Bootstrap Critical Values for Testing Homogeneity of Covariance Matrices

Ji Zhang and Dennis D. Boos

Institute of Statistics Mimeo Series No. 1949

June 1989

ABSTRACT

Bartlett's modified likelihood ratio statistic Λ is often suggested in multivariate analysis for testing equality of covariance matrices. Unfortunately, the χ^2-approximation to the null distribution of $-2 \log \Lambda$ is only useful when the data is very close to the normal distribution. This paper presents a pooled bootstrap procedure which replaces the χ^2-approximation and makes Bartlett's statistic a useful tool for data analysis.

KEY WORDS : Bartlett's test; Discriminant analysis; Equal covariance matrices; Multivariate analysis; Resampling; Test validity.

1. INTRODUCTION

The assumption of equal covariance matrices arises in numerous multivariate statistical analyses. One very important application is in discriminant analysis where the assumption is needed to justify the use of Fisher's linear discriminant function (LDF). If the covariance matrices are not equal, then we might prefer the quadratic discriminant function (QDF), which uses separate covariance estimates instead of the pooled estimate (Gnanadesikan and Kettenring, 1989). The LDF, however, is much simpler to use than the QDF and is more powerful when covariance matrices are approximately equal.

Thus there is some interest in testing for homogeneity of covariance matrices when sampling from k independent multivariate distributions. The standard test statistic is Bartlett's modified likelihood ratio statistic:

$$\Lambda = \frac{\left(\prod_{i=1}^{k} |A_i|^{(n_i-1)/2}\right)(N - k)^{p(N-k)/2}}{|A|^{(N-k)/2} \prod_{i=1}^{k} (n_i - 1)^{p(n_i-1)/2}} , \tag{1.1}$$

where $A_i = \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)(X_{ij} - \overline{X}_i)'$, $\overline{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$, $A = \sum_{i=1}^{k} A_i$ and $N = \sum_{i=1}^{k} n_i$,

for k samples of independent $p \times 1$ random vectors $\left\{X_{i1} , \dots , X_{in_i} , \ i = 1, \dots, k\right\}$.
Under the null hypothesis of equal covariance matrices and *multivariate normality* (Muirhead, 1982, pp. 298 – 309)

$$-2 \log \Lambda \ \xrightarrow{d} \ \chi^2_{p(p+1)(k-1)/2} , \tag{1.2}$$

as $\min(n_1, \ldots, n_k) \rightarrow +\infty$ with $\frac{n_i}{N} \rightarrow \lambda_i \in (0, 1)$ for $i = 1, \ldots, k$.

Although the statistic Λ is derived under the assumption of multivariate normal distributions, it is certainly a reasonable statistic for comparing covariance matrices even when the distributions are not normal. The asymptotic chi-squared distribution in (1.2), however, is very sensitive to the normality assumption. Even small deviations from normality can upset this asymptotic distribution and lead to *very liberal* tests if critical values based on (1.2) are used (see, e.g. , Table 1). If the normality assumption is replaced by the assumption of a common elliptical distribution with kurtosis parameter κ, then κ can be estimated and an asymptotically valid test based on (1.1) can be constructed (Muirhead, 1982, pp. 331). The elliptical distribution assumption, however, is still quite strong and hard to verify in small samples.

This paper shows how to use a bootstrap procedure to estimate critical values for (1.1) and other test statistics when the data cannot be assumed to be normal. The procedure is based on pooling samples after substracting means and is similar to a procedure proposed by Boos and Brownie (1989) for the univariate case. In Section 2 we present the bootstrap method and some asymptotic results to justify its use. Details of the proofs are given in the Appendix. Section 3 reports on some Monte Carlo work which shows that the procedure works well in sample sizes as low as $n_i = 20$. Finally, in Section 4 we illustrate the bootstrap approach in the context of a two sample discriminant analysis problem which partially motivated the research.

We conclude this section by noting that the popular statistical package, SAS, has a procedure DISCRIM with option POOL = TEST which tests equality of covariance matrices using (1.1) and (1.2) and then uses the QDF in place of the LDF if the test rejects. The practical result is that if the data have approximately equal covariance matrices but are not normally distributed, then the suboptimal QDF will be used much more often then it should be. The example in Section 4 further illustrates this point.

2. ASYMPTOTIC PROPERTIES OF THE BOOTSTRAP PROCEDURE

Let $\left\{X_{i1}, \dots, X_{in_i}, \ i = 1, \dots, k\right\}$ be k samples of independent $p\times 1$ random vectors, where X_{ij} has mean vector μ_i, covariance matrix Σ_i, and distribution function $G_i(x)$. Our results are for the null hypothesis

$$H_0: \ \Sigma_1 = \Sigma_2 = \dots = \Sigma_k \ \text{ and } \ \mu_4(G_1) = \mu_4(G_2) = \dots = \mu_4(G_k), \tag{2.1}$$

where the fourth moment $\mu_4(G_i)$ is defined below. The extra assumption on the fourth moments is used to pool distributional information from each sample. Note that the assumption that the $G_i(x)$ are from the same elliptical family with equal Σ_i is much stronger than (2.1).

