Journal of Mathematics and Statistics 7 (4): 332-342, 2011
ISSN 1549-3644

A Test for Two-Sample Repeated Measures Designs: Effect of High-Dimensional Data

Boonyarit Choopradit and Samruam Chongcharoen
School of Applied Statistics,
National Institute of Development Administration,
Bangkapi District, Bangkok, 10240 Thailand

Abstract: Problem statement: High-dimensional repeated measures data are increasingly encountered in various areas of modern science since classical multivariate statistics, e.g,. Hotelling's T^2, are not well defined in the case of high-dimensional data. **Approach:** In this study, the test statistics with no specific form of covariance matrix for analyzing high-dimensional two-sample repeated measures designs with common equal covariance are proposed. The asymptotic distributions of the proposed test statistics also were derived. **Results:** A simulation study exposes the approximated Type I errors in the null case very well even though the number of subjects of each sample as small as 10. Numerical simulations study indicates that the proposed test have good power. Application of the new tests is demonstrated using data from the body-weight of male Wistar rats example. **Conclusion:** The proposed test statistics have an asymptotically distributed as standard normal distributions, under the null hypothesis. Simulation studies show that these test statistics still accurately control Type I error and have quite good power for any the covariance matrix pattern considered.

Key words: Hypothesis test, repeated measures design, high-dimensional data, type I error, two sample, covariance matrix, simulation study, asymptotically distributed, asymptotic distributions

INTRODUCTION

Repeated measurements across time (occasion, drug or treatment) on the same subject (e.g., patient, animal, cell culture, block) are frequently observed in several scientific fields, for example in medicine, pharmaceutical, agriculture, industrial and social sciences (Zangiabadi and Ahrari, 2005; Sharifi *et al*., 2008; Arbabi *et al*., 2009; Tu and Koh, 2010; Mosallanejad *et al*., 2011). This type of data commonly called univariate repeated measures data (Davis, 2002). A design contain such a data is called repeated measures design. A main advantage of this design is that test results are more powerful since subjects serve as their own controls and then variability among the subjects due to individual differences is eliminated. The simplest repeated measures design is when data is collected as a sequence of equally spaced points in time. There exist several methods of analyzing repeated measurement data. For an overview see Everitt (1995) and Keselman *et al*. (2001). However, these studies pertain only to the case when the dimension of repeated measurements is not exceeds the number of subjects.

Let $x_{1k} = (x_{1k1}, x_{1k2}, \ldots, x_{1kp})^T$ and $x_{2l} = (x_{2l1}, x_{2l2}, \ldots, x_{2lp})^T$ be multivariate random vectors each of p repeated observations measured (times) on kth subject in sample 1 and lth subject in sample 2, respectively, where the samples are drawn independently from two populations (groups) and $k = 1, 2, \ldots, n_1$, $l = 1, 2, \ldots, n_2$. Suppose $E(x_{1k}) = \mu_1$, $E(x_{2l}) = \mu_2$, $Var(x_{1k}) = \Sigma_1$ and $Var(x_{2l}) = \Sigma_2$, where μ_i is the population mean vector and Σ_i is the population covariance matrix of group i for $i = 1, 2$. Then $x = \left(x_{1k}^T, x_{2l}^T\right)^T$ is the vector of all observations from kth subject in sample 1 and lth subject in sample 2, with $E(x) = \mu = \left(\mu_1^T, \mu_2^T\right)^T$ and $Var(x) = \Sigma_1 \oplus \Sigma_2$. The corresponding sample estimators are $\bar{x} = \left(\bar{x}_{1\bullet}^T, \bar{x}_{2\bullet}^T\right)^T$ and $\hat{\Sigma}_1 \oplus \hat{\Sigma}_2$, respectively, where $\bar{x}_{i\bullet} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij} = \left(\bar{x}_{i\bullet 1}, \bar{x}_{i\bullet 2}, \ldots, \bar{x}_{i\bullet p}\right)^T$ is the vector of means of ith sample with $E\left(\bar{x}_{i\bullet}\right) = \mu_i$, $Var\left(\bar{x}_{i\bullet}\right) = \left(1/n_i\right)\Sigma_i$ and $\hat{\Sigma}_i = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{i\bullet}\right)\left(x_{ij} - \bar{x}_{i\bullet}\right)^T$ is the sample covariance matrix in group i for $j = 1, 2, \ldots, n_i$, $i = 1, 2$.

Corresponding Author: Boonyarit Choopradit, School of Applied Statistics, National Institute of Development Administration, Bangkapi District, Bangkok, 10240, Thailand

The notation $\oplus$ is the direct sum and a sample from each population with sample sizes denoted by n_1 and n_2.

Traditionally, if data with fixed dimension are normally distributed, people would use the famous classical Hotelling's T^2 test for two-sample which is defined as $T^2 = \frac{n_1 n_2}{n_1+n_2}(\bar{x}_{1\cdot} - \bar{x}_{2\cdot})^T \hat{\Sigma}^{-1}(\bar{x}_{1\cdot} - \bar{x}_{2\cdot})$, where $\bar{x}_{i\cdot}$ is the ith sample mean vector as defined above, for i = 1, 2 and the pooled sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ be defined by $\hat{\boldsymbol{\Sigma}} = \frac{1}{n_1+n_2-2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{i\cdot})(x_{ij} - \bar{x}_{i\cdot})^T$. It is well known that under the null hypothesis, $\frac{n-p+1}{np}T^2$ has a central F-distribution (Davis, 2002; Anderson, 2003) with p and n-p+1 degrees of freedom where $n = n_1 + n_2 - 2$. Nowadays, however, modern society demands that the use of subjects (e.g., animals, patients) in scientific experiments must be adequately controlled and reduced. This position is enforced by ethical committees who authorize or deny permission for these kind of experiments. As a consequence the statistician sometimes has to work with very small subjects, but the large dimension of repeated measurements on each subject. Unfortunately, if n<p, called high-dimensional data, the Hotelling's T^2 test is not well defined because the sample covariance matrix becomes singular.

Recently, there has been some interest in investigating the behavior of high dimensionality n<p that can be found in the literature. A non-exact test for two sample case when n<p was first proposed by Dempster (1958). Many works have been published on hypothesis testing for means when both p and n go to infinity with the ratio p/n must remain bounded, Bai and Saranadasa (1996), Fujikoshi *et al*. (2004); Srivastava and Fujikoshi (2006); Srivastava (2007; 2009); Schott (2007) and Srivastava and Du (2008). In addition, when sample covariance matrix does not have an inverse, Chongcharoen (2011) proposed one way to modify a sample covariance matrix. Yahya *et al*. (2011) proposed approach for feature selection in high dimensional data. For analysis of the one-sample high-dimensional repeated measures designs, see Choopradit and Chongcharoen (2011). However, the mainstream attempt in this study interests in the analysis of two-sample repeated measures designs for high-dimensional data. In an influential work, Ahmad (2008) proposed the modified version of the ANOVA-type statistic for analysis of two-sample repeated measures designs when the data are multivariate normal and the dimension can be large compared to the sample size using a modified Box's approximation (Box, 1954) based on quadratic and bilinear forms.

