COMPUTATIONAL STATISTICS & DATA ANALYSIS

ELSEVIER Computational Statistics & Data Analysis 22 (1996) 235–249

Rank-based comparisons of treatments with a control for repeated measures designs

Yuh-Ing Chen[a,]*, Sheng-Shu Cheng[a], Hong-Long Wang[b]

[a]*Graduate Institute of Statistics, National Central University, Chung-Li, Taiwan 320, Taiwan, ROC*

[b]*Graduate School of Statistics, National Chung-Shin University, Taipei, Taiwan, ROC*

Received August 1994; revised July 1995

Abstract

In this paper the problem of comparing several treatments with a control in a one-way repeated measures design is considered. Multiple testing procedures based on rank transformation data are proposed for determining which treatments are more effective than the control. The results of a Monte Carlo level and power study are presented.

Keywords: Monte Carlo study; Rank transformation data; Repeated measured design

1. Introduction

Let $X_i^t = (X_{i0}, X_{i1}, \ldots, X_{ik})$, $i = 1, \ldots, n$, be a random sample from a continuous $(k + 1)$-variate distribution with distribution function F and covariance matrix $\Sigma = (\sigma_{ij})$. The setting in which the X_{ij} is the response for the ith experimental unit receiving the jth treatment ($j = 0$ denotes the control) is generally referred to as the one-way repeated measures design. When F is a normal distribution function and the corresponding covariance matrix Σ satisfies $\sigma_{ij} = \tau^2\delta_{ij} + \beta_i + \beta_j$, where $\delta_{ij} = 1$, if $i = j$, and 0 otherwise, which is commonly referred to be a spherical matrix (see, for instance, Huynh and Feldt, 1970), the procedure based on the ANOVA F statistic is usually employed for testing the equality of the $(k + 1)$ treatments (see, for example, Crowder and Hand, 1990). Note that, under the assumption of compound

*Corresponding author.

SSDI 0167-9473(95)00051-8

symmetry, that is, $\Sigma = \sigma^2[(1 - \pi)I + \pi\mathbf{1}\mathbf{1}^t]$ with $-1/k < \pi < 1$, where I is an identy matrix and $\mathbf{1}$ is a vector of ones, these repeated measures can be expressed as exchangeable random variables when the treatments and the control are equally effective. From this point of view, Agresti and Pendergast (1986) considered rank tests for detecting treatment effects based on a single ranking of the entire sample which are related to the one proposed by Koch (1969) and the rank analog of the ANOVA F statistic suggested by Iman et al. (1984), respectively. Kepner and Robinson (1988) later provided a theoretic support for the use of these statistics in the one-way repeated measures design. Ernst and Kepner (1993) further investigated the performance of the rank tests for repeated measures designs via a Monte Carlo study.

In comparing several treatments with a control, however, procedures that are able to decide which treatments (if any) are better than the control would be more preferred. To this end, Wang (1992), based on the sample average vector of the repeated measures, proposed a multiple comparison procedure for comparing k treatments with a control when the normally distributed repeated measures satisfy the sphericity condition. However, there are very limited practical situations in which the normal assumptions is tenable. Moreover, the central limit theorem assures that the mean vector is approximately normal only for sufficiently large sample sizes. Sometimes there are technical or economic reasons for taking only a few repeated observations and, hence, one cannot rely on the central limit theorem for normality. In this case, non-parametric procedures which provide practical alternatives for comparing several treatments with a control in the one-way repeated measures design would be needed.

In Section 2 we discuss previously proposed testing procedures. In Section 3 we consider rank-based multiple comparisons procedures for determining the treatments which are more effective than the control. In Section 4 a numerical example of studying the lens strength on the visual acuity presented in Crowder and Hand (1990) is illustrated. In Section 5 we describe the method of conducting the Monte Carlo study investigation of the relative level and power performances of the competing multiple testing procedures considered in this paper. In Section 6 we present and discuss the simulation results.

2. The previous work

Suppose that the independent random vectors $\boldsymbol{X}_i$ are identically distributed to a $(k + 1)$-variate normal distribution with the mean vector $\boldsymbol{\mu}^t = (\mu_0, \mu_1, \ldots, \mu_k)$ and the covariance matrix Σ. Let

$$\bar{X}_{i.} = \sum_{j=0}^{k} X_{ij}/(k + 1),$$

$$\bar{X}_{.j} = \sum_{i=1}^{n} X_{ij}/n,$$

$$\bar{X}_{..} = \sum_{i=1}^{n} \sum_{j=0}^{k} X_{ij}/[n(k + 1)].$$

Assume that Σ is spherical in the sense that $\mathrm{Var}(X_{ij} - X_{ij'})$ remains constant for and i and $j \neq j'$. Wang (1992) proposed to claim $\mu_j > \mu_0$ if

$$W\colon \sqrt{n}(\bar{X}_{.j} - \bar{X}_{.0})/\sqrt{2\mathrm{MSAB}} \geq t(\alpha; k, k(n-1), 0.5), \tag{1}$$

where

$$\mathrm{MSAB} = \sum_{i=1}^{n} \sum_{j=0}^{k} (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2/[k(n-1)],$$

and $t(\alpha; k, k(n-1), 0.5)$ is the upper αth percentile of the maximum component of a k-variate equicorrelated t-distribution with $k(n-1)$ degrees of freedom and the common correlation coefficient 0.5 which has been tabulated in Dunnett (1964). When Σ is not spherical, however, the level performance of Wang's procedure tends to be anti-conservative. Let λ_j, $j = 1, \ldots, k+1$, be the eigenvalues of $\Sigma(I - \mathbf{1}\mathbf{1}^t/(k+1))$, where, again, I is an identy matrix and $\mathbf{1}$ is a vector of ones. Since the λ's being constant is the necessary and sufficient condition for Σ being spherical, Greenhouse and Geisser (1959) defined a measured of departure from the spherificity to be

$$\varepsilon = \left(\sum_{j=1}^{k} \lambda_j\right)^2 \Bigg/ \left(k \sum_{j=1}^{k} \lambda_j^2\right), \tag{2}$$

which is between (including) $1/k$ and 1. The estimation of the unknown constant ε has been extensively discussed by Greenhouse and Geisser (1959), Huynh and Feldt (1970) among others. Since ε is less than 1 when Σ is not spherical, Wang further suggested to replace the critical value $t(\alpha; k, k(n-1), 0.5)$ by $t(\alpha; k, k(n-1)\hat{\varepsilon}, 0.5)$ in the multiple comparison procedure, where $\hat{\varepsilon}$ is an estimate of ε.

