Songklanakarin J. Sci. Technol.
32 (3), 321-326, May - Jun. 2010

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

Original Article

Scale invariant for one-sided multivariate likelihood ratio tests

Samruam Chongcharoen*

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

Received 12 May 2009; Accepted 25 June 2009

Abstract

Suppose $X_1, X_2, ..., X_n$ is a random sample from $N_p(\theta, V)$ distribution. Consider $H_0 : \theta_1 = \theta_2 = ... = \theta_p = 0$ and $H_1 : \theta_i \geq 0$ for $i = 1, 2, ..., p$, let $H_1 - H_0$ denote the hypothesis that H_1 holds but H_0 does not, and let ~ H_0 denote the hypothesis that H_0 does not hold. Because the likelihood ratio test (LRT) of H_0 versus $H_1 - H_0$ is complicated, several ad hoc tests have been proposed. Tang, Gnecco and Geller (1989) proposed an approximate LRT, Follmann (1996) suggested rejecting H_0 if the usual test of H_0 versus ~ H_0 rejects H_0 with significance level 2α and a weighted sum of the sample means is positive, and Chongcharoen, Singh and Wright (2002) modified Follmann's test to include information about the correlation structure in the sum of the sample means. Chongcharoen and Wright (2007, 2006) give versions of the Tang-Gnecco-Geller tests and Follmann-type tests, respectively, with invariance properties. With LRT's scale invariant desired property, we investigate its powers by using Monte Carlo techniques and compare them with the tests which we recommend in Chongcharoen and Wright (2007, 2006).

Keywords: Follmann's test; likelihood ratio tests; modified Follmann's test; Tang-Gnecco-Geller 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 $i = 1, 2, ..., p$.

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 (LRT) 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. Tang *et al.* (1989) proposed an approximate likelihood ratio test, and Follmann (1996) proposed one-sided modifications of Hotelling's T^2 tests of H_0 versus ~ H_0 that are easier to implement. Using exact computations and Monte Carlo methods, Chongcharoen *et al.* (2002) 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. For a partially known covariance matrix, they compared the powers of these tests with Kudo's test replaced by Shorack's test.

Chongcharoen and Wright (2007; 2006) studied versions of the Tang-Gnecco-Geller test, Follmann's test and the modified Follmann's test that are permutation and scale invariant. Because the Boyett-Shuster test does not require the assumption of normality and are quite complicated, we do not consider further.

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

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, sample covariance and the unbiased sample covariance are

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

It is well known that S_x and $\hat{S}$ are 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 θ = 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 $w_i = n_i / \sigma_i^2$ 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.

It is clear that for most matched-pair designs, one wants the test to be invariant under changes in the units of measurement for any or all of the response variables as well as changes in the order of the response variables. The likelihood function and the constraint region, H_1, are invariant under permutations of the indices of the response variables. For under scale changes for the response variables, the LRTs test statistic, Kudo's test statistic, Shorack's test statistic and Perlman's test statistic, are shown below:

When covariance matrix V is known, Kudo's test statistic for testing H_0 versus $H_1 - H_0$ rejects H_0 for large value of

$$\bar{\chi}_{01}^2 = n\bar{X}^* V^{-1} \bar{X}^* \quad (1.3)$$

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 $\bar{\chi}_{01}^2$ is given by Robertson *et al.* (1988, pp.219-220) Theorem 4.6.1, i.e. for any real t,

$$P(\bar{\chi}_{01}^2 \geq t) = \sum_{j=0}^{p} Q(j, p; V) P(\chi_j^2 \geq t) \quad (1.4)$$

where the weights $Q(j, p; V), j = 0, 1, 2, ..., p$, are nonnegative and sum to one and χ_j^2 is a chi-squared variable with j degrees of freedom ($\chi_0^2 \equiv 0$). The weights $Q(j, p; V)$, $j = 0, 1, 2, ..., p$, are called level probabilities and can be computed using the FORTRAN programs by Bohrer and Chow (1978) and Sun (1988) for $p < 10$.

For considering scale invariant property of Kudo's test statistic, we let $M_0 = diag(1/\sqrt{V_{i,i}})$, then the transformation $Y = M_0 X$ with $R_0 = M_0 V M_0$ has Kudo's test statistic as $n\bar{Y}^* R_0^{-1} \bar{Y}^* = n\bar{X}^* V^{-1} \bar{X}^* = \bar{\chi}_{01}^2$ where $\bar{Y}^*$ can be computed as $\bar{X}^*$.