In this study, we introduce test statistics for testing interaction and time effects which follow a multivariate approach to repeated measures for analyzing high-dimensional two-sample repeated measures designs. The test statistics asymptotically follows a standard normal distribution and is not affected by even a large the dimension of repeated measurements, p. The test statistics are derived with no specific form of covariance matrix but under equal population covariance assumption.

The modified ANOVA-type statistic: Here and in the following, for any natural number p, I_p denotes the p×p identity matrix of dimension p, $1_p = (1,\ldots,1)^T_{p\times 1}$ a p-dimensional column vector of ones, $J_p = 1_p 1_p^T$ a square matrix of ones and $P_p = I_p - p^{-1} J_p$ is the centering matrix. Note that P_p is a projection matrix. Usually, the null hypothesis to be tested is H_0: $H\mu = 0$. The matrix H can be formulated in distinctive settings depending on the objectives of the experimental research. As similar null hypothesis given in Brunner *et al*. (1999) and Ahmad (2008), we can also write H_0: $G\mu = 0$ where $G = H^T (HH^T)^- H$ is the general hypothesis matrix whereas $(HH^T)^-$ denoting a generalized inverse of HH^T and G is a projection matrix. We note that $G\mu = 0$ if and only if $H\mu = 0$. For the two-sample repeated measures designs (Davis, 2002), the situation when repeated measurements at p time points are obtained from two independent groups of subjects is considered. Let $x_{ij} = (x_{ij1}, x_{ij2,\ldots,} x_{ijp})^T$ denote the vector of observations from the jth subject in group i for j = 1, 2, …, n_i, i = 1, 2. The matrix G can be formulated to test any appropriately hypotheses as $G^{AB} = P_2 \otimes P_p$ (interaction effect hypothesis), $G^A = P_2 \otimes \frac{1}{p}J_p$ (group effect hypothesis) and $G^B = \frac{1}{2}J_2 \otimes P_p$ (time effect hypothesis).

It is clear from the hypotheses above that the high dimensionality influences only interaction effect hypothesis and time effect hypothesis whereas group effect hypothesis is a univariate hypothesis of the means of two independently samples from normal populations and does not depend on p. Group effect hypothesis, therefore, is not of our main interest because it can be tested using the usual t-test for two independent samples (Rencher, 2002; Johnson and Wichern, 2002). When n<p, $\hat{\Sigma}$ is singular. Any test statistic involved the inversion of $\hat{\Sigma}$ will not exist. Ahmad (2008) proposed the modified ANOVA-type statistic under covariance matrices of two groups are the same, using quadratic and symmetric bilinear forms based on Box's approximation (Box, 1954) as followed.

The interaction effect: The hypothesis of no interaction effect, $H_0^{AB}:(P_2 \otimes P_p)\mu = 0$, can also be written as $H_0^{AB}: P_p(\mu_1 - \mu_2) = 0$. Therefore, the generating matrix without any loss of generality for the interaction effect is G = P_p. Let x_{1k} and x_{2l} defined above with $x_{1k} \sim N_p(\mu_1, \Sigma_1)$ and $x_{2l} \sim N_p(\mu_2, \Sigma_2)$ where the samples are drawn independently from two populations. Define the differences x_{1k} - x_{2l}, for all k = 1, 2, …, n_1, l = 1, 2, …, n_2, where $E(x_{1k} - x_{2l}) = \mu_1 - \mu_2$ and $Var(x_{1k} - x_{2l}) = \Sigma_1 + \Sigma_2$. Under $H_0^{AB}: G(\mu_1 - \mu_2) = 0$, $E[G(x_{1k} - x_{2l})] = E\left[G(\bar{x}_{1.} - \bar{x}_{2.})\right] = 0$, $Var[G(x_{1k} - x_{2l})] = G(\Sigma_1 + \Sigma_2)G$ and $Var\left[G(\bar{x}_{1.} - \bar{x}_{2.})\right] = G\left(\frac{1}{n_1}\Sigma_1 + \frac{1}{n_2}\Sigma_2\right)G$. Denote $Q_{kl}^{AB} = (x_{1k} - x_{2l})^T G (x_{1k} - x_{2l})$ and $Q_{rs}^{AB} = (x_{1r} - x_{2s})^T G (x_{1r} - x_{2s})$ are the quadratic forms, $Q_{klrs}^{AB} = (x_{1k} - x_{2l})^T G (x_{1r} - x_{2s})$ is the symmetric bilinear form. For assuming that the two groups have common covariance matrix Σ, that is $\Sigma_1 = \Sigma_2 = \Sigma$, the estimators:

$$\frac{1}{n_1 n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^{AB},\quad \frac{1}{n_1 n_2 (n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{kl}^{AB} Q_{rs}^{AB}$$

$$\text{and } \frac{1}{n_1 n_2 (n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{klrs}^{AB^2},$$

are unbiased and consistent estimators of 2tr(GΣG), $4[tr(G\Sigma G)]^2$ and $4tr(G\Sigma G)^2$, respectively. For testing H_0^{AB}, the modified ANOVA-type statistic, T_{MATS}^{AB}, is given by:

$$T_{MATS}^{AB} = \frac{2 n_1^2 n_2^2 \tilde{f}^{AB}}{N \sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^{AB}} (\bar{x}_{1.} - \bar{x}_{2.})^T G (\bar{x}_{1.} - \bar{x}_{2.})$$

where, N = $n_1 + n_2$. Under H_0^{AB} the modified ANOVA-type statistic T_{MATS}^{AB} has asymptotically a central $\chi^2_{f^{AB}}$ distribution with f^{AB} degree of freedom which is estimated by:

$$\tilde{f}^{AB} = \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{kl}^{AB} Q_{rs}^{AB} \Bigg/ \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{klrs}^{AB^2}$$

