Songklanakarin J. Sci. Technol.
31 (3), 351-359, May - Jun. 2009

Songklanakarin Journal
of Science and Technology
http://rdo.psu.ac.th/sjst

Original Article

Powers of some one-sided multivariate tests with unknown population covariance matrix

Samruam Chongcharoen*

School of Applied Statistics,
National Institute of Development Administration,
Bang Kapi, Bangkok, 10240 Thailand.

Received 16 October 2008; Accepted 13 May 2009

Abstract

For a multivariate normal population, Kudo (1963), Shorack (1967) and Perlman (1969) derived the likelihood ratio tests of the null hypothesis that the mean vector is zero with a one-sided alternative for a known covariance matrix, a partially known covariance matrix and a completely unknown covariance matrix, respectively. Because these tests may be tedious to use, Tang, Gnecco and Geller (1989) developed approximate likelihood ratio tests and Follmann (1996) proposed one-sided modifications of the usual omnibus chi-squared test and Hotelling's T^2 test. Also, we consider a modification of Follmann's test (the new test) to include information of off diagonal of covariance matrix , which adjusts for possibly unequal variances. For the non-normal population, Boyett and Shuster (1977) proposed a nonparametric one-sided test and we use their technique to develop nonparametric versions of Perlman's test, Follmann's test, the new test and the Tang-Gnecco-Geller test. Following Chongcharoen, Singh and Wright (2002), who considered known and partially known covariance matrices, we study the powers of these one-sided tests for an unknown covariance matrix using Monte Carlo techniques and make recommendations concerning their use.

Keywords: Follmann's test,Kudo's test, Perlman's test, Simple order, Simple tree order, Tang-Gnecco-Geller test and Boyett and Shuster test

1. Introduction

Suppose one uses a matched-pair design to compare the multivariate responses of two treatments. If the responses are p dimensional and $\theta = (\theta_1, \theta_2, ..., \theta_p)'$ is the difference, treatment one minus treatment two, of the mean responses, then one may test the null hypothesis, $H_0: \theta_1 = \theta_2 =, ..., = \theta_p = 0$, to determine if there is a difference in the two treatments. Furthermore, if one believes that for each coordinate, the mean responses for treatment one are at least as large as those for treatment two, then the alternative can be constrained by $H_1: \theta_i \geq 0$ for i = 1, 2, …, p.

*Corresponding author.
Email address: samruam@as.nida.ac.th

Based on a random sample from the normal distribution with mean θ and covariance matrix *V*, Kudo (1963), Shorack (1967) and Perlman (1969) derived the likelihood ratio test of H_0 versus H_1 - H_0 for the cases in which *V* is known, known up to a multiplicative constant and completely unknown, respectively. Because the likelihood ratio tests with restricted alternatives are complicated to use, Tang, Gnecco and Geller (1989) proposed an approximate likelihood ratio test, and Follmann (1996) proposed one-sided modifications of the usual χ^2 and Hotelling's T^2 tests of H_0 versus ~ H_0 that are easier to implement. Using exact computations and Monte Carlo methods, Chongcharoen, Singh and Wright (1998) compared the performance of Kudo's test, Follmann's test, a new test, which is a modification of Follmann's test, the permutation test of Boyett and Shuster and the Tang-Gnecco-Geller test for a known covariance matrix, and for a partially known covariance matrix, they compared the powers of these tests with Kudo's test replaced by Shorack's test.

Because situations with a completely unknown covariance matrix occur frequently in practice, it is important to study that case, too. In his Ph.D. dissertation, Chongcharoen studied the power of these one-sided tests for unknown covariance matrices with equal variances. Those results are summarized here, and unequal variances are considered, as well as tests obtained by combining the Boyett-Shuster technique with Follmann's test, the new test, Perlman's test and the Tang-Gnecco-Geller test.

Throughout this paper, we suppose that $X_1, X_2, ..., X_n$ is a random sample from a p-dimensional multivariate normal distribution with unknown mean $\theta = (\theta_1, \theta_2, ..., \theta_p)'$ and unknown positive definite covariance matrix *V*. We consider testing the null hypothesis $H_0 : \theta = 0$ versus $H_1 - H_0$ where $H_1 : \theta \in \Omega_p$ and $\Omega_p = \{x : x_i \geq 0$ for i = 1,2, …, p$\}$ is the p-dimensional nonnegative orthant. The sample mean and covariance are

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n} \text{ and } S_x = \sum_{i=1}^{n} \frac{(X_i - \bar{X})(X_i - \bar{X})'}{n},$$

and it is well known that S_x is positive definite with probability one.

The hypotheses H_0 and H_1 also arise in the one-way analysis of variance when the means are known to satisfy an order restriction. For observations which come from k normal populations whose means are known to satisfy a simple ordering, i.e. H_S: $\mu_1 \leq \mu_2 \leq, ..., \leq \mu_k$, Bartholomew (1959a, 1959b, 1961) derived the likelihood ratio test of $\mu_1 = \mu_2 =, ..., = \mu_k$ with the alternative restricted by H_S for the cases of known variances and variances known up to a multiplicative constant. Suppose the observations are Y_{ij} for j =1, 2, …, n_i, and i =1, 2, …, k, and the sample means are $\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_k$. With known variances, $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$, Kudo (1963) noted that for p = k-1, $X_i = \bar{Y}_{i+1} - \bar{Y}_i$ for i = 1, 2, …, p, $X = (X_1, X_2, ..., X_p)'$ and $\theta = E(X)$, the hypotheses on μ are equivalent to H_0 and H_1 above, and Bartholomew's and Kudo's tests are equivalent. With $w_i = n_i / \sigma_i^2$ for i = 1, 2, …, k, the correlation matrix for X satisfies

$$\rho_{i,i+1} = -\sqrt{\frac{w_i w_{i+2}}{(w_i + w_{i+1})(w_{i+1} + w_{i+2})}} \text{ for } i = 1, 2, ..., p-1$$

and $\rho_{ij} = 0$ for $|i - j| \geq 2$. (1.1)

If the weights are equal, i.e. $w_1 = w_2 = ... = w_k$, the correlation matrix in (1.1) is denoted by R_S.

