2.38), (2.41) is identical to (2.6).

Now that the matrices $\Sigma^{(\alpha)}$ each have been reduced to the form $\sigma_\alpha I_p$, a direct extension of the Pitman-Sugiura-Nagao method, as in Giri (1973), yields the final inequality (2.7). We sketch the details. Put $\tau_\alpha = \sigma_\alpha/\sigma_1$, $1 \leq \alpha \leq k$, so that

$$\begin{aligned} 1 - \pi_1(\sigma_1 I_p, \cdots, \sigma_k I_p) &= 1 - \pi_1(I_p, \tau_2 I_p, \cdots, \tau_k I_p) \\ (2.42)\qquad &= M \cdot \textstyle\int_{\{\lambda_1 < c_1\}}\left[\Pi_{\alpha=1}^{k}\tau_\alpha^{-pn_\alpha/2}|S^{(\alpha)}|^{(n_\alpha - p - 1)/2}\right] \\ &\quad \times \exp\left\{-\tfrac{1}{2}\mathrm{tr}\left[S^{(1)} + \Sigma_{\alpha=2}^{k}\tau_\alpha^{-1}S^{(\alpha)}\right]\right\}\Pi_{\alpha=1}^{k}\, dS^{(\alpha)}, \end{aligned}$$

where M is a positive constant. Make the transformation $U_\alpha = (S^{(1)})^{-\frac{1}{2}}S^{(\alpha)}(S^{(1)})^{-\frac{1}{2}}$, $2 \leq \alpha \leq k$, as in Theorem 3.1 of Olkin and Rubin (1964) (the Jacobian is $|S^{(1)}|^{(p+1)(k-1)/2}$), integrate out $S^{(1)}$, and then make the transformation $V_\alpha = \tau_\alpha^{-1}U_\alpha$, $2 \leq \alpha \leq k$. This yields

$$(2.43)\quad 1 - \pi_1(I_p, \tau_2 I_p, \cdots, \tau_k I_p) = M'\textstyle\int_{\{f(\tau V) > c'\}} f(V)\Pi_{\alpha=2}^{k}\dfrac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

where $c' = c_1^{-\frac{1}{2}}$, M' is another positive constant, and

$$(2.44)\quad \begin{aligned} f(V) &\equiv f(V_2, \cdots, V_k) \equiv |I_p + \Sigma_{\alpha=2}^{k}V_\alpha|^{-n/2}\Pi_{\alpha=2}^{k}|V_\alpha|^{n_\alpha/2}, \\ f(\tau V) &\equiv f(\tau_2 V_2, \cdots, \tau_k V_k). \end{aligned}$$

Hence

$$\pi_1(I_p, \tau_2 I_p, \cdots, \tau_k I_p) - \pi_1(I_p, \cdots, I_p)$$

$$= M'\left[\int_{\{f(V) > c' > f(\tau V)\}} - \int_{\{f(\tau V) > c' > f(V)\}}\right] \cdot f(V) \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

$$(2.45) \qquad \geqslant c'M'\left[\int_{\{f(V) > c' > f(\tau V)\}} - \int_{\{f(\tau V) > c' > f(V)\}}\right] \cdot \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

$$= c'M'\left[\int_{\{f(V) > c'\}} - \int_{\{f(\tau V) > c'\}}\right] \cdot \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

$$= 0,$$

since the measure $\Pi |V_\alpha|^{-(p+1)/2}\, dV_\alpha$ is invariant under the transformation $V_\alpha \to \tau_\alpha V_\alpha$. The second-to-last equality in (2.45) requires that

$$(2.46) \qquad \int_{\{f(V) > c', f(\tau V) > c'\}} \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}} < \infty,$$

which is true since this integral is bounded above by

$$(2.47) \qquad \frac{1}{c'} \int_{\{f(V) > c'\}} f(V) \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}} = \frac{1}{c'M'}\left[1 - \pi_1(I_p, \cdots, I_p)\right].$$

Thus, (2.7) is established, and the proof of Theorem 2.1 is complete.

REMARK 2.5. Giri (1973) attempted to use the argument of the preceding paragraph to combine (2.6) and (2.7), i.e., to prove in a single step that

$$\pi_1(D^{(1)}, \cdots, D^{(k)}) \geqslant \pi_1(I_p, \cdots, I_p)$$

when the $D^{(\alpha)}$ are diagonal. However, when $k \geqslant 3$ the step from (2.42) to (2.43) is not valid if the matrices $\tau_\alpha I_p$ are replaced by $D^{(\alpha)}$, since in general

$$\operatorname{tr} D^{-1} S^{\frac{1}{2}} U S^{\frac{1}{2}} \neq \operatorname{tr} D^{-1} US.$$

(In Giri (1973), page 59, the quantities P' and P on the second line of his expression (2.7) should appear in the reverse order. When this correction is made, the subsequent equality is no longer valid.) When $k = 2$, though, one can take $D^{(1)} = I_p$ without loss of generality, and the step from (2.42) to (2.43) can be accomplished—see Sugiura and Nagao (1968).