The time effect: The hypothesis of no time effect, $H_0^B: \left(\frac{1}{2}J_2 \otimes P_p\right)\mu = 0$, can also be written as $H_0^B: P_p(\mu_1 + \mu_2) = 0$. As similar in the interaction effect, therefore, the generating matrix without any loss of generality for the time effect is G = P_p. Let x_{1k} and x_{2l} defined above with $x_{1k} \sim N_p(\mu_1, \Sigma_1)$ and $x_{2l} \sim N_p(\mu_2, \Sigma_2)$ where the samples are drawn independently from two populations. Define the sums $x_{1k} + x_{2l}$, for all k = 1, 2, …, n_1, l = 1, 2, …, n_2, where $E(x_{1k} + x_{2l}) = \mu_1 + \mu_2$ and $Var(x_{1k} + x_{2l}) = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$. Under $H_0^B: G(\mu_1 + \mu_2) = 0$, $E[G(x_{1k} + x_{2l})] = E\left[G(\bar{x}_{1.} + \bar{x}_{2.})\right] = 0$, $Var[G(x_{1k} + x_{2l})] = G(\Sigma_1 + \Sigma_2)G$ and $Var\left[G(\bar{x}_{1.} + \bar{x}_{2.})\right] = G\left(\frac{1}{n_1}\Sigma_1 + \frac{1}{n_2}\Sigma_2\right)G$. Denote $Q_{kl}^{B} = (x_{1k} + x_{2l})^T G (x_{1k} + x_{2l})$ and $Q_{rs}^{B} = (x_{1r} + x_{2s})^T G (x_{1r} + x_{2s})$ are the quadratic forms, $Q_{klrs}^{B} = (x_{1k} + x_{2l})^T G (x_{1r} + x_{2s})$ is the symmetric bilinear form. For assuming that the two groups have common covariance matrix Σ, that is $\Sigma_1 = \Sigma_2 = \Sigma$, the estimators:

$$\frac{1}{n_1 n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^{B},\quad \frac{1}{n_1 n_2 (n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{kl}^{B} Q_{rs}^{B}$$

$$\text{and } \frac{1}{n_1 n_2 (n_1-1)(n_2-1)} \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{klrs}^{B^2},$$

are unbiased and consistent estimators of 2tr(GΣG), $4[tr(G\Sigma G)]^2$ and $4tr(G\Sigma G)^2$, respectively. For testing H_0^B, the modified ANOVA-type statistic, T_{MATS}^{B}, is given by:

$$T_{MATS}^{B} = \frac{2 n_1^2 n_2^2 \tilde{f}^{B}}{N \sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^{B}} (\bar{x}_{1.} + \bar{x}_{2.})^T G (\bar{x}_{1.} + \bar{x}_{2.})$$

where N = $n_1 + n_2$. Under H_0^B the modified ANOVA-type statistic T_{MATS}^{B} has asymptotically a central $\chi^2_{f^B}$ distribution with f^B degree of freedom which is estimated by:

$$\tilde{f}^{B} = \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{kl}^{B} Q_{rs}^{B} \Bigg/ \underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{klrs}^{B^2}.$$

METERIALS AND METHODS

A proposed test: When n<p, only the first n eigenvalues of $\hat{\Sigma}$ will be non-zero and the smallest

eigenvalues will tend to zero pretty quickly as the dimensionality grows. Consequently, many of the classical techniques, e.g., the Hotelling's T^2 test, encounter a mathematical barrier and becomes inapplicable since in the case $\hat{\Sigma}$ is degenerate. Hence, we look for tests which do not require the nonsingularity of $\hat{\Sigma}$.

We will derive the asymptotic distributions of the test statistics for two-sample repeated measures design when data are high dimension (n<p) under general conditions only that Σ is a positive definite covariance matrix (denoted as $\Sigma>0$). The test statistics considered under covariance matrices of two groups are equal and both p and n go to infinity. We should note that the limiting ratio p/n is allowed to be greater than one. Consequently, these tests can be used when n<p. From hypotheses presented as before the formulations of the test statistics for interaction effect involves differences of the vector x_{1k} and x_{2l}, but those for time effect involves sums of the vector x_{1k} and x_{2l}. Finally, the asymptotic distributions of these test statistics for the interaction effect and time effect are same since covariance of differences and sums remains the same. Hence, the details in the following section will be shown only for the interaction effect while for the time effect follows the same pattern.

The interaction effect: For testing the null hypothesis of no interaction effect as above we first consider the following result.

Theorem 1: For assuming $\Sigma_1 = \Sigma_2 = \Sigma$, let $x_{hj} \sim N_p(0,\Sigma)$, $\Sigma>0$, where j = 1, 2, …, n_h, h = 1, 2 and let G be a matrix defined the same as above. Then $\frac{n_1 n_2}{N}\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right)^T G\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right) \overset{d}{=} \sum_{i=1}^{p}\lambda_i C_i$, where $N = n_1 + n_2$, λ_i are the eigen values of $G\Sigma$ and $C_i \sim \chi_1^2$, i = 1, 2, …, p, are independent.

Proof. Under $H_0^{AB}: G(\mu_1 - \mu_2) = 0$, for assuming $\Sigma_1 = \Sigma_2 = \Sigma$, notice that $\sqrt{\frac{n_1 n_2}{N}} G\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right) \sim N_p\left(0, G\Sigma G\right)$. We set $\sqrt{\frac{n_1 n_2}{N}} G\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right) \equiv Q \sim N_p\left(0, G\Sigma G\right)$. Let v_1, v_2, …, v_p and λ_1, λ_2, …, λ_p are the corresponding orthonormal eigenvectors and eigenvalues of $G\Sigma$. It follows that:

$$Q = \sum_{i=1}^{p}\xi_i v_i$$

where, $\xi_i = Q^T v_i$, i = 1, 2, …, p, which are independent and we have $E(\xi_i) = 0$ and $Var(\xi_i) = \lambda_i > 0$. Since Q is a normal random vector, we have $\xi_i/\sqrt{\lambda_i}$ iid ~ N(0,1) for i = 1, 2, …, p.

Due to the orthonormality of the eigenvectors v_i, i = 1, 2, …, p, then we have:

$$Q^T Q = \sum_{i=1}^{p}\xi_i^2$$

Since $\xi_i/\sqrt{\lambda_i}$ iid ~ N(0,1), then $\xi_i^2/\lambda_i \sim \chi_1^2$. Hence $\xi_i^2 \overset{d}{=} \lambda_i C_i$, where $C_i \sim \chi_1^2$, i = 1, 2, …, p, are independent. It follows that $\sum_{i=1}^{p}\xi_i^2 \overset{d}{=} \sum_{i=1}^{p}\lambda_i C_i$, as desired.

Since the quantities tr($G\Sigma G$), tr$(G\Sigma G)^2$ and $\lambda_{max} = \max_{1\le i\le p}\lambda_i$ vary with p. To derive the asymptotic null distribution in the following theorem, we impose the following regular assumptions:

(A) $tr\left(G\Sigma G\right)^2/p \to c_1 \in (0,\infty)$ as $p \to \infty$;

(B) $\lambda_{max}/\sqrt{p} \to 0$ as $p \to \infty$;

(C) $n_1/N \to c_2 \in (0,1)$.