Also, Bartholomew considered an arbitrary partial order restriction, which includes the simple tree order, i.e. H_T: $\mu_1 \leq \mu_j$ for j = 2, 3, …, k. For this ordering, one takes differences, $X_i = \bar{Y}_{i+1} - \bar{Y}_1$ for i=1, 2, …, p, and with p = k-1 and w_i as above, the correlation matrix of $X = (X_1, X_2, ..., X_p)'$ satisfies

$$\rho_{i,j} = \sqrt{\frac{w_{i+1} w_{j+1}}{(w_{i+1} + w_1)(w_{j+1} + w_1)}} \text{ for } 1 \leq i \neq j \leq p. \quad (1.2)$$

If the weights are equal, i.e. $w_1 = w_2 = ... = w_k$, the correlation matrix in (1.2) is denoted by R_T. We compare the powers of the proposed tests for several correlation matrices including R_S and R_T.

2. Perlman's Test

Perlman (1969) showed that the likelihood ratio test (LRT) of H_0 versus $H_1 - H_0$ rejects for large value of

$$U = \frac{\bar{X}^{*\prime} S_x^{-1} \bar{X}^*}{1 + (\bar{X} - \bar{X}^*)' S_x^{-1} (\bar{X} - \bar{X}^*)}, \quad (2.1)$$

where $\bar{X}^*$ is the restricted maximum likelihood estimate of θ under H_1. In particular, $\bar{X}^*$ minimizes $(\bar{X} - \theta)' S_x^{-1} (\bar{X} - \theta)$ subject to $\theta \in \Omega_p$, and can be computed using a quadratic programming routine such as QPROG in IMSL. The null hypothesis distribution of U is given by the following: for any real number t,

$$P(U \geq t) = \sum_{j=0}^{p} Q(j, p; V)\, P(\chi_j^2 / \chi_{n-p}^2 \geq t), \quad (2.2)$$

where χ_q^2 is a chi-squared random variable with q degrees of freedom ($\chi_0^2 \equiv 0$) and χ_j^2 and χ_{n-p}^2 are independent. The weights Q(j, p; V), j = 0,1, …, p, are called **level probabilities**, are nonnegative, sum to one and can be computed using the FORTRAN programs by Bohrer and Chow (1978) and Sun (1988) for p<10. Perlman (1969) obtained the maximum of (2.2) over all positive definite *V*. However, using this maximum makes the test too conservative. Following Lei *et al.* (1995), we approximate Q(j, p; *V*) by using S_x in place of *V*.

3. Follmann's Test

For *V* completely unknown, Follmann (1996) used the unbiased sample covariance, $\hat{S} = n S_x / (n-1)$, in place of *V*. Then Follman's statistics are

$$F^s = n \bar{X}' \hat{S}^{-1} \bar{X} \text{ and } \sum_{j=1}^{p} \bar{X}_j. \quad (3.1)$$

With $F_{p,n-p}$ the F-distribution with degrees of freedom p and n-p, under H_0,

$$F^s = n \bar{X}' \hat{S}^{-1} \bar{X} \sim \frac{(n-1)p}{n-p} F_{p,n-p},$$

cf. Anderson (1984). Follmann's test rejects H_0 if

$$F^s > \frac{(n-1)p}{n-p} F_{2\alpha;p,n-p} \text{ and } \sum_{j=1}^{p} \bar{X}_j > 0,$$

where $F_{2\alpha;p,n-p}$ is the 1-2α^{th} quantile of the central F-distribution with p and n-p degrees of freedom. After his Theorem 2.1, Follmann (1996) noted that the significance level of this test is α.

4. The New Test

For V known, Chongcharoen, Singh and Wright (2002) proposed a new test, which is Follmann's test applied to $Z_i = V^{-1/2} X_i$. However, Chongcharoen and Wright (2002) recommend Follmann's test applied to

$W_i = BX_i$ where $B_{ii} = 1/\sqrt{V_{ii}}$ and $B_{ij} = 0$ for $1 \le i \ne j \le p$.

when V is completely unknown, consider a known positive definite p×p matrix A and $\eta = A\theta$. Follmann's test of $\eta = 0$ versus $\eta \ge 0$ is based on A $X_i \sim N(\eta, AVA')$. Now

$$\sum_{i=1}^{n} \frac{(AX_i - A\overline{X})(AX_i - A\overline{X})'}{n-1} = A\hat{S}A'.$$

Follmann's statistics based on $A\overline{X}$ are

$$n(A\overline{X})'(A\hat{S}A')^{-1}(A\overline{X}) = n\overline{X}'\hat{S}^{-1}\overline{X} \sim \frac{(n-1)p}{n-p} F_{p,n-p} \text{ and } \sum_{j=1}^{p}(A\overline{X})_j.$$

If one replaces A with $\hat{B}$ where $\hat{B}_{ii} = 1/\sqrt{\hat{S}_{ii}}$ and $\hat{B}_{ij} = 0$ for $1 \le i \ne j \le p$, then the new test in this setting is based on

$$G^S = n\,\overline{X}'\,\hat{S}^{-1}\,\overline{X} \text{ and } \sum_{j=1}^{p} (\hat{B}\overline{X})_j, \tag{4.1}$$

and H_0 is rejected with approximate significance level α if

$$G^S > \frac{(n-1)p}{n-p} F_{2\alpha;p,n-p} \text{ and } \sum_{j=1}^{p} (\hat{B}\overline{X})_j > 0.$$

The significance level is approximate because this choice of A depends on the data. We investigate the accuracy of this approximation by Monte Carlo techniques in section 11.