The above proof of the inequality (2.7) gives no indication of the behavior of $\pi_1(\sigma_1 I_p, \cdots, \sigma_k I_p)$ other than that its minimum occurs when $\sigma_1 = \cdots = \sigma_k$. The proof may be modified, however, to obtain the stronger result given in Lemma 2.6 below, which extends a univariate result of Carter and Srivastava (1977). This result is a monotonicity property which provides some information as to the relative distances of alternatives of the form $(\sigma_1 I_p, \cdots, \sigma_k I_p)$ from the null hypothesis $\sigma_1 = \cdots = \sigma_k$ (with respect to the modified LRT).

LEMMA 2.6. *Assume that $\sigma_i \neq \sigma_j$ for at least one pair i, j. If* $\max\{\sigma_1, \cdots, \sigma_\beta\} \leqslant \min\{\sigma_{\beta+1}, \cdots, \sigma_k\}$, *where* $1 \leqslant \beta \leqslant k-1$, *then*

$$\psi(\lambda) \equiv \pi_1(\sigma_1 I_p, \cdots, \sigma_\beta I_p, \lambda\sigma_{\beta+1} I_p, \cdots, \lambda\sigma_k I_p)$$

is strictly increasing in λ for $\lambda \geqslant 1$.

PROOF. It suffices to show that $\psi(\lambda) > \psi(1)$ when $\lambda > 1$. Without loss of generality assume that $\sigma_1 = \max\{\sigma_1, \cdots, \sigma_\beta\}$, and put $\tau_\alpha = \sigma_\alpha/\sigma_1$ as before. Then, as in (2.45),

$$(2.48) \quad \psi(\lambda) - \psi(1)$$
$$= M'\left[\int_{\{f(\tau V) > c' > f(\lambda\tau V)\}} - \int_{\{f(\lambda\tau V) > c' > f(\tau V)\}}\right] \cdot f(V) \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

where $f(V)$, $f(\tau V)$ are given by (2.44) and where

$$f(\lambda\tau V) \equiv f(\tau_2 V_2, \cdots, \tau_\beta V_\beta, \lambda\tau_{\beta+1} V_{\beta+1}, \cdots, \lambda\tau_k V_k).$$

Since

$$(2.49) \qquad f(\tau V) \geqslant c' \Leftrightarrow f(V) \geqslant c'\tau' \frac{|I_p + \Sigma_{\alpha=2}^k \tau_\alpha V_\alpha|^{n/2}}{|I_p + \Sigma_{\alpha=2}^k V_\alpha|^{n/2}}$$

where $\tau' = \Pi_{\alpha=2}^k \tau_\alpha^{-(pn_\alpha)/2}$, it follows that

$$\psi(\lambda) - \psi(1) \geqslant c'\tau' M'\left[\int_{\{f(\tau V) > c' > f(\lambda\tau V)\}} - \int_{\{f(\lambda\tau V) > c' > f(\tau V)\}}\right]$$
$$\cdot \frac{|I_p + \Sigma_{\alpha=2}^k \tau_\alpha V_\alpha|^{n/2}}{|I_p + \Sigma_{\alpha=2}^k V_\alpha|^{n/2}} \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

$$(2.50) \qquad = c'\tau' M'\left[\int_{\{f(\tau V) > c'\}} - \int_{\{f(\lambda\tau V) > c'\}}\right]$$
$$\cdot \frac{|I_p + \Sigma_{\alpha=2}^k \tau_\alpha V_\alpha|^{n/2}}{|I_p + \Sigma_{\alpha=2}^k V_\alpha|^{n/2}} \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

$$= c'\tau' M' \int_{\{f(\tau V) > c'\}} \left[\frac{|I_p + \Sigma_{\alpha=2}^k \tau_\alpha V_\alpha|^{n/2}}{|I_p + \Sigma_{\alpha=2}^k V_\alpha|^{n/2}} \right.$$
$$\left. - \frac{|I_p + \Sigma_{\alpha=2}^\beta \tau_\alpha V_\alpha + \lambda^{-1}\Sigma_{\alpha=\beta+1}^k \tau_\alpha V_\alpha|^{n/2}}{|I_p + \Sigma_{\alpha=2}^\beta V_\alpha + \lambda^{-1}\Sigma_{\alpha=\beta+1}^k V_\alpha|^{n/2}} \right] \Pi_{\alpha=2}^k \frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}$$

(if $\beta = 1$, the sums $\Sigma_{\alpha=2}^\beta$ are vacuous). The last equality in (2.50) is obtained by making the transformation $V_{\beta+1} \to \lambda^{-1} V_{\beta+1}, \cdots, V_k \to \lambda^{-1} V_k$ in the second integral of the preceding expression, while the validity of the first equality in (2.50) follows from an argument similar to the one involving (2.46) and (2.47). Finally, since $\tau_2 \leqslant 1, \cdots, \tau_\beta \leqslant 1$ and $\tau_{\beta+1} \geqslant 1, \cdots, \tau_k \geqslant 1$ with at least one strict inequality, the integrand on the last line of (2.50) is strictly positive by Lemma 2.7 which follows. This completes the proof of Lemma 2.6.

