DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 15 | Issue 2 Article 19

11-1-2016

Comparison of Some Multivariate Nonparametric Tests in Profile Analysis to Repeated Measurements

Mehrdad Vossoughi
Shiraz University of Medical Sciences, Shiraz, Iran, vossoughim@sums.ac.ir

Shila Shahvali
Fars Planning and Budget Organization, Shiraz, Iran

Erfan Sadeghi
Isfahan University of Medical Sciences, Isfahan, Iran

Recommended Citation
Vossoughi, Mehrdad; Shahvali, Shila; and Sadeghi, Erfan (2016) "Comparison of Some Multivariate Nonparametric Tests in Profile Analysis to Repeated Measurements," *Journal of Modern Applied Statistical Methods*: Vol. 15: Iss. 2, Article 19.
DOI: 10.22237/jmasm/1478002620
Available at: http://digitalcommons.wayne.edu/jmasm/vol15/iss2/19

Comparison of Some Multivariate Nonparametric Tests in Profile Analysis to Repeated Measurements

Cover Page Footnote
Acknowledgements: The authors thank Dr. Shahram Hamedani of the Center for Research Improvement of the School of Dentistry for improving the use of English in the manuscript.

Journal of Modern Applied Statistical Methods
November 2016, Vol. 15, No. 2, 273-292.
doi: 10.22237/jmasm/1478002620

ISSN 1538 − 9472

Comparison of Some Multivariate Nonparametric Tests in Profile Analysis to Repeated Measurements

Mehrdad Vossoughi
Shiraz University of Medical Sciences
Shiraz, Iran

Shila Shahvali
Bureau of Statistics and Information, Fars Planning and Budget Organization
Shiraz, Iran

Erfan Sadeghi
Isfahan University of Medical Sciences
Isfahan, Iran

Through Monte Carlo simulations, the performance of six multivariate nonparametric tests for testing the hypothesis of parallelism in profile analysis was studied. In conclusion, the tests based on ranks were as efficient as Hotelling's T^2 under multivariate normal distribution. For the heavy tailed distribution, the tests based on signs performed best.

Keywords: Monte Carlo simulation, multivariate, nonparametric, profile analysis, heavy tailed

Introduction

Research in many areas of application frequently involves repeated measurements in which response from each experimental unit is measured repeatedly over different occasions such as time points. The linear mixed model to repeated measurements (Laird & Ware, 1982; Ware, 1985) was developed to analyze incomplete and unbalanced data. However, the performance of this complex approach is highly sensitive to the choice of model for mean function and correlation structure for errors (Littell, Pendergast, & Natarajan, 2000; Park, Park, & Davis, 2001; Vossoughi, Ayatollahi, Towhidi, & Ketabchi, 2012). Although several nonparametric methods have been developed for non-normal responses (Azzalini & Bowman, 1991; Singer, Poleto, & Rosa, 2004; Wernecke & Kalb, 1999; Wernecke & Kaufmann, 2000), model building and software implementation of these methods are extremely complicated.

Due to these difficulties, investigators are often interested in using the traditional approaches especially when the circumstances are controlled for

Dr. Vossoughi is an Assistant Professor in the Department of Dental Public Health. Email him at: vossoughim@sums.ac.ir.

obtaining complete data. In this context, the profile analysis method using MANOVA tests makes no assumption regarding the correlation structure and trend of mean model and hence is widely used. Nevertheless, the MANOVA tests perform poorly when the distribution of errors much deviates from multivariate normal (Davis, 1980, 1982; Everitt, 1979; Olson, 1974; Um & Randles, 1998).

Bhapkar (1984) and Sen (1984) discussed asymptotically distribution-free analogous of profile analysis. Multivariate extensions of Kruskal-Wallis and Brown-Mood median tests based on marginal ranks and signs were discussed in Puri and Sen (1971) but suffer from a lack of invariance with respect to affine transformations. Several authors provided detailed descriptions of affine invariant and non-invariant competitors based on spatial signs and ranks (Hettmansperger, Möttönen & Oja, 1998; Hettmansperger & Oja, 1994; Möttönen & Oja, 1995; Oja, 1999; Oja & Randles, 2004). The asymptotic efficiency of multivariate spatial sign and rank tests were studied by Möttönen, Oja, and Tienari (1997), Möttönen, Hettmansperger, Oja, and Tienari (1998), Nordhausen, Oja, and Tyler (2006) and Oja and Randles (2004). The theory and software implementation of affine invariant/non-invariant spatial sign and rank tests were well described by Oja (2010).

The aim of this study is to compare the performance of six nonparametric multivariate multi-sample tests with Hotelling's T^2in profile analysis for repeated measurements. For this propose, Monte Carlo simulations based on broad spectrum of scenarios are used to study the empirical type I error rates and powers of the tests in testing the hypothesis of parallelism. Affine/non-affine invariant multivariate generalizations of multi-sample tests are compared based on spatial scores discussed in Oja (2010, Ch. 11) and multivariate generalization of multi-sample tests based on marginal scores discussed in Chapter 5 of Puri and Sen (1971).

Although the test of group main effect or hypothesis that the two groups are at the same level can also be assessed using multivariate multi-sample procedures, it was not included in the simulations for three priori reasons. First, rather than testing the general multivariate hypothesis $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_k$ to assess group main effect, summarizing the response vector of each subjects using its individual mean and then applying univariate tests is generally implemented in a parametric profile analysis (Davis, 2002; Rencher, 1995). Second, the performance of Hotelling's T^2 and its nonparametric counterparts were studied to test above general hypothesis (Möttönen et al., 1998; Nordhausen et al., 2006; Um & Randles, 1998). Finally, group main effect has no direct interpretation in the presence of significant interaction and hence is not the primary hypothesis of interest in profile analysis.

Although the Monte Carlo comparison of methods for the analysis of repeated measurements has been an active area of research (Bhapkar & Patterson, 1978; Marcucci, 1986; Mendoza, Toothaker, & Nicewander, 1974; Park et al., 2001; Schwertman, Flynn, Stein, & Schenk, 1985; Schwertman, Fridshal, & Magrey, 1981), this study has been designed to examine some different aspects. First, the performances of recent nonparametric tests based on spatial signs and ranks considered here have not yet been studied in the area of profile analysis. Second, the effect of various correlation structures for errors has not included by most of the previous literature on this subject. Finally, the performance of the non-invariant tests under various transformation matrices widely used in the profile analysis are examined.