We define the resampling space R_S to be

$$R_S = \left\{ X_{ij} - \hat{\mu}_i, \ j=1, \dots, n_i, \ i=1, \dots, k \right\}, \tag{2.2}$$

where $\hat{\mu}_i$ is a location estimator with the translation property

$$\hat{\mu}_i(Y + c\cdot e_{n_i}') = \hat{\mu}_i(Y) + c,$$

where $Y = (Y_1, Y_2, \dots, Y_{n_i})$ is a $p\times n_i$ data matrix, e_{n_i} is a $n_i\times 1$ vector of 1's, and c is a $p\times 1$ vector of constants. Usually we let $\hat{\mu}_i = \overline{X}_i$, the ith sample mean vector, but the theorems allow $\hat{\mu}_i$ to be a vector of sample medians, etc. Our bootstrap method is to draw $\left\{ X_{i1}^*, \dots, X_{in_i}^*, \ i = 1, \dots, k \right\}$ with replacement from R_S. Equivalently, we may view these bootstrap random samples as iid samples from the distribution function

$$G_N(x) = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} I(X_{ij} - \hat{\mu}_i \le x), \tag{2.3}$$

where I is the indicator function and $\le$ means elementwise. The practical implementation of this approach and some small sample results are described in Section 3. Further motivation and a discussion of philosophical issues are given in Boos and Brownie (1989, Section 2).

In this section we want to give an asymptotic justification for resampling from (2.2). In particular we shall show for a class of statistics including Bartlett's $-2 \log \Lambda$ that the bootstrap distribution based on resampling from (2.2) converges with probability one to the same distribution as the true limit distribution under (2.1). All proofs are given in the Appendix.

We will first give a straightforward result (Theorem 1) on the limit distribution of quadratic-type statistics and then follow it with the bootstrap results (Theorem 2). Corollary 1 and Corollary 2 will be specific for Bartlett's statistic. The minimal notation we need here is as follows:

i) I_p is the $p \times p$ identity matrix, and e_p is a $p \times 1$ vector of 1's.

ii) For a $p \times p$ symmetric matrix M with elements m_{ij}, let $\text{uvec}\{M\} = M^u$ be a $p_1 \times 1$ vector formed from the elements in the upper triangular half of M, including the diagonal elements, where $p_1 = \frac{p(p+1)}{2}$,

$$M^u = (m_{11}, m_{12}, \ldots, m_{1p}, m_{22}, \ldots, m_{2p}, \ldots, m_{p-1\ p-1}, m_{p-1\ p}, m_{pp})'.$$

iii) For a random $p \times 1$ vector $X = (x_1, x_2, \ldots, x_p)'$ with mean μ, covariance matrix Σ, and finite fourth moments, let μ_4 and β_2 be symmetric $p_1 \times p_1$ matrices:

$$\mu_4 = \mathcal{E}\Big([\text{uvec}\{(X - \mu)(X - \mu)'\}][\text{uvec}\{(X - \mu)(X - \mu)'\}]'\Big), \text{ and}$$

$$\beta_2 = \mathcal{E}\Big([\text{uvec}\{\Sigma^{-\frac{1}{2}}(X - \mu)(X - \mu)'\Sigma^{-\frac{1}{2}}\}][\text{uvec}\{\Sigma^{-\frac{1}{2}}(X - \mu)(X - \mu)'\Sigma^{-\frac{1}{2}}\}]'\Big),$$

where $\Sigma^{-\frac{1}{2}}$ is the symmetric square root decomposition matrix of Σ^{-1} such that $\Sigma^{-1} = \Sigma^{-\frac{1}{2}} \cdot \Sigma^{-\frac{1}{2}}$.

iv) For a scalar b, let $H(b) = \text{Diag}[\text{uvec}\{(b - 1)I_p + e_p e_p'\}]$, a $p_1 \times p_1$ matrix.

v) Finally, let $Q(S_1, S_2, \ldots, S_k)$ be a function of the $p \times p$ symmetric positive definite matrices $S_1, S_2, \ldots, S_k$ such that

Q1) $Q(BS_1B', BS_2B', \ldots, BS_kB') = Q(S_1, S_2, \ldots, S_k)$, for any non-singular $p \times p$ matrix B ;

Q2) $Q(I_p , I_p , \dots , I_p) = 0$;

Q3) $\frac{\partial Q}{\partial V}(S_1 , S_2 , \dots , S_k) = 0$ at $(S_1 , S_2 , \dots , S_k) = (I_p , I_p , \dots , I_p)$,
where $V = \left((S_1^u)' , \dots , (S_k^u)'\right)'$;

Q4) Q has continuous second partial derivatives, and

$$A = \frac{1}{2} \frac{\partial^2 Q}{\partial V' \partial V}(S_1 , S_2 , \dots , S_k) \text{ at } (S_1 , S_2 , \dots , S_k) = (I_p , I_p , \dots , I_p) ,$$

a $kp_1 \times kp_1$ matrix, is not the zero matrix.

Theorem 1 below is a general result for Q functions satisfying Q1 – Q4 . The Q function associated with Bartlett's $-2 \log \Lambda$ in (1.2) is

$$Q(S_1 , S_2 , \dots , S_k) = \sum_{i=1}^{k} \lambda_i \log[\frac{|S|}{|S_i|}] ,$$

where $S = \sum_{i=1}^{k} \lambda_i S_i$, $\lambda_i \in (0, 1)$ for $i = 1, \dots, k$, $\sum_{i=1}^{k} \lambda_i = 1$. Other interesting Q functions are from the likelihood ratio tests of H_0 versus H_a : the covariance matrices are proportional and H_0 versus H_a : the correlation matrices are equal (see Manly and Rayner, 1987). Zhang (1989) verifies that all three of these Q functions satisfy Q1 – Q4 above.