LEMMA 2.7. *Let U_1, U_2, W_1, W_2 be $p \times p$ symmetric positive definite matrices. Assume that $U_2 - U_1$ and $W_1 - W_2$ are positive semidefinite and at least one difference is not the zero matrix. Then*

$$\varphi^*(t) \equiv \frac{|U_1 + tW_1|}{|U_2 + tW_2|}$$

is strictly increasing in t for $t \geqslant 0$.

PROOF. Set $A_j = U_j^{-\frac{1}{2}} W_j U_j^{-\frac{1}{2}}$, $j = 1, 2$, where $U_j^{\frac{1}{2}}$ is the symmetric positive definite square root of U_j. Then $\varphi^*(t)$ is proportional to

$$\varphi(t) \equiv \frac{|I + tA_1|}{|I + tA_2|} = \frac{\Pi_{i=1}^p[1 + tch_i(A_1)]}{\Pi_{i=1}^p[1 + tch_i(A_2)]},$$

where $ch_1(A) \geqslant \cdots \geqslant ch_p(A)$ denote the ordered characteristic roots of A. Since $W_1 - W_2$ and $U_1^{-1} - U_2^{-1}$ are positive semidefinite and at least one is nonzero, two applications of the Courant-Fischer theorem (cf. Bellman (1970), Theorem 3, page 117) show that

$$ch_i(A_1) \geqslant ch_i(A_2), \qquad 1 \leqslant i \leqslant p, \tag{2.51}$$

with at least one inequality strict. From this the result follows easily.

3. Unbiasedness of the LRT for H_2. The LRT for testing H_2, the hypothesis of equality of k normal populations, rejects H_2 if

$$\lambda_2 \equiv \lambda_2(\bar{X}^{(1)}, \cdots, \bar{X}^{(k)}; S^{(1)}, \cdots, S^{(k)}) \equiv \frac{|S + T|^N}{\Pi_\alpha |S^{(\alpha)}|^{N_\alpha}} > c_2, \tag{3.1}$$

where

$$\begin{aligned} S &= \Sigma_{\alpha=1}^k S^{(\alpha)} \\ T &= \Sigma_{\alpha=1}^k N_\alpha(\bar{X}^{(\alpha)} - \bar{X}^{(+)})(\bar{X}^{(\alpha)} - \bar{X}^{(+)})', \\ X^{(+)} &= \frac{1}{N}\Sigma_{\alpha=1}^k N_\alpha \bar{X}^{(\alpha)}; \end{aligned} \tag{3.2}$$

cf. Anderson (1958, Section 10.3). Anderson suggested a modified LRT for H_2 based on $|S + T|^n / \Pi |S^{(\alpha)}|^{n_\alpha}$. In this section, however, we show that it is the LRT itself, not the modified LRT, that is unbiased for testing H_2.

Denote the power function of the test (3.1) by

$$\begin{aligned} \pi_2 &\equiv \pi_2(\mu^{(1)}, \cdots, \mu^{(k)}; \Sigma^{(1)}, \cdots, \Sigma^{(k)}) \\ &= P_{\mu^{(1)}, \ldots, \mu^{(k)}; \Sigma^{(1)}, \ldots, \Sigma^{(k)}}[\lambda_2 > c_\alpha]. \end{aligned} \tag{3.3}$$

Both λ_2 and π_2 are invariant under nonsingular affine transformations, i.e.,

$$\begin{aligned} \lambda_2(\bar{X}^{(1)}, \cdots, \bar{X}^{(k)}; S^{(1)}, \cdots, S^{(k)}) = \lambda_2(A\bar{X}^{(1)} + b, \cdots, A\bar{X}^{(k)} + b; \\ AS^{(1)}A', \cdots, AS^{(k)}A') \end{aligned} \tag{3.4}$$

$$(3.5)\quad \pi_2(\mu^{(1)}, \cdots, \mu^{(k)}; \Sigma^{(1)}, \cdots, \Sigma^{(k)}) = \pi_2(A\mu^{(1)} + b, \cdots, A\mu^{(k)} + b; A\Sigma^{(1)}A', \cdots, A\Sigma^{(k)}A')$$

for all $p \times p$ nonsingular matrices A and all $p \times 1$ vectors b.