Methodology

Parametric profile analysis

The structure of profile analysis for the analysis of repeated measurements is now considered. Suppose that repeated measurements have been taken from k groups of subjects at p occasions. Let $\mathbf{y}_{ij} = (y_{ij1},\ldots, y_{ijp})^{\mathrm{T}}$ represent the response vector from the j^{th} subject in group i for $j = 1,\ldots, n_k$, $i = 1,\ldots, k$. The profile analysis model is

$$\mathbf{y}_{ij} = \boldsymbol{\mu}_i + \varepsilon_{ij}, \tag{1}$$

where the vector $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1},\ldots, \varepsilon_{ijp})^T$ is the vector of errors for the j^{th} subject in group I and $\boldsymbol{\mu}_i = (\mu_{i1},\ldots, \mu_{ip})^T$ is the population mean vector for the i^{th} group. Error vectors are assumed to be independent and normally distributed with mean vector $\mathbf{0}$ and common covariance matrix $\boldsymbol{\Sigma}$.

Arguably, in the presence of group × occasion interaction, the tests of main effects are confounded. Therefore, the primary aim in the profile analysis is to test the hypothesis of parallelism of k group profiles. The test of the hypothesis can be constructed as

$$\mathrm{H}_0 = \mathbf{C}\boldsymbol{\mu}_1 = \ldots = \mathbf{C}\boldsymbol{\mu}_k \text{ or } \boldsymbol{\mu}_1^* = \ldots = \boldsymbol{\mu}_k^*, \tag{2}$$

where $\boldsymbol{\mu}_h^*$ is the mean of transformed observations, $\mathbf{y}_{ij}^* = \mathbf{C}\mathbf{y}_{ij}$. Here, $\mathbf{C}$ is a p-1 × p transformation matrix with rank p-1 satisfying $\mathbf{C}\mathbf{1} = \mathbf{0}$, where $\mathbf{1}$ is the unit matrix. For instance, when p = 3, three widely-used matrices are:

$\mathbf{C}_1$: Mean difference	$\mathbf{C}_2$: Adjacent difference	$\mathbf{C}_3$: Last-value difference
$\frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$

For example, the analogous hypothesis of parallelism for $k = 2$ and the transformation matrix $\mathbf{C}_2$ is

$$H_0 : \begin{pmatrix} \mu_{12} - \mu_{11} \\ \mu_{13} - \mu_{12} \\ \vdots \\ \mu_{1p} - \mu_{1,p-1} \end{pmatrix} = \begin{pmatrix} \mu_{22} - \mu_{21} \\ \mu_{23} - \mu_{22} \\ \vdots \\ \mu_{2p} - \mu_{2,p-1} \end{pmatrix}. \tag{3}$$

Then, one-way multivariate analysis of variance (MANOVA) test statistics such as Wilk's Λ (if $k > 2$) or Hotelling's T^2 (if $k = 2$) can be used to assess the equality of mean vectors of transformed variables $\mathbf{y}_{ij}^*$ or equivalently hypothesis of parallelism. Similarly, nonparametric multivariate tests can be applied on the transformed observations to assess the equality of population locations when the underlying distribution deviates from normality.

Nonparametric counterparts of MANOVA tests

A brief overview of six nonparametric multivariate multi-sample tests used for profile analysis in the Monte Carlo simulations are now considered. The focus is primarily on recent methods that are supplied in standard statistical software packages. Here, we assume the p-dimensional data vectors are generated independently using model

$$\mathbf{y}_{ij} = \boldsymbol{\theta}_i + \boldsymbol{\varepsilon}_{ij}, \tag{4}$$

where $\boldsymbol{\theta}_i$ denotes the p-dimensional location vector for group i which is not necessary the corresponding mean vector and $\boldsymbol{\varepsilon}_{ij}$ is the vector of errors from an elliptical multivariate distribution with location vector $\mathbf{0}$ and scatter matrix $\boldsymbol{\Sigma}$. When measurements are not normally distributed, nonparametric multi-sample multivariate tests can be employed to test the hypothesis of no group × occasions interaction effect as

$$H_0 = \boldsymbol{\theta}_1^* = \ldots = \boldsymbol{\theta}_k^*, \tag{5}$$

where $\boldsymbol{\theta}_i^*$ indicates the location vector of transformed variables from group i.

Tests based on spatial signs

The test statistic based on spatial signs for testing H_0 is

$$Q = \sum_{i=1}^{c} \left\{ n_i \bar{\mathbf{U}}_i^{*\prime} \bar{\mathbf{U}}_i^* \right\} \tag{6}$$

where $\bar{U}_i^*$ denotes the sample mean vector of spatial signs transformed using inner centering and outer standardization. Although the test is location invariant, it is not affine invariant; that is the condition $Q(\mathbf{AY}) = Q(\mathbf{Y})$ is not satisfied for every nonsingular matrix $\mathbf{A}$ with rank p.

The affine invariant test statistic is

$$Q = p \sum_{i=1}^{c} \left\{ n_i \bar{\mathbf{U}}_i^{*\prime} \bar{\mathbf{U}}_i^* \right\} \tag{7}$$

where, here, $\bar{\mathbf{U}}_i^*$ is the sample mean vector of spatial signs transformed using inner centering and inner standardization.

The test statistics are multivariate generalizations of two- and several-sample Mood's median test and are asymptotically distributed as $\chi^2_{(c-1)p}$ when H_0 is true. The spatial sign tests are denoted by *SS* and *SSI* for the non invariant and invariant versions in the simulations, respectively. See Oja (2010) regarding the theory and software implementation of spatial sign and rank tests.

Tests based on spatial ranks

The constructions of tests based on spatial ranks are essentially the same as the spatial sign cases, with the difference that $\bar{\mathbf{U}}_i^*$'s are replaced by the corresponding sample mean vector of transformed spatial ranks, $\bar{\mathbf{R}}_i^*$. Due to the fact that the spatial ranks are naturally centered, one needs only to standardize them using outer or inner approaches to construct non affine or affine invariant versions of test statistic. The test statistics using outer and inner standardization are in the form of

$$Q = \sum_{i=1}^{c} \left\{ n_i \bar{\boldsymbol{R}}_i^{*\prime} \bar{\boldsymbol{R}}_i^* \right\} \tag{8}$$

and

$$Q = p \sum_{i=1}^{c} \left\{ n_i \bar{\boldsymbol{R}}_i^{*\prime} \bar{\boldsymbol{R}}_i^* \right\}, \tag{9}$$

respectively. The asymptotic null distribution of both test statistics is $\chi^2_{(c-1)p}$. The non invariant and affine invariant spatial rank tests are denoted by *SR* and *SRI* in the simulations, respectively.

