Ann Inst Stat Math (2019) 71:327–364
https://doi.org/10.1007/s10463-017-0643-8

Testing in nonparametric ANCOVA model based on ridit reliability functional

Debajit Chatterjee[1] · **Uttam Bandyopadhyay**[2]

Received: 19 January 2017 / Revised: 18 December 2017 / Published online: 6 February 2018

Abstract In the spirit of Bross (Biometrics 14:18–38, 1958), this paper considers ridit reliability functionals to develop test procedures for the equality of $\mathrm{K}(>2)$ treatment effects in nonparametric analysis of covariance (ANCOVA) model with d covariates based on two different methods. The procedures are asymptotically distribution free and are not based on the assumption that the distribution functions (d.f.'s) of the response variable and the associated covariates are continuous. By means of simulation study, the proposed methods are compared with the methods provided by Tsangari and Akritas (J Multivar Anal 88:298–319, 2004) and Bathke and Brunner (Recent advances and trends in nonparametric statistics, Elsevier, Amsterdam, 2003) under ANCOVA in terms of type I error rate and power.

Keywords Asymptotic distribution · Nonparametric ANCOVA model · Ridit · U-statistic · Nadaraya–Watson weight · Bandwidth

1 Introduction

Many statistical studies involve analysis in which response variables correspond to two or more treatments in the presence of one or more auxiliary variables termed as

✉ Uttam Bandyopadhyay
ubstat@caluniv.ac.in

Debajit Chatterjee
debajit1.chatterjee@gmail.com

1 Department of Statistics, Maulana Azad College, 8, Rafi Ahmed Kidwai Rd, Taltala, Kolkata 700013, India

2 Department of Statistics, University of Calcutta, 35 Ballygunge Circular Road, Kolkata 700019, India

covariates. In some cases, the problem is to judge whether the treatment effects are significantly different when the covariate effects are eliminated. A standard practice to handle the problem is to assume suitable linear model and normality of the error components. Besides, many authors (see, for example, Grigoletto and Akritas 1999; Thas et al. 2012; Neve and Thas 2015) meet the problem under semiparametric model in which normality assumption is relaxed by setting a covariate-influenced linear or nonlinear model for the response variables. Further, dropping the parametric part in the semiparametric model, tests for homogeneity of covariate-eliminated response distributions are obtained under fully nonparametric heteroscedastic model with one covariate (see, for example, Dette and Neumeyer 2001; Akritas and Keilegom 2001; Munk et al. 2007). In addition, the use of ANCOVA model under fully nonparametric setup can be found in the works of Akritas et al. (2000), Tsangari and Akritas (2004), Wang and Akritas (2006) among others. There is also an alternative approach (Bathke and Brunner 2003) that keeps analogy between parametric and nonparametric ANCOVA models.

Bross Irwin (1958) introduces ridit as a reliability functional for comparing ordinal-scale responses. In this context, for two independent variables U and V, a reliability functional can be defined by

$$R = P(U < V) + \frac{1}{2}P(U = V). \tag{1}$$

There are several interpretations for the above functional, like Wilcoxon functional, concordance probability, stress strength and nonparametric treatment effect. Some authors use the concept of ridit to frame tests for comparing two or more independent samples under nonparametric setup (see, for example, Akritas et al. 1997; Brunner and Munzel 2000; Brunner and Puri 2001; Konietschke et al. 2012; Fischer et al. 2014; Bandyopadhyay and Chatterjee 2015; Friedrich et al. 2017). The advantages of using such ridit analysis are that it helps to construct tests more precisely in accordance with the structure of the null and alternative hypotheses (see, for example, Terpstra and Magel 2003) and it need not require the existence of moments of the response variables. In the present study, we develop tests for the covariate eliminating treatment effects under nonparametric setup. Here, we follow two different approaches, conditional distribution approach (Tsangari and Akritas 2004) and an alternative approach (Bathke and Brunner 2003) based on marginal distributions to construct new test procedures with the help of reliability functionals that generalize (1).

The content of the paper is as follows. The setup and the homogeneity hypothesis under various ridit reliability functionals are described in Sect. 2. Section 3 contains the proposed asymptotically distribution-free (ADF) tests and the test by Tsangari and Akritas (2004) as competitor. Section 4 develops some alternative tests using the nonparametric model provided in Bathke and Brunner (2003). Simulated values of type I error rate (empirical level) and power for various test procedures, proposed and competitor, are obtained in Sect. 5. A data study, considered in Sect. 6, illustrates the use of the different tests. The paper concludes in Sect. 7, followed by some technical details in "Appendix."