THEOREM 3.1. *The* LRT (3.1) *for* H_2 *is unbiased,* i.e.,

$$\pi_2(\mu^{(1)}, \cdots, \mu^{(k)}; \Sigma^{(1)}, \cdots, \Sigma^{(k)}) \geq \pi_2(0, \cdots, 0; I_p, \cdots, I_p)$$

for all $\mu^{(1)}, \cdots, \mu^{(k)}, \Sigma^{(1)}, \cdots, \Sigma^{(k)}$.

The proof of Theorem 3.1 follows the same pattern as that of Theorem 2.1, but there are several nontrivial differences. First, an analog of Lemma 2.2 will be established. For $0 \leq t \leq 1$, define Σ_t as in (2.12). Also, for any $p \times 1$ vector μ partitioned as $\mu = (\mu_1, \mu_{(2)}')'$ with $\mu_{(2)} : (p-1) \times 1$, define $\mu_t = (t\mu_1, \mu_{(2)}')'$.

LEMMA 3.2. *For fixed* $\mu^{(1)}, \cdots, \mu^{(k)}, \Sigma^{(1)}, \cdots, \Sigma^{(k)}$, *the power function*

$$(3.6)\quad \pi_2(\mu_t^{(1)}, \cdots, \mu_t^{(k)}; \Sigma_t^{(1)}, \cdots, \Sigma_t^{(k)})$$

is increasing in t, $0 \leq t \leq 1$.

PROOF. Write λ_2 as

$$(3.7)\quad \lambda_2 = \frac{|S_{22} + T_{22}|^N[(S+T)_{11\cdot 2}]^N}{\Pi_\alpha |S_{22}^{(\alpha)}|^{N_\alpha}(S_{11\cdot 2}^{(\alpha)})^{N_\alpha}},$$

where the partitioning is in accordance with (2.8) and (2.9). The $p \times p$ matrix T can be written as

$$(3.8)\quad T = \bar{X}B\bar{X}' = \begin{bmatrix} \bar{X}_1 B \bar{X}_1' & \bar{X}_1 B \bar{X}_2' \\ \bar{X}_2 B \bar{X}_1' & \bar{X}_2 B \bar{X}_2' \end{bmatrix},$$

where

$$(3.9)\quad \bar{X} = (\bar{X}^{(1)}, \cdots, \bar{X}^{(k)}) \equiv \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix}$$

with $\bar{X}_1$: $1 \times k$ and $\bar{X}_2$: $(p-1) \times k$, and where

$$(3.10)\quad B = D_N - N^{-1}\tilde{N}\tilde{N}' : k \times k$$

with $D_N = \text{diag}(N_1, \cdots, N_k)$ and $\tilde{N} = (N_1, \cdots, N_k)'$. Therefore,

$$(3.11)\quad (S+T)_{11\cdot 2} = \Sigma_{\alpha=1}^k S_{11\cdot 2}^{(\alpha)} + \Delta + \bar{X}_1 E \bar{X}_1',$$

where

$$(3.12)\quad E = B - B\bar{X}_2'(\bar{X}_2 B \bar{X}_2')^- \bar{X}_2 B,$$

$$(3.13)\quad \Delta = \left[\Sigma_{\alpha=1}^k S_{12}^{(\alpha)}(S_{22}^{(\alpha)})^{-1}S_{21}^{(\alpha)} + \bar{X}_1 B \bar{X}_2'(\bar{X}_2 B \bar{X}_2')^- \bar{X}_2 B \bar{X}_1'\right] - (\Sigma_{\alpha=1}^k S_{12}^{(\alpha)} + \bar{X}_1 B \bar{X}_2')(\Sigma_{\alpha=1}^k S_{22}^{(\alpha)} + \bar{X}_2 B \bar{X}_2')^{-1}(\Sigma_{\alpha=1}^k S_{21}^{(\alpha)} + \bar{X}_2 B \bar{X}_1')$$

and $(\bar{X}_2 B \bar{X}_2')^-$ is the Moore-Penrose generalized inverse of $\bar{X}_2 B \bar{X}_2'$.

Now define $S^{(k+1)} = \bar{X}B\bar{X}'$ and write $(S_{22}^{(k+1)})^{-\frac{1}{2}}$ to denote $[(S_{22}^{(k+1)})^{-}]^{\frac{1}{2}}$. From the properties of a generalized inverse and the argument at the top of page 523 of Rao (1973) it follows that

$$S_{22}^{(k+1)}(S_{22}^{(k+1)})^{-}S_{22}^{(k+1)} = (S_{22}^{(k+1)})$$
$$S_{22}^{(k+1)}(S_{22}^{(k+1)})^{-}S_{21}^{(k+1)} = S_{21}^{(k+1)} \tag{3.14}$$
$$(S_{22}^{(k+1)})^{\frac{1}{2}}(S_{22}^{(k+1)})^{-\frac{1}{2}}S_{21}^{(k+1)} = S_{21}^{(k+1)}.$$