Tests based on marginal ranks and signs

The multivariate multi-sample rank sum test compares the difference between the sample average rank vector $\bar{\mathbf{r}}_i$ and the combined-data average rank vector $\bar{\mathbf{r}}_.$ as

$$L_R = \sum_{i=1}^{c} n_i \left(\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_. \right)' \mathbf{V}^{-1} \left(\bar{\mathbf{r}}_i - \bar{\mathbf{r}}_. \right). \tag{10}$$

The test reduces to the Kruskal-Wallis test when $p = 1$ and to Wilcoxon-Mann-Whitney test when $p = 1$ and $c = 2$.

The multivariate multi-sample median test uses the corresponding average vectors based on sample signs (computed regarding combined-data median vector) to test the null hypothesis as

$$L_S = \sum_{i=1}^{c} n_i \left(\bar{\mathbf{s}}_i - \bar{\mathbf{s}}_. \right)' \mathbf{V}^{-1} \left(\bar{\mathbf{s}}_i - \bar{\mathbf{s}}_. \right). \tag{11}$$

Write $\mathbf{V}$ to denote the sample covariance matrix of marginal ranks and signs in L_R and L_S, respectively. The asymptotic null distribution of both statistics is

$\chi^2_{(c-1)p}$. The multivariate multi-sample location tests based on the vector of marginal signs and ranks were discussed in detail by Puri and Sen (1971).

The marginal sign and rank tests are denoted by *MS* and *MR*, respectively, in the simulation.

Simulation study

The structure of a Monte Carlo study used to investigate the performances of tests according to empirical type I error rates and powers is now discussed. The profile model (4) with two groups ($k = 2$), number of measurements $p = 4$, 8 and sample sizes $n = 10, 20$ and 30 for each of the two samples was considered. The performances of MANOVA test (here Hotelling's T^2 since $k = 2$) and the six nonparametric counterparts in testing the hypothesis of parallelism were compared under various scenarios. In the simulations, Hotelling's T^2 test was denoted by T^2.

Consider three types of correlation structures for errors; compound symmetry (CS) with $\rho = 0.2$, first-order autoregressive (AR1) with $\rho = 0.5$, and an unstructured model (UN). The UN structure considered here was an arbitrary $p \times p$ correlation matrix producing a positive definite covariance matrix. Errors were generated from multivariate t with 3 degrees of freedom (denoted by t (3)) as a heavy-tailed distribution and multivariate normal distribution with mean vector $\mathbf{0}$ and variances 3 for above correlation structures. Therefore the two distributions had the same mean vector and covariance matrix and differ only by degrees of heaviness of their tails. The MANOVA tests have been shown to have low powers when the underlying distribution is heavy-tailed, in particular (see e.g. (Somorčík, 2006). The reason is that the sample mean vector and covariance matrix would not provide proper estimates of location and variation under the presence of outliers (see, e.g. Um & Randles, 1998).

Throughout the simulations, $\boldsymbol{\theta}_1$ was considered to be a zero vector. To compute the empirical type I error rates, data were simulated under the hypothesis of parallelism, $H_0 : \mathbf{C}\boldsymbol{\theta}_1 = \mathbf{C}\boldsymbol{\theta}_2$, when $\boldsymbol{\theta}_2$ was also considered to be a zero vector. However, the hypothesis of interaction or $H_1 : \mathbf{C}\boldsymbol{\theta}_1 \neq \mathbf{C}\boldsymbol{\theta}_2$ was simulated when $\boldsymbol{\theta}_2 = (0, 1, 1, 0)^T$ and $(0, 0, 1, 1, 1, 1, 0, 0)^T$ for $p = 4$ and 8, respectively; so that the empirical powers were computed. Also considered are the three transformation matrices $\mathbf{C}_1$ to $\mathbf{C}_3$ presented above to evaluate the robustness of non affine-invariant tests.

For each combination of above scenarios, 1000 replications were carried out and significance level was considered to be 0.05. All simulations performed using

R 3.0.1 (R Development Core Team, 2013). In this respect, the multi-sample tests implemented using the R packages MNM, ICSNP, and Hotelling. Multivariate normal and *t* data were generated using the R packages MASS and mvtnorm, respectively.

Results

Displayed in Tables 1 and 2 are empirical type I error rates of the tests for errors generated from multivariate normal and multivariate $t(3)$ distributions, respectively. Each value is the proportion of 1000 replications for which the hypothesis of parallelism or null hypothesis was incorrectly rejected. In general, all tests preserved the nominal 5 percent level under all scenarios. However, for $p = 8$ and smaller sample size $n = 10$, the type I error rates of nonparametric tests were smaller than those of parametric one.

Displayed in Table 3 are empirical powers of the test for multivariate normal distribution. Each power value computed as the proportion of 1000 replications for which the hypothesis of parallelism was correctly rejected. In summary, among the tests, the affine invariant and non-invariant tests based on spatial ranks as well as test based on marginal ranks reached a power level fully close to that of Hotelling T^2 in which the differences were considerably negligible for all correlation structures. However, for the smaller sample size $n = 10$ and larger number of replication $p = 8$, the amount of difference somewhat increased. The test based on marginal signs performed unsatisfactorily; that is its powers were much lower than those of other test statistics for all correlation structures and transformation matrices. Interestingly, for all transformation matrices, the competitor based on spatial signs dominated the test based on marginal sign and was comparable to the best tests in the multivariate normal case. The empirical power trends of tests for multivariate normal distribution are visualized in Figure 1.

Shown in Table 4 are empirical powers of the test for data generated from multivariate $t(3)$ as a heavy tailed distribution. The results showed that the tests based on spatial signs and ranks and tests based on marginal ranks fully dominated Hotelling's T^2 for larger sample sizes $n = 20$ and 30 and any given correlation structure. For a fixed sample size, the amount of superiority somewhat decreased as p increased. In summary, the tests based on spatial signs yielded the greater values than the counterparts based on spatial and marginal ranks. Note that for a fixed p, the larger the size of sample, the greater the amount of difference in power levels. However, the performance of marginal sign test

Table 1. The empirical type I error rates of tests under multivariate normal distribution.