These assumptions are similar to those imposed by Bai and Saranadasa (1996) and Srivastava and Du (2008) for the study of their two-sample testing procedures and for the study of their testing mean vector in one- and two-sample cases, respectively.

The following theorem establishes the asymptotic normality of our test statistic for testing the hypothesis:

$$H_0^{AB}: G(\mu_1 - \mu_2) = 0$$

Theorem 2: For $n_1, n_2, p \to \infty$ and assuming $\Sigma_1=\Sigma_2=\Sigma$, let $x_{ij} \sim N_p(\mu_i,\Sigma)$, $\Sigma>0$, where j = 1, 2, …, n_i, i = 1, 2. Under $H_0^{AB}: G(\mu_1 - \mu_2) = 0$ and the assumption (A)-(C) are satisfied, then:

$$T_*^{AB} = \frac{\frac{2n_1 n_2}{N}\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right)^T G\left(\bar{x}_{1\cdot} - \bar{x}_{2\cdot}\right) - \frac{1}{n_1 n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^{AB}}{\sqrt{\frac{2}{n_1 n_2 (n_1-1)(n_2-1)}\underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\ne r,\ l\ne s} Q_{klrs}^{AB^2}}} \to N(0,1)$$

in distribution.

Proof: For assuming $\Sigma_1=\Sigma_2=\Sigma$, then $\bar{x}_{1\cdot} - \bar{x}_{2\cdot} \sim N_p\left(\mu_1 - \mu_2, \frac{N}{n_1 n_2}\Sigma\right)$. Under the null hypothesis

$H_0^{AB}: G(\mu_1-\mu_2)=0$ and G is a projection matrix, obviously, $E\left[\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})\right]=2tr(G\Sigma G)$. Using cumulant generating function of quadratic and bilinear forms (Searle, 1997) and moment-cumulant relationships (Rao, 2001) we obtained $E\left[\frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}Q_{kl}^{AB}\right]=2tr(G\Sigma G)$ and $Var\left[\frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}Q_{kl}^{AB}\right]=\frac{2(N+2)}{n_1n_2}tr(G\Sigma G)^2$. Consider the statistic:

$$R_n=\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})-\frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}Q_{kl}^{AB}$$

Therefore, clearly under H_0^{AB} we have $E(R_n) = 0$. Ahmad (2008) [Theorem 3.5] shown that $Var\left[\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})\right]=8tr(G\Sigma G)^2$ and $Cov\left[\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot}),\frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}Q_{kl}^{AB}\right]=\frac{4N}{n_1n_2}tr(G\Sigma G)^2$.

It follows that $Var(R_n)\simeq 8tr(G\Sigma G)^2$ for large n_1, n_2. One can see that $\left[\frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}Q_{kl}^{AB}-2tr(G\Sigma G)\right]\Big/\sqrt{8tr(G\Sigma G)^2}\to 0$. Therefore, we need only show that

$$\left\{\begin{array}{l}\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})-E\\ \left[\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})\right]\end{array}\right\}\Big/\sqrt{8tr(G\Sigma G)^2}\to N(0,1)$$

in distribution.

Recall Theorem 1 since $\lambda_i C_i \overset{d}{=} \lambda_i z_i^2, z_i \text{ iid} \sim N(0,1)$ and note that $tr(G\Sigma G)=tr(G\Sigma)=\sum_{i=1}^p\lambda_i$ we may rewrite:

$$\frac{\frac{2n_1n_2}{N}(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}-\bar{x}_{2\cdot})-2tr(G\Sigma G)}{\sqrt{8tr(G\Sigma G)^2}} \overset{d}{=} \frac{\sum_{i=1}^p\lambda_i(z_i^2-1)}{\sqrt{2tr(G\Sigma G)^2}} \qquad (A.1)$$

Now we show that (A.1) is asymptotically distributed as N (0, 1) by using Lyapounov's theorem. We usually note that $tr(G\Sigma G)^2=\sum_{i=1}^p\lambda_i^2$. Let $S=\sum_{i=1}^p\lambda_i(z_i^2-1)$. Then $E(S) = 0$ and $Var(S) = 2tr(G\Sigma G)^2$. Therefore, as $n_1,n_2,p\to\infty$ and under assumption (A)-(C) are satisfied, it then follows that:

$$\frac{E\sum_{i=1}^p\left|\lambda_i(z_i^2-1)\right|^3}{\left(\sqrt{2tr(G\Sigma G)^2}\right)^3}=\frac{\sum_{i=1}^p\lambda_i^3E\left|z_1^2-1\right|^3}{\left(\sqrt{2tr(G\Sigma G)^2}\right)^3} \le \frac{\left(\lambda_{max}/\sqrt{p}\right)E\left|z_1^2-1\right|^3}{2^{3/2}\sqrt{tr(G\Sigma G)^2/p}}\to 0.$$

And Lyapounov's condition is satisfied. By Lyapounov's Central Limit Theorem, the expression (A.1) tends to N(0,1) in distribution.

To complete the construction of our test statistic, we need only find a ratio-consistent estimator of 4tr $(G\Sigma G)^2$. A natural estimator of $4tr(G\Sigma G)^2$ is $4tr(G\hat{\Sigma}G)^2$. However, $4tr(G\hat{\Sigma}G)^2$ is generally neither unbiased nor ratio-consistent. It is usual to verify that $\frac{1}{n_1n_2(n_1-1)(n_2-1)}\underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s}Q_{klrs}^{AB^2}$, is an unbiased and ratio-consistent estimator of 4tr $(G\Sigma G)^2$. Hence the theorem is proved.

Due to Theorem 2 the test with an α level of significance will rejects H_0^{AB} if $T_*^{AB}>z_{1-\alpha}$ where $z_{1-\alpha}$ is the $100(1-\alpha)\%$ quantile of N(0,1). Note that, the approximating asymptotic null distribution of T_*^{AB} is similarly to the test obtained in Bai and Saranadasa (1996); Srivastava and Du (2008) and Zhang and Xu (2009) when both p and n go to infinity.

The time effect: We consider testing the null hypothesis of no time effect as above. As mentioned above, the proof of the following theorem can be obtained similar that of Theorem 1, thus we present it without proof.