Thus, Δ again can be written as in (2.19), where now the range of summation is $1, \cdots, k+1$ and where $(S_{22}^{(k+1)})^{-}$ is substituted for $(S_{22}^{(k+1)})^{-1}$, so

$$\Delta = YPY', \tag{3.15}$$

where now

$$Y = (Y_1, \cdots, Y_{k+1}) \qquad : 1 \times (k+1)(p-1)$$
$$Y_\alpha = S_{12}^{(\alpha)}(S_{22}^{(\alpha)})^{-\frac{1}{2}} \qquad : 1 \times (p-1)$$
$$P = I_{(k+1)(p-1)} - W'(WW')^{-1}W \qquad : (k+1)(p-1) \times (k+1)(p-1)$$
$$W = (W_1, \cdots, W_{k+1}) \qquad : (p-1) \times (k+1)(p-1)$$
$$W_\alpha = (S_{22}^{(\alpha)})^{\frac{1}{2}} \qquad : (p-1) \times (p-1).$$

Since $Y_{k+1} = \bar{X}_1 B\bar{X}_2'(\bar{X}_2 B\bar{X}_2')^{-\frac{1}{2}}$, (3.15) shows that with $\bar{X}_2$ and $\{S_{22}^{(\alpha)} | 1 \leqslant \alpha \leqslant k\}$ held fixed, $\Delta + \bar{X}_1 B\bar{X}_1'$ is a positive semidefinite quadratic form in $\bar{X}_1$ and $Y_\alpha \equiv S_{12}^{(\alpha)}(S_{22}^{(\alpha)})^{-\frac{1}{2}}$, $1 \leqslant \alpha \leqslant k$. When the population mean vectors and covariance matrices are given by $\mu_t^{(\alpha)}$ and $\Sigma_t^{(\alpha)}$, $1 \leqslant \alpha \leqslant k$, the conditional distribution of $(Y_1, \cdots, Y_k)$ given $\bar{X}_2$ and $\{S_{22}^{(\alpha)} | 1 \leqslant \alpha \leqslant k\}$ is again (2.22), while the conditional distribution of $\bar{X}_1$ is

$$\bar{X}_1 | \bar{X}_2 \sim N_k(tF, \Omega), \tag{3.16}$$

where

(3.17)

$$F = \left(\mu_1^{(1)} + \Sigma_{12}^{(1)}(\Sigma_{22}^{(1)})^{-1}(\bar{X}_2^{(1)} - \mu_{(2)}^{(1)}), \cdots, \mu_1^{(k)} + \Sigma_{12}^{(k)}(\Sigma_{22}^{(k)})^{-1}(\bar{X}_2^{(k)} - \mu_{(2)}^{(k)})\right) \qquad : 1 \times k$$

$$\Omega = \operatorname{diag}(N_1^{-1}\Sigma_{11\cdot 2}^{(1)}, \cdots, N_k^{-1}\Sigma_{11\cdot 2}^{(k)}) \qquad : k \times k.$$

Thus, the conditional distribution of $(Y_1, \cdots, Y_k, \bar{X}_1)$ given $\bar{X}_2$ and $S_{22}^{(\alpha)}$, $1 \leqslant \alpha \leqslant k$, is of the form $N(t\Phi^*, \Lambda^*)$, where Φ^* and Λ^* depend on $\bar{X}_2$ and $S_{22}^{(\alpha)}$, $\mu^{(\alpha)}$, $\Sigma^{(\alpha)}$, $1 \leqslant \alpha \leqslant k$, but not on t. Therefore, by Anderson's Theorem, for every constant c the conditional probability

$$P_{\mu_t^{(1)}, \ldots, \mu_t^{(k)}; \Sigma_t^{(1)}, \ldots, \Sigma_t^{(k)}}\left[\Delta + \bar{X}_1 E\bar{X}_1' > c | \bar{X}_2, S_{22}^{(\alpha)}; 1 \leqslant \alpha \leqslant k\right] \tag{3.18}$$

is increasing in t for $t \geqslant 0$. Lemma 3.2 now follows readily.

By an argument similar to that which led to (2.25), Lemma 3.2 implies that

$$\pi_2(\mu^{(1)}, \cdots, \mu^{(k)}; \Sigma^{(1)}, \cdots, \Sigma^{(k)}) \tag{3.19}$$

$$\geqslant \pi_2\left(\begin{pmatrix} 0 \\ \mu_{(2)}^{(1)} \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \mu_{(2)}^{(k)} \end{pmatrix}; \begin{pmatrix} \Sigma_{11\cdot 2}^{(1)} & 0 \\ 0 & \Sigma_{22}^{(1)} \end{pmatrix}, \cdots, \begin{pmatrix} \Sigma_{11\cdot 2}^{(k)} & 0 \\ 0 & \Sigma_{22}^{(k)} \end{pmatrix}\right)$$

$$\geqslant \pi_2(0, \cdots, 0; D^{(1)}, \cdots, D^{(k)})$$

$$\equiv \pi_2^0(D^{(1)}, \cdots, D^{(k)}),$$

where $D^{(\alpha)}$ is the diagonal matrix defined by (2.26) and where $\mu_{(i)}^{(\alpha)}$ denotes $(\mu_i^{(\alpha)}, \cdots, \mu_p^{(\alpha)})'$. (Note that the proof of (3.19) requires p steps, whereas the proof of (2.5) required only $p - 1$ steps.) Thus, the analog of (2.5) has been established.