			Correlation structure								
			CS			AR1			UN		
	Matrix **C**	Test	10	20	30	10	20	30	10	20	30
$p = 4$		T^2	047	044	048	061	047	040	044	049	045
		SRI	046	050	051	057	047	041	045	050	046
		SSI	044	040	044	053	045	043	045	052	044
	$\mathbf{C}_1$:	*SR*	045	048	050	056	046	041	047	049	043
		SS	047	040	045	060	045	044	055	047	037
		MR	048	045	049	048	037	038	054	053	039
		MS	038	046	041	048	037	041	035	040	044
	$\mathbf{C}_2$:	*SR*	046	046	052	059	048	040	045	047	041
		SS	049	047	045	053	045	041	052	049	036
		MR	051	047	045	050	046	043	046	047	049
		MS	040	041	044	055	040	038	042	050	049
	$\mathbf{C}_3$:	*SR*	047	046	050	054	050	042	044	053	047
		SS	047	038	043	050	045	043	046	046	043
		MR	042	040	047	057	048	038	043	040	046
		MS	042	035	040	049	036	036	043	047	049
$p = 8$		T^2	050	051	050	047	052	050	036	054	042
		SRI	018	043	042	016	042	043	015	042	037
		SSI	022	043	045	019	047	046	017	041	035
	$\mathbf{C}_1$:	*SR*	023	045	045	016	040	043	015	043	040
		SS	020	046	046	017	044	044	015	044	034
		MR	015	038	048	017	033	045	015	046	043
		MS	021	045	044	018	036	047	017	044	038
	$\mathbf{C}_2$:	*SR*	021	045	041	016	042	042	016	040	040
		SS	023	046	045	015	045	040	016	040	042
		MR	020	040	048	023	045	047	014	033	039
		MS	023	042	038	016	031	036	011	038	036
	$\mathbf{C}_3$:	*SR*	023	046	048	017	045	045	019	037	038
		SS	020	040	043	016	043	045	012	040	038
		MR	020	040	042	018	032	042	015	041	041
		MS	020	044	040	017	039	037	014	036	033

*Note: The entries within table correspond to empirical type I error rates multiplied by 1000.

was unsatisfactory since it was just as efficient as Hotelling T^2 for some specific choices of **C**. Surprisingly; even the permutation procedure provided no additional gain in efficiency for Hotelling's T^2 under the heavy tailed distribution and hence not reported here.

Table 2. The empirical type I error rates of tests under multivariate *t* distribution.

			Correlation structure								
			CS			AR1			UN		
	Matrix **C**	Test	10	20	30	10	20	30	10	20	30
$p = 4$		T^2	040	036	043	034	045	047	047	046	038
		SRI	051	049	050	048	056	047	057	054	046
		SSI	047	054	051	044	055	047	046	051	061
	$\mathbf{C}_1$:	*SR*	052	046	056	046	056	050	054	052	050
		SS	051	049	060	048	052	050	047	053	054
		MR	046	043	049	047	046	050	048	054	041
		MS	050	049	049	046	057	057	046	051	039
	$\mathbf{C}_2$:	*SR*	048	048	051	047	055	048	053	053	051
		SS	045	052	050	050	052	046	050	054	057
		MR	048	048	055	043	047	050	052	056	052
		MS	049	044	049	040	048	062	050	045	043
	$\mathbf{C}_3$:	*SR*	046	050	046	048	057	050	054	055	050
		SS	048	058	044	053	051	042	044	054	062
		MR	042	055	051	042	047	047	050	055	055
		MS	043	047	046	049	049	041	047	055	059
$p = 8$		T^2	044	033	033	037	034	051	034	028	044
		SRI	022	036	029	013	031	051	021	031	041
		SSI	019	036	038	019	040	044	031	034	046
	$\mathbf{C}_1$:	*SR*	019	034	027	010	030	049	020	030	045
		SS	015	031	039	019	043	041	018	035	049
		MR	018	028	038	014	026	029	021	030	045
		MS	020	038	038	015	035	038	013	029	036
	$\mathbf{C}_2$:	*SR*	018	035	027	011	032	053	019	029	042
		SS	017	039	038	017	041	046	025	032	041
		MR	021	037	032	019	030	044	014	033	047
		MS	015	037	039	015	034	045	018	031	049
	$\mathbf{C}_3$:	*SR*	020	027	026	012	033	046	019	028	042
		SS	014	030	040	012	039	044	020	033	047
		MR	012	022	022	013	042	048	018	032	038
		MS	020	037	038	012	038	043	018	026	044

*Note: The entries within table correspond to empirical type I error rates multiplied by 1000.

Although not reported in the tables, additional simulations demonstrated that the superiority of nonparametric tests was not attained until n reached 15. Figure 2 shows the empirical power trends of tests for the heavy tailed distribution.

Table 3. The empirical powers of tests under multivariate normal distribution.

			Correlation structure								
			CS			AR1			UN		
	Matrix **C**	Test	10	20	30	10	20	30	10	20	30
$p = 4$		T^2	178	323	487	226	470	669	183	366	585
		SRI	172	325	478	220	450	659	181	361	562
		SSI	166	283	409	216	400	585	176	323	510
	$\mathbf{C_1}$:	*SR*	172	326	483	218	450	649	186	359	561
		SS	157	283	416	198	402	582	168	313	490
		MR	154	311	460	210	422	631	170	323	528
		MS	113	174	288	131	282	402	121	187	294
	$\mathbf{C_2}$:	*SR*	169	332	481	218	455	658	183	367	567
		SS	170	292	419	198	412	598	175	329	500
		MR	153	308	471	210	439	650	171	357	541
		MS	100	185	307	131	276	460	110	226	314
	$\mathbf{C_3}$:	*SR*	170	326	473	216	450	656	180	367	554
		SS	172	272	409	185	409	578	167	315	502
		MR	162	306	458	188	435	617	175	365	551
		MS	117	194	291	138	267	372	128	222	349
$p = 8$		T^2	175	421	655	154	388	613	154	398	626
		SRI	90	382	619	82	354	560	78	355	594
		SSI	86	359	577	76	338	535	85	344	567
	$\mathbf{C_1}$:	*SR*	92	386	618	85	353	564	83	360	605
		SS	101	378	592	95	344	522	93	353	578
		MR	72	357	592	62	333	558	70	310	575
		MS	55	191	332	47	193	333	46	186	351
	$\mathbf{C_2}$:	*SR*	80	376	612	84	355	566	81	349	590
		SS	78	341	579	87	344	535	75	316	553
		MR	65	316	556	72	326	537	72	286	532
		MS	47	117	216	45	198	307	54	123	236
	$\mathbf{C_3}$:	*SR*	84	378	611	78	353	567	86	346	587
		SS	86	351	569	72	336	503	80	321	539
		MR	70	342	567	59	309	517	67	306	546
		MS	60	165	302	48	165	247	60	157	274