The permutation and scale invariant statistic of the Tang-Gnecco-Geller test which was recommended in Chongcharoen and Wright (2007), it is G_{0S} that given $a \vee b$ denoting the maximum of a and b and let

$$Z = \sqrt{n} R_0^{-\frac{1}{2}} M_0 \bar{X} \quad \text{and} \quad g(z) = \sum_{i=1}^{p} (z_i \vee 0)^2.$$

H_0 is rejected for the large values of g(z) with the null distribution for any real number t

$$P(g(z) \geq t) = \sum_{i=1}^{p} (C_i^p / 2^p) P(\chi_i^2 \geq t)$$

where C_i^p is the number of combinations of p things taken i at a time, χ_i^2 a chi-square variable with i degrees of freedom and $\chi_0^2 = 0$.

In Chongcharoen and Wright (2006), the recommended permutation and scale invariant statistic of the Follman-type test is N_{0T} that reject null hypothesis if

$\chi^2 = n\bar{X}'V^{-1}\bar{X} > \chi^2_{2\alpha,p}$ and $U_{0T} > 0$ with $U_{0T} = \sum_{i=1}^{p}\sqrt{(V^{-1})_{i,i}}\bar{X}_i$

where $\chi^2_{2\alpha,p}$ is the $(1-2\alpha)^{th}$ quantile of the chi-square distribution with p degrees of freedom.

When covariance matrix $V = \sigma^2 V_0$ is partially known, V_0 known, Shorack's test statistic for testing H_0 versus $H_1 - H_0$ rejects H_0 for large value of

$$\bar{E}^2 = \frac{n\underline{\bar{X}}^{*\prime}V_0^{-1}\underline{\bar{X}}^*}{\sum_{i=1}^{n}\underline{X}_i'V_0^{-1}\underline{X}_i} \tag{1.5}$$

where $\underline{\bar{X}}^*$ solves

$$\min_{\underline{\bar{X}}^*\in\Omega_p}(\underline{\bar{X}} - \underline{\bar{X}}^*)'V_0^{-1}(\underline{\bar{X}} - \underline{\bar{X}}^*) \tag{1.6}$$

The null hypothesis distribution of $\bar{E}^2$ is given by Robertson *et al.* (1988, pp.221) Theorem 4.6.2, i.e. for any real t,

$$P(\bar{E}^2 \ge t) = \sum_{j=0}^{p} Q(j,p;V_0)P(B_{j/2,(np-j)/2} \ge t) \tag{1.7}$$

where the weights $Q(j,p;V_0), j = 0,1,2,....,p$, level probability, are the same as the covariance known case and $B(a,b)$ is a random variable having a beta distribution with parameter a and b $(B(0,b) \equiv 0)$.

For considering scale invariant property of Shorack's test statistic, we let $M_1 = diag(1/\sqrt{(V_0)_{i,i}})$, then the transformation $Y = M_1 X$ with $R_1 = M_1 V_0 M_1$ has Shorack's test statistic as

$$\frac{n\underline{\bar{Y}}^{*\prime}R_1^{-1}\underline{\bar{Y}}^*}{\sum_{i=1}^{n}\underline{Y}_i'R_1^{-1}\underline{Y}_i} = \frac{n\underline{\bar{X}}^{*\prime}V_0^{-1}\underline{\bar{X}}^*}{\sum_{i=1}^{n}\underline{X}_i'V_0^{-1}\underline{X}_i} = \bar{E}^2 \tag{1.8}$$

where $\bar{Y}^*$ can be computed as $\bar{X}^*$.

In Chongcharoen and Wright (2007), the recommended permutation and scale invariant statistic of the Tang-Gnecco-Geller test is G_{1S} that H_0 is rejected for large value of

$G_{1S} = \frac{g(z)}{\sum_{j=1}^{n} X_j'V_0^{-1}X_j}$ with $Z = \sqrt{n}R_1^{-\frac{1}{2}}M_1\bar{X}$ and

$$g(z) = \sum_{i=1}^{p}(z_i \vee 0)^2 .$$

The null distribution of G_{1S} is given by the following: for any real number t,

$$P(G_{1S} \ge t) = \sum_{i=1}^{p}(C_i^p/2^p)P(B_{i/2,(np-i)/2} \ge t)$$

where C_i^p is the number of combinations of p things taken i at a time; $B_{a,b}$ a beta random variable and $B_{0,b} \equiv 0$.