Let R_{ij} be the rank of X_{ij} among the $N = n(k+1)$ observations and set

$$\bar{R}_{i.} = \sum_{j=0}^{k} R_{ij}/(k+1), \qquad \bar{R}_{.j} = \sum_{i=1}^{n} R_{ij}/n, \qquad \bar{R}_{..} = (N+1)/2.$$

Note that, under the assumption of compound symmetry that the components of $\boldsymbol{X}_i$ are equally correlated repeated measures on the ith experimental unit, the null hypothesis, denoted by H_0, of no treatment effects can be expressed as

$$\mathrm{H}_0^*\colon F(x_0, x_1, \ldots, x_k) = F(x_{\pi_0}, x_{\pi_1}, \ldots, x_{\pi_k})$$

for all $\boldsymbol{x}^t = (x_0, x_1, \ldots, x_k)$ and all permutations $(\pi_0, \pi_1, \ldots, \pi_k)$ of $(0, 1, \ldots, k)$. Agresti and Pendergast (1986) then obtained that, when H_0^* is true, $\mathrm{Cov}(R_{ij}, R_{ij'}) = \rho$ for all $j \neq j'$, and $\mathrm{Cov}(R_{ij}, R_{i'j'}) = \lambda$ for all j and j' with $i \neq i'$. Note that both ρ and λ depend on n, the number of observation vectors. Let $\sigma^2 = \mathrm{Var}(R_{ij}) = (N^2 - 1)/12$. They also found

$$\mathrm{Var}(\bar{R}_{.j}) = [1 + (n-1)\lambda]\sigma^2/n,$$

and

$$\mathrm{Cov}(\bar{R}_{.j}, \bar{R}_{j'}) = [\rho + (n-1)\lambda]\sigma^2/n$$

for $j, j' = 0, 1, \ldots, k$ and $j \neq j'$. Since $\mathrm{Var}(\sum_{j=0}^{k} \bar{R}_{.j}) = 0$ implies $\lambda = -(1 + k\rho)/[(k+1)(n-1)]$, the two equations stated above can be rewritten respectively, as

$$\mathrm{Var}(\bar{R}_{.j}) = k\sigma^2(1-\rho)/N,$$

and

$$\mathrm{Cov}(\bar{R}_{.j}, \bar{R}_{.j'}) = -\sigma^2(1-\rho)/N.$$

Agresti and Pendergast then conjectured that the limiting distribution of the random variable

$$n \sum_{j=0}^{k} [\bar{R}_{.j} - (N+1)/2]^2/[\sigma^2(1-\rho)]$$

is a χ^2-distribution with k degrees of freedom, denoted by χ_k^2, provided that the limiting distribution of the random vector $\bar{\boldsymbol{R}}^t = (\bar{R}_{.1}, \ldots, \bar{R}_{.k})$ is a k-variate normal distribution. Kepner and Robinson (1988) latter showed that this conjecture holds when H_0^* is true and proved that the two estimators of $\sigma^2(1-\rho)$ raised by Agresti and Pendergast are both consistent, namely,

$$\mathrm{RMSE} = \sum_{i=1}^{n} \sum_{j=0}^{k} (R_{ij} - \bar{R}_{i.})^2/(nk), \tag{3}$$

$$\mathrm{RMSAB} = \sum_{i=1}^{n} \sum_{j=0}^{k} [R_{ij} - \bar{R}_{i.} - \bar{R}_{.j} + (N+1)/2]^2/[k(n-1)]. \tag{4}$$

Finally, for testing of H_0^*, Kepner and Robinson suggested to use either the Koch's (1969) statistic

$$\mathrm{RT}_1 = \frac{n \sum_{j=0}^{k} [\bar{R}_{.j} - (N+1)/2]^2/k}{\mathrm{RMSE}}$$

or the rank transformation statistic proposed by Iman et al. (1984)

$$\mathrm{RT}_2 = \frac{n \sum_{j=0}^{k} [\bar{R}_{.j} - (N+1)/2]^2/k}{\mathrm{RMSAB}}$$

compared to their limiting χ_k^2/k-distribution or to an F-distribution with k and $k(n-1)$ degrees of freedom in the spirit of Iman and Davenport (1980). Ernst and Kepner (1993) further conducted a Monte Carlo study to investigate the level and power performances of some competing tests for detecting the treatment effects. According to their simulation results, the test based on RT_2 compared to an F-distribution maintains a reasonable level and has a nice power performance for non-normal distributions.

3. The proposed multiple test

Following the results in Kepner and Robinson (1988), we obtain that, under H_0^*, the limiting distribution of the random vector $\{\sqrt{n}/\sqrt{2\sigma^2(1-\rho)}\}(\bar{R}_{.1}-\bar{R}_{.0},\ldots,\bar{R}_{.k}-\bar{R}_{.0})$ is a k-variate normal distribution with mean $\mathbf{0}$ and covariance matrix $\Sigma=(\sigma_{ij})$, where $\sigma_{ii}=1$ and $\sigma_{ij}=1/2$ for $i,\ j=1,\ldots,k$ and $i\neq j$. Therefore, the limiting distribution of the random variable

$$\max_{1\le j\le k}\left[\frac{\sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})}{\sqrt{2\sigma^2(1-\rho)}}\right]$$

is the same as that of the maximum of k equally correlated standard normal variates with common correlation 0.5, denoted by $z(k,\,0.5)$. For the form of the distribution $z(k,\,0.5)$, see, for example, Gupta (1963).