5. An Approximate Likelihood Ratio Test

Tang, Gnecco and Geller (1989) proposed an approximate likelihood ratio test for V known. With

$$Z = \sqrt{n}\; V^{-1/2}\,\overline{X} \sim N(\sqrt{n}V^{-1/2}\theta, I) \text{ and } T = \sum_{j=1}^{p}(Z_j \vee 0)^2,$$

where $Z_j \vee 0$ denotes the maximum of Z_j and 0, H_0 is rejected if T is too large. The null hypothesis distribution of this test statistic is given for any real number t by

$$P(T \ge t) = \sum_{j=0}^{p}(C_j^p / 2^p)\, P(\chi_j^2 \ge t),$$

where C_j^p denotes the number of combinations of p things taken j at a time. They gave the critical values of the test for p up to 10 in their paper.

When V is completely unknown, they suggested that one could simply replace V by $\hat{S}$ in the covariance known case. However, in the Monte Carlo study it was found that with their critical values, the significance levels of this test can be quite large. For instance, with p=3 and a target level of $\alpha = .05$, the estimated significance level for this test with $V = R_S$ is 0.242 for n=6 and 0.086 for n=20, and the corresponding values for $V = R_T$ are 0.259 for n=6 and 0.085 for n=20.

Thus we adjusted by letting

$$Z = \sqrt{n}\,\hat{S}^{-1/2}\,\overline{X} \quad \text{and}$$

$$T^* = \sum_{j=1}^{p} \max(Z_j, 0)^2. \tag{5.1}$$

Under the null hypothesis,

$$Z'Z = n\overline{X}'\,\hat{S}^{-1}\,\overline{X} \sim \frac{(n-1)p}{n-p} F_{p,n-p},$$

and combining this result with the null hypothesis distribution given in Tang *et al.* (1989), suggests the following approximation: under H_0 for any real number t,

$$P(T^* \ge t) \cong \sum_{j=0}^{p} (C_j^p / 2^p)\, P\left[F_{j,n-p} \ge t\frac{(n-p)}{j(n-1)}\right]. \tag{5.2}$$

The accuracy of this approximation is studied in subsection 11.

6. Nonparametric One-Sided Tests in Multivariate Analysis

Boyett and Shuster (1977) proposed a nonparametric multivariate one-sided test. For a matched pair design, X_i is the difference, treatment minus placebo, for i = 1, 2, …, n. We consider the null hypothesis that the responses for the treatment and placebo are interchangeable, which means that $\theta = 0$, where $\theta = E(X_i)$, provided this mean exists. If the treatment is believed to have mean responses at least as large as the placebo, one may want to test H_0 versus H_1 - H_0, where these hypotheses are defined as in section 1. Let

$$E_1 = \{(c_1X_1', \ldots, c_nX_n')' : c_i = 1 \text{ or } -1\}.$$

Under the null hypothesis that treatment and placebo are equivalent, conditionally on the set E_1, $X_d = (X_1', X_2', \cdots, X_n')'$ has a uniform distribution over E_1. Let

$$\overline{X} = (\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_p)' \text{ and } (S_j^*)^2 = \sum_{i=1}^{n} \frac{X_{ij}^2}{n-1} \quad j = 1, 2, \ldots, p,$$

then the t statistic corresponding to the jth component is given by

$$t_j = \frac{\sqrt{n}\,\overline{X}_j}{\sqrt{(S_j^*)^2 - \dfrac{n\overline{X}_j^2}{n-1}}} \tag{6.1}$$

and $t_{max}(X_d) = \max\{t_1, t_2, \ldots, t_p\}$. For $y \in E_1$, let $t_{max}(y)$ be calculated as above and

$$J(y) = \begin{cases} 1 \text{ if } t_{max}(y) \geq t_{max}(X_d) \\ 0 \text{ if } t_{max}(y) < t_{max}(X_d), \end{cases}$$

then the significance level for this randomized test is

$$\alpha = \sum_{y \in E_1} \frac{J(y)}{2^n}. \qquad (6.2)$$

7. Boyett–Shuster–Perlman Test

This test is a combination of the Boyett-Shuster technique and Perlman's test obtained by simply replacing the t statistic with Perlman's statistic. Let $X_1, X_2, \ldots, X_n$, X_d and E_1 be defined as in section 6. Thus, Perlman's statistic is

$$BS^{(P)}(X_d) = \frac{\bar{X}^{*\prime} S_X^{-1} \bar{X}^*}{1 + (\bar{X} - \bar{X}^*)' S_X^{-1} (\bar{X} - \bar{X}^*)} \qquad (7.1)$$

where $\bar{X}^*$ is the restricted maximum likelihood estimate of θ under H_1. For $y \in E_1$, let $BS^{(P)}(y)$ be calculated as $BS^{(P)}(X_d)$, and let

$$J_P(y) = \begin{cases} 1 \text{ if } BS^{(P)}(y) \geq BS^{(P)}(X_d) \\ 0 \text{ if } BS^{(P)}(y) < BS^{(P)}(X_d). \end{cases}$$

The significance level for this randomized test is

$$\alpha = \sum_{y \in E_1} \frac{J_P(y)}{2^n}. \qquad (7.2)$$

8. Boyett–Shuster–Follmann Test

With $X_1, X_2, \ldots, X_n$, X_d and E_1 be defined as in sections 6 and 7, Follmann's statistic may be written as

$$BS^{(F)}(X_d) = n\bar{X}'\hat{S}^{-1}\bar{X} I_{\left(\sum_{J=1}^{p} \bar{X}_j > 0\right)},$$

where I denotes the indicator function. For $y \in E_1$, let $BS^{(F)}(y)$ be calculated as $BS^{(F)}(X_d)$ above, and let

$$J_F(y) = \begin{cases} 1 \text{ if } BS^{(F)}(y) \geq BS^{(F)}(X_d) \\ 0 \text{ if } BS^{(F)}(y) < BS^{(F)}(X_d). \end{cases}$$