Theorem 1. Let $\left\{X_{i1} , \dots , X_{in_i} , i = 1, \dots, k\right\}$ be k samples of independent $p \times 1$ random vectors, where X_{ij} has mean vector μ_i , covariance matrix Σ_i , and distribution function $G_i(x)$. Suppose that $\mu_4(G_i) < \infty$ for $i = 1, \dots, k$, and H_0 of (2.1) holds. Then if $Q(S_1 , \dots , S_k)$ satisfies Q1 – Q4 , where $S_1 , \dots , S_k$ are the k sample covariances, and $\min(n_1 , \dots , n_k) \to +\infty$ with $\frac{n_i}{N} \to \lambda_i \in (0, 1)$ for $i = 1, \dots, k$, we have that

$$NQ(S_1, \ldots, S_k) \xrightarrow{d} Z'\mathbf{A}Z,$$

where $\mathbf{A}$ is defined in Q4, and Z is a $kp_1 \times 1$ multivariate normal random vector with mean vector 0 and covariance matrix $\mathrm{Diag}[\lambda_i^{-1},\ i = 1, \ldots, k] \otimes (\beta_2 - I_p^4)$, where $I_p^4 = \mathrm{uvec}\{I_p\} \cdot (\mathrm{uvec}\{I_p\})'$, and $\beta_2 = \beta_2(G_1) = \ldots = \beta_2(G_k)$.

Corollary 1. If the conditions of Theorem 1 hold, then

$$-2 \log \Lambda \xrightarrow{d} Z'\mathbf{A}Z,$$

where Λ is given by (1.1), Z is defined in Theorem 1, and $\mathbf{A} = (\mathrm{Diag}[\lambda] - \lambda\lambda') \otimes H(\frac{1}{2})$, with $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)'$, and $H(\frac{1}{2})$ is defined in iv) of the notation.

Remarks about *Corollary 1.*

i) Here we describe $Z'\mathbf{A}Z$ more completely. If we let

$$\Sigma_0^{\frac{1}{2}} = \left(\mathrm{Diag}[\lambda_i^{-\frac{1}{2}},\ i = 1, \ldots, k]\right) \otimes H(2^{\frac{1}{2}}),$$

then the $kp_1 \times 1$ random vector $U = \Sigma_0^{\frac{1}{2}}\mathbf{A}Z$ has a multivariate normal distribution with mean vector 0 and covariance matrix $\Sigma_u = [\ I_k - \lambda^{\frac{1}{2}}\lambda^{\frac{1}{2}\prime}\ ] \otimes \Sigma_{u1}$, where $\lambda^{\frac{1}{2}} = (\lambda_1^{\frac{1}{2}}, \ldots, \lambda_k^{\frac{1}{2}})'$, and $\Sigma_{u1} = [\ H(2^{-\frac{1}{2}})\beta_2 H(2^{-\frac{1}{2}}) - \frac{1}{2}I_p^4\ ]$. Note that $[\ I_k - \lambda^{\frac{1}{2}}\lambda^{\frac{1}{2}\prime}\ ]$ is an idempotent matrix with rank $k - 1$. Thus if we denote by f_1 the rank of Σ_{u1}, we have that

$$-2 \log \Lambda \xrightarrow{d} Z'\mathbf{A}Z \stackrel{d}{=} U'U \stackrel{d}{=} \sum_{i=1}^{f_1} \alpha_i \chi^2_{k-1}(i),$$

where α_i are the non-zero latent roots of Σ_{u1}, and the $\chi^2_{k-1}(i)$ are independent χ^2_{k-1} variables.

ii) When the X_{ij} are all from the same elliptical distribution with kurtosis parameter κ (Muirhead, 1982, pp. 40 – 41), then under H_0, $\Sigma^{-\frac{1}{2}}(X_{ij} - \mu_i)$ has the same elliptical distribution with covariance matrix I_p , $\beta_2 - I_p^4 = (1+\kappa)H(2) + \kappa I_p^4$, and $\Sigma_{u1} = (1+\kappa)I_{p_1} + \frac{\kappa}{2}I_p^4$. Thus

$$-2 \log \Lambda/(1+\kappa) \xrightarrow{d} Z'AZ/(1+\kappa) \stackrel{d}{=} U'U/(1+\kappa) \stackrel{d}{=} \sum_{i=1}^{f_1} \alpha_i^e \chi_{k-1}^2(i) ,$$

where α_i^e are the non-zero latent roots of $\Sigma_{u1}/(1+\kappa)$. Note that in this case, $f_1 = p_1 = p(p+1)/2 = \text{rank}(\Sigma_{u1}/(1+\kappa))$. If $k = 2$ this result is the same as Muirhead's Theorem 8.2.18 (1982, pp. 329 – 331), where his matrix V is of order p^2 as a result of counting all the elements in the sample covariance matrix.

iii) When the X_{ij} are all from multivariate normal distributions, then the kurtosis $\kappa = 0$ and all the latent roots of $\Sigma_{u1}/(1+\kappa)$ are equal to one. Hence we get the classical result

$$-2 \log \Lambda \xrightarrow{d} Z'AZ \stackrel{d}{=} \sum_{i=1}^{f_1} \chi_{k-1}^2(i) = \chi_f^2 ,$$

where $f = f_1 \cdot (k - 1) = p(p+1)(k - 1)/2$.

Our next result is an analog of Theorem 1 for $Q(S_1^*, \ldots, S_k^*)$, where $S_1^*, \ldots, S_k^*$ are sample covariance matrices of bootstrap samples from $G_N(x)$ of (2.3). Here the null hypothesis H_0: $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_k$ need not hold since our bootstrap method forces a null distribution in the bootstrap resampling space where we always have that H_0^*: $\Sigma_1^* = \Sigma_2^* = \ldots = \Sigma_k^* = \Sigma(G_N)$. Since the bootstrap distribution of a statistic is random, depending on the data in some fashion, we use the notation $\xrightarrow{d^*}$ a.s. to denote "convergence in distribution almost surely" for such random distributions.