Theorem 3: For assuming $\Sigma_1=\Sigma_2=\Sigma$, let $x_{hj}\sim N_p(0,\Sigma)$, $\Sigma>0$, where j = 1, 2, …, n_h, h = 1, 2 and let G be a matrix defined the same as above. Then $\frac{n_1n_2}{N}(\bar{x}_{1\cdot}+\bar{x}_{2\cdot})^T G(\bar{x}_{1\cdot}+\bar{x}_{2\cdot})\overset{d}{=}\sum_{i=1}^p\lambda_iC_i$, where $N = n_1 + n_2$, λ_i are the eigenvalues of $G\Sigma$ and $C_i\sim\chi_1^2$, i = 1, 2, …, p, are independent.

The following theorem establishes the asymptotic normality of our test statistic for testing the hypothesis $H_0^B : G(\mu_1 + \mu_2) = 0$.

Theorem 4: For $n_1, n_2, p \to \infty$ and assuming $\Sigma_1=\Sigma_2=\Sigma$, let $x_{ij} \sim N_p(\mu_i,\Sigma)$, $\Sigma>0$, where j = 1, 2, …, n_i, i = 1, 2. Under $H_0^B : G(\mu_1 + \mu_2) = 0$ and the assumption (A)-(C) are satisfied, then:

$$T_*^B = \frac{\frac{2n_1n_2}{N}\left(\bar{x}_{1\cdot} + \bar{x}_{2\cdot}\right)^T G\left(\bar{x}_{1\cdot} + \bar{x}_{2\cdot}\right) - \frac{1}{n_1n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} Q_{kl}^B}{\sqrt{\frac{2}{n_1n_2(n_1-1)(n_2-1)}\underbrace{\sum_{k=1}^{n_1}\sum_{l=1}^{n_2}\sum_{r=1}^{n_1}\sum_{s=1}^{n_2}}_{k\neq r,\ l\neq s} Q_{klrs}^{B^2}}} \to N(0,1)$$

in distribution

The proof of Theorem 4 is similar to that of Theorem 2. Due to Theorem 4 the test with an α level of significance will rejects H_0^B if $T_*^B > z_{1-\alpha}$ where $z_{1-\alpha}$ is the $100(1-\alpha)\%$ quantile of N(0,1).

Simulation study: We assess the effectiveness of the proposed two-sample tests for high dimensional data by means of a Monte Carlo simulation study. Simulation results were obtained so as to assess the accuracy of the asymptotic standard normal distribution in approximating the actual null distributions of T_*^{AB} and T_*^B. If the distribution is derived correctly for the proposed test statistics, then we would expect that the estimated Type I errors should be close to the nominal significance level setting. We also estimate empirical powers of our proposed test statistics. We begin with the description of the parameter selection for our simulation.

Parameter selection: For Type I error simulation study which were designed to evaluate the performance of the proposed two-sample test for repeated measures designs with high-dimensional data, we used Monte Carlo technique for 5,000 iterations with setting a nominal significance level of $\alpha = 0.05$. We then took n_1, n_2 = 10, 15, 20, 30 and p = 30, 50, 70, 100. The upper and lower limits were calculated according to $0.05 \pm 3\sqrt{0.05\times 0.95/5000}$ = (0.041, 0.059), i.e., three standard errors around the nominal significance level of 0.05. Thus, any estimated Type I error rate falling within these limits is not significantly different from the nominal significance level of 0.05. The estimated type I error rates are given as in Table 1.

For the empirical power computations, we chose G as an appropriated testing matrix defined above, $\mu_1 = (0,0,\ldots,0)_{p\times 1}^T$ and for the nonzero $\mu_2 = \eta_i(u_1,u_2,\ldots,u_p)_{p\times 1}^T$ where $u_j = j/p$, j = 1, 2, …, p and η_i is the ith element of the vector of constants, η = 0(0.2) 1.4. The Table 2 and 3 give empirical powers of the proposed test T_*^{AB} (interaction effect) and of the proposed test T_*^B (time effect) for 5,000 iterations of simulations study with setting a nominal significance level of $\alpha = 0.05$, p = 50, 70, 100 and for sample sizes n_1 = 10, n_2 = 20 and n_1 = 20, n_2 = 30 respectively.

Here, the four different covariance patterns are considered: (a) simple (SIM) pattern, (b) compound symmetry (CS) pattern, (c) unstructured (UN) pattern and (d) heterogeneous compound symmetry (CSH) pattern. A simple covariance pattern is defined as $\Sigma = \sigma^2 I$, where I denotes the p×p dimensional identity matrix with $\sigma^2 > 0$. A compound symmetry covariance pattern is defined as $\Sigma = \sigma^2 I + \kappa J$, where $J = 11^T$ denotes the p×p dimensional matrix of 1s and κ is appropriate constant. The unstructured covariance pattern refers to the SAS PROC MIXED unstructured pattern (UN). And a heterogeneous compound symmetry covariance pattern is defined as $\Sigma = (\sigma_{ij})_{i,j=1}^p$, where $\sigma_{ij} = \sigma_i^2 > 0$ (i = j) and $\sigma_{ij} = \sigma_i\sigma_j\rho$ $(i \neq j)$, ρ is the correlation parameter satisfying $|\rho| < 1$. We set $\sigma^2 = 1$ for SIM, $\sigma^2 = \kappa = 1$ for CS, σ_{ij} iid ~ Unif(1,2) (if i = j) and $\rho_{ij} = (j-1)/p^2$ (if $i < j$) for UN and σ_{ij} iid ~ Unif(2,3) (if i = j), ρ = 0.5 for CSH, where i, j = 1, 2, …, p. The multivariate normal random vectors were generated using IMSL subroutine RNMVN of FORTRAN.

RESULTS AND DOSCUSSION

Simulation result: Table 1 show the closeness of estimated Type I errors with nominal significance level setting at α = 0.05. As can be seen from the Table 1, the estimated Type I errors of the test statistic T_*^{AB} for interaction effect and the test statistic T_*^B for time effect are close to the nominal significance level setting at α = 0.05 reasonably well in all cases considered including small sample size as 10. This is shown that the proposed tests are reasonable tests. Moreover, we note that the accuracy of Type I error control is not affected by changing the covariance pattern and by the increasing the dimension p.