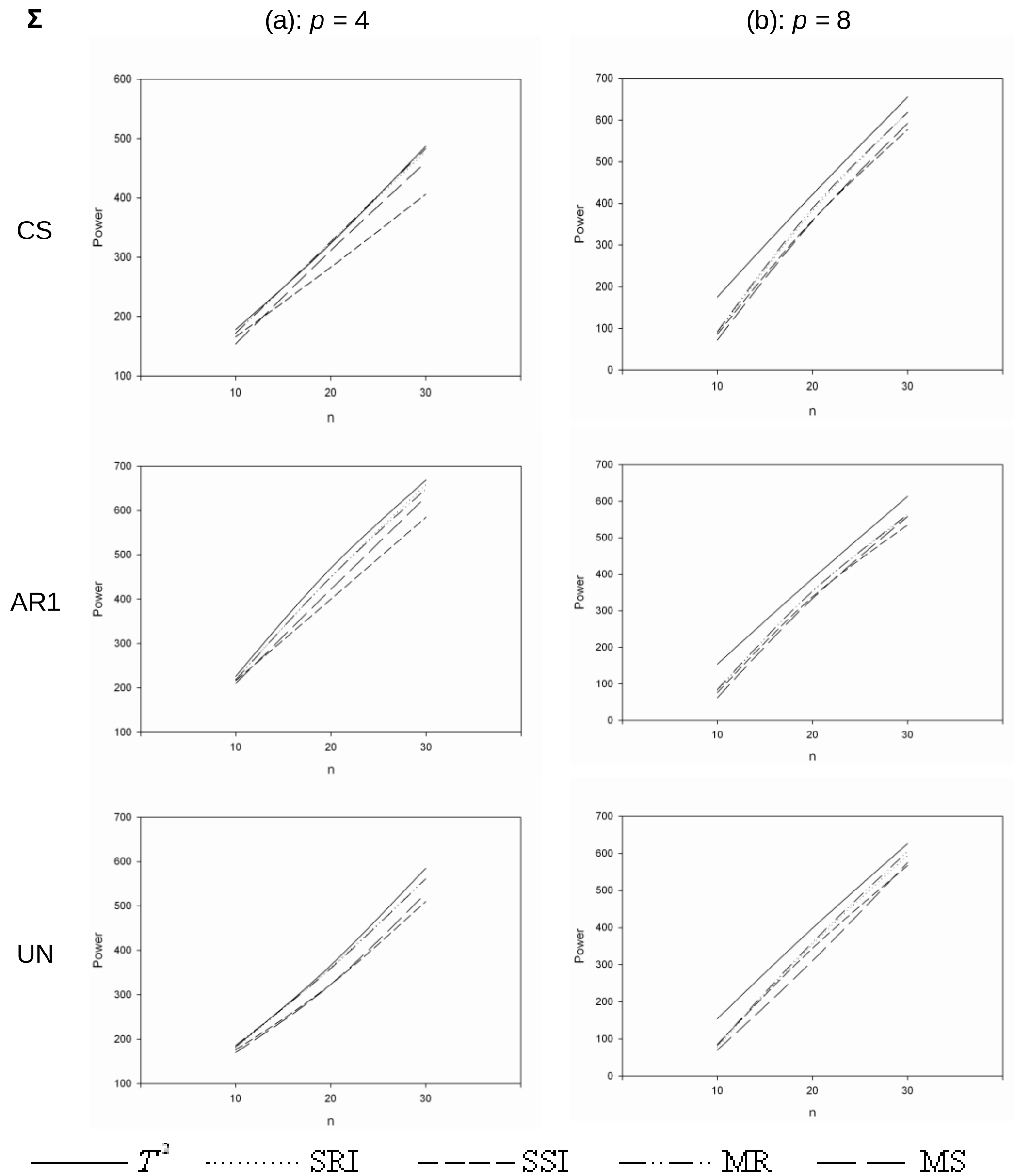
*Note: The entries within table correspond to empirical powers multiplied by 1000.

Table 4. The empirical powers of tests under multivariate *t* distribution.

			Correlation structure								
			CS			AR1			UN		
	Matrix **C**	Test	10	20	30	10	20	30	10	20	30
$p = 4$		T^2	261	480	653	344	607	779	273	527	691
		SRI	294	623	822	397	736	919	321	668	872
		SSI	309	654	856	427	782	946	336	700	900
	$\mathbf{C}_1$:	*SR*	293	620	823	395	728	920	314	667	871
		SS	317	648	840	414	771	940	328	689	889
		MR	264	587	788	370	713	915	263	601	816
		MS	180	410	582	287	511	721	176	391	602
	$\mathbf{C}_2$:	*SR*	293	616	818	402	735	918	322	674	872
		SS	315	653	846	424	782	946	342	691	891
		MR	276	591	802	380	720	922	286	643	846
		MS	180	401	636	263	537	795	179	387	661
	$\mathbf{C}_3$:	*SR*	287	635	826	397	733	918	326	670	863
		SS	303	658	841	401	756	939	343	695	893
		MR	278	591	793	359	711	903	294	630	847
		MS	219	432	642	263	544	759	212	495	723
$p = 8$		T^2	317	632	822	267	590	773	296	607	799
		SRI	194	709	917	169	661	893	184	669	902
		SSI	196	761	953	170	716	934	180	731	940
	$\mathbf{C}_1$:	*SR*	201	717	920	175	676	900	195	682	911
		SS	233	776	954	205	729	931	212	743	945
		MR	176	681	893	145	638	868	166	653	878
		MS	108	444	718	112	405	695	113	439	722
	$\mathbf{C}_2$:	*SR*	200	709	916	172	678	899	183	669	906
		SS	190	760	947	195	730	931	172	719	939
		MR	151	607	870	149	635	866	149	578	840
		MS	101	268	504	099	378	632	102	257	496
	$\mathbf{C}_3$:	*SR*	200	711	917	161	667	901	178	672	902
		SS	218	765	946	176	693	923	197	718	930
		MR	171	634	890	135	601	836	154	608	859
		MS	119	414	694	100	349	606	126	379	645