It was observed, in Kepner and Robinson (1988), that both the estimators, RMSE and RMSAB stated in (3) and (4), provide consistent estimators of $\sigma^2(1-\rho)$. Slutsky's theorem then implies that the limiting distribution of the two statistics,

$$\max_{1\le j\le k}\left[\frac{\sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})}{\sqrt{2\mathrm{RMSE}}}\right],$$

$$\max_{1\le j\le k}\left[\frac{\sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})}{\sqrt{2\mathrm{RMSAB}}}\right],$$

is also the distribution $z(k,\,0.5)$. Hence, we consider to claim that the jth treatment is better than the control if

$$\mathrm{RMT}_1\colon\ \sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})/\sqrt{2\mathrm{RMSE}}\ge z(\alpha;k,0.5),\quad j=1,\ldots,k, \tag{5}$$

$$\mathrm{RMT}_2\colon\ \sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})/\sqrt{2\mathrm{RMSAB}}\ge z(\alpha;k,0.5),\quad j=1,\ldots,k, \tag{6}$$

where $z(\alpha;\,k,\,0.5)$ is the upper αth percentile of $z(k,\,0.5)$ which has been extensively tabulated in Gupta (1963). However, according to the simulation results in Ernst and Kepner (1993), two more multiple testing procedures utilizing the statistics in (5) and (6), respectively, but different critical value, namely, $t(\alpha;k,k(n-1),0.5)$ as stated in (1), are obtained which suggest to claim that the jth treatment is better than the control if

$$\mathrm{RMT}_1^*\colon\ \sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})/\sqrt{2\mathrm{RMSE}}\ge t(\alpha;k,k(n-1),0.5),\quad j=1,\ldots,k, \tag{7}$$

$$\mathrm{RMT}_2^*\colon\ \sqrt{n}(\bar{R}_{.j}-\bar{R}_{.0})/\sqrt{2\mathrm{RMSAB}}\ge t(\alpha;k,k(n-1),0.5),\quad j=1,\ldots,k, \tag{8}$$

Note that, if the assumption of compound symmetry does not hold, the null hypothesis H_0 may not be expressed as H_0^*. In this case, as we will see from the simulation results in Section 6, the proposed multiple procedure, RMT_2^*, tends to be anti-conservative in the level performance. To determine which treatments are more effective than the control in such a setting, we consider a modified procedure analogous to the parametric adjustment employed by Wang (1992). For simplicity, however, we use the smallest value of ε, $1/k$, and then modify the procedure by comparing its test statistic with $t(\alpha; k, n-1)$.

4. An example

To investigate the effect of the lens strength on the visual acuity, the response times of the eyes each through lenses of powers 6/6, 6/18, 6/36 and 6/60 to a stimulus (a light flash) were measured, where, for example, the power 6/36 indicates that the magnification is such that the eye will perceive as being at 6 ft an object actually positioned at a distance of 36 ft. The data in Table 1 is the time lag (milliseconds) between the stimulus and the electrical response at the back of the cortex. [These data correspond to the left eye visual acuity with varying lens strength as given in Table 3.2 of Crowder and Hand (1990).]
We calculate the following statistics based on the original data:

$$\bar{X}_{1.} = 118.75,\quad \bar{X}_{2.} = 112.25,\quad \bar{X}_{3.} = 118.75,\quad \bar{X}_{4.} = 114,\quad \bar{X}_{5.} = 114.75,$$

$$\bar{X}_{6.} = 111,\quad \bar{X}_{7.} = 110.75,\quad \bar{X}_{.0} = 113.86,\quad \bar{X}_{.1} = 114.57,$$

$$\bar{X}_{.2} = 111.14,\quad \bar{X}_{.3} = 117.71,\quad \bar{X}_{..} = 114.32.$$

It can be computed that MSAB $= 23.16$ and hence

$$\sqrt{n}(\bar{X}_{.1} - \bar{X}_{.0})/\sqrt{2\mathrm{MSAB}} = 0.276,$$

$$\sqrt{n}(\bar{X}_{.2} - \bar{X}_{.0})/\sqrt{2\mathrm{MSAB}} = -1.048,$$

$$\sqrt{n}(\bar{X}_{.3} - \bar{X}_{.0})/\sqrt{2\mathrm{MSAB}} = 1.489.$$

We observe, from Dunnett (1964), that $t(0.10; 3, 18, 0.5) = 1.82$. Therefore, Wang's procedure leads to claim that, under level $\alpha = 0.10$, there is no effect of the lens strength on the visual acuity. Now, we calculate the following statistics based on the rank transformation data:

$$\bar{R}_{1.} = 21.875,\quad \bar{R}_{2.} = 8.5,\quad \bar{R}_{3.} = 23.75,\quad \bar{R}_{4.} = 11.75,\quad \bar{R}_{5.} = 13.125,$$

$$\bar{R}_{6.} = 14.25,\quad \bar{R}_{7.} = 8.25,\quad \bar{R}_{.0} = 12.36,\quad \bar{R}_{.1} = 14.00,$$

$$\bar{R}_{.2} = 11.85,\quad \bar{R}_{.3} = 19.79\quad \bar{R}_{..} = 14.50$$

It can be computed that RMSAB $= 35.08$ and thus

$$\sqrt{n}(\bar{R}_{.1} - \bar{R}_{.0})/\sqrt{2\mathrm{RMSAB}} = 0.518,$$

$$\sqrt{n}(\bar{R}_{.2} - \bar{R}_{.0})/\sqrt{2\mathrm{RMSAB}} = -0.158,$$

$$\sqrt{n}(\bar{R}_{.3} - \bar{R}_{.0})/\sqrt{2\mathrm{RMSAB}} = 2.344.$$

Table 1
Visual acuity with varying lens strength

Subject	6/6	6/18	6/36	6/60
1	116	119	116	124
2	110	110	114	115
3	117	118	120	120
4	112	116	115	113
5	113	114	114	118
6	119	115	94	116
7	110	110	105	118