Theorem 2. Let $\left\{X_{i1}, \ldots, X_{in_i}, i = 1, \ldots, k\right\}$ be k samples of independent $p \times 1$ random vectors, where X_{ij} has mean vector μ_i , covariance matrix Σ_i , and distribution function $G_i(x)$ with finite fourth moment. Suppose that $S_1^*, \ldots, S_k^*$ are k sample covariances based on the iid bootstrap samples $\left\{X_{i1}^*, \ldots, X_{in_i}^*, i = 1, \ldots, k\right\}$ drawn from $G_N(x)$ of (2.3). If Q satisfies Q1 – Q4 and

$\min(n_1, \ldots, n_k) \to +\infty$ with $\frac{n_i}{N} \to \lambda_i \in (0, 1)$ and $\hat{\mu}_i \overset{a.s.}{\to} \mu_{i\infty}$ for $i = 1, \ldots, k$, then

$$NQ(S_1^*, \ldots, S_k^*) \overset{d^*}{\to} Z'AZ \quad \text{a.s.},$$

where A is defined in Q4, and Z is a $kp_1 \times 1$ multivariate normal random vector with mean vector 0 and covariance matrix $\text{Diag}[\lambda_i^{-1}, i = 1, \ldots, k] \otimes (\bar{\beta}_2(G) - I_p^4)$, and

$$G(x) = \sum_{i=1}^{k} \lambda_i G_i(x + \mu_{i\infty}) . \tag{2.4}$$

When (2.1) holds, the result in Theorem 2 only changes by the fact that $\beta_2(G) = \beta_2$, where β_2 is the common value of $\beta_2(G_1), \ldots, \beta_2(G_k)$, provided that $\mu_i - \mu_{i\infty} = d$ for $i = 1, \ldots, k$. In Corollary 2 we assume (2.1) and concentrate on Bartlett's statistic (1.1).

Corollary 2. If the conditions of Theorem 2 and (2.1) both hold, and $\mu_i - \mu_{i\infty} = d$ for $i = 1, \ldots, k$, then

$$-2 \log \Lambda^* \overset{d^*}{\to} Z'AZ \quad \text{a.s.},$$

where Λ^* is (1.1) when data is generated from (2.3), and Z and A are as in Corollary 1.

Therefore from Corollary 1 and Corollary 2 we get the result that the probability of rejecting (2.1) converges to α when using Bartlett's statistic $L = -2 \log \Lambda$ as a test statistic and level α critical values are chosen from the bootstrap distribution via sampling from $G_N(x)$ of (2.3). In other words, the resulting test procedure is asymptotically valid. In the next section we check the validity in finite samples for several situations. Note also that under the alternative hypothesis H_a, L is asymptotically normal while $L^* \overset{d^*}{\to} Z'AZ$ a.s. continues to hold by Theorem 2. Thus we have that $P(L \geq \text{Bootstrap } \alpha \text{ critical value} \mid H_a) \to 1$ as $\min(n_1, \ldots, n_k) \to +\infty$ (see Boos and Brownie, 1989, pp.72).

3. MONTE CARLO SIMULATIONS

To carry out our bootstrap procedure for Bartlett's statistic $L = -2 \log \Lambda$ we generate B sets of k independent samples by sampling with replacement from R_S of (2.2) with $\hat{\mu}_i = \overline{X}_i$, and calculate L for each set. We label these $L_1^*, \ldots, L_B^*$ and note that their empirical distribution is the bootstrap estimate of the null distribution of L. Since H_0 is rejected for large L, the bootstrap level α critical value for L is the $(1-\alpha)^{th}$ percentile of the L^* distribution, and the bootstrap p value $\hat{p}_B$ is the proportion of the L_i^* that are at least as large as L_0, the value of L based on the original sample.

In this research, we investigated the properties of our bootstrap procedure under three different families of distributions:

i) The multivariate normal distribution — $MN(\mu, \Sigma)$;

ii) The multivariate Student's t distribution with 5 degrees of freedom — $MT(5; \mu, \Sigma)$;

iii) The contaminated normal distribution — $CN(2, 0.9; 0, 1)$.

$MT(5; \mu, \Sigma)$ is generated as $MN(\mu, \Sigma)$ divided by $(\chi^2_{(5)}/5)^{\frac{1}{2}}$, and $CN(2, 0.9; 0, 1)$ is generated by letting each component be a $N(0, 1)$ with probability 0.9 and a $\chi^2_{(2)}$ with probability 0.1, where the components are independent of each other. Note that all three distributions have finite fourth moments. MT is elliptically distributed with its kurtosis being larger than that of MN, and CN is non-elliptically distributed with its kurtosis even larger.

In all simulations, 1000 independent sets of Monte Carlo random samples were generated and p values were computed from our bootstrap procedure and from the χ^2-approximation of (1.2). Actually, we multiplied $L = -2 \log \Lambda$ by a correction γ and used $-2\gamma \log \Lambda$ when calculating the p value from the χ^2-approximation of (1.2) (Mardia, Kent, and Bibby, 1982, pp. 140), where

TABLE 1. Simulations Under H_0 for Bartlett's $-2 \log \Lambda$ for $k = 2$ Populations