Therefore, the significance level for this randomized test is

$$\alpha = \sum_{y \in E_1} \frac{J_F(y)}{2^n}.$$

9. Boyett–Shuster–New Test

With $X_1, X_2, \ldots, X_n$, X_d and E_1 defined as in the last three sections, the statistic for the new test may be written

$$BS^{(N)}(X_d) = n\bar{X}'\hat{S}^{-1}\bar{X} I\left(\sum_{j=1}^{p} (\hat{B}\bar{X})_j > 0\right).$$

For $y \in E_1$, let $BS^{(N)}(y)$ be calculated as $BS^{(N)}(X_d)$ and

$$J_N(y) = \begin{cases} 1 \text{ if } BS^{(N)}(y) \geq BS^{(N)}(X_d) \\ 0 \text{ if } BS^{(N)}(y) < BS^{(N)}(X_d). \end{cases}$$

The significance level α can be computed as in the last three tests.

10. Boyett-Shuster-Tang-Gnecco-Geller Test

We use the T* statistic from the Tang-Gnecco-Geller test in place of the t statistic in the Boyett-Shuster test. Hence, the test statistic for this test is

$$BS^{(T)}(X_d) = \sum_{j=1}^{p} \max(Z_j, 0)^2$$

where $Z = \sqrt{n}\, \hat{S}^{-1/2} \bar{X}$. With $y \in E_1$ and

$$J_T(y) = \begin{cases} 1 \text{ if } BS^{(T)}(y) \geq BS^{(T)}(X_d) \\ 0 \text{ if } BS^{(T)}(y) < BS^{(T)}(X_d), \end{cases}$$

the significance level α can be computed as for the four tests above.

11. Power Comparisons

For p=3 and 6, the performances of these nine tests are studied by Monte Carlo techniques for multivariate normal distributions and R_S and R_T, that is for the simple order and the simple tree order correlations with equal weights and k=4 and 7 (p = k-1) as well as some other forms of correlation structures. Recall, R_S and R_T are given in (1.1) and (1.2), respectively. Also, we consider several variance patterns including increasing variances, decreasing variances, V-shaped variances and inverted V-shaped variances. We discuss the relationship between symmetry properties of the power functions and variance patterns below.

In addition to the correlation matrices in (1.1) and (1.2), which we denote by $R_{p,1}$ and $R_{p,2}$, we also consider the following correlation matrices $R = [\rho_{ij}]_{pxp}$:

$R_{3,3}$ ($R_{6,3}$) with ρ_{ij} = -0.4 (-0.1) for $1 \leq i \neq j \leq p$,

$R_{3,4}$ with $\rho_{12} = \rho_{23}$ = -0.4 and ρ_{13} = 0.4,

$R_{3,5}$ with $\rho_{12} = \rho_{23}$ = 0.4 and ρ_{13} = -0.4, and

$R_{6,4}$ with $\rho_{12}=\rho_{14}=\rho_{25}=\rho_{26}=\rho_{35}=\rho_{36}=\rho_{45}=\rho_{46}=-0.4$ and the other ρ_{ij}=0.4 for $i \neq j$.

Let a be a positive constant, [x] the largest integer less than and equal to x and σ_j the standard deviation of the jth component of X, i.e. $\sqrt{(V_{jj})}$ for j = 1, 2,…, p. The form for increasing variances is

$$\sigma_j = 1 + a(j - 1)$$

with a=p for p=3 and 6, the form for V-shaped variances is

$$\sigma_j = 1 + a\left(\left[\frac{p+1}{2}\right] - j\right) \text{for } j \le \frac{p+1}{2} \text{ and}$$

$$\sigma_j = 1 + a\left(\left[\frac{p+1}{2}\right] + j - p - 1\right) \text{for } j > \frac{p+1}{2}$$

with a=3 for p=3 and a=5 for p=6, and the form for inverted V-shaped variances is

$$\sigma_j = 1 + a(j-1) \text{ for } j \le \frac{p+1}{2} \text{ and}$$

$$\sigma_j = 1 + a(p-j) \text{ for } j > \frac{p+1}{2}$$

with a=4 for p=3 and a=5 for p=6. We obtain the co variances from the relation

$$V_{ij} = \sigma_i \sigma_j \rho_{ij} \text{ for } 1 \le i,j \le p.$$

The power functions of the tests considered here have the following symmetry property: the power at (μ, V) is the same as at $(P\mu, PVP')$ where P is a $p \times p$ permutation matrix. Thus, the variance patterns considered here provide information about other patterns. Even if one pattern is a permutation of another pattern, such as the V-shaped and inverted V-shaped patterns with p=6, which are described above, considering the second pattern may give information about correlation structures or mean vectors that were not considered with the first. However, to save some space, decreasing variance patterns are not discussed.

Also, all of the tests except Follmann's test and the Tang-Gnecco-Geller test are scale invariant in the following sense: if $A = diag(a_1, a_2, ..., a_p)$ with $a_i > 0$ for $1 \le i \le p$, then the power function at (μ, V) is the same as at $(A\mu, AVA')$. Thus, considering unequal variances, provides information about the power of these tests for equal variances at the new alternatives, $A\mu$, by taking $a_i = 1/\sigma_i$.