To establish the analog of (2.6), note that the function

$$\pi_2^0(\Sigma^{(1)}, \cdots, \Sigma^{(k)}) \equiv \pi_2(0, \cdots, 0; \Sigma^{(1)}, \cdots, \Sigma^{(k)}) \tag{3.20}$$

satisfies the same properties as $\pi_1(\Sigma^{(1)}, \cdots, \Sigma^{(k)})$ in Section 2, in particular properties (2.3), (2.16), and therefore (2.28). Thus, if we define $\tilde{\pi}_2^0$ in the same manner as $\tilde{\pi}_1$, i.e.,

$$\tilde{\pi}_2^0(\delta^{(1)}, \cdots, \delta^{(k)}) = \pi_2^0(D^{(1)}, \cdots, D^{(k)}) \tag{3.21}$$

(cf. (2.30) and (2.31)), then $\tilde{\pi}_2^0$ satisfies the same properties as $\tilde{\pi}_1$, in particular (2.35), (2.39), (2.40), and (2.41). Therefore π_2^0 satisfies

$$\pi_2^0(D^{(1)}, \cdots, D^{(k)}) \geqslant \pi_2^0(\sigma_1 I_p, \cdots, \sigma_k I_p), \tag{3.22}$$

where $\sigma_\alpha = |D^{(\alpha)}|^{1/p} = |\Sigma^{(\alpha)}|^{1/p}$, which is the analog of (2.6). (We remark that if $k = 2$, then $\tilde{\pi}_2^0(0, \delta)$ satisfies (2.37) and hence π_2^0 satisfies Proposition 2.4, i.e., $\pi_2^0(\Sigma^{(1)}, \Sigma^{(2)})$ is a Schur-convex function of the logarithms of the characteristic roots of $(\Sigma^{(1)})^{-1}\Sigma^{(2)}$.)

We now apply the Pitman-Sugiura-Nagao-Giri method to obtain the analog of (2.7) for π_2^0. The presence of $\bar{X}$ in the statistic λ_2 requires some preliminary manipulations. Define

$$\hat{X} = \bar{X} D_N^{-\frac{1}{2}}$$

$$\hat{B} = D_N^{-\frac{1}{2}} B D_N^{-\frac{1}{2}} \equiv I_k - N^{-1}\hat{N}\hat{N}',$$

where $\hat{N} = D_N^{-\frac{1}{2}}\tilde{N} = (N_1^{\frac{1}{2}}, \cdots, N_k^{\frac{1}{2}})'$. The matrix $\hat{B}$ is a symmetric idempotent matrix of rank $k - 1$, so there exists a $k \times k$ orthogonal matrix ψ such that

$$\hat{B} = \psi\begin{pmatrix} I_{k-1} & 0 \\ 0 & 0 \end{pmatrix}\psi' \equiv \psi_1\psi_1',$$

where $\psi = (\psi_1, \psi_2)$ with ψ_1: $k \times (k - 1)$ and ψ_2: $k \times 1$. For later use, note that

$\psi_1'\psi_1 = I_{k-1}$, and that we can choose ψ such that $\psi_2 = (\hat{N}'\hat{N})^{-\frac{1}{2}}\hat{N}$. Lastly, define

$$Z = \hat{X}\psi = (\hat{X}\psi_1, \hat{X}\psi_2) \equiv (Z_1, Z_2),$$

so that

$$T \equiv \bar{X}B\bar{X}' = \hat{X}\hat{B}\hat{X}' = Z_1Z_1'.$$

When $\mu^{(\alpha)} = 0$ and $\Sigma^{(\alpha)} = \sigma_\alpha I_p$, $1 \leq \alpha \leq k$, the p rows of the $p \times (k-1)$ random matrix Z_1 are independent and identically distributed, each having the nonsingular normal distribution $N_{k-1}(0, \psi_1'D_\sigma\psi_1)$, where $D_\sigma = \text{diag}(\sigma_1, \cdots, \sigma_k)$. Thus the density of Z_1 is given by

$$(3.23) \qquad (2\pi)^{-p(k-1)/2}|\psi_1'D_\sigma\psi_1|^{-p/2}\exp\{-\tfrac{1}{2}\text{tr}\, Z_1(\psi_1'D_\sigma\psi_1)^{-1}Z_1'\}.$$