	DISTRIBUTIONS					
Nominal	Normal (μ, Σ)		Multi-T $(5; \mu, \Sigma)$		Contaminated Normal	
Level	χ^2	Bootstrap	χ^2	Bootstrap	χ^2	Bootstrap
	$n_1 = 20$, $n_2 = 20$, $p = 2$ dimensions					
0.05	0.059	0.046	0.231	0.045	0.315	0.050
0.10	0.100	0.095	0.327	0.098	0.403	0.112
	$n_1 = 20$, $n_2 = 40$, $p = 2$ dimensions					
0.05	0.057	0.055	0.230	0.054	0.335	0.064
0.10	0.108	0.104	0.323	0.109	0.425	0.133
	$n_1 = 20$, $n_2 = 20$, $p = 5$ dimensions					
0.05	0.056	0.012	0.374	0.023	0.390	0.019
0.10	0.110	0.055	0.496	0.067	0.490	0.077
	$n_1 = 20$, $n_2 = 40$, $p = 5$ dimensions					
0.05	0.045	0.024	0.430	0.030	0.397	0.031
0.10	0.086	0.055	0.531	0.077	0.510	0.092

NOTE: Entries are the proportion of rejections in 1000 Monte Carlo replications, using a 3 stage procedure in the bootstrap simulations (see the text for details).

TABLE 2. Simulations Under H_0 for Bartlett's $-2 \log \Lambda$ for $k = 6$ Populations

Nominal Level	DISTRIBUTIONS: Normal (μ, Σ) χ^2	Normal (μ, Σ) Bootstrap	Multi-T $(5; \mu, \Sigma)$ χ^2	Multi-T $(5; \mu, \Sigma)$ Bootstrap	Contaminated Normal χ^2	Contaminated Normal Bootstrap
$n_1 = n_2 = \dots = n_6 = 20$, $p = 2$ dimensions						
0.05	0.042	0.031	0.511	0.031	0.700	0.056
0.10	0.090	0.081	0.628	0.102	0.767	0.124
$n_1 = n_2 = 20$, $n_3 = n_4 = 30$, $n_5 = n_6 = 40$, $p = 2$ dimensions						
0.05	0.046	0.044	0.627	0.036	0.802	0.049
0.10	0.082	0.079	0.723	0.095	0.864	0.110
$n_1 = n_2 = \dots = n_6 = 20$, $p = 5$ dimensions						
0.05	0.061	0.045	0.887	0.061	0.853	0.038
0.10	0.109	0.087	0.936	0.135	0.908	0.110
$n_1 = n_2 = 20$, $n_3 = n_4 = 30$, $n_5 = n_6 = 40$, $p = 5$ dimensions						
0.05	0.041	0.032	0.905	0.049	0.919	0.044
0.10	0.089	0.066	0.945	0.121	0.948	0.114

NOTE: Entries are the proportion of rejections in 1000 Monte Carlo replications, using a 3 stage procedure in the bootstrap simulations (see the text for details).

$$\gamma = 1 - \frac{2p^2 + 3p - 1}{6(p + 1)(k - 1)}\Big(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{N-k}\Big). \tag{3.1}$$

Rejection of H_0 at $\alpha = 0.05$ or $\alpha = 0.10$ means that a p value is less than 0.05 or a p value is less than 0.10, respectively. To calculate the bootstrap p value, we used a three-stage sequential procedure: a) start with $B = 100$; b) if $\hat{p}_B > 0.25$ then stop; else take 300 (100 for $k = 6$) more replications and use all $B = 400$ ($B = 200$ for $k = 6$) replications to compute a new $\hat{p}_B$; c) if $\hat{p}_B > 0.15$ then stop; else take 300 more replications and use all $B = 700$ ($B = 500$ for $k = 6$) replications to compute the final $\hat{p}_B$. The sequential procedure was used because of the large cpu time required. All simulations were run under SAS/IML release 5.18 on IBM cpu model 3090 computers using random generators RANUNI , RANNOR , and RANGAM.

We simulated four cases:

i) $k = 2$, $p = 2$, sample size = (20, 20), (20, 40);

ii) $k = 2$, $p = 5$, sample size = (20, 20), (20, 40);

iii) $k = 6$, $p = 2$, sample size = (20, 20, 20, 20, 20, 20), (20, 20, 30, 30, 40, 40);

iv) $k = 6$, $p = 5$, sample size = (20, 20, 20, 20, 20, 20), (20, 20, 30, 30, 40, 40).

The results are displayed in Table 1 and Table 2. A brief summary is as follows.

i) Although the bootstrap procedure is a little conservative in the normal case (MN), it is much better in the MT and CN cases, while the χ^2-approximation performs very badly in both of these latter two cases.

ii) For the sample sizes studied the bootstrap procedure performs better when the number of populations k increases, and it performs worse when dimension p increases.

iii) A further simulation (not displayed) using the larger sample sizes (30, 30) and (30, 50)

TABLE 3. Power Study for Bartlett's $-2 \log \Lambda$ for $k = 2$ Populations, $p = 2$ Dimensions, and $n_1 = n_2 = 20$

Nominal Level	MN(0, V) and MN(0, I) $\chi^2_{(3)}$	MN(0, V) and MN(0, I) Bootstrap	MT(5; 0, V) and MT(5; 0, I) $\chi^2_{(3)}$	MT(5; 0, V) and MT(5; 0, I) Bootstrap
0.05	0.762	0.642	0.763	0.487
0.05-adjusted	0.735	0.657	0.499	0.525
0.10	0.852	0.812	0.829	0.649
0.10-adjusted	0.851	0.817	0.599	0.658

Nominal Level	MN(0, C) and MN(0, I) $\chi^2_{(3)}$	MN(0, C) and MN(0, I) Bootstrap	MT(5; 0, C) and MT(5; 0, I) $\chi^2_{(3)}$	MT(5; 0, C) and MT(5; 0, I) Bootstrap
0.05	0.264	0.233	0.436	0.155
0.05-adjusted	0.244	0.243	0.164	0.177
0.10	0.386	0.372	0.533	0.294
0.10-adjusted	0.386	0.381	0.264	0.300

NOTE: Entries are the proportion of rejections in 1000 Monte Carlo replications, using a 3 stage procedure in the bootstrap simulations (see the text for details).

for $k = 2$ and $p = 5$ at the normal distribution helped confirm the convergence of the critical values.