Hence, we conclude, at the 10% significance level, that the lens of power 6/60 results in less visual acuity than that of power 6/6. Note that, using the sample covariance matrix in computing ε in (2), we obtain the Greenhouse and Geisser's estimator of ε which is 0.428. For simplicity, we use the smallest value of ε, namely, $\frac{1}{3}$, to obtain the modified critical value $t(0.10; 3, 6, 0.5) = 2.02$. (In fact, under the sample correlation structure, the approximate level of the modified testing procedure obtained from a simulation study based on 5000 replications is 0.0878.) It is obvious that our conclusion still holds.

5. Methodology

We conducted a Monte Carlo study to examine the relative levels and powers of Wang's (1992) procedure and the multiple tests suggested in this paper for comparing several treatments with a control in a one-way repeated measures design. We considered $k = 3$ and 4 treatments with $n = 10$, 20 and 30 observations in the level study and $n = 10$ and 20 in the power study. For each of these settings, multivariate normal, multivariate t with 10 degrees of freedom (d.f.), multivariate Cauchy (i.e. multivariate t with 1 d.f.) and multivariate exponential distributions were considered as the underlying distributions. For the definitions of multivariate normal, multivariate t and multivariate Cauchy, see, for example, Fang et al. (1990). Note that multivariate t with 10 d.f. represents the symmetric and moderately heavy-tailed distribution, multivariate Cauchy represents the symmetric and heavy-tailed distribution and multivariate exponential represents the asymmetric distribution.

This Monte Carlo study was implemented on a VAX 9320 computer at National Central University and all programmings were done in FORTRAN 77. The International Mathematical and Statistical Libraries (IMSL) routine RNMVN was used to generate multivariate normal with zero mean vector and covariance matrix Σ, denoted by $\boldsymbol{Z}$. The IMSL routine RNCHI was employed to generate the chi-squared with v d.f. variates, denoted by U. The multivariate t variates were then formulated by $\boldsymbol{Z}/\sqrt{U/v}$. Moreover, the algorithm provided by Sim (1993) was employed to generate the appropriate multivariate exponential variates. Note that, in generating multivariate normal, t and cauchy variates, the common correlation $\rho_{jj'} = 0.2$ and 0.8 and unequal correlation $\rho_{jj'} = 0.5^{|j-j'|}$ were considered for the $\mathbf{Z}$. The three different correlation structures were also used for the multivariate exponential variates. In the level study, the multivariate normal (t, Cauchy, exponential) distribution with standard normal (t, Cauchy, exponential) marginal distributions was considered. In the power study, we used the multivariate normal (t, Cauchy, exponential) distribution with various values of location parameters, denoted by $\theta_0, \theta_1, \dots, \theta_k$, and the designated treatment effects configurations correspond to values of $\theta_{i0} = \theta_i - \theta_0$ for $i = 1, \dots, k$.

The experiment-wise error rate (proportion of experiments with at least one treatment erroneously declared more effective than the control) was utilized to evaluate the level performances of the multiple test procedures under consideration. The experiment-wise power (probability of correctly detecting at least one treatment which is better than the control) and the comparison-wise power (probability of correctly detecting all the treatments which are better than the control) were employed to assess the power performances of the testing procedures. The results of the level study are presented in Table 3 and those of the power study are reported in

Table 2
Summary statistics for judging the adequacy of the simulation
(a) Multivariate normal

	$\theta_j = 0,\ \ j = 0, 1, 2, 3$											
	$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.2, & j \neq j' \end{cases}$				$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.8, & j \neq j' \end{cases}$				$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.5^{\lvert j-j' \rvert}, & j \neq j' \end{cases}$			
$n = 10$	$\theta_j = 0,\ \ j = 0, 1, 2, 3$											
$\rho_{jj'}, j \leq j'$	0.995	0.194	0.200	0.193	0.995	0.798	0.794	0.797	1.008	0.500	0.255	0.153
		0.993	0.198	0.198		0.998	0.794	0.800		0.996	0.503	0.198
			0.999	0.192			0.990	0.794			1.009	0.504
				0.997				0.996				1.002
$n = 20$												
$\rho_{jj'}, j \leq j'$	0.992	0.200	0.204	0.203	1.005	0.806	0.807	0.807	0.993	0.495	0.247	0.125
		0.991	0.196	0.198		1.007	0.808	0.808		0.995	0.497	0.249
			1.003	0.202			1.007	0.808			0.998	0.497
				1.001				1.008				0.996
$n = 30$												
$\rho_{jj'}, j \leq j'$	1.001	0.194	0.201	0.201	0.991	0.797	0.795	0.794	1.001	0.503	0.252	0.123
		0.997	0.202	0.204		1.002	0.798	0.798		1.003	0.504	0.250
			1.003	0.202			0.996	0.796			1.005	0.501
				0.998				0.994				1.000