For fixed dimension p, we constrained the ratio of maximum variance to minimum variance to be the same for each pattern of variances considered. We consider mean vectors of the form, $\theta = c\upsilon$ with c a constant and υ a vector. We refer to the vector υ as a direction and choose c so that the usual F test has power equal to 0.70 provided $\upsilon \neq 0$. For example, for p=3 we consider directions (0,0,0), (0,0,1), (0,1,0), (1,0,0) (0,1,1), (1,0,1), (1,1,0) and (1,1,1). We used 10,000 iterations and recorded the proportion of rejections for these tests. All of these tests are conducted using the level of significance α = .05. The proportion of rejections for Perlman's test, Follmann's test, the new test, the Tang-Gnecco-Geller test, the Boyett Shuster test, the Boyett-Shuster-Perlman test, the Boyett-Shuster-Follmann test, the Boyett-Shuster-New test and the Boyett-Shuster-Tang-Gnecco-Geller test are denoted by

$$\hat{\pi}_P, \hat{\pi}_F, \hat{\pi}_N, \hat{\pi}_T, \hat{\pi}_{BS}, \hat{\pi}_{BSP}, \hat{\pi}_{BSF}, \hat{\pi}_{BSN},$$

and $\hat{\pi}_{BST}$, respectively

Some of 108 tables, which are the results from the power computations, are given here. For all nine tests, every V and n considered with both p=3 and p=6, the Monte Carlo power estimates at the null hypothesis are close to 0.05. The maximum difference is 0.005 and the Monte Carlo power estimates of the non-null powers of Hotelling's T^2 were between 0.6901 and 0.7183. Before considering the four variance patterns, we note the following:

- With n≥20, for both p=3 and 6, every variance pattern and every correlation matrix considered, the power of the Boyett-Shuster test was less than 0.7 (the power of Hotelling's T^2, which is an omnibus test) for some alternative considered here. There are some apparent exceptions. For instance, with inverted V-shaped variances, p=6, $R_{6,2}$, and n=20, the minimum power of the Boyett-Shuster test given in the table is 0.738. However, because this test is scale invariant, alternatives considered for the case of equal variances, p=6, $R_{6,2}$ and n=20 also provide alternatives for inverted V-shaped variances, and some of those alternatives had power as low as 0.662. The same is true of the other apparent exceptions. Thus, we do not consider the Boyett-Shuster test further for n≥20.
- For the case of equal variances, Follmann's test tends to have slightly larger powers than the new test. In particular, the differences, $\hat{\pi}_F - \hat{\pi}_N$, for p=3 and p=6, each correlation matrix and each sample size, range from -0.003 to 0.016, their median is 0 and their mean is 0.0014. Thus, if the variances are equal, then the loss in power due to using the new test in place of Follmann's test is not large. On the other hand, the same difference for all the cases considered here ranges from -0.313 to 0.117, has median -0.047 and mean -0.074. One might like to determine when Follmann's test is preferred over the new test. However, the largest difference in these powers, 0.117, occurs for p=6, $R_{6,2}$, increasing variances, n=20 and μ = c(0,0,0,0,0,1), but for the same p, V, n and μ = c(1,1,1,1,1, 1), the difference is -0.130. So, in this case, the choice between these two tests depends on the mean vector. Hence, if one is concerned that the variances may be unequal, we recommend the new test. We do not consider Follmann's test further.
- For every p, V and alternative considered with n≥ 20, Perlman's test and its Boyett-Shuster version, i.e. the Boyett-Shuster-Perlman test, have essentially the same powers. The same is true of Follmann's test, the new test, and the Tang-Gnecco-Geller test. In particular, for all such cases,

$$-0.02 \le \hat{\pi}_P - \hat{\pi}_{BSP}, \hat{\pi}_F - \hat{\pi}_{BSF}, \hat{\pi}_N - \hat{\pi}_{BSN}, \hat{\pi}_T - \hat{\pi}_{BST} \le 0.012$$

and these differences have median and mean about -0.001. While the Boyett-Shuster versions are more complicated to use, they should be considered if the normality assumption were in question. For p=3 and n=6, using the Boyett-Shuster version of these tests could result in a significant loss in power for some directions. The loss is more severe for

Perlman's test and the Tang-Gnecco-Geller test and ranges from 0.003 to 0.129 with median and mean about 0.065. For Follmann's test and the new test, the loss ranges from 0.003 to 0.058 with median and mean about 0.05. For p=6 and n=10, these losses are not as severe, possibly because n is larger. For p=6 and n=10, the typical loss if one uses the Boyett-Shuster-Perlman test in place of the Perlman test is about 0.02, and in the cases of Follmann's test, the new test and the Tang-Gnecco-Geller test, the typical loss is less than 0.01.

Equal variances: With p=3 and 6 and the various correlation matrices considered, Tables 1-3 give the power estimates for the nine tests in the case of equal variances. In comparing two tests, we are looking for one that performs well over the entire alternative region. So we consider the minimum of the powers over the alternatives considered. If

Table 1. Numerical comparisons for power of the tests for simple order with equal weights, p=3, n=6

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0)	0.000	0.053	0.051	0.049	0.050	0.050	0.046	0.050	0.047	0.047	0.047	0.048
(1,1,1)	0.747	**0.930**	0.890	0.890	0.890	0.905	0.577	0.817	0.838	0.838	0.837	0.850
(1,0,1)	1.182	0.900	0.889	0.888	0.889	0.902	0.775	0.801	0.836	0.836	0.836	0.846
(0,1,1)	1.008	0.904	0.890	**0.889**	**0.889**	**0.903**	0.662	0.808	**0.843**	0.842	0.842	0.847
(1,1,0)	1.008	0.900	**0.881**	**0.880**	0.880	0.895	0.667	0.803	0.830	0.829	0.829	0.843
(1,0,0)	1.930	0.856	0.880	0.869	0.878	0.886	0.894	0.792	0.830	0.819	0.828	0.830
(0,1,0)	1.671	0.872	0.888	**0.884**	0.884	0.898	0.819	0.791	0.836	0.832	0.832	0.840
(0,0,1)	1.930	0.864	**0.892**	0.881	0.890	0.896	0.898	0.805	0.836	0.826	0.834	0.840
Min.		0.856	0.880	0.869	0.878	0.886	0.577	0.791	0.830	0.819	0.828	0.830
Average		0.889	0.887	0.883	0.886	0.898	0.756	0.802	0.836	0.832	0.834	0.842