We also did a power study for $k = 2$ (populations) and $p = 2$ (dimensions) under the normal distribution and the multivariate t_5 distribution. Here we studied two alternatives: a) with one covariance $V = \text{Diag}[2, 4]$, and the other one the identity matrix; b) with one covariance $C = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$, and the other one the identity matrix. Note that the deviation from the identity matrix is small in the second case. The results of the power study are shown in Table 3.

Two estimates of power are given in Table 3. The odd-numbered rows are the percentage of rejections when using the $\chi^2_{(3)}$ and bootstrap critical values. Since the $\chi^2_{(3)}$ produces very liberal tests when sampling from nonnormal distributions, some adjustment must be made to compare powers when sampling from the multivariate t_5 distribution. The even-numbered rows are obtained by using as critical values the 5th percentile of the empirical distribution of p values under H_0 obtained in constructing Table 1. Since these latter estimates involve two sources of error, they are more variable than the entries in the odd-numbered rows (which have standard deviation bounded by 0.016), and they have some bias.

Table 3 shows that the bootstrap loses a little power at the normal distribution but seems to gain in adjusted power at the multivariate t_5 distribution when compared to the $\chi^2_{(3)}$ procedure. This latter gain in power, however, seems unlikely and may be due to variability or bias in the adjusted estimates.

Tables 1 – 3 suggest that the bootstrap procedure holds its level well under H_0 and loses very little power under H_a, while the χ^2-approximation performs very badly for nonnormal distributions. Therefore, we feel that the bootstrap procedure can be highly recommended for data analysis.

4. EXAMPLE

The data in Table 4 is from a study of the effect of soil and bush characteristics on the presence or absence of blueberry maggots. The goal of the study was to give practical recommendations on where to expect blueberry bushes to be infested with maggots. We thank Gwen Pearson and the NCSU Entomology Department for the consulting interaction and for the use of the data.

The data was collected from numerous locations in North Carolina where blueberry maggots were found (27 sites, INFEST = 1) and from locations where they were absent (29 sites, INFEST = 0). Many variables were measured but preliminary analysis reduced the set to HT = average bush height, RAD = square root of the average bush radius, and CLAY = percent clay in the soil. We decided to use discriminant analysis.

When the SAS procedure DISCRIM with option POOL = TEST is run on the Table 4 data, the Bartlett's test for equal covariance matrices rejects equality with $p = 0.014$ and the SAS program selects the QDF method which separately estimates covariance matrices. The output shows this QDF approach then results in 10 misclassifications out of 56 (7 of the 27 INFEST = 1 classified as INFEST = 0 and 3 of the 29 INFEST = 0 classified as INFEST = 1).

Our bootstrap approach based on Bartlett's test with B = 4000 bootstrap replications gave the p value $p = 0.136$ for testing equality of the covariance matrices. Skewness and kurtosis multivariate normality tests (Mardia, et al., 1982, pp. 148 – 149) resulted in $p = 0.18$ for skewness and $p = 0.76$ for kurtosis for INFEST = 0 and $p = 0.000$ for skewness and $p = 0.002$ for kurtosis for INFEST = 1. Thus, the INFEST = 1 group is clearly not normally distributed which helps explain the difference in p values when using Bartlett's test with $\chi^2_{(6)}$ and bootstrap critical values.

TABLE 4. Data Set Used in the Example

OBS.	HT	RAD	CLAY	INFEST	OBS.	HT	RAD	CLAY	INFEST
1	80	7.4162	0.4	0	30	140	8.3666	1.3	1
2	95	7.0711	0.8	0	31	160	6.7082	1.7	1
3	50	6.3246	1.2	0	32	120	8.3666	0.4	1
4	60	5.0000	0.8	0	33	130	7.7460	3.6	1
5	80	7.0711	0.8	0	34	170	8.6603	1.3	1
6	90	6.7082	0.9	0	35	160	8.3666	1.2	1
7	90	6.3246	1.2	0	36	114	6.1644	0.8	1
8	65	6.3246	1.3	0	37	160	8.6603	1.6	1
9	70	6.7082	1.2	0	38	130	8.3666	1.2	1
10	110	7.0711	0.8	0	39	125	7.7460	1.2	1
11	140	7.7460	1.2	0	40	140	8.3666	0.8	1
12	110	8.9443	0.8	0	41	142	6.1644	1.2	1
13	130	8.3666	1.3	0	42	175	7.7460	0.8	1
14	100	8.3666	1.7	0	43	130	8.0623	1.7	1
15	118	5.9161	0.0	0	44	120	8.3666	1.7	1
16	120	7.7460	0.8	0	45	125	8.0623	2.1	1
17	116	5.1962	0.8	0	46	130	7.7460	1.6	1
18	134	5.8310	0.8	0	47	120	8.9443	1.2	1
19	140	8.6603	1.7	0	48	180	8.9443	2.1	1
20	100	5.6569	0.8	0	49	135	8.3666	2.5	1
21	120	5.9161	0.8	0	50	130	7.7460	1.2	1
22	122	5.2915	0.8	0	51	140	8.3666	1.6	1
23	115	6.3246	0.0	0	52	160	8.3666	2.5	1
24	130	7.0711	1.3	0	53	170	10.0000	0.8	1
25	105	6.3246	0.8	0	54	180	11.4018	1.2	1
26	100	8.3666	1.2	0	55	160	13.4164	0.8	1
27	150	8.9443	2.5	0	56	240	10.0000	0.0	1
28	102	7.0711	1.2	0					
29	129.5	5.9582	1.2	0					

Finally, we ran the Table 4 data through SAS DISCRIM with the POOL = YES option. The resulting LDF approach has only 7 misclassifications (3 of the 27 INFEST = 1 classified as INFEST = 0 and 4 of the 29 INFEST = 0 classified as INFEST = 1) and is much easier to interpret and explain. The difference between 7 and 10 misclassifications is not "proof" but it is consistent with the lack of robustness to nonnormality of the QDF mentioned in Gnanadesikan and Kettenring (1989, pp. 42).