In Chongcharoen and Wright (2006), the recommended permutation and scale invariant statistic of the Follman-type test is N_{1T} that reject null hypothesis if

$$F = \frac{n(n-1)\bar{X}'V_0^{-1}\bar{X}}{\sum_{j=1}^{n}(X_j - \bar{X})'V_0^{-1}(X_j - \bar{X})} > F_{2\alpha;p,(n-1)p}$$

and $U_{1T} > 0$ with $U_{1T} = \sum_{i=1}^{p}\sqrt{(V_0^{-1})_{i,i}}\bar{X}_i$

where $F_{2\alpha;p,(n-1)p}$ is the $(1-2\alpha)^{th}$ quantile of the F distribution with p and $(n-1)p$ degrees of freedom.

When covariance matrix V is completely unknown, Perlman's test statistic for testing H_0 versus $H_1 - H_0$ rejects H_0 for large value of

$$U = \frac{\underline{\bar{X}}^{*\prime}S_x^{-1}\underline{\bar{X}}^*}{1 + (\underline{\bar{X}} - \underline{\bar{X}}^*)'S_x^{-1}(\underline{\bar{X}} - \underline{\bar{X}}^*)}, \tag{1.9}$$

where $\underline{\bar{X}}^*$ is the restricted maximum likelihood estimate of θ under H_1. In particular, $\underline{\bar{X}}^*$ 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,

$$P(U \ge t) = \sum_{j=0}^{p} Q(j,p;V)\, P(\chi_j^2/\chi_{n-p}^2 \ge t), \tag{1.10}$$

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 $Q(j,p;V), j = 0,1,2,....,p$, level probabilities, are the same as the covariance known case. Perlman (1969) obtained the maximum of (1.10) 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.

Again for considering scale invariant property of Perlman's test statistic, we let $M_2 = diag(1/\sqrt{\hat{S}_{i,i}})$, then the transformation $Y = M_2 X$ with $R_2 = M_2 \hat{S} M_2$, $S_y = M_2 S_x M_2$ has Perlman's test statistic as:

$$\frac{\underline{\bar{Y}}^{*\prime} S_y^{-1} \underline{\bar{Y}}^*}{1+(\underline{\bar{Y}}-\underline{\bar{Y}}^*)' S_y^{-1} (\underline{\bar{Y}}-\underline{\bar{Y}}^*)} = \frac{\underline{\bar{X}}^{*\prime} S_x^{-1} \underline{\bar{X}}^*}{1+(\underline{\bar{X}}-\underline{\bar{X}}^*)' S_x^{-1} (\underline{\bar{X}}-\underline{\bar{X}}^*)} = U \tag{1.11}$$

where $\bar{Y}^*$ can be computed as $\bar{X}^*$.

In Chongcharoen and Wright (2007), the recommended permutation and scale invariant statistic of the Tang-Gnecco-Geller test is G_{2S} that H_0 is rejected for large value of $G_{2S} = g(Z)$ with $Z = \sqrt{n} R_2^{-\frac{1}{2}} M_2 \bar{X}$ and $g(z) = \sum_{i=1}^{p} (z_i \vee 0)^2$.

The null distribution of G_{2S} is given by the following: for any real number t

$$P(G_{2S} \geq t) = \sum_{i=1}^{p} (C_i^p / 2^p) P(F_{j,(n-p)} \geq t\frac{(n-p)}{(j(n-1))})$$

where C_i^p is the number of combinations of p things taken i at a time, $F_{j,n-p}$ the random variable of the F distribution with j and $n-p$ degrees of freedom.

In Chongcharoen and Wright (2006), the recommended permutation and scale invariant statistic of the Follman-type test is N_{2T} that reject null hypothesis if

$$F_2 = n\bar{X}'\hat{S}^{-1}\bar{X} > \frac{(n-1)p}{n-p} F_{2\alpha;p,n-p} \text{ and}$$

$$U_{2T} > 0 \text{ with } U_{2T} = \sum_{i=1}^{p} \sqrt{(\hat{S}^{-1})_{i,i}}\, \bar{X}_i .$$

where $F_{2\alpha;p,n-p}$ the $(1-2\alpha)^{th}$ quantile of the F distribution with p and $n-p$ degrees of freedom.