Now let $\tau_\alpha = \sigma_\alpha/\sigma_1$, $1 \leq \alpha \leq k$, and let $D_\tau = \text{diag}(1, \tau_2, \cdots, \tau_k)$. Then

$$\begin{aligned} &1 - \pi_2^0(\sigma_1 I_p, \cdots, \sigma_k I_p) = 1 - \pi_2^0(I_p, \tau_2 I_p, \cdots, \tau_k I_p) \\ (3.24) \qquad &= M \cdot \int_{\{\lambda_2 < c_2\}} |\psi_1'D_\tau\psi_1|^{-p/2}\left[\Pi_{\alpha=1}^k \tau_\alpha^{-pn_\alpha/2}|S^{(\alpha)}|^{(n_\alpha-p-1)/2}\right] \\ &\cdot\exp\{-\tfrac{1}{2}\text{tr}\left[\Sigma_{\alpha=1}^k\tau_\alpha^{-1}S^{(\alpha)} + Z_1(\psi_1'D_\tau\psi_1)^{-1}Z_1'\right]\}\, dZ_1\Pi_{\alpha=1}^k\, dS^{(\alpha)}. \end{aligned}$$

Make the transformation $R = (S^{(1)})^{-\frac{1}{2}}Z_1$, $U_\alpha = (S^{(1)})^{-\frac{1}{2}}S^{(\alpha)}(S^{(1)})^{-\frac{1}{2}}$, $2 \leq \alpha \leq k$, and integrate out $S^{(1)}$. The Jacobian is $|S^{(1)}|^{(p+2)(k-1)/2}$, so (3.24) becomes

(3.25)

$$M'\int_{\{f(U_2, \cdots, U_k, R) > c'\}}|\psi_1'D_\tau\psi_1|^{-p/2}\cdot|I_p + \Sigma_{\alpha=2}^k\tau_\alpha^{-1}U_\alpha + R(\psi_1'D_\tau\psi_1)^{-1}R'|^{-(n+k-1)/2}$$
$$\left[\Pi_{\alpha=2}^k\tau_\alpha^{-pn_\alpha/2}|U_\alpha|^{(n_\alpha-p-1)/2}\, dU_\alpha\right] dR,$$

where $c' = c_2^{-(N-1)/2N}$ and

$$(3.26) \quad f(U_2, \cdots, U_k, R) = \left[|I_p + \Sigma_{\alpha=2}^k U_\alpha + RR'|^{-N/2}\Pi_{\alpha=2}^k|U_\alpha|^{N_\alpha/2}\right]^{(N-1)/N}$$

(note that $n + k - 1 = N - 1$). Now let $Q = R(\psi_1'D_\tau\psi_1)^{-\frac{1}{2}}$ and $V_\alpha = \tau_\alpha^{-1}U_\alpha$, $2 \leq \alpha \leq k$, so that

$$\begin{aligned} (3.27) \quad &1 - \pi_2^0(I_p, \tau_2 I_p, \cdots, \tau_k I_p) \\ &= M \cdot \int_{\{f(\tau V, \tau Q) > c'\}} f(V, Q)\left[\Pi_{\alpha=2}^k|V_\alpha|^{(N_\alpha-N)/2N}\right]\left[\Pi_{\alpha=2}^k\frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}\right] dQ, \end{aligned}$$

where

$$\begin{aligned} (3.28) \qquad &f(V, Q) \equiv f(V_2, \cdots, V_k, Q) \\ &f(\tau V, \tau Q) \equiv f\left(\tau_2 V_2, \cdots, \tau_k V_k, Q(\psi_1'D_\tau\psi_1)^{\frac{1}{2}}\right). \end{aligned}$$

Hence, arguing as in (2.45) and (2.50),

$$\begin{aligned} &\pi_2^0(I_p, \tau_2 I_p, \cdots, \tau_k I_p) - \pi_2^0(I_p, \cdots, I_p) \\ &\qquad = M'\left[\int_{\{f(V,Q) > c' > f(\tau V, \tau Q)\}} - \int_{\{f(\tau V, \tau Q) > c' > f(V,Q)\}}\right] \\ &\qquad \cdot f(V, Q)\Pi_{\alpha=2}^k|V_\alpha|^{(N_\alpha-N)/2N}\Pi_{\alpha=2}^k\frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}\, dQ \end{aligned}$$

(3.29)

$$\geq c'M'\left[\int_{\{f(V,Q)>c'\}} - \int_{\{f(\tau V,\tau Q)>c'\}}\right]\Pi_{\alpha=2}^{k}|V_\alpha|^{(N_\alpha-N)/2N}\Pi_{\alpha=2}^{k}\frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}\,dQ$$

$$= c'M'\int_{\{f(V,Q)>c'\}}\left[1-\left(\frac{\Pi_{\alpha=2}^{k}\tau_\alpha^{(N-N_\alpha)/N}}{|\psi_1'D_\tau\psi_1|}\right)^{p/2}\right]$$

$$\cdot\,\Pi_{\alpha=2}^{k}|V_\alpha|^{(N_\alpha-N)/2N}\Pi_{\alpha=2}^{k}\frac{dV_\alpha}{|V_\alpha|^{(p+1)/2}}\,dQ.$$