The discriminant analysis above was run under SAS/STAT release 6.03 on an IBM PC – AT.

5. APPENDIX

Here we list some propositions and brief proofs of the Section 2 results. First we define some more notation.

For $p \times p$ matrix M and any symmetric $p \times p$ matrix S, there exists a $p_1 \times p_1$ matrix denoted by M_q (for "quadratic") with its elements being quadratic cross product of elements from M such that $(MSM')^u = M_q S^u$ for any S, and hence

$$\beta_2 = \mathcal{E}\Big([\mathrm{uvec}\{\Sigma^{-\frac{1}{2}}(X - \mu)(X - \mu)'\Sigma^{-\frac{1}{2}}\}][\mathrm{uvec}\{\Sigma^{-\frac{1}{2}}(X - \mu)(X - \mu)'\Sigma^{-\frac{1}{2}}\}]'\Big)$$

$$= (\Sigma^{-\frac{1}{2}})_q \cdot \mu_4 \cdot \Big((\Sigma^{-\frac{1}{2}})_q\Big)'.$$

For symmetric $p \times p$ matrix M we also define M^4 to be the symmetric $p_1 \times p_1$ matrix $M^4 = M^u(M^u)'$.

Proposition 1. Let $\Big\{X_{i1}, \ldots, X_{in_i}, \ i = 1, \ldots, k\Big\}$ be k samples of independent $p \times 1$ random vectors, where X_{ij} has mean vector μ_i, covariance matrix Σ_i, and distribution function $G_i(x)$ with finite fourth moment. Let S_i, $i = 1, \ldots, k$, be the sample covariance matrices. Then if

$\min(n_1, \ldots, n_k) \rightarrow +\infty$ with $\frac{n_i}{N} \rightarrow \lambda_i \in (0, 1)$ for $i = 1, \ldots, k$,

$$N^{\frac{1}{2}}[(S_1^u - \Sigma_1^u)', \ldots, (S_k^u - \Sigma_k^u)']' \xrightarrow{d} Z,$$

where Z is a multivariate normal random vector with mean vector 0 and covariance matrix

$\mathrm{Diag}[\lambda_1^{-1}(\mu_4(G_1) - \Sigma_1^4), \ldots, \lambda_k^{-1}(\mu_4(G_k) - \Sigma_k^4)]$.

Proof: By the central limit theorem we have that $n_i^{\frac{1}{2}}(S_i^u - \Sigma_i^u) \xrightarrow{d} MN\Big(0, \mu_4(G_i) - \Sigma_i^4\Big)$. Thus since $\frac{n_i}{N} \to \lambda_i$, we have by Slutsky's Theorem that

$$N^{\frac{1}{2}}(S_i^u - \Sigma_i^u) = N^{\frac{1}{2}} n_i^{-\frac{1}{2}} \cdot n_i^{\frac{1}{2}}(S_i^u - \Sigma_i^u) \xrightarrow{d} MN\Big(0 , \lambda_i^{-1}(\mu_4(G_i) - \Sigma_i^4)\Big) .$$

Since $S_1 , \dots , S_k$ are independent, the conclusion of joint convergence then follows.

Proposition 2. Let $Q(S_1 , S_2 , \dots , S_k) = \sum_{i=1}^{k} \lambda_i \log[\frac{|S|}{|S_i|}]$, where $S = \sum_{i=1}^{k} \lambda_i S_i$, $\lambda_i \in (0, 1)$ for $i = 1, \dots, k$, $\sum_{i=1}^{k} \lambda_i = 1$, and $S_1 , \dots , S_k$ are $p \times p$ symmetric positive definite matrices. Then Q , as function of $V = \Big((S_1^u)' , \dots , (S_k^u)'\Big)'$, satisfies properties Q1 – Q4.

Proof: See Zhang (1989).

Proof of *Theorem 1*: By Q1 we have $NQ(S_1 , \dots , S_k) = NQ(\Sigma^{-\frac{1}{2}} S_1 \Sigma^{-\frac{1}{2}} , \dots , \Sigma^{-\frac{1}{2}} S_k \Sigma^{-\frac{1}{2}})$. Then under (2.1) with Σ and μ_4 denoting the common values of Σ_i and $\mu_4(G_i)$, we have by Theorem 3.3B of Serfling (1980, pp. 124)

$$NQ(S_1 , \dots , S_k) \xrightarrow{d} Z'AZ ,$$

where $Z \stackrel{d}{=} MN\Big(0 , \mathrm{Diag}[\lambda_i^{-1}(\Sigma^{-\frac{1}{2}})_q \cdot (\mu_4 - \Sigma^4) \cdot \big((\Sigma^{-\frac{1}{2}})_q\big)' , i = 1, \dots, k]\Big)$.