(b) Multivariate exponential

	$\theta_j = 0, \quad j = 0, 1, 2, 3$											
	$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.2, & j \neq j' \end{cases}$				$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.8, & j \neq j' \end{cases}$				$\rho_{jj'} = \begin{cases} 1, & j = j' \\ 0.5^{\lvert j - j' \rvert}, & j \neq j' \end{cases}$			
$n = 10$												
θ_j	1.003	1.006	1.007	1.002	1.004	1.003	1.008	1.010	1.003	1.005	1.005	1.011
$\rho_{jj'}, j \le j'$	1.012	0.204	0.206	0.212	1.013	0.809	0.806	0.800	1.000	0.497	0.253	0.117
		1.001	0.214	0.214		1.009	0.815	0.806		1.008	0.504	0.248
			1.012	0.205			1.012	0.811			1.015	0.509
				1.012				1.003				1.013
$n = 20$												
θ_j	1.000	1.002	1.005	1.004	1.000	0.996	1.005	1.008	0.999	0.996	0.998	1.012
$\rho_{jj'}, j \le j'$	0.998	0.198	0.198	0.209	1.004	0.802	0.806	0.808	0.994	0.491	0.247	0.120
		1.002	0.205	0.203		0.998	0.807	0.807		0.992	0.498	0.254
			1.009	0.199			1.006	0.812			0.996	0.506
				1.008				1.008				1.009
$n = 30$												
θ_j	1.003	1.002	1.003	1002	1.000	1.000	1.007	1.003	0.998	0.999	0.998	1.005
$\rho_{jj'}, j \le j'$	0.999	0.201	0.197	0.203	1.007	0.809	0.808	0.804	0.995	0.497	0.247	0.126
		1.006	0.203	0.208		1.009	0.805	0.808		0.997	0.500	0.259
			1.008	0.199			1.004	0.806			0.998	0.495
				1.007				1.006				1.000

Tables 4 and 5. Since, in each case, we used 5000 replications in obtaining the estimated error rate or power under the nominal level $\alpha = 0.05$, we are guaranteed a standard error not greater than 0.0031 for estimating the experiment-wise error rate. We then indicate, by + (−) signs, whenever the estimated error rate is two or more standard errors above (below) 0.05.

6. Results

6.1. Adequacy of the data generation

To assess the adequacy of the data generation, we computed, based on 5000 replications, the average mean vector and average correlation coefficients of the generated data from $(k + 1)$-dimensional normal or exponential distribution with

Table 3
Experiment-wise error rate estimates for $\alpha = 0.05$

			$k = 3$					$k = 4$				
Distribution	n	τ	RMT_1	RMT_2	RMT_1^*	RMT_2^*	W	RMT_1	RMT_2	RMT_1^*	RMT_2^*	W
Multivariate normal	10	0.2	0.048	0.060+	0.036−	0.048	0.048	0.049	0.059+	0.038−	0.049	0.048
		0.8	0.053	0.065+	0.041−	0.053	0.053	0.047	0.060+	0.039−	0.050	0.052
		0.5[a]	0.056	0.068+	0.046	0.059+	0.057+	0.055	0.065+	0.047	0.059+	0.057+
	20	0.2	0.050	0.056	0.045	0.052	0.048	0.049	0.054	0.045	0.048	0.052
		0.8	0.050	0.056	0.046	0.051	0.056	0.050	0.056	0.047	0.052	0.051
		0.5[a]	0.060+	0.064+	0.056	0.060+	0.063+	0.058+	0.059+	0.050	0.057+	0.058+
	30	0.2	0.054	0.057+	0.051	0.054	0.054	0.053	0.057+	0.050	0.053	0.054
		0.8	0.050	0.054	0.045	0.050	0.046	0.047	0.050	0.044	0.048	0.047
		0.5[a]	0.062+	0.065+	0.057+	0.061+	0.061+	0.058+	0.062+	0.057+	0.059+	0.060+
Multivariate t with 10 d.f.	10	0.2	0.050	0.061+	0.038−	0.053	0.054	0.050	0.060+	0.040−	0.052	0.051
		0.8	0.048	0.060+	0.037−	0.048	0.052	0.055	0.065+	0.045	0.056	0.057+
		0.5[a]	0.059+	0.072+	0.049	0.061+	0.061+	0.063+	0.073+	0.051	0.061+	0.060+
	20	0.2	0.049	0.054	0.044	0.049	0.050	0.047	0.052	0.041−	0.047	0.047
		0.8	0.053	0.059+	0.047	0.056	0.050	0.048	0.052	0.043−	0.048	0.047
		0.5[a]	0.058+	0.064+	0.055	0.059+	0.058+	0.058+	0.063+	0.051	0.059+	0.057+
	30	0.2	0.052	0.055	0.048	0.052	0.052	0.049	0.052	0.045	0.048	0.049
		0.8	0.054	0.057+	0.050	0.054	0.051	0.051	0.054	0.048	0.050	0.047
		0.5[a]	0.058+	0.062+	0.053	0.058+	0.060+	0.063+	0.066+	0.060+	0.063+	0.064+
Multivariate Cauchy	10	0.2	0.046	0.055	0.036−	0.047	0.031−	0.046	0.060+	0.037−	0.049	0.030−
		0.8	0.047	0.057+	0.034−	0.046	0.034−	0.044	0.057+	0.035−	0.046	0.033−
		0.5[a]	0.058+	0.069+	0.045	0.057+	0.033−	0.055	0.064+	0.045	0.058+	0.036−
	20	0.2	0.050	0.055	0.044	0.049	0.029−	0.045	0.050	0.038−	0.045	0.031−
		0.8	0.053	0.059+	0.048	0.054	0.034−	0.049	0.053	0.047	0.049	0.030−
		0.5[a]	0.058+	0.063+	0.051	0.057+	0.036−	0.061+	0.066+	0.055	0.062+	0.036−
	30	0.2	0.051	0.048	0.053	0.051	0.032−	0.049	0.053	0.047	0.049	0.030−
		0.8	0.050	0.047	0.053	0.050	0.032−	0.044	0.046	0.040−	0.044	0.031−
		0.5[a]	0.059+	0.055	0.063+	0.059+	0.040−	0.057+	0.058+	0.053	0.057+	0.038−
Multivariate exponential	10	0.2	0.045	0.060+	0.035−	0.047	0.047	0.048	0.060+	0.038−	0.051	0.056
		0.8	0.046	0.058+	0.036−	0.045	0.036−	0.045	0.057+	0.037−	0.045	0.040−
		0.5[a]	0.061+	0.072+	0.050	0.062+	0.060+	0.062+	0.074+	0.051	0.064+	0.067+
	20	0.2	0.044	0.049	0.039−	0.044	0.046	0.049	0.053	0.043−	0.049	0.050
		0.8	0.045	0.050	0.041−	0.045	0.042−	0.046	0.049	0.041−	0.045	0.043−
		0.5[a]	0.060+	0.064+	0.054	0.060+	0.059+	0.071+	0.077+	0.068+	0.072+	0.067+
	30	0.2	0.048	0.053	0.046	0.048	0.050	0.048	0.052	0.045	0.049	0.050
		0.8	0.053	0.056	0.048	0.053	0.050	0.052	0.054	0.049	0.052	0.047
		0.5[a]	0.062+	0.065+	0.058+	0.062+	0.059+	0.076+	0.080+	0.073+	0.076+	0.075+