Table 1: Simulated Type I error of T_*^{AB} and T_*^{B} under the null hypothesis for four different covariance matrix pattern applied at nominal significance level $\alpha = 0.05$

			Interaction effect				Time effect			
n_1	n_2	p	SIM	CS	UN	CSH	SIM	CS	UN	CSH
10	10	30	0.047	0.048	0.045	0.046	0.044	0.044	0.044	0.044
		50	0.046	0.045	0.046	0.045	0.047	0.046	0.049	0.048
		70	0.044	0.044	0.046	0.044	0.045	0.045	0.042	0.042
		100	0.043	0.041	0.041	0.041	0.046	0.043	0.046	0.047
10	20	30	0.047	0.046	0.050	0.050	0.051	0.052	0.052	0.052
		50	0.048	0.047	0.048	0.049	0.049	0.047	0.045	0.046
		70	0.050	0.050	0.049	0.052	0.048	0.047	0.048	0.047
		100	0.047	0.048	0.048	0.050	0.050	0.050	0.050	0.050
10	30	50	0.050	0.052	0.051	0.051	0.047	0.048	0.050	0.047
		70	0.048	0.047	0.048	0.047	0.051	0.053	0.052	0.053
		100	0.046	0.049	0.050	0.047	0.045	0.044	0.044	0.042
15	20	50	0.050	0.052	0.049	0.051	0.048	0.047	0.050	0.048
		70	0.052	0.052	0.052	0.053	0.047	0.047	0.046	0.047
		100	0.052	0.052	0.052	0.052	0.049	0.049	0.051	0.050
15	30	50	0.049	0.048	0.050	0.048	0.053	0.054	0.054	0.056
		70	0.048	0.048	0.047	0.049	0.047	0.050	0.045	0.047
		100	0.049	0.051	0.049	0.050	0.045	0.045	0.045	0.045
20	20	50	0.052	0.054	0.056	0.055	0.055	0.057	0.057	0.057
		70	0.050	0.050	0.053	0.052	0.048	0.047	0.049	0.049
		100	0.052	0.052	0.053	0.052	0.054	0.055	0.051	0.054
20	30	50	0.053	0.054	0.056	0.055	0.052	0.051	0.053	0.052
		70	0.052	0.050	0.052	0.050	0.052	0.052	0.050	0.053
		100	0.049	0.049	0.050	0.051	0.052	0.050	0.051	0.050

Table 2: Empirical powers of T_*^{AB} (interaction effect) under the alternative hypothesis for four different covariance matrix pattern applied at nominal significance level $\alpha = 0.05$

		$n_1 = 10, n_2 = 20$			$n_1 = 20, n_2 = 30$		
COV	η	p = 50	p = 70	p = 100	p = 50	p = 70	p = 100
SIM	0.0	0.048	0.050	0.047	0.053	0.052	0.049
	0.2	0.076	0.092	0.095	0.125	0.150	0.163
	0.4	0.232	0.282	0.372	0.508	0.617	0.732
	0.6	0.603	0.713	0.843	0.933	0.978	0.997
	0.8	0.917	0.973	0.996	0.999	1.000	1.000
	1.0	0.995	1.000	1.000	1.000	1.000	1.000
	1.2	1.000	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
CS	0.0	0.047	0.050	0.048	0.054	0.050	0.049
	0.2	0.079	0.089	0.094	0.127	0.147	0.156
	0.4	0.228	0.288	0.362	0.500	0.605	0.724
	0.6	0.595	0.721	0.833	0.936	0.978	0.996
	0.8	0.916	0.978	0.994	0.999	1.000	1.000
	1.0	0.996	1.000	1.000	1.000	1.000	1.000
	1.2	1.000	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
UN	0.0	0.048	0.049	0.048	0.056	0.052	0.050
	0.2	0.067	0.073	0.074	0.098	0.108	0.117
	0.4	0.148	0.179	0.221	0.311	0.388	0.467
	0.6	0.368	0.461	0.582	0.751	0.860	0.944
	0.8	0.708	0.830	0.925	0.975	0.997	1.000
	1.0	0.938	0.987	0.998	1.000	1.000	1.000
	1.2	0.995	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
CSH	0.0	0.049	0.052	0.050	0.055	0.050	0.051
	0.2	0.070	0.079	0.082	0.111	0.125	0.130
	0.4	0.175	0.218	0.272	0.383	0.468	0.574
	0.6	0.453	0.565	0.694	0.836	0.927	0.980
	0.8	0.804	0.906	0.971	0.993	1.000	1.000
	1.0	0.973	0.997	1.000	1.000	1.000	1.000
	1.2	0.999	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000

Table 3:Empirical powers of T_*^{AB} (time effect) under the alternative hypothesis for four different covariance matrix pattern applied at nominal significance level $\alpha = 0.05$

		$n_1 = 10, n_2 = 20$			$n_1 = 20, n_2 = 30$		
COV	η	p = 50	p = 70	p = 100	p = 50	p = 70	p = 100
SIM	0.0	0.049	0.048	0.050	0.052	0.052	0.052
	0.2	0.077	0.086	0.093	0.124	0.135	0.159
	0.4	0.249	0.286	0.351	0.497	0.598	0.715
	0.6	0.604	0.725	0.832	0.935	0.977	0.997
	0.8	0.918	0.972	0.995	0.999	1.000	1.000
	1.0	0.995	0.999	1.000	1.000	1.000	1.000
	1.2	1.000	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
CS	0.0	0.047	0.047	0.050	0.051	0.052	0.050
	0.2	0.080	0.085	0.095	0.121	0.137	0.160
	0.4	0.241	0.285	0.355	0.498	0.603	0.722
	0.6	0.609	0.714	0.826	0.932	0.975	0.996
	0.8	0.921	0.972	0.996	0.999	1.000	1.000
	1.0	0.996	1.000	1.000	1.000	1.000	1.000
	1.2	1.000	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
UN	0.0	0.045	0.048	0.050	0.053	0.050	0.051
	0.2	0.069	0.069	0.076	0.094	0.098	0.112
	0.4	0.156	0.176	0.217	0.304	0.380	0.467
	0.6	0.377	0.464	0.571	0.742	0.846	0.936
	0.8	0.717	0.833	0.921	0.979	0.997	1.000
	1.0	0.937	0.984	0.998	1.000	1.000	1.000
	1.2	0.996	0.999	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000
CSH	0.0	0.046	0.047	0.050	0.052	0.053	0.050
	0.2	0.074	0.076	0.081	0.106	0.111	0.130
	0.4	0.183	0.218	0.269	0.375	0.460	0.574
	0.6	0.463	0.558	0.678	0.835	0.918	0.977
	0.8	0.809	0.906	0.967	0.992	0.999	1.000
	1.0	0.972	0.994	1.000	1.000	1.000	1.000
	1.2	0.999	1.000	1.000	1.000	1.000	1.000
	1.4	1.000	1.000	1.000	1.000	1.000	1.000

The corresponding empirical power curves of the test statistic T_*^{AB} for the interaction effect with the aforementioned four different covariance patterns are summarized in Table 2 and Fig. 1. To make it clear, the four covariance patterns are listed separately. We observe that the empirical power functions of T_*^{AB} have a steep ascending slope and is very high for each sample size groups and all covariance patterns considered. The corresponding empirical power of T_*^{AB} test give good power when $\eta \geq 0.6$ for $n_1 = 10$ and $n_2 = 20$ and$\eta \geq 0.4$ for $n_1 = 20$ and $n_2 = 30$. As expected, the power of T_*^{AB} test increases for increasing both p and η when fixed the sample sizes for all cases considered as well. In addition, the power of T_*^{AB} test also increases for increasing the sample sizes when both p and η are fixed. Moreover, the empirical power is also unaffected by changing the covariance pattern. In a separate simulation where the test statistic T_*^{B} were considered for the time effect as displayed in Table 3 and Fig. 2, the results obtained similar as interaction effect which reported above.