2. Power Comparisons

For $p = 3$ and 6, we compare the performances of the LRT with the tests which we recommend in Chongcharoen and Wright (2007, 2006) 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 that R_S and R_T are given in (1.1) and (1.2), respectively, which we denote by $R_{p,1}$ and $R_{p,2}$, we also consider the following correlation matrices $R = \left(\rho_{ij}\right)_{pxp}$:

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

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

$R_{3,5}$ with $\rho_{12} = \rho_{23} = 0.4$ and $\rho_{13} = -0.4$, 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 $\rho_{ij} = 0.4$ for $i \neq j$

Sample sizes considered here are $n = 6$, 20, and 100 for $p = 3$ and $n = 10$, 20, and 100, for $p = 6$. We consider mean vectors of the form, $\theta = c\upsilon$ with c a constant and υ a vector.

The vector υ is called the direction and we choose c so that the usual F-test has power equal to 0.70 provided υ is non-null, i.e., $\upsilon \neq 0$. We consider directions of the form $(\upsilon_1, \ldots, \upsilon_p)'$ with $\upsilon_i = 0$ or 1 for $1 \leq i \leq p$. With 10,000 iterations, the proportion of times each test rejects the null hypothesis is recorded. Throughout, the level of significance is $\alpha = 0.05$. Let

N_{0T} = the new permutation and scale invariant test for known variance

G_{0S} = the Tang-Gnecco-Geller permutation and scale invariant test for known variance

N_{1T} = the new permutation and scale invariant test for partially known variance

G_{1S} = the Tang-Gnecco-Geller permutation and scale invariant test for partially known variance

N_{2T} = the new permutation and scale invariant test for unknown variance

G_{2S} = the Tang-Gnecco-Geller permutation and scale invariant test for unknown variance

We consider three tests here, i.e. version of Tang-Gnecco-Geller test and the new test which we recommend on Chongcharoen and Wright (2007, 2006) and a version of LRT. For all of these tests, all n, and all the correlation structures considered, the power estimates under the null hypothesis range from 0.045 to 0.053. The Monte Carlo power estimates of the non-null powers of Hotelling's T^2 were between 0.689 and 0.716. For a given correlation matrix, we estimate power of three tests over the 2^p-1 non-null directions which 4 of 63 estimates power tables given Table 1-4.

For variance V known, both $p = 3$ and 6, the LRT (Kudo's test) is the best overall tests for every correlation matrix considered, the minimum power and averages power ranges 0.750 (0.733) to 0.858 (0.880) and 0.770 (0.761) to 0.873 (0.917) respectively for $p = 3$ (6). G_{0S} is the second best test for this case but it has minimum power less than 0.7 for $R_{3,5}, R_{6,2}(R_T)$ and $R_{6,4}$. N_{0T} has minimum power less than 0.7 for $R_{3,2}(R_T), R_{3,5}$ and it has very bad power for $R_{6,4}$. So we recommend Kudo's test over other two tests for this variance case.

For V known up to a multiplicative constant, the LRT (Shorack's test) has highest powers of overall tests for every correlation matrix in both $p = 3$ and 6 and every n considered. The minimum power and average power ranges 0.753 (0.736) to 0.898 (0.896) and 0.771 (0.762) to 0.911 (0.930)

Table 1. Estimates of the power of the tests when V unknown for simple order correlation with equal weights, R_S, $p = 3$, $n = 6$.

Direction	c	Perl	N_{2T}	$G\text{-}_{2S}$
(0,0,0)	0.000	0.053	0.050	0.050
(1,1,1)	0.747	0.930	0.890	0.905
(1,0,1)	1.182	0.900	0.889	0.902
(0,1,1)	1.008	0.904	0.890	0.902
(1,1,0)	1.008	0.900	0.880	0.895
(1,0,0)	1.930	0.856	0.877	0.887
(0,1,0)	1.671	0.872	0.884	0.896
(0,0,1)	1.930	0.864	0.888	0.894
Min.		0.856	0.877	0.887
Average		0.889	0.885	0.897

Table 2. Estimate of the power of the tests when V unknown for simple tree order correlation with equal weights, R_T, $p = 3$, $n = 6$.