[a] $\rho_{jj'} = 0.5^{|j-j'|}$.
+ (−): At least two standard error above (below) $\alpha = 0.05$.

sample sizes $n = 10$, 20 and 30, respectively. The adequacy of the data generation for $k = 3$ and $k = 4$ is quite similar. Therefore, we only summarized the results for $k = 3$ in Table 2. By comparing these summarized statistics with their theoretical counterparts, the simulated data seem to possess approximately the desired distributional properties.

6.2. Comparison of testing procedures

When the repeated measures have a common intervariable correlation coefficient, it is evident, upon examination of Table 3, that both the testing procedures, RMT_1 (Eq. (5)) and RMT_2^* (Eq. (8)), reasonably maintain their levels. In this case, the testing procedure, RMT_1^* (Eq. (7)) tends to be conservative in holding its level, while the level performance of RMT_2 (Eq. (6)) is anti-conservative, especially, for the case of small sample size corresponding to $n = 10$. When the intervariable correlation coefficients are unequal, however, all the testing procedures mentioned above tend to be anti-conservative. Therefore, in the power comparison, we simply considered the testing procedures, RMT_1 and RMT_2^* for the case of equal correlation.

Wang's procedure, W (Eq. (1)), holds its level quite well when the repeated measures are distributed to the equally correlated multivariate t with 10 d.f. or

Table 4(a)
Experiment-wise power estimates for $\alpha = 0.05$ and $k = 3$

					$n = 10$			$n = 20$		
Distribution	τ	θ_{10}	θ_{20}	θ_{30}	RMT_1	RMT_2^*	W	RMT_1	RMT_2^*	W
Multivariate normal	0.2	0.0	0.0	0.4	0.138	0.148	0.151	0.230	0.241	0.255
		0.0	0.2	0.4	0.161	0.163	0.168	0.260	0.269	0.283
		0.2	0.4	0.4	0.221	0.224	0.230	0.378	0.380	0.397
	0.8	0.0	0.0	0.4	0.359	0.393	0.451	0.692	0.715	0.782
		0.0	0.2	0.4	0.395	0.423	0.478	0.720	0.736	0.797
		0.2	0.4	0.4	0.553	0.557	0.606	0.851	0.854	0.900
Multivariate t with 10 d.f.	0.2	0.0	0.0	0.4	0.128	0.136	0.141	0.209	0.216	0.215
		0.0	0.2	0.4	0.146	0.150	0.157	0.237	0.242	0.241
		0.2	0.4	0.4	0.211	0.214	0.212	0.290	0.296	0.300
	0.8	0.0	0.0	0.4	0.327	0.355	0.386	0.615	0.640	0.665
		0.0	0.2	0.4	0.365	0.390	0.413	0.646	0.663	0.686
		0.2	0.4	0.4	0.514	0.521	0.549	0.792	0.794	0.812
Multivariate Cauchy	0.2	0.0	0.0	0.4	0.082	0.086	0.044	0.116	0.120	0.046
		0.0	0.2	0.4	0.093	0.097	0.050	0.137	0.140	0.052
		0.2	0.4	0.4	0.129	0.133	0.063	0.189	0.191	0.068
	0.8	0.0	0.0	0.4	0.178	0.189	0.086	0.276	0.285	0.089
		0.0	0.2	0.4	0.204	0.212	0.097	0.309	0.304	0.101
		0.2	0.4	0.4	0.292	0.292	0.130	0.415	0.417	0.139
Multivariate exponential	0.2	0.0	0.0	0.4	0.237	0.254	0.176	0.472	0.494	0.287
		0.0	0.2	0.4	0.279	0.293	0.200	0.516	0.531	0.321
		0.2	0.4	0.4	0.398	0.400	0.277	0.651	0.653	0.440
	0.8	0.0	0.0	0.4	0.697	0.750	0.598	0.933	0.946	0.804
		0.0	0.2	0.4	0.753	0.786	0.609	0.947	0.955	0.815
		0.2	0.4	0.4	0.852	0.857	0.705	0.981	0.981	0.893