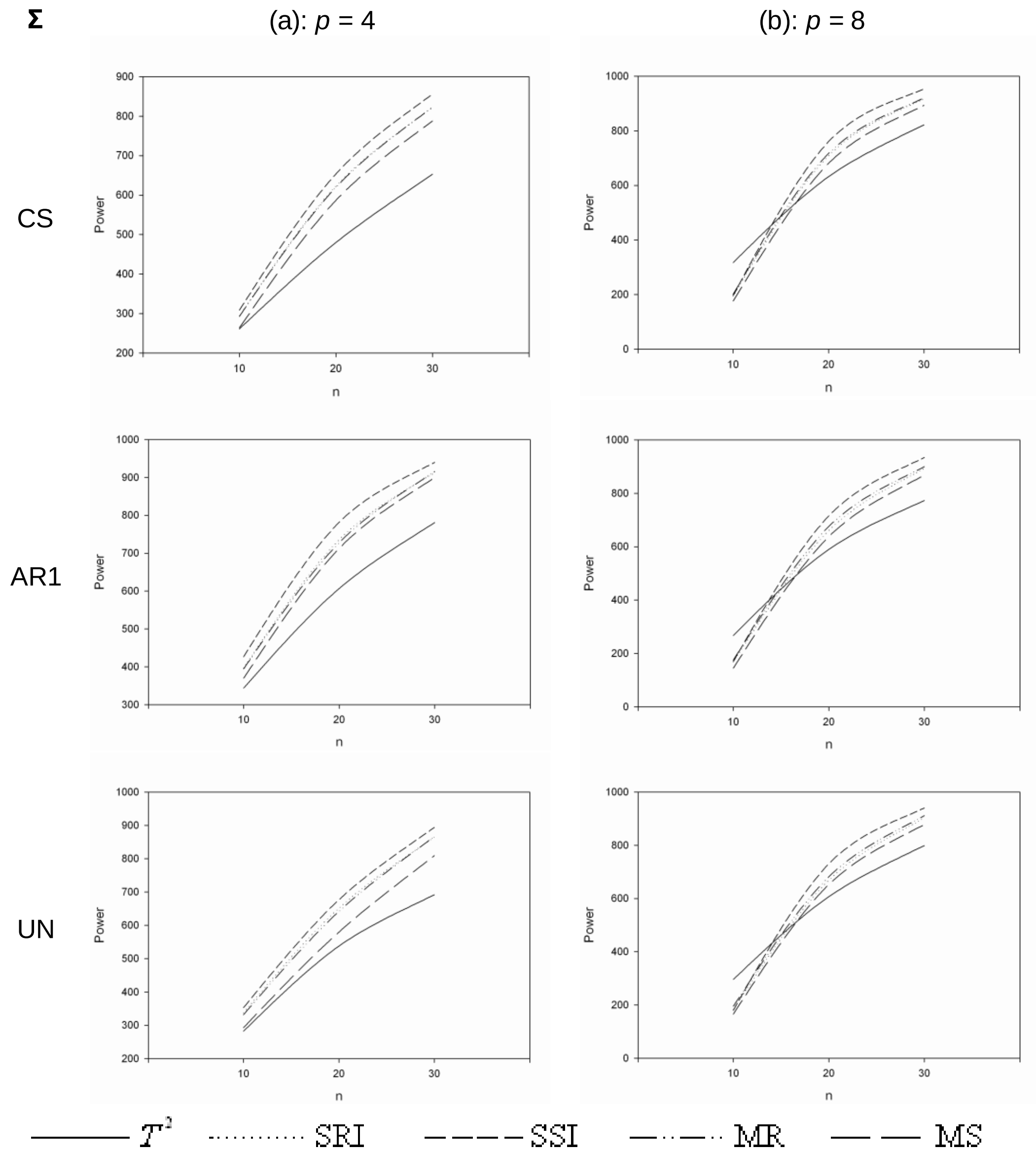
*Note: The entries within table correspond to empirical powers multiplied by 1000.

Except for the test based on marginal sign, the performances of other non invariant tests were relatively robust with respect to different choices of transformation matrix **C** to test parallelism. There was not a unique choice for **C** which corresponded to the best performance of the tests. Figure 3 illustrates the degree of stability in power values for the 4 non-invariant tests for the three transformation matrices $\mathbf{C}_1$ - $\mathbf{C}_3$ when $n = 30$.



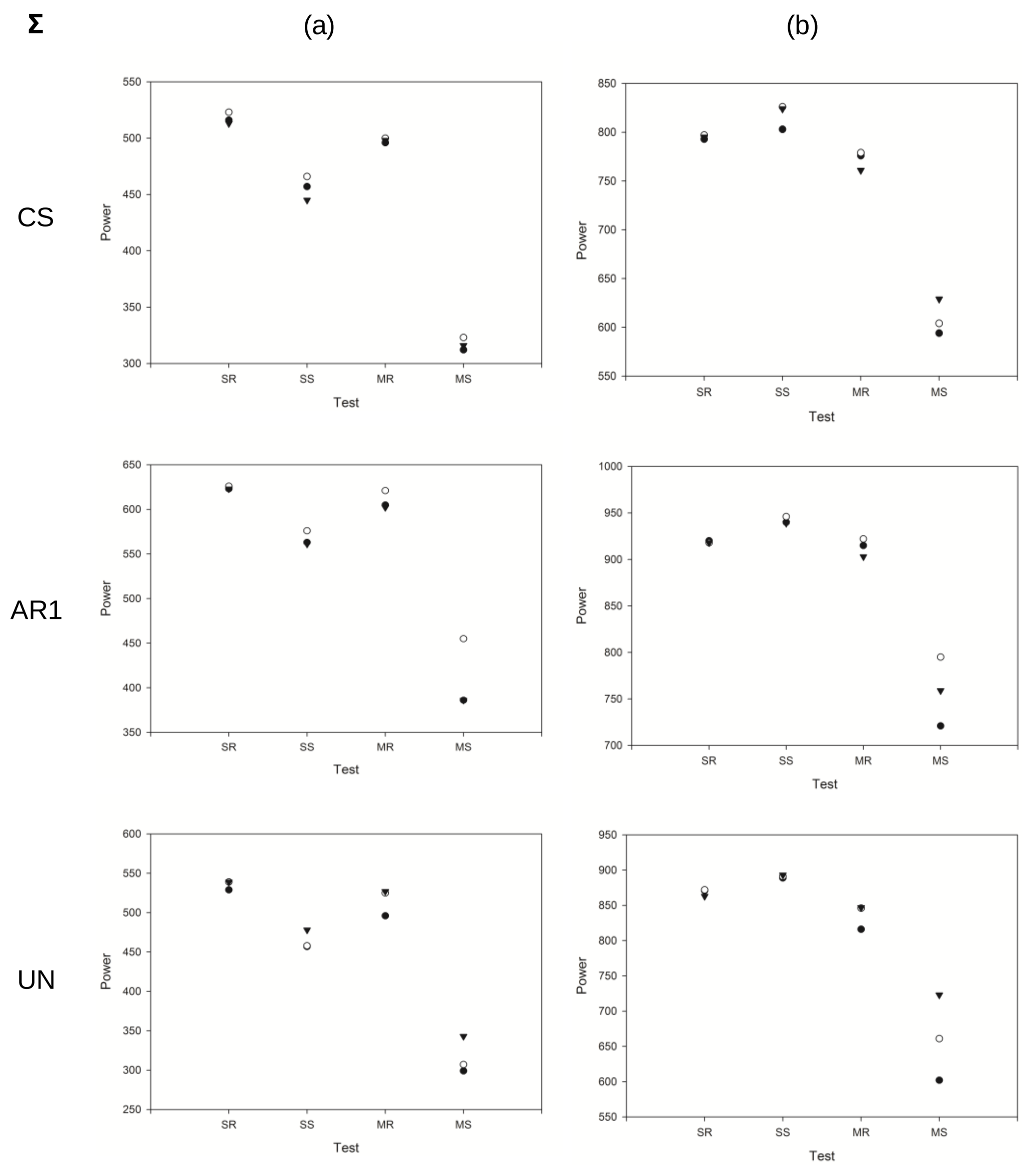
*Note: For purpose of better illustration, the powers of non-invariant tests are displayed only for the matrix $\mathbf{C}_2$.