2 Setup and hypothesis using conditional distribution approach

The present work is on comparison among more than two treatments, $1, 2, \ldots, K(> 2)$, when the subjects under consideration are not homogeneous. Let $(Y_k, \mathbf{X}_k)$ be the variables associated with a subject corresponding to treatment k, where Y_k is the real-valued response variable and $\mathbf{X}_k = \left(X_k^{(1)}, X_k^{(2)}, \ldots, X_k^{(d)}\right)^T$ is the associated d-component covariate. Let $F_{k\mathbf{x}}$ be the conditional d.f. of Y_k when $\mathbf{X}_k = \mathbf{x}$ and G_k be the marginal d.f. of $\mathbf{X}_k$. Here, we do not require to assume that the d.f.'s $F_{k\mathbf{x}}$ and G_k are continuous. That is, the procedures are valid for discrete distributions also. In particular, when $F_{k\mathbf{x}}$ is continuous, we need to assume that the derivative, $F'_{k\mathbf{x}}(y)$, of $F_{k\mathbf{x}}(y)$ exists and is continuous in a suitable domain of y. Let $\{(\mathbf{X}_{kj}, Y_{kj}),\ j = 1, 2, \ldots, n_k\}$ be the observations from n_k subjects who receive treatment k, $k = 1, 2, \ldots, K$. We write $N = \sum_{k=1}^{K} n_k$, the total number of subjects receiving all the treatments. In this paper, the relative treatment effect for treatment k, when the covariate matrix $\mathfrak{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K)$ is realized at $\mathfrak{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K)$, is defined by a suitable functional, denoted by $\tau_k(\mathfrak{x}) = \tau_k(\theta(\mathfrak{x}))$, $k = 1, 2, \ldots, K$, where $\theta(\mathfrak{x}) = (F_{1\mathbf{x}_1}, F_{2\mathbf{x}_2}, \ldots, F_{K\mathbf{x}_K})^T$. Here, $\tau_k(\mathfrak{x})$ is used to provide relative comparison for the treatment k among the treatment groups through their conditional distributions when the covariates are realized at $\mathfrak{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K)$. Therefore, $\tau_k(\mathfrak{x})$ is influenced by the treatments as well as by the covariate levels. This two-way measure can be converted into one-way treatment comparison by adjusting the covariate levels. One approach would be to integrate $\tau_k(\mathfrak{x})$ with respect to all the covariates according to a common d.f. This resultant measure can then be interpreted as covariate-eliminated treatment effect for treatment k relative to others $(k = 1, 2, \ldots, K)$. That means, if $G(\mathbf{x})$ denotes the overall d.f. associated with the covariate data $\{\mathbf{X}_{kj},\ j = 1, 2, \ldots, n_k,\ k = 1, 2, \ldots, K\}$, the covariate-eliminated treatment effect for the treatment k, denoted by $\tau_{k\cdot}$, can be represented by

$$\tau_{k\cdot} = \int \tau_k(\mathfrak{x}) \prod_{l=1}^{K} dG(\mathbf{x}_l), \quad k = 1, 2, \ldots, K.$$

Clearly, $\tau_{k\cdot}$ depends on the choice of G. While choosing G, it is to be kept in mind that G should represent universal population with respect to the covariate. A particular choice of G may be $\sum_{k=1}^{K} \lambda_k G_k$, where $\lambda_k > 0,\ k = 1, 2, \ldots, K$ with $\sum_{k=1}^{K} \lambda_k = 1$, or simply $\frac{1}{K} \sum_{k=1}^{K} G_k$. This will ensure that $\tau_{k\cdot}$ $(k = 1, 2, \ldots, K)$ are independent of the sample sizes and hence, as in Brunner et al. (2017), they are comparable irrespective of the observed sample sizes. With this background, we concentrate on the equality of the treatment effects after eliminating the covariate effects. Formally, such problem can be written as $\{\tau_{1\cdot} = \tau_{2\cdot} = \cdots = \tau_{K\cdot}\}$, which is equivalent to the problem of testing the null hypothesis

$$H_0 : C\boldsymbol{\tau}_{\cdot} = 0 \tag{2}$$

against an alternative H_1, where $\boldsymbol{\tau}. = (\tau_1., \tau_2., \ldots, \tau_K.)^T$ is the vector of covariate-eliminated treatment effects and C is a contrast matrix of order $(K-1) \times K$ with $\text{rank}(C) = K-1$ of the following form

$$C = \begin{pmatrix} 1 & -1 & 0 & 0 \cdots & 0 \\ 1 & 0 & -1 & 0 \cdots & 0 \\ \vdots & & & & \\ 1 & 0 & 0 & 0 \cdots & -1 \end{pmatrix}.$$