Analysis of the body-weight of male Wistar rats data: The data and the experimental description for a motivating example to deal with the high-dimensional data is reported in Brunner *et al*. (2002). The body-weight of male Wistar rats was observed over a period of 22 weeks to assess the toxicity of a drug. A group of ten animals was given a placebo, while a second group of ten animals was given a high dose of the drug. The main question to be addressed is whether the body-weights of the two test groups differ in their evolution over time. For this data, we have $n_i = 10$, i = 1, 2 and p = 22.

We get test values of $T_*^{AB} = -0.6273$ with a corresponding p-value is 0.7348 and $T_*^{B} = 6.2078$ with a corresponding p-value < 0.0001. We conclude that a p-value highly significant for time effect but a p-value accepting the null hypothesis of no interaction effect.

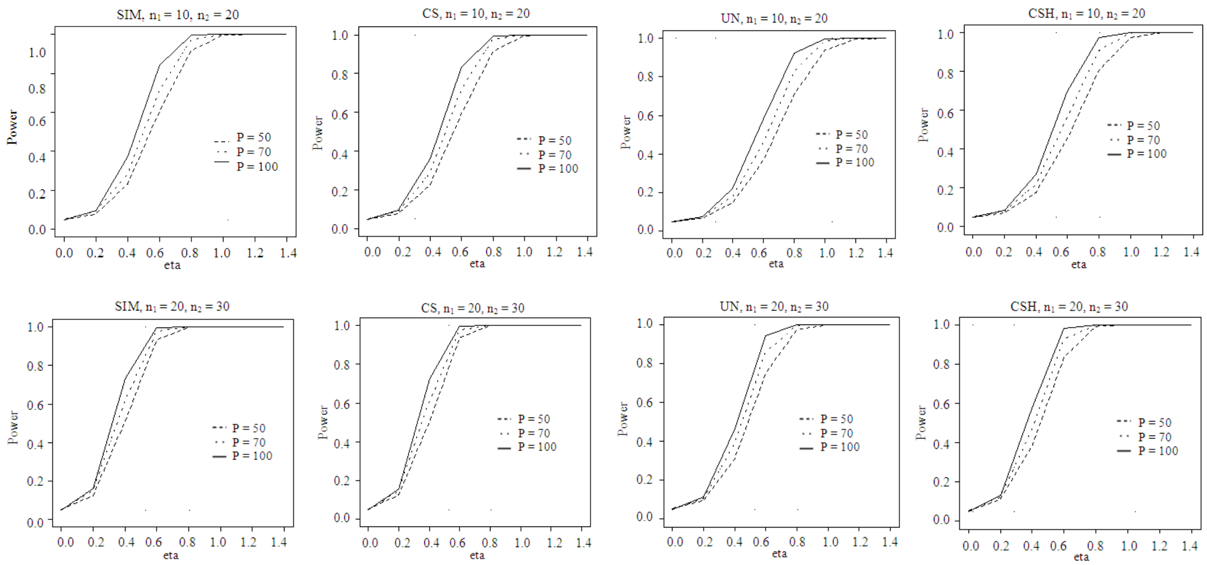

Fig. 1: Empirical power curves for T_*^{AB} (interaction effect) with n_1 = 10, n_2 = 20 and n_1 = 20, n_2 = 30

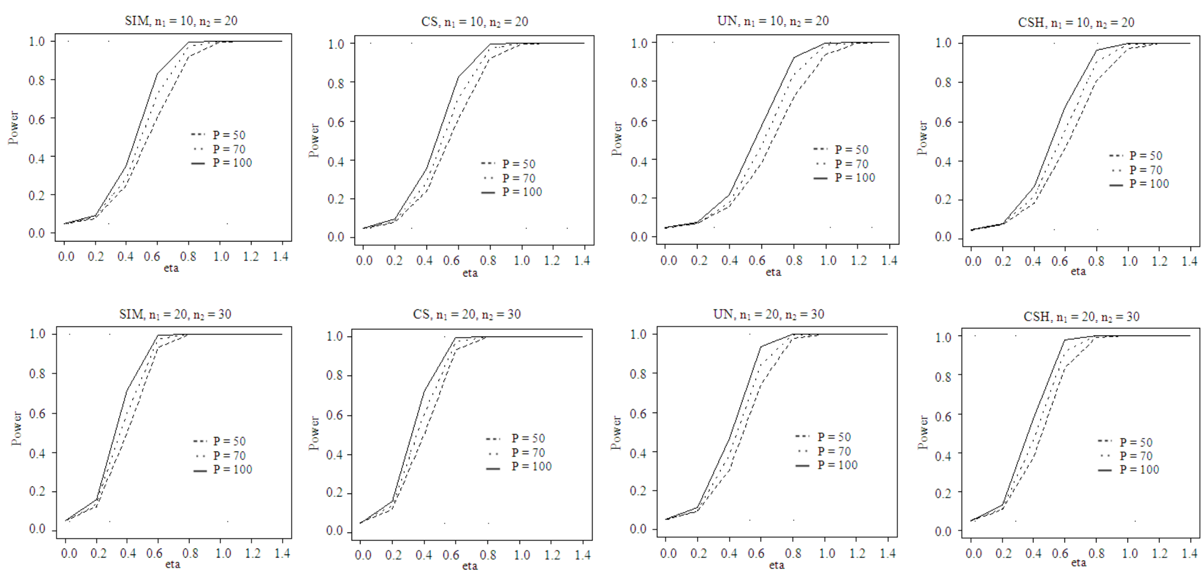

Fig. 2: Empirical power curves for T_*^{B} (time effect) with n_1 = 10, n_2 = 20 and n_1 = 20, n_2 = 30

CONCLUSION

In this study, we developed test statistics for analyzing high-dimensional two-sample repeated measures designs when the data are multivariate normal. We began by highlighting the previous work in the literature describing test for the hypothesis of no interaction and no time effects. We proposed test statistics T_*^{AB} for no interaction effect and T_*^{B} for no time effect which have an asymptotically distributed as standard normal distributions, under the null hypothesis, with common covariance Σ. One of the main advantages of these statistics is that they can be used for both unstructured and factorially structured repeated measures designs when the underlying hypothesis matrix G is appropriately defined. Monte Carlo simulation studies in this study show that the general behavior of these test statistics with asymptotic standard normal distribution still accurately control Type I error and have quite good power for any the covariance matrix pattern considered with a moderate sample size and any large dimension p. The strong support is provided in the simulation results. In our study, the power simulation results suggest that it may be assumed that the quality of the proposed test statistics, T_*^{AB} and T_*^{B}, are maintained even if the dimension is in the thousands and it may also be applied for the microarray data analysis.