Table 4(b)
Experiment-wise power estimates for $\alpha = 0.05$ and $k = 4$

						$n = 10$			$n = 20$		
Distribution	τ	θ_{10}	θ_{20}	θ_{30}	θ_{40}	RMT_1	RMT_2^*	W	RMT_1	RMT_2^*	W
Multivariate normal	0.2	0.0	0.0	0.0	0.4	0.123	0.133	0.139	0.226	0.235	0.248
		0.0	0.0	0.4	0.4	0.175	0.190	0.198	0.332	0.344	0.356
		0.2	0.2	0.4	0.4	0.205	0.207	0.224	0.365	0.367	0.380
	0.8	0.0	0.0	0.0	0.4	0.346	0.379	0.433	0.644	0.667	0.735
		0.0	0.0	0.4	0.4	0.479	0.529	0.582	0.788	0.819	0.865
		0.2	0.2	0.4	0.4	0.546	0.556	0.600	0.827	0.830	0.870
Multivariate t with 10 d.f.	0.2	0.0	0.0	0.0	0.4	0.110	0.116	0.120	0.188	0.195	0.193
		0.0	0.0	0.4	0.4	0.163	0.176	0.174	0.279	0.290	0.290
		0.2	0.2	0.4	0.4	0.195	0.198	0.193	0.315	0.317	0.313
	0.8	0.0	0.0	0.0	0.4	0.312	0.341	0.362	0.592	0.614	0.642
		0.0	0.0	0.4	0.4	0.428	0.475	0.498	0.747	0.778	0.789
		0.2	0.2	0.4	0.4	0.486	0.492	0.521	0.789	0.792	0.797
Multivariate Cauchy	0.2	0.0	0.0	0.0	0.4	0.073	0.078	0.042	0.102	0.102	0.044
		0.0	0.0	0.4	0.4	0.100	0.105	0.052	0.148	0.152	0.055
		0.2	0.2	0.4	0.4	0.119	0.121	0.061	0.171	0.173	0.063
	0.8	0.0	0.0	0.0	0.4	0.156	0.165	0.074	0.249	0.257	0.079
		0.0	0.0	0.4	0.4	0.222	0.244	0.108	0.357	0.372	0.110
		0.2	0.2	0.4	0.4	0.263	0.267	0.123	0.393	0.396	0.123
Multivariate exponential	0.2	0.0	0.0	0.0	0.4	0.229	0.246	0.178	0.469	0.486	0.288
		0.0	0.0	0.4	0.4	0.318	0.346	0.252	0.577	0.607	0.388
		0.2	0.2	0.4	0.4	0.393	0.399	0.278	0.641	0.643	0.418
	0.8	0.0	0.0	0.0	0.4	0.689	0.730	0.571	0.929	0.940	0.783
		0.0	0.0	0.4	0.4	0.775	0.835	0.692	0.963	0.973	0.879
		0.2	0.2	0.4	0.4	0.855	0.865	0.696	0.976	0.978	0.881

multivariate normal. For the setting where the repeated measures are distributed to the multivariate exponential with common correlation 0.8, the level performance of W tends to be conservative unless the sample size is large about 30. Wang's procedure also has an inflated error rate when the intervariable correlation coefficients are unequal for all distributions except the case of multivariate cauchy where the error rate is already relatively conservative when the coefficients are equal.

The power estimates in Tables 4 and 5 show that RMT_2^* is slightly better than RMT_1 in comparing several treatments with a control in one-way repeated measures designs. When the multivariate distribution is normal or t with 10 d.f., RMT_2^* is slightly less powerful than the Wang's procedure W. When the multivariate distribution is exponential, however, Wang's procedure performs poorly. In this case, both the procedures, RMT_1 and RMT_2^*, have better power performances than W. Moreover, for the multivariate Cauchy distribution, although it does not seem to be fair to compare directly the power performances of RMT_2^* and Wang's procedure since the estimated level of RMT_2^* is roughly 1.5 (less than 2, anyway) times of that of W, the estimated power of RMT_2^* is 2 to 3 times of that of W. The increase in power compared to the favored level indicates, however, that the power performance of RMT_2^* is better than that of Wang's procedure.

As a direct consequence of simulation results, we recommend to use the rank-based multiple testing procedure RMT_2^* when the assumption of compound symmetry is tenable for two reasons. First, the procedure RMT_2^* has a reasonable level performance across a variety of distributions, while the Wang's procedure W does not hold its level for either a symmetric and heavy-tailed distribution or an asymmetric distribution with small sample size about 20. Second, the procedure RMT_2^* performs better in power than the Wang's procedure W for an asymmetric or a symmetric and heavy-tailed distribution and it can be regarded as a valid competitor to W for a normal or a symmetric and moderately heavy-tailed distribution.

Table 5(a)
Comparison-wise power estimates for $\alpha = 0.05$ and $k = 3$

					$n = 10$			$n = 20$		
Distribution	τ	θ_{10}	θ_{20}	θ_{30}	RMT_1	RMT_2^*	W	RMT_1	RMT_2^*	W
Multivariate normal	0.2	0.0	0.0	0.4	0.052	0.058	0.058	0.086	0.091	0.095
		0.0	0.2	0.4	0.063	0.068	0.070	0.106	0.113	0.118
		0.2	0.4	0.4	0.096	0.104	0.107	0.178	0.187	0.195
	0.8	0.0	0.0	0.4	0.127	0.144	0.163	0.239	0.252	0.275
		0.0	0.2	0.4	0.157	0.180	0.203	0.302	0.326	0.356
		0.2	0.4	0.4	0.276	0.304	0.340	0.524	0.551	0.602
Multivariate t with 10 d.f.	0.2	0.0	0.0	0.4	0.049	0.055	0.056	0.077	0.082	0.081
		0.0	0.2	0.4	0.059	0.065	0.057	0.099	0.104	0.106
		0.2	0.4	0.4	0.090	0.100	0.098	0.159	0.165	0.164
	0.8	0.0	0.0	0.4	0.117	0.131	0.142	0.214	0.227	0.235
		0.0	0.2	0.4	0.145	0.165	0.176	0.268	0.289	0.300
		0.2	0.4	0.4	0.250	0.277	0.298	0.474	0.499	0.516
Multivariate Cauchy	0.2	0.0	0.0	0.4	0.031	0.036	0.017	0.043	0.046	0.017
		0.0	0.2	0.4	0.035	0.039	0.019	0.054	0.056	0.020
		0.2	0.4	0.4	0.052	0.058	0.025	0.081	0.085	0.027
	0.8	0.0	0.0	0.4	0.064	0.071	0.032	0.099	0.105	0.033
		0.0	0.2	0.4	0.080	0.088	0.039	0.122	0.129	0.038
		0.2	0.4	0.4	0.130	0.141	0.059	0.206	0.218	0.060
Multivariate exponential	0.2	0.0	0.0	0.4	0.087	0.097	0.064	0.168	0.180	0.104
		0.0	0.2	0.4	0.116	0.131	0.078	0.220	0.237	0.130
		0.2	0.4	0.4	0.201	0.219	0.125	0.376	0.393	0.215
	0.8	0.0	0.0	0.4	0.236	0.264	0.211	0.316	0.330	0.280
		0.0	0.2	0.4	0.319	0.385	0.281	0.458	0.509	0.391
		0.2	0.4	0.4	0.544	0.613	0.467	0.741	0.790	0.642