Again, if the rows of C are represented by $\{\mathbf{c}_l^T, l = 1, 2, \ldots, K-1\}$, it follows that

$$H_0 \Leftrightarrow \bigcap_{l=1}^{K-1} H_0(l),$$

where $H_0(l) : \mathbf{c}_l^T \boldsymbol{\tau}. = 0, l = 1, 2, \ldots, K-1.$

It is to be noted that under the assumption mentioned above, $\tau_k.$ can also be considered as a functional $\tau_k(\boldsymbol{\theta}.)$, where $\boldsymbol{\theta}. = (F_1., F_2., \ldots, F_K.)^T$ with

$$F_k.(y) = \int F_{k\mathbf{x}_k}(y) dG(\mathbf{x}_k), \quad k = 1, 2, \ldots, K. \tag{3}$$

The equality of the covariate-eliminated d.f.'s implies equality of the covariate-eliminated treatment effects, but the converse may not be true. Thus, the present problem of testing is more general than that considered by Akritas et al. (2000) and Tsangari and Akritas (2004). Moreover, for testing H_0 against an order-restricted alternative H_1, it is possible to construct a test using suitably chosen functionals. However, in this paper, tests are obtained for H_0 against all alternatives H_1 from three sets of ridit reliability functionals that are used by Brunner and Puri (2001) (referred to as BP), Bandyopadhyay and De (2011) (referred to as BD) and Bandyopadhyay and Chatterjee (2015) (referred to as BC). Conditional versions of such functionals are obtained in the following way.

2.1 BP-type functional

Here, the functional

$$p_{kk'}(\mathbf{x}_k, \mathbf{x}_{k'}) = P(Y_k > Y_{k'} | \mathbf{x}_k, \mathbf{x}_{k'}) + \frac{1}{2} P(Y_k = Y_{k'} | \mathbf{x}_k, \mathbf{x}_{k'}) \tag{4}$$

is used as a measure for comparing the conditional distribution of Y_k given $\mathbf{X}_k = \mathbf{x}_k$ with that of $Y_{k'}$ given $\mathbf{X}_{k'} = \mathbf{x}_{k'}$, $k, k' = 1, 2, \ldots, K$. Writing

$$F^0_{k\mathbf{x}_k}(y) = P(Y_k > y | \mathbf{x}_k) + \frac{1}{2} P(Y_k = y | \mathbf{x}_k),$$

(4) reduces to

$$p_{kk'}(x_k, x_{k'}) = \int F^0_{kx_k}(y)\, dF_{k'x_{k'}}(y),\ \ k, k' = 1, 2, \ldots, K$$

with $p_{kk}(x_k, x_k) = \frac{1}{2}$. Hence, a covariate-eliminated measure for comparing two distributions can be obtained as

$$\begin{aligned} p_{kk'} &= \int\int p_{kk'}(x_k, x_{k'})\, dG(x_k)\, dG(x_{k'}) \\ &= \int F^0_{k\cdot}(y)\, dF_{k'\cdot}(y), \end{aligned}$$

where

$$F^0_{k\cdot}(y) = \int F^0_{kx_k}(y) dG(x_k).$$

Further, the functional

$$p_k(\mathfrak{x}) = \frac{1}{K} \sum_{k'=1}^{K} p_{kk'}(x_k, x_{k'})$$

can be interpreted as the relative treatment effect for treatment k when the covariate $\mathfrak{X}$ is realized at $\mathfrak{x} = (x_1, x_2, \ldots, x_K)^T$. Hence, eliminating the effect of covariate, the main effect due to treatment k is defined by

$$p_{k\cdot} = \int p_k(\mathfrak{x}) \prod_{l=1}^{K} dG(x_l) = \frac{1}{K}\left[\sum_{k'(\neq k)=1}^{K} p_{kk'} + \frac{1}{2}\right],\ \ k = 1, 2, \ldots, K, \qquad (5)$$

which, being dependent on an universal d.f. G, is free from sample sizes (Brunner et al. 2017). Moreover, as discussed in (3), $p_{k\cdot}$ depends on $\theta_\cdot$, $k = 1, 2, \ldots, K$.