ACKNOWLEDGEMENT

We are grateful to the Commission on Higher Education, Thailand, for financial support through a grant fund under the Strategic Scholarships Fellowships Frontier Research Networks.

REFERENCES

Ahmad, M.R., 2008. Analysis of High Dimensional Repeated Measures Designs: The One- and Two-Sample Test Statistics. Doctoral dissertation, Gerog-August University, Göttingen, Germany.

Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis. 3rd Edn. Wiley, New Jersey, ISBN: 9780471360919, pp: 721.

Arbabi, M., M. Sadeghi and C. Anyakora, 2009. Phenanthrene contaminated soil biotreatment using slurry phase bioreactor. Am. J. Environ. Sci., 5: 223-229. DOI: 10.3844/ajessp.2009.223.229

Bai, Z. and H. Saranadasa, 1996. Effect of high dimension: By an example of a two sample problem. Statistica Sinica, 6: 311-329.

Box, G.E.P., 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems, I: Effect of inequality of variance in the one-way classification. Ann. Math. Statist., 25: 290-302.

Brunner, E., S. Domhof and F. Langer, 2002. Nonparametric Analysis of Longitudinal Data in Factorial Experiments. 1st Edn., Wiley, New York, ISBN: 9780471441663, pp: 261.

Brunner, E., U. Munzel and M.L. Puri, 1999. Rank score tests in factorial designs with repeated measures. J. Multivariate Anal. 70: 286-317. DOI: 10.1006/jmva.1999.1821

Chongcharoen, S., 2011. Inversion of covariance matrix for high dimensional data. J. Math. Stat., 7: 227-229. DOI: 10.3844/jmssp.2011.227.229

Choopradit, B. and S. Chongcharoen, 2011. A test for one-sample repeated measures designs: Effect of high-dimensional data. J. Applied Sci., 11: 3285-3292. DOI: 10.3923/jas.2011.3285.3292

Davis, C.S, 2002. Statistical Methods for the Analysis of Repeated Measurements. 1st Edn., Springer, New York, ISBN: 9780387953700, pp: 415.

Dempster, A.P., 1958. A high dimensional two sample significance test. Ann. Math. Stat., 29: 995-1010.

Everitt, B.S., 1995. The analysis of repeated measures: A practical review with examples. Statistician, 44: 113-135.

Fujikoshi, Y., T. Himeno and H. Wakaki, 2004. Asymptotic results of a high dimensional MANOVA test and power comparison when the dimension is large. J. Japan Statis. Soc. 34: 19-26.

Johnson, R.A. and D.W. Wichern, 2002. Applied Multivariate Statistical Analysis. 5th Edn., Prentice Hall, Upper Saddle River, ISBN: 0130925535, pp: 767.

Keselman, H.J., J. Algina and R.K. Kowalchuk, 2001. The analysis of repeated measures designs: A review. Br. J. Math. Statist. Psychol., 54: 1-20.

Mosallanejad, B., R. Avizeh and H. Najafzadeh, 2011. Evaluation of prophylactic and therapeutic effects of silymarin on diazepam-induced hepatotoxicity in cats. Am. J. Applied Sci., 8: 848-853. DOI: 10.3844/ajassp.2011.848.853

Rao, C.R., 2001. Linear Statistical Inference and Its Applications. 2nd Edn. Wiley, New York, ISBN: 9780471218753, pp: 625.

Rencher, A.C., 2002. Methods of Multivariate Analysis. 2nd Edn. John Wiley and Sons, New York, ISBN: 9780471418894, pp: 708.

Schott, J.R., 2007. Some high dimensional tests for a one-way MANOVA. J. Multivariate Anal., 98: 1825-1839. DOI: 10.1016/j.jmva.2006.11.007

Searle, S.R., 1997. Linear Models. 3rd Edn., Wiley, New York, ISBN: 9780471184997, pp: 532.

Sharifi, D., D. Kazemi and A. Veshkini, 2008. Ultrasonographic evaluation of transcutaneous electrical neural stimulation on the repair of severed superficial digital flexor tendon in horses. Am. J. Anim. Vet. Sci., 3: 73-77. DOI: 10.3844/ajavsp.2008.73.77

Srivastava, M.S. and M. Du, 2008. A test for mean vector with fewer observations than the dimension. J. Multivariate Anal., 99: 386-402. DOI: 10.1016/j.jmva.2006.11.002

Srivastava, M.S. and Y. Fujikoshi, 2006. Multivariate analysis of variance with fewer observations than the dimension. J. Multivariate Anal., 97: 1927-1940. DOI: 10.1016/j.jmva.2005.08.010

Srivastava, M.S., 2007. Multivariate theory for analyzing high dimensional data. J. Japan Statist. Soc., 37: 53-86.

Srivastava, M.S., 2009. A test for the mean vector with fewer observations than the dimension under non-normality. J. Multivariate Anal., 100: 518-532. DOI: 10.1016/j.jmva.2008.06.006

Tu, C. and W.Y. Koh, 2010. Using bioequivalence approach to assess toxicity effect of herbal products, TyrelTM and RumbionTM. Am. J. Pharmacol. Toxicol., 5: 109-111. DOI: 10.3844/ajptsp.2010.109.111

Yahya, A.A., A. Osman, A.R. Ramli and A. Balola, 2011. Feature selection for high dimensional data: An evolutionary filter approach. J. Comput. Sci., 7: 800-820. DOI: 10.3844/jcssp.2011.800.820

Zangiabadi, N. and M.N. Ahrari, 2005. Electro-diagnostic and clinical changes of peripheral neuropathies in bam earthquake victims. Am. J. Environ. Sci., 1: 206-208. DOI: 10.3844/ajessp.2005.206.208

Zhang, J.T. and J.F. Xu, 2009. On the k-sample Behrens-Fisher problem for high-dimensional data. Sci. China Ser. A., 52: 1285-1304. DOI: 10.1007/s11425-009-0091-x