Direction	c	Perl	N_{2T}	$G\text{-}_{2S}$
(0,0,0)	0.000	0.048	0.050	0.049
(1,1,1)	1.929	0.826	0.880	0.895
(1,0,1)	1.671	0.797	0.885	0.857
(0,1,1)	1.671	0.799	0.886	0.859
(1,1,0)	1.671	0.797	0.882	0.856
(1,0,0)	1.929	0.763	0.843	0.845
(0,1,0)	1.929	0.768	0.850	0.854
(0,0,1)	1.929	0.774	0.854	0.854
Min.		0.763	0.843	0.845
Average		0.789	0.869	0.860

Table 3. Estimates of the power of the tests when V unknown for simple order correlation with equal weights, R_S, $p = 6$, $n = 10$. (6 of 63 directions shown)

Direction	c	Perl	N_{2T}	$G\text{-}_{2S}$
(0,0,0,0,0,0)	0.000	0.051	0.050	0.052
(1,1,1,1,1,1)	0.273	0.954	0.866	0.910
(0,1,1,1,1,1)	0.302	0.948	0.866	0.907
(0,0,1,1,1,1)	0.364	0.936	0.865	0.905
(0,0,0,1,1,1)	0.485	0.916	0.864	0.902
(0,0,0,0,1,1)	0.749	0.890	0.871	0.902
(0,0,0,0,0,1)	1.560	0.849	0.864	0.893
Min.		0.849	0.859	0.893
Average		0.919	0.867	0.906

Table 4: Estimates of the power of the tests when V unknown for simple tree order correlation with equal weights, R_T, $p = 6$, $n = 10$. (6 of 63 directions shown)

Direction	c	Perl	N_{2T}	$G\text{-}_{2S}$
(0,0,0,0,0,0)	0.000	0.051	0.051	0.049
(1,1,1,1,1,1)	1.560	0.809	0.874	0.912
(0,1,1,1,1,1)	1.209	0.788	0.873	0.850
(0,0,1,1,1,1)	1.103	0.768	0.864	0.821
(0,0,0,1,1,1)	1.103	0.752	0.859	0.814
(0,0,0,0,1,1)	1.209	0.733	0.815	0.813
(0,0,0,0,0,1)	1.560	0.724	0.744	0.835
Min.		0.722	0.741	0.809
Average		0.754	0.843	0.823

respectively for $p = 3$ (6). For $p = 3$, all n considered, G_{1S} has minimum power less than that of usual F-test for correlation matrices with the number of positive elements larger than those of negative elements only for direction (0,1,0). It has overall powers second best to LRT. N_{1T} also has bad powers for R_T when $n = 20,100$ and for $R_{3,5}$ when $n = 100$. Both cases are on direction $(\upsilon_1, \upsilon_2, \upsilon_3)'$ with only one $\upsilon_i = 1; i = 1,2,3$. For $p = 6$ with all possible 63 directions of all n and every correlation matrix considered, both G_{1S} and N_{1T} have very bad powers on R_T and $R_{6,4}$. It can be noted that these two tests are affected by correlation matrices with positive elements. Clearly, for this partially known variance case, the LRT gives the best powers for every direction considered, so we recommend it for all p .

For variance V completely unknown, with $p = 3$ and small n all three tests, LRT (Perlman's test), N_{2T} and G_{2S} have essentially high power for every correlation matrix considered. That is, Perlman's test , N_{2T} and G_{2S} have minimum power ranges 0.763, 0.843, 0.813 to 0.881, 0.882, 0.895 and have averages powers ranges 0.789, 0.869, 0.860 to 0.908, 0.888, 0.901 respectively. Since they do not have essentially difference in power, so we recommend all these three tests for any correlation cases here. When sample size n increases to moderate or large, the power of G_{2S} at direction (0,1,0) has power smaller than that of Hotelling's $T\text{-}^2$ and this also happens for N_{2T} only for large n . So for some protection when the sample size is medium or large , we recommend Perlman's test over other two tests. For $p = 6$ and small n , all three tests considered give high power for almost every correlation matrix except for $R_{6,4}$, G_{2S} has smaller power than that of usual F-test only at direction (0,0,1,0,0,0) and N_{2T} also has smaller power than that of usual F-test in several directions. For correlation matrices R_S, R_T and $R_{6,3}$ on average of minimum powers and average powers G_{2S} will be ranked as the best test with N_{2T} as the second best test and the third is Perlman's test. When sample size n increases to moderate or large, both N_{2T} and G_{2S} have some power smaller than 0.7 for R_T and $R_{6,4}$, so we recommend Perlman's test over these two tests for these cases.