Since $(\Sigma^{-\frac{1}{2}})_q \cdot \Sigma^4 \cdot \big((\Sigma^{-\frac{1}{2}})_q\big)' = \mathrm{uvec}\{\Sigma^{-\frac{1}{2}} \, \Sigma \, \Sigma^{-\frac{1}{2}}\}[\mathrm{uvec}\{\Sigma^{-\frac{1}{2}} \, \Sigma \, \Sigma^{-\frac{1}{2}}\}]' = I_p^4$ and

$(\Sigma^{-\frac{1}{2}})_q \cdot \mu_4 \cdot \big((\Sigma^{-\frac{1}{2}})_q\big)' = \beta_2$,

$$Z \stackrel{d}{=} MN\Big(0 , \mathrm{Diag}[\lambda_i^{-1} , \; i = 1, \dots, k] \otimes (\beta_2 - I_p^4)\Big) .$$

Proof of *Corollary 1:*

$$-2 \log \Lambda = \sum_{i=1}^{k} (n_i - 1) \log[\frac{|A|}{|A_i|}] + \sum_{i=1}^{k} p(n_i - 1) \log[\frac{n_i - 1}{N - k}]$$

$$= \sum_{i=1}^{k} (n_i - 1) \log[\frac{\left| \sum_{j=1}^{k} (\frac{n_j - 1}{N-k}) S_j \right|}{|S_i|}] .$$

Then using the Q function from Proposition 2 we have by Taylor expansion and the fact that

$S_i^u - \Sigma^u = O_P(N^{-\frac{1}{2}})$

$$-2 \log \Lambda - NQ(S_1 , \dots , S_k)$$

$$= \sum_{i=1}^{k} (n_i - 1) \log[\frac{\left| \sum_{j=1}^{k} (\frac{n_j - 1}{N-k}) S_j \right|}{|S_i|}] - \sum_{i=1}^{k} N\lambda_i \log[\frac{\left| \sum_{j=1}^{k} \lambda_j S_j \right|}{|S_i|}] \xrightarrow{P} 0 .$$

Therefore by Slutsky's Theorem and Theorem 1,

$$-2 \log \Lambda \xrightarrow{d} Z'AZ .$$

Proposition 3. Let $\left\{X_{i1} , \dots , X_{in_i} , \ i = 1, \dots, k\right\}$ be k samples of independent $p\times 1$ random vectors, where X_{ij} has mean vector μ_i , covariance matrix Σ_i , and distribution function $G_i(x)$ with finite fourth moment. Let G_N be defined as in (2.3), G be defined as in (2.4) , and suppose that $\hat{\mu}_i$ has the translation property with $\hat{\mu}_i \xrightarrow{a.s.} \mu_{i\infty}$ for $i = 1, \dots, k$. Then if $\min(n_1 , \dots , n_k) \to +\infty$ with $\frac{n_i}{N} \to \lambda_i \in (0, 1)$ for $i = 1, \dots, k$, we have that

i) $\mu(G_N) \to \mu(G)$ a.s. ;

ii) $\Sigma(G_N) \to \Sigma(G)$ a.s. and hence $\Sigma^4(G_N) \to \Sigma^4(G)$ a.s. ;

iii) $\mu_4(G_N) \to \mu_4(G)$ a.s.

Proof: The results follow by applying the strong law of large numbers.

Proposition 4. Let $\{X_{i1}, \dots, X_{in_i}, \; i = 1, \dots, k\}$ be k samples of independent $p \times 1$ random vectors, with distribution function $G_i(x)$ and finite fourth moment. Suppose S_i^*, $i = 1, \dots, k$, are k sample covariances based on the iid bootstrap samples $\{X_{i1}^*, \dots, X_{in_i}^*, \; i = 1, \dots, k\}$ drawn from G_N of (2.3). Then, as $\min(n_1, \dots, n_k) \to +\infty$ with $\frac{n_i}{N} \to \lambda_i \in (0, 1)$ and $\hat{\mu}_i \overset{a.s.}{\to} \mu_{i\infty}$ for $i = 1, \dots, k$, we have

$$N^{\frac{1}{2}}[\text{uvec}\{S_i^*\} - \Sigma^u(G)] \overset{d^*}{\to} MN\Big(0 \, , \, \lambda_i^{-1}[\mu_4(G) - \Sigma^4(G)]\Big) \text{ a.s. for } i = 1, \dots, k,$$

where $G(x)$ is defined as in (2.4).

Proof: The proof is similar to the proof of Lemma 1 of Boos and Brownie (1989).

Proofs of Theorem 2 and Corollary 2 are similar to the proofs of Theorem 1 and Corollary 1. Full details are contained in Zhang (1989).

REFERENCES

Boos , D.D. , and Brownie , C. (1989), "Bootstrap Methods for Testing Homogeneity of Variances," *Technometrics*, 31, 69−82.

Gnanadesikan , R. , and Kettenring , J.R. (1989), "Discriminant Analysis and Clustering," *Statistical Science*, 4, 34−69.

Manly , B.F.J. , and Rayner , J.C.W. (1987), "The Comparison of Sample Covariance Matrices Using Likelihood Ratio Tests," *Biometrika*, 74, 841−847.

Mardia , K.V. , Kent , J.T. , and Bibby , J.M. (1982), "*Multivariate Analysis*," New York : Academic Press, Inc.

Muirhead , R.J. (1982), "*Aspects of Multivariate Statistical Theory*," New York : John Wiley & Sons, Inc.

Serfling , R.J. (1980), "*Approximation Theorems of Mathematical Statistics*," New York : John Wiley & Sons, Inc.

Zhang , J. (1989), "*Bootstrap Methods for Tests about Covariance Matrices*," Ph.D Dissertation, Department of Statistics, North Carolina State University.