Table 2. Numerical comparisons for power of the tests for simple tree order with equal weights, p=3, n=6

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0)	0.000	0.048	0.050	0.050	0.050	0.049	0.048	0.045	0.045	0.045	0.450	0.046
(1,1,1)	1.930	0.826	0.880	0.880	0.880	0.895	**0.995**	0.723	0.830	0.830	0.830	0.840
(1,0,1)	1.671	0.797	0.886	0.873	0.885	0.849	0.938	0.732	0.838	0.825	0.837	0.784
(0,1,1)	1.671	0.800	0.887	0.873	0.886	0.850	0.935	0.732	0.836	0.822	0.835	0.789
(1,1,0)	1.671	0.797	0.884	0.868	0.883	0.848	0.940	0.728	0.829	0.814	0.828	0.783
(1,0,0)	1.930	0.763	0.858	0.784	0.847	0.835	0.898	0.726	0.814	0.741	0.803	0.772
(0,1,0)	1.930	0.768	0.867	0.792	0.856	0.842	0.901	0.725	0.815	0.742	0.804	0.771
(0,0,1)	1.930	0.774	0.868	0.793	0.855	0.844	0.904	0.735	0.814	0.741	0.802	0.786
Min.		0.763	0.858	0.784	0.847	0.835	0.898	0.723	0.814	0.741	0.802	0.771
Average		0.789	0.876	0.838	0.870	0.852	0.930	0.729	0.825	0.788	0.820	0.789

Table 3. Numerical comparisons for power of the tests for simple order with equal weights when p=6 and n=10

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0,0,0,0)	0.000	0.051	0.049	0.049	0.050	0.052	0.053	0.049	0.052	0.052	0.050	0.053
(1,1,1,1,1,1)	0.273	0.954	0.866	0.866	0.866	0.910	0.268	0.912	0.864	0.870	0.859	0.910
(0,1,1,1,1,1)	0.302	0.948	0.866	0.866	0.866	0.907	0.270	0.904	0.865	0.871	0.861	0.908
0,0,1,1,1,1)	0.364	0.936	0.865	0.865	0.865	0.905	0.308	0.892	0.862	0.869	0.860	0.901
(0,0,0,1,1,1)	0.485	0.916	0.865	0.864	0.865	0.903	0.398	0.879	0.865	0.872	0.863	0.899
(0,0,0,0,1,1)	0.750	0.890	0.873	0.866	0.873	0.902	0.622	0.864	0.872	0.868	0.869	0.898
(0,0,0,0,0,1)	1.560	0.849	0.869	0.842	0.869	0.893	0.959	0.840	0.868	0.846	0.864	0.885
Min.		0.849	0.865	0.842	0.865	0.893	0.268	0.840	0.862	0.846	0.859	0.885
Average		0.916	0.867	0.862	0.867	0.903	0.471	0.882	0.866	0.866	0.863	0.900

the minimum powers for two tests are close, then we consider their average powers. The tables give minimum and average powers. The results are summarized below.

- Only Perlman's test has power greater than Hotelling's T^2 for every p, *V*, n and every direction considered.
- For small n, i.e. n=6 if p=3 and n=10 if p=6, the Boyett-Shuster test is preferred for the simple tree correlation matrix ($R_{3,2}$ and $R_{6,2}$), and in fact this is true for all variance patterns. Thus, for small n, we recommend the Boyett-Shuster test if all of the correlations seem to be positive. For $R_{3,1}$, $R_{6,1}$,$R_{3,3}$ and $R_{6,3}$, the Tang-Gnecco-Geller test is preferred. Hence, for small n, we recommend the Tang-Gnecco-Geller test if none of the correlations seem to be positive. For $R_{3,4}$, $R_{3,5}$ and $R_{6,4}$, the new test is preferred. Therefore, for small n and a mixture of positive and negative correlations, we recommend the new test.
- For moderate n (n=20) and p=3, we recommend Perlman's test. Perlman test and the Tang-Gnecco-Geller test have similar powers for $R_{3,2}$ and $R_{3,4}$ and Perlman's test is preferred for $R_{3,1}$,$R_{3,3}$ and $R_{3,5}$. In fact, for $R_{3,5}$, the power of the Tang-Gnecco-Geller test is less than Hotelling's T^2 for some alternatives considered. For n=20 and p=6, Perlman test and the Tang-Gnecco-Geller test have similar powers for $R_{6,1}$, The Tang-Gnecco-Geller test is preferred for $R_{6,2}$ and $R_{6,3}$, and Perlman's test is preferred for $R_{6,4}$. As with p=3, for $R_{6,4}$, the power of the Tang-Gnecco-Geller test is less than Hotelling's T^2 for some alternatives considered. Thus, for p=6 and n=20, we recommend the Tang-Gnecco-Geller test if the non-zero correlations are of the same sign and Perlman's test if their signs are mixed. For p=6, n=30 and equal variances, the power estimates for these tests were obtained. For n=30, Perlman's test is preferred except for $R_{6,3}$, and in that case, the Tang-Gnecco-Geller test has slightly larger minimum power, but Perlman's test has slighly larger average power. Hence, for p=6 and n=30, we recommend Perlman's test.
- For large n, we recommend Perlman's test. It is the preferred test except for for $R_{6,3}$, and as with n=30, the Tang-Gnecco-Geller test has slightly larger minimum power, but Perlman's test has slighly larger average power.

Unequal variances: For p=3 and 6 and the various correlation matrices and variance patterns considered, Tables 4-6 give the power estimates for the nine tests. We now attempt to summarize these results for the cases with unequal variances.