In summary, when V known and partially known LRT (Kudo's test and Shorack's test) have the best overall powers over the other two tests for all p and n considered, we recommend Kudo's test for V known and Shorack's test for V partially known. For V unknown , $p = 3$ and small

n, we recommend all three tests, i.e. Perlman's tests, N_{2T}, and G_{2S} because they give essentially equal high powers. When n is moderate ($n = 20$) or large ($n = 100$), we recommend Perlman's test. For $p = 6$ and small n we recommend G_{2S} except for V with no negative and positive correlations. If n is moderate or large, we recommend Perlman's test.

3. Two examples for illustration

In this section, we apply the proposed test and the N_{2T} and G_{2S} tests to the two examples of the data from the matched-pair design and the one-way analysis of variance, (Rencher A.C. 2002, p135 and p227). The data sets are described next.

1) The matched-pair data: Two types of coating for resistance to corrosion, 15 pieces of pipe were coated with each type of coating. Two pipes, one with each type of coating, were buried together and left for the same duration of time at 15 different locations, providing a natural pairing of observations. Corrosion for two coatings were measure by two variables $(p = 2)$. The tests that coating 2 is better than coating 1 are conducted.

2) The one-way analysis of variance data: The measurements in a dental study on boys from ages 8 to 14 $(p = 3)$ were reported by Potthoff and Roy (1964). We want to test that this measurement has the growth curve.

The results are shown in Table 5. For the matched-pair data the $p-values$ of the three tests equal to 1.1045e-02, 2.4196e-02 and 9.7431e-03 respectively. Thus, all the three tests lead to rejection of the hypothesis that coating 2 has equality the same as coating 1. The $p-values$ of the three tests for the one-way analysis of variance data are 9.6860e-06, 4.7028e-05 and 1.3925e-05 respectively and also lead to the rejection of hypothesis that the latter measurements in the dental study do not increase with ages.

Acknowledgement

The author would like to express sincere thanks to Professor F.T. Wright of the University of Missouri-Columbia, USA, for his invaluable suggestions.

Table 5. Observed *p-values* for testing the equality of the mean difference for two data.

	LRT	N_{2T}	G_{2S}
The matched-pair data:			
Statistic	0.7728	5.0231	10.8189
Average sum		1.1981>0	
p-values	1.1045e-02	2.4196e-02	9.7431e-03
The one-way analysis of variance data:			
Statistic	5.1971	20.8813	73.0846
Average sum		3.0852>0	
p-values	9.6860e-06	4.7028e-05	1.3925e-05

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

Chongcharoen, S. and Wright, F.T. 2007. Permutation and scale invariant one-sided approximate likelihood ratio tests. Probability and Statistics letter, 77(8), 774-781.

Chongcharoen, S. and Wright, F.T. 2006. Scale invariant Follman-type tests. International Journal of Statistics and Management Systems. (Submitted.)

Chongcharoen, S., Singh, B. and Wright, F. T. 2002. Powers of some one-sided multivariate tests with the population covariance matrix known up to a multiplicative constant. Journal of Statistical Planning and Inference, 107, 103-121.Chongcharoen, Samruam,1998. One-sided Multivariate Tests. Ph.D. dissertation, University of Missouri-Columbia.

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

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

Lei, X., Peng, Y.B. and Wright, F. T. 1995. Testing homogeneity of normal means with a simply ordered alternative and dependent observations. Computational Statistics & Data Analysis, 20;173-183.

Perlman, M. D. 1969. One-sided problems in multivariate analysis. Annals Mathematical Statistics, 40, 549-567.

Potthoff, R. F., and Roy, S. N. 1964. A Generalized Multivariate Analysis of Variance Model Useful Espectially for Growth Curve Problems, Biometrika, 51, 313-326.

Rencher, A.C. 2002. Methods of Multivariate Analysis. 2nd edition, John Wiley & Sons, Inc. publication.

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

Shorack, G. R. 1967. Testing against ordered alternatives in model I analysis of variance: Normal theory and nonparametrics. Annals 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., Gnecco, C. and Geller, N. L. 1989. An approximate likelihood ratio test for a normal mean vector with nonnegative components with application to clinical trials. Biometrika, 76, 577-583.