Figure 1. The empirical powers of tests under multivariate normal distribution.



*Note: For purpose of better illustration, the powers of non-invariant tests are displayed only for the matrix $\mathbf{C}_2$.

Figure 2. The empirical powers of tests under multivariate t (3) distribution.



symbol key: ●: **C$_1$**, ○: **C$_2$**, and ▼: **C$_3$**

Figure 3. The empirical powers of non-invariant tests for *n* = 30 for various transformation matrices under multivariate normal (a) and *t* (b) distributions

Conclusion

The results of the study revealed that the tests based on spatial and marginal ranks could serve as efficient tools for profile analysis since they performed notably better than Hotelling's T^2 for the heavy tailed distribution and were as efficient as it under normality. Similar results reported in simulation studies by Nordhausen et al. (2006) and Möttönen et al. (1998) only in the context of two sample comparison of locations for normal and t distributions. Interestingly, even for moderate tailed t distributions, the tests based on ranks were superior to Hotelling's T^2 in both studies. Um and Randles (1998) also reported that the multi sample extensions of multivariate rank tests proposed by Randles and Peters (1990) were more efficient than Lawly-Hotelling's $\mathbf{U}$ for light-tailed and heavy-tailed distributions. However, the results revealed that when there was sufficient evidence to conclude that the underlying distribution was heavy tailed, the tests based on spatial signs were the best choices to profile analysis. Similarly, this aspect was reported in the study by Nordhausen et al. (2006) and for a different sign test by Um and Randles (1998). It should also be noted that above studies conducted in areas not involving repeated measurements and various correlation structures for errors. The simulations also illustrated that when the number of replication was large (here $p = 8$) the mentioned nonparametric tests outperformed Hotelling's T^2 only for larger sample sizes ($n \geq 10$). The panel (b) of Figure 2 illustrated this issue for which Hotelling's T^2 performed slightly better than any nonparametric counter parts for $p = 8$ even if the underlying distribution was heavy-tailed. The effect of sample size relative to the number of measurements has been not reported yet and hence further research in this area is necessary.

In the context of two sample comparison (as our study), Hotelling's T^2 and all the MANOVA tests (Wilks' $\mathbf{\Lambda}$, Pilla's $\mathbf{V}$, Lawley-Hotelling's $\mathbf{U}$ and Roy's $\boldsymbol{\theta}$) are functions of each other and give equivalent results; see Rencher (1998). The power of the MANOVA tests has been compared by several authors. However, they are asymptotically equivalent for sufficient sample sizes (Olson, 1974). Therefore it is implied the nonparametric alternatives can be confidently applied in place of MANOVA tests in profile analysis regardless of the nature of underlying distribution. Park et al. (2001) investigated the performance of profile analysis using Hotelling's T^2 and mixed model approach to test group and interaction effects. Also, Vossoughi et al. (2012) compared the performance of profile analysis, linear mixed model and summary measure approach in repeated measurements generated from a linear mixed model setting. Similarly, both studies showed that the profile analysis preserved the nominal significance level and performed relatively robust to the underlying correlation structure but provided less power values than the competitors, in general.

Marcucci (1986) demonstrated that profile analysis using Hotelling's $T^2$and exclusively univariate split-plot analysis with d.f. adjustments gave type I error rates closest to the nominal level, but not one of which was most powerful along various correlation structures and patterns of means. The interested reader is also referred to Schwertman et al. (1985), Boik (1991) and Davidson (1972) for further assessment on this issue.

Thought not reported here, we conducted additional simulations for a variety type of the location trend over occasions such as linear trend as $\boldsymbol{\theta}_2 = (0.25, 0.5, 0.75, 1)'$. The larger number of measurements $p = 8$ and sample size $n = 50$ in each group were also considered. However, the similar results were yielded and hence not further included in the study.

In conclusion, the findings implied that the use of some nonparametric multivariate tests in place of the parametric counterparts can considerably improve the result of profile analysis for heavy-tailed distributions. Accordingly, the tests based on spatial and marginal ranks are severe competitors for parametric tests in profile analysis since they performed as well as Hotelling's T^2 under multivariate normal distribution and dominated it under heavy-tailed distribution. Moreover, the simulation results revealed that the tests based on spatial signs under heavy tailed distributions, were more efficient than the MANOVA tests for the analysis of repeated measurements.

References

Azzalini, A., & Bowman, A. (1991). Nonparametric regression methods for repeated measurements. In G. Roussas (Ed.), *Nonparametric functional estimation and related topics* (pp. 377-387). Netherlands: Springer.

Bhapkar, V. P. (1984). Univariate and multivariate multisample location and scale tests. In P. R. Krishnaiah & P. K. Sen (Eds.), *Nonparametric methods* (pp. 31-62). Oxford: Elsevier.

Bhapkar, V. P., & Patterson, K.W. (1978). A Monte Carlo study of some multivariate nonparametric statistics for profile analysis of several samples. *Journal of Statistical Computation and Simulation*, *6*(3/4), 223-237. doi: 10.1080/00949657808810191

Boik, R. J. (1991). Scheffés mixed model for multivariate repeated measures: A relative efficiency evaluation. *Communications in Statistics - Theory and Methods, 20*(4), 1233-1255. doi: 10.1080/03610929108830562

Davidson, M. L. (1972). Univariate versus multivariate tests in repeated-measures experiments. *Psychological Bulletin, 77*(6), 446-452. doi: 10.1037/h0032674

Davis, A. W. (1980). On the effects of moderate multivariate nonnormality on Wilks's likelihood ratio criterion. *Biometrika, 67*(2), 419-427. doi: 10.1093/biomet/67.2.419

Davis, A. W. (1982). On the effects of moderate multivariate nonnormality on Roy's largest root test. *Journal of the American Statistical Association, 77*(380), 896-900. doi: 10.1080/01621459.1982.10477904

Davis, C. S. (2002). *Statistical methods for the analysis of repeated measurements*. New York, NY: Springer-Verlag.