- For small n (n=6 for p=3 and n=10 for p=6), as with equal variances, the Boyett Shuster test is preferred for the simple tree correlation matrices ($R_{3,2}$ and $R_{6,2}$). For the cases with no positive correlations ($R_{3,1}$, $R_{3.3}$, $R_{6,1}$ and $R_{6,3}$), the Tang-Gnecco-Geller test is preferred. Actually, for $R_{6,3}$ it appears from the tables for increasing variances, V-shaped variances and inverted V-shaped variances that the Boyett-Shuster test has larger minimum power than the Tang-Gnecco-Geller test. However, because the powers of the Boyett-Shuster test and the new test are scale invariant, the alternatives for the case of equal variances can be transformed to alternatives for these variance patterns. Thus, when these alternatives are included, the minimum powers for the Boyett-Shuster test and the new test are 0.798 and 0.864, respectively. So the new test would be preferred over the Boyett-Shuster test for this correlation matrix. Furthermore, for this correlation matrix and all four variance patterns, the Tang-Gnecco-Geller test has slightly better powers than the new test. For $R_{3,4}$, $R_{3,5}$ and $R_{6,4}$, we recommend the new test. For $R_{6,4}$, Perlman's test has similar powers, but it is more complicated to use. Also, for $R_{3,4}$, it appears in the tables for increasing, V-shaped and inverted V-shaped variances that the Boyett-Shuster test would be preferred. But, as above, when the alternatives derived from the case of equal variances are included, the minimum powers for the Boyett-Shuster test and the new test are 0.854 and 0.876. For that reason, we also have recommended the new test for $R_{3,4}$. Thus, as in the case of equal variances and small sample sizes, we recommend the Boyett-Shuster test if all the correlations are positive, the Tang-Gnecco-Geller test if there are no positive correlations and the new test if the are both positive and negative correlations.
- For moderate n (n=20) and p=3, we recommend Perlman's test. One could use the Tang-Gnecco-Geller test for $R_{3,1}$ and $R_{3,4}$, but it is desirable to be able to recommend one test for all covariance matrices. For p=6, Perlman's test is preferred for $R_{6,1}$, $R_{6,2}$ and $R_{6,4}$, and in fact, Perlman's test is the only one with minimum power greater than that of Hotelling's T^2 for $R_{6,4}$. For $R_{6,2}$, it appears from the table for inverted V-shaped variances that the Tang-Gnecco-Geller test is preferred. However, all of these tests are permutation invariant, and including the alternatives for V-shaped variances, the minimum powers for the Tang-Gnecco-Geller test and Perlman's test are 0.663 and 0.736, respectively. For $R_{6,3}$, the Tang-Gnecco-Geller test is preferred for all of the variance patterns considered. Thus, for moderate n and possibly unequal variances, we recommend Perlman's test except for larger p (p=6) and all correlations negative, and in the latter case, we recommend the Tang-Gnecco-Geller test. It should be noted that for equal variances, the Tang-Gnecco-Geller test performed at least as well as Perlman's test for $R_{6,1}$ and $R_{6,2}$, but for unequal variances, Perlman's test is preferred for these two correlation matrices. Also, based on the power estimates for p=6 with equal variances, we conjecture that Perlman's test could be used for all covariance matrices with n=30.
- As in the case of equal variances, for large n, we recommend Perlman's test. For $R_{6,3}$, the Tang-Gnecco-Geller test has powers like Perlman's test and could be used with no loss in power. For $R_{6,2}$ and inverted V-shaped variances, it appears that the Tang-Gnecco-Geller test would be preferred to Perlman's test. However, if one includes the alternatives derived from V-shaped variances, the minimum powers of the Tang-Gnecco-Geller test and Perlman's test are 0.627 and 0.735, respectively.

Table 4. Numerical comparisons for power of the tests for the simple order correlation and increasing covariance structure at a=2 with p=3 and n=6

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0)	0.000	0.053	0.054	0.050	0.050	0.050	0.046	0.050	0.051	0.047	0.047	0.048
(1,1,1)	1.417	**0.921**	0.882	0.885	0.886	0.899	0.733	0.820	0.831	0.834	0.834	0.842
(1,0,1)	1.782	0.892	0.870	0.882	0.883	0.894	0.860	0.807	0.818	0.829	0.830	0.837
(0,1,1)	3.666	**0.903**	0.892	0.890	**0.891**	**0.905**	0.670	0.809	0.839	0.837	0.838	0.848
(1,1,0)	1.529	0.898	0.855	0.881	0.881	0.896	0.774	0.810	0.803	0.828	0.828	0.836
(1,0,0)	1.930	0.856	0.771	0.876	0.878	0.886	0.894	0.792	0.726	0.826	0.828	0.833
(0,1,0)	5.014	0.872	0.887	0.882	0.884	0.897	0.819	0.791	0.835	0.830	0.832	0.843
(0,0,1)	9.649	0.864	0.892	0.870	0.890	0.892	**0.898**	0.805	0.836	0.815	0.834	0.836
Min.		0.856	0.771	0.870	0.878	0.886	0.670	0.791	0.726	0.815	0.828	0.833
Average		0.887	0.864	0.881	0.885	0.896	0.807	0.805	0.813	0.828	0.832	0.839

Table 5. Numerical comparisons for power of the tests for the simple tree order correlation and increasing covariance structure at a=2 with p=3 and n=6

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0)	0.000	0.048	0.050	0.050	0.050	0.049	0.048	0.045	0.046	0.046	0.045	0.048
(1,1,1)	2.227	0.831	0.871	0.883	0.883	0.895	0.969	0.754	0.821	0.832	0.832	0.838
(1,0,1)	2.027	0.797	0.797	0.876	0.870	0.873	0.925	0.744	0.756	0.830	0.825	0.821
(0,1,1)	5.909	0.800	0.883	0.735	0.882	0.779	0.938	0.733	0.834	0.689	0.832	0.694
(1,1,0)	2.047	0.795	0.796	0.877	0.874	0.878	0.934	0.736	0.752	0.828	0.825	0.820
(1,0,0)	1.930	0.763	0.649	0.869	0.847	0.872	0.898	0.726	0.616	0.825	0.803	0.819
(0,1,0)	5.789	0.768	0.862	0.701	0.856	0.824	0.901	0.725	0.810	0.653	0.804	0.747
(0,0,1)	9.649	0.774	0.891	0.550	0.855	0.767	0.904	0.735	0.835	0.509	0.802	0.681
Min.		0.763	0.649	0.550	0.847	0.767	0.898	0.725	0.616	0.509	0.802	0.681
Average		0.790	0.821	0.784	0.867	0.841	0.924	0.736	0.775	0.738	0.818	0.774