Therefore, the proof will be complete if it can be shown that

$$|\psi_1'D_\tau\psi_1| \geq \Pi_{\alpha=1}^{k}\tau_\alpha^{(N-N_\alpha)/N} \tag{3.30}$$

for all positive $\tau_2, \cdots, \tau_k$ (recall that $\tau_1 = 1$). Since $\psi \equiv (\psi_1, \psi_2)$ is orthogonal, we have

$$\Pi_{\alpha=1}^{k}\tau_\alpha = |D_\tau| = |\psi' D_\tau \psi| = \frac{|\psi'_1 D_\tau \psi_1|}{\psi_2' D_\tau^{-1}\psi_2}. \tag{3.31}$$

Thus, since $\psi_2 = (\Sigma_{\alpha=1}^{k}N_\alpha)^{-\frac{1}{2}}\hat{N} \equiv N^{-\frac{1}{2}}\hat{N}$,

$$|\psi_1'D_\tau\psi_1| = \left[\Sigma_{\alpha=1}^{k}\left(\frac{N_\alpha}{N}\right)\tau_\alpha^{-1}\right](\Pi_{\alpha=1}^{k}\tau_\alpha), \tag{3.32}$$

so the desired inequality (3.30) is equivalent to

$$\Sigma_{\alpha=1}^{k}\left(\frac{N_\alpha}{N}\right)\nu_\alpha \geq \Pi_{\alpha=1}^{k}\nu_\alpha^{N_\alpha/N}, \tag{3.33}$$

where $\nu_\alpha = \tau_\alpha^{-1}$, which is true by the concavity of $h(x) \equiv \log x$. Therefore (3.29) is nonnegative, which establishes the analog of (2.7) for π_2^0. The proof of Theorem 3.1 is complete.

Acknowledgment. A referee deserves thanks for carefully reading the manuscript and suggesting improvements in its presentation.

REFERENCES

[1] ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176.

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

[3] BELLMAN, R. (1970). *Introduction to Matrix Analysis*, 2nd edition. McGraw-Hill, New York.

[4] BROWN, G. W. (1939). On the power of the L_1 test for equality of several variances. *Ann. Math. Statist.* **10** 119–128.

[5] CARTER, E. M. and SRIVASTAVA, M. S. (1977). Monotonicity of the power functions of modified likelihood ratio criterion for the homogeneity of variances and of the sphericity test. *J. Multivariate Anal.* **7** 229–233.

[6] COHEN, A. and STRAWDERMAN, W. (1971). Unbiasedness of tests for homogeneity of variances. *Ann. Math. Statist.* **42** 355–360.

[7] DAS GUPTA, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. *Ann. Math. Statist.* **40** 697–701.

[8] DAS GUPTA, S. and GIRI, N. (1973). Properties of tests concerning covariance matrices of normal distributions. *Ann. Statist.* **1** 1222–1224.

[9] GIRI, N. (1973). On a class of unbiased tests for the equality of k covariance matrices. In *Multivariate Statistical Inference*, (D. G. Kabe and R. P. Gupta, eds.) North Holland, Amsterdam and American Elsevier, New York.

[10] HARDY, G. H., LITTLEWOOD, J. E., and POLYA, G. (1952). *Inequalities*, 2nd edition. Cambridge University Press, Cambridge.

[11] OLKIN, I. and RUBIN, H. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.

[12] PERLMAN, M. D. (1980). Tests for variances and covariance matrices: unbiasedness and monotonicity of power. Unpublished manuscript.

[13] PITMAN, E. J. G. (1939). Tests of hypotheses concerning location and scale parameters. *Biometrika* **31** 200–215.

[14] RAMACHANDRAN, K. V. (1958). A test of variances. *J. Amer. Statist. Assoc.* **53** 741–747.

[15] RAO, C. R. (1973). *Linear Statistical Inference*, 2nd edition. Wiley, New York.

[16] SRIVASTAVA, M. S., KHATRI, C. G. and CARTER, E. M. (1978). On monotonicity of the modified likelihood ratio test for the equality of two covariances. *J. Multivariate Anal.* **8** 262–267.

[17] SUGIURA, N. and NAGAO, N. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Ann. Math. Statist.* **39** 1686–1692.

[18] SUGIURA, N. and NAGAO, N. (1969). On Bartlett's test and Lehmann's test for homogeneity of variances. *Ann. Math. Statist.* **40** 2018–2032.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195