Everitt, B. S. (1979). A Monte Carlo investigation of the robustness of Hotelling's one-and two-sample T^2 tests. *Journal of the American Statistical Association, 74*(365), 48-51. doi: 10.1080/01621459.1979.10481606

Hettmansperger, T. P., Möttönen, J. & Oja, H. (1998). Affine invariant multivariate rank tests for several samples. *Statistica Sinica, 8*(3), 785-800.

Hettmansperger, T. P., & Oja, H. (1994). Affine invariant multivariate multisample sign tests. *Journal of the Royal Statistical Society. Series B (Methodological), 56*(1), 235-249.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38*(4), 963-974. doi: 10.2307/2529876

Littell, R. C., Pendergast, J., & Natarajan, R. (2000). Tutorial in biostatistics: Modelling covariance structure in the analysis of repeated measures data. *Statistics in Medicine, 19*, 1793-1819.

Marcucci, M. (1986). A comparison of the power of some tests for repeated measurements. *Journal of Statistical Computation and Simulation, 26*(1/2), 37-53. doi: 10.1080/00949658608810948

Mendoza, J. L., Toothaker, L. E., & Nicewander, W. A. (1974). A Monte Carlo comparison of the univariate and multivariate methods for the groups by trials repeated measures design. *Multivariate Behavioral Research, 9*(2), 165-177. doi: 10.1207/s15327906mbr0902_3

Möttönen, J., Hettmansperger, T. P., Oja, H., & Tienari, J. (1998). On the efficiency of affine invariant multivariate rank tests. *Journal of Multivariate Analysis, 66*(1), 118-132. doi: 10.1006/jmva.1998.1740

Möttönen, J., & Oja, H. (1995). Multivariate spatial sign and rank methods. *Journal of Nonparametric Statistics*, *5*(2), 201-213. doi: 10.1080/10485259508832643

Möttönen, J., Oja, H., & Tienari, J. (1997). On the efficiency of multivariate spatial sign and rank tests. *The Annals of Statistics*, *25*(2), 542-552. doi: 10.1214/aos/1031833663

Nordhausen, K., Oja, H., & Tyler, D. E. (2006). On the efficiency of invariant multivariate sign and rank tests. In E. P. Liski, J. Isotalo, J. Niemelä, S. Puntanen & G. P. H. Styan (Eds.), *Festschrift for Tarmo Pukkila on his 60th birthday* (pp.217-231). Tampere, Finland: Dept. of Mathematics, Statistics and Philosophy, University of Tampere.

Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates: A review. *Scandinavian Journal of Statistics*, *26*(3), 319-343. doi: 10.1111/1467-9469.00152

Oja, H. (2010). *Multivariate nonparametric methods with R: An approach based on spatial signs and ranks*. New York, NY: Springer.

Oja, H., & Randles, R. H. (2004). Multivariate nonparametric tests. *Statistical Science*, *19*(4), 598-605. doi: 10.1214/088342304000000558

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association*, *69*(348), 894-908.

Park, T., Park, J.-K., & Davis, C. S. (2001). Effects of covariance model assumptions on hypothesis tests for repeated measurements: Analysis of ovarian hormone data and pituitary-pteryomaxillary distance data. *Statistics in Medicine, 20*(16), 2441-2453. doi: 10.1002/sim.859

Puri, M. L., & Sen, P. K. (1971). *Nonparametric methods in multivariate analysis*. New York, NY: Wiley.

R Development Core Team. (2013). R: A language and environment for statistical computing. Reference index (Version 3.0.1). Vienna, Austria: R Foundation for Statistical Computing.

Randles, R. H., & Peters, D. (1990). Multivariate rank tests for the two-sample location problem. *Communications in Statistics - Theory and Methods, 19*(11), 4225-4238. doi: 10.1080/03610929008830439

Rencher, A. C. (1995). *Methods of multivariate analysis*. New York, NY: Wiley.

Rencher, A. C. (1998). *Multivariate statistical inference and applications*. New York, NY: Wiley.

Schwertman, N. C., Flynn, W., Stein, S., & Schenk, K. L. (1985). A Monte Carlo study of alternative procedures for testing the hypothesis of parallelism for complete and incomplete growth curve data. *Journal of Statistical Computation and Simulation*, *21*(1), 1-37. doi: 10.1080/00949658508810794

Schwertman, N. C., Fridshal, D., & Magrey, J. M. (1981). On the analysis of incomplete growth curve data, a Monte Carlo study of two nonparametric procedures. *Communications in Statistics-Simulation and Computation*, *10*(1), 51-66. doi: 10.1080/03610918108812192

Sen, P. K. (1984). Nonparametric procedures for some miscellaneous problems. In P. R. Krishnaiah & P. K. Sen (Eds.), *Nonparametric methods* (pp. 699-739). Oxford: Elsevier.

Singer, J. M., Poleto, F. Z., & Rosa, P. (2004). Parametric and nonparametric analyses of repeated ordinal categorical data. *Biometrical Journal*, *46*(4), 460-473. doi: 10.1002/bimj.200310045

Somorčík, J. (2006). Tests using spatial median. *Austrian Journal of Statistics*, *35*(2/3), 331-338. doi: 10.17713/ajs.v35i2&3.380

Um, Y., & Randles, R. H. (1998). Nonparametric tests for the multivariate multi-sample location problem. *Statistica Sinica*, *8*(3), 801-812.

Vossoughi, M., Ayatollahi, S. M. T., Towhidi, M., & Ketabchi, F. (2012). On summary measure analysis of linear trend repeated measures data: performance comparison with two competing methods. *BMC Medical Research Methodology*, *12*(33), 1-12. doi: 10.1186/1471-2288-12-33

Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. *American Statistician*, *39*(2), 95-101.

Wernecke, K.-D., & Kalb, G. (1999). On nonparametric repeated measures analysis of variance. In W. Gaul & H. Locarek-Junge (Eds.), *Classification in the information age* (pp. 596-604). Berlin: Springer.

Wernecke, K.-D., & Kaufmann, J. (2000). A new approach on nonparametric repeated measures analysis of variance. In R. Decker & W. Gaul (Eds.), *Classification and information processing at the turn of the millennium* (pp. 430-437). Berlin: Springer.