Table 6. Numerical comparisons for power of the tests for the simple order correlation and increasing covariance structure at a=2 with p=6 and n=10

Direct	c	Perl	Foll	ONew	New	Tang	BS	BSP	BSF	OBSN	BSN	BST
(0,0,0,0,0,0)	0.000	0.051	0.049	0.051	0.050	0.053	0.053	0.049	0.050	0.051	0.050	0.050
(1,1,1,1,1,1)	0.905	**0.939**	0.853	0.870	0.870	0.907	0.588	0.905	0.849	0.866	0.866	0.901
(0,1,1,1,1,1)	1.643	0.945	0.870	0.870	0.870	0.909	0.315	0.902	0.867	0.867	0.867	0.904
(0,0,1,1,1,1)	2.550	0.934	0.862	0.861	0.862	0.903	0.317	0.892	0.858	0.857	0.857	0.896
(0,0,0,1,1,1)	4.082	0.914	0.865	0.864	0.864	0.899	0.393	0.876	0.859	0.856	0.859	0.89
(0,0,0,0,1,1)	7.301	0.889	0.871	0.859	0.870	0.894	0.607	0.863	0.868	0.856	0.868	0.884
(0,0,0,0,0,1)	17.170	0.849	0.869	0.833	0.869	0.881	0.959	0.840	0.864	0.828	0.864	0.869
Min.		0.849	0.853	0.833	0.862	0.881	0.315	0.840	0.849	0.828	0.857	0.869
Average		0.912	0.865	0.860	0.868	0.899	0.530	0.880	0.861	0.855	0.864	0.891

In summary, Perlman's test has the best overall powers of the nine tests. If n is large (n =100), we recommend Perlman's test. If n is moderate (n=20), we recommend Perlman's test except for p=6 with all non-zero correlations of the same sign. For moderate n, p=6 and all correlations negative, we recommend the Tang-Gnecco-Geller test. For moderate n, p=6 and all other covariance matrices, we recommend the Tang-Gnecco-Geller test if the variances are nearly equal and Perlman's test if the variances are not equal. If n is small, we recommend the Boyett-Shuster test when *V* has all correlations positive, the Tang-Gnecco-Geller test when V has no positive correlations and the new test when V has negative and positive correlations.

Acknowledgment

Professor Dr., F.T. Wright, Department of Statistics, University of Missouri-Columbia, for his great suggestions and writing corrections in this paper. This paper is supported by the Thailand Research Fund.

References

Bartholomew, D.J. 1959a. A test of homogeneity for ordered alternatives. Biometrika, 46, 46-48.

Bartholomew, D.J. 1959a. A test of homogeneity for ordered alternatives II. Biometrika, 46, 328-335.

Bartholomew, D.J. 1961. A test of homogeneity of means under restricted alternatives (with discussion). Journal of the Royal Statistical Society, Series B, 23, 239-281.

Bohrer, R. and Chow, W. 1978. Weights for one-sided multivariate inferences. Applied Statistics, 24, 380-384.

Boyett, J.M. and Shuster, J.J. 1977. Nonparametric one-sided tests in multivariate analysis with medical applications. Journal of the American Statistics Association, 72, 665-668.

Chongcharoen, S. 1998. One-sided Multivariate Tests. Ph.D. dissertation, University of Missouri-Columbia.

Chongcharoen, S., Singh, B. and Wright, F.T. 2002.Powers of some one-sided multivariate tests with the population covariance matrix known up to multiplicative constant. Journal of Statistical Planning and Inference, 107, 103-121.

Conaway, M., Pillers, C., Robertson, T., and Sconing, J. 1991. A circular cone test for testing homogeneity against a simple tree order. Canadian Journal of Statistics, 19, 283-296.

Follmann, D. 1996. A simple multivariate test for one-sided alternatives. Journal of the American Statistics Association, 91, 854-861.

Johnson, R.A. and D.W. Wichern. 1992. Applied Multivariate Statistical Analysis, 3rd Edition. Simon & Schuster.

Kudo, A. 1963. A multivariate analogue of the one-sided test. Biometrika, 50, 403-418.

Rao, C.R. 1973. Linear Statistical Inferences and Its Applications, Second Edition. John Wiley & Sons, New York.

Robertson, T., F.T. Wright, and R.L. Dykstra. 1988. Order Restricted Statistical Inference. John Wiley & Sons, New York.

Sen, P.K. 1984. Sub hypotheses testing against restricted alternatives for the Cox regression model. Journal of Statistical Planning Inference, 10, 31-42.

Shorack, G.R. 1967. Testing against ordered alternatives in model I analysis of variance: Normal theory and nonparametric. Annals of Mathematical Statistics, 38, 1740-1753.

Sun, H.J. 1988. A FORTRAN subroutine for computing normal orthant probability. Communications in Statistics-Simulation and Computation, 17, 1097-1111.

Tang, D., C. Gnecco and N.L. Geller. 1989. An approximate likelihood ratio test for a Normal mean vector with nonnegative components with application to clinical trials. Biometrika, 76, 577-83.

Tang, D. and Lin, S.P. 1997. An approximate likelihood ratio test for comparing several treatments to a control. Journal of the American Statistics Association, 92, 1155-1162.