Table 5(b)
Comparison-wise power estimates for $\alpha = 0.05$ and $k = 4$

						$n = 10$			$n = 20$		
Distribution	τ	θ_{10}	θ_{20}	θ_{30}	θ_{40}	RMT_1	RMT_2^*	W	RMT_1	RMT_2^*	W
Multivariate normal	0.2	0.0	0.0	0.0	0.4	0.036	0.041	0.041	0.064	0.067	0.072
		0.0	0.0	0.4	0.4	0.055	0.063	0.066	0.110	0.118	0.123
		0.2	0.2	0.4	0.4	0.070	0.076	0.082	0.138	0.145	0.150
	0.8	0.0	0.0	0.0	0.4	0.093	0.105	0.119	0.169	0.178	0.195
		0.0	0.0	0.4	0.4	0.164	0.195	0.219	0.316	0.345	0.376
		0.2	0.2	0.4	0.4	0.218	0.242	0.271	0.411	0.438	0.482
Multivariate t with 10 d.f.	0.2	0.0	0.0	0.0	0.4	0.032	0.037	0.037	0.053	0.056	0.056
		0.0	0.0	0.4	0.4	0.052	0.060	0.058	0.092	0.099	0.098
		0.2	0.2	0.4	0.4	0.066	0.073	0.072	0.118	0.122	0.121
	0.8	0.0	0.0	0.0	0.4	0.085	0.097	0.103	0.155	0.164	0.171
		0.0	0.0	0.4	0.4	0.146	0.176	0.186	0.290	0.317	0.329
		0.2	0.2	0.4	0.4	0.195	0.217	0.213	0.374	0.395	0.415
Multivariate Cauchy	0.2	0.0	0.0	0.0	0.4	0.022	0.025	0.012	0.030	0.031	0.013
		0.0	0.0	0.4	0.4	0.032	0.035	0.016	0.046	0.049	0.017
		0.2	0.2	0.4	0.4	0.040	0.043	0.020	0.058	0.061	0.021
	0.8	0.0	0.0	0.0	0.4	0.043	0.047	0.021	0.068	0.072	0.023
		0.0	0.0	0.4	0.4	0.071	0.083	0.033	0.121	0.131	0.035
		0.2	0.2	0.4	0.4	0.093	0.102	0.041	0.155	0.163	0.043
Multivariate exponential	0.2	0.0	0.0	0.0	0.4	0.065	0.073	0.050	0.128	0.135	0.078
		0.0	0.0	0.4	0.4	0.113	0.133	0.080	0.226	0.248	0.133
		0.2	0.2	0.4	0.4	0.162	0.181	0.098	0.305	0.321	0.165
	0.8	0.0	0.0	0.0	0.4	0.175	0.195	0.154	0.237	0.248	0.208
		0.0	0.0	0.4	0.4	0.325	0.377	0.293	0.457	0.478	0.396
		0.2	0.2	0.4	0.4	0.462	0.539	0.391	0.651	0.707	0.541

Acknowledgements

We thank the referee for many constructive suggestions which improved the presentation of this paper. The work of the first author was supported by the National Science Council of Taiwan under the contract number NSC-84-2121-M008-007.

References

Agresti, A. and J. Pendergast, Comparing mean ranks for repeated measures data, *Comm. Statist. Theory Methods*, **15** (1986) 1417–1433.

Crowder, M.J. and D.J. Hand, *Analysis of repeated measures* (Chapman and Hall, London, 1990).

Dunnett, C.W., New tables for multiple comparisons with a control, *Biometrics*, **20** (1964) 482–491.

Ernst, M.D. and J.L. Kepner, A Monte Carlo study of rank tests for repeated measures design, *Comm. Statist. Simulation Comput.*, **22** (1993) 671–678.
Fang, K.T., S. Koltz and K.W. Ng, *Symmetric multivariate and related distributions* (Chapman and Hall, New York, 1990).
Greenhouse, S.W. and S. Geisser, On the methods in the analysis of profile data, *Psychometrika*, **24** (1959) 95–112.
Gupta, S.S., Probability integrals of multivariate normal and multivariate *t*, *Ann. Math. Statist.*, **28** (1963) 792–828.
Hochberg, Y. and A.C. Tamhane, *Multiple comparison procedures* (Wiley, New York, 1987).
Huynh, H. and L.S. Feldt, Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions. *J. Amer. Statist. Assoc.*, **65** (1970) 1582–1589.
Iman, R.L. and J.M. Davenport, Approximations of the critical region of the Friedman statistic, *Comm. Statist. Theory Methods*, **A9** (1980) 571–595.
Iman, R.L., S.C. Hora and W.J. Conover, Comparison of asymptotically distribution-free procedures for the analysis of complete blocks, *J. Amer. Statist. Assoc.*, **79** (1984) 674–685.
Kepner, J.L. and D.H. Robinson, Nonparametric methods for detecting treatment effects in repeated measures design, *J. Amer. Statist. Assoc.*, **83** (1988) 456–461.
Koch, G.G., Some aspects of the statistical analysis of split plot experiments in completely randomized layouts. *J. Amer. Statist. Assoc.*, **64** (1969) 485–505.
Sim, C.H., Generation of poisson and gamma random vectors with given marginal and covariance matrix. *J. Statist. Comput. Simulation*, **47** (1993) 1–10.
Wang, H.L., Multiple comparison procedures in repeated measures design, Doctoral thesis (Dept. of Statistics, The Ohio State University, Columbus, Ohio, 1992).