2.2 BD-type functional

Here, the relative treatment effect due to treatment *k*, when the covariate $\mathfrak{X}$ is realized at $\mathfrak{x}$, is defined by the ridit reliability functional

$$\begin{aligned} R_k(\mathfrak{x}) =& P[Y_k > \max(Y_m, m = 1, 2, \ldots, K; m \neq k)|\mathfrak{x}] \\ &+ \frac{1}{2} \sum_{1 \le k_1 \le K, k_1 \neq k} P[Y_k = Y_{k_1} \\ &> \max(Y_m, m = 1, 2, \ldots, K; m \neq k, k_1)|\mathfrak{x}] \end{aligned}$$

$$+\frac{1}{3}\sum_{1\le k_1<k_2\le K,\ k_1,k_2\ne k}\sum P[Y_k=Y_{k_1}=Y_{k_2}$$
$$>\max(Y_m, m=1,2,\ldots,K; m\ne k,k_1,k_2)|\mathfrak{x}]$$
$$+\cdots+\frac{1}{K}P[Y_1=Y_2=\cdots=Y_K|\mathfrak{x}],\ \ k=1,2,\ldots,K,$$

which, by eliminating the effect of covariate, gives

$$R_{k\cdot}=\int R_k(\mathfrak{x})\prod_{l=1}^{K}dG(x_l)$$

and is interpreted as the main effect due to treatment k, $k=1,2,\ldots,K$, and is independent of sample sizes. Moreover, after some routine steps, it is not difficult to express $R_{k\cdot}$ as a function of $\theta_{\cdot}$, $k=1,2,\ldots,K$.

2.3 BC-type functional

Similarly as above, the relative treatment effect due to treatment k, when the covariate $\mathfrak{X}$ is realized at $\mathfrak{x}$, is defined by the ridit reliability functional

$$R_k^*(\mathfrak{x})=P[Y_k<\min(Y_m, m=1,2,\ldots,K; m\ne k)|\mathfrak{x}]$$
$$+\frac{1}{2}\sum_{1\le k_1\le K,k_1\ne k}P[Y_k=Y_{k_1}$$
$$<\min(Y_m, m=1,2,\ldots,K; m\ne k,k_1)|\mathfrak{x}]$$
$$+\frac{1}{3}\sum_{1\le k_1<k_2\le K,\ k_1,k_2\ne k}\sum P[Y_k=Y_{k_1}=Y_{k_2}$$
$$<\min(Y_m, m=1,2,\ldots,K; m\ne k,k_1,k_2)|\mathfrak{x}]$$
$$+\cdots+\frac{1}{K}P[Y_1=Y_2=\cdots=Y_K|\mathfrak{x}],\ \ k=1,2,\ldots,K,$$

which, by eliminating the effect of covariate, gives

$$R_{k\cdot}^*=\int R_k^*(\mathfrak{x})\prod_{l=1}^{K}dG(x_l)$$

and is interpreted as the main effect due to treatment k according to the given set of ridit reliability functional, $k=1,2,\ldots,K$, and is independent of sample sizes. By the same argument as above, $R_{k\cdot}^*$ is dependent on $\theta_{\cdot}$, $k=1,2,\ldots,K$.

Here, it is important to compare the covariate-eliminated treatment effect for treatment k with that for treatment k' $(k\ne k')$. For example, suppose $F_{k\cdot}$ is stochastically larger than $F_{k'\cdot}$ (that is, treatment k is more effective than treatment k'). Then, for

BD (BC)-type functional, it follows from Bandyopadhyay and Chatterjee (2015) that $R_{k\cdot} > R_{k'\cdot}$ ($R^*_{k\cdot} < R^*_{k'\cdot}$). Further, with the same argument, it is not difficult to verify $p_{k\cdot} > p_{k'\cdot}$. In particular, when $F_{1\cdot} = F_{2\cdot} = \cdots = F_{K\cdot}$, it follows that $R_{k\cdot} = R^*_{k\cdot} = \frac{1}{K}$ and $p_{k\cdot} = \frac{1}{2}$, $k = 1, 2, \ldots, K$.

Next, under BP-type model, the testing problem becomes

$$H_{0p} : C\mathbf{p}. = 0 \;\text{ against }\; H_{1p} : C\mathbf{p}. \neq 0,$$

whereas, under BD-type model, the problem becomes

$$H_{0R} : C\mathbf{R}. = 0 \;\text{ against }\; H_{1R} : C\mathbf{R}. \neq 0,$$

and finally, under BC-type model, the problem becomes

$$H_{0R^*} : C\mathbf{R}.^* = 0 \;\text{ against }\; H_{1R^*} : C\mathbf{R}.^* \neq 0,$$

where $\mathbf{p}. = (p_{1\cdot}, p_{2\cdot}, \ldots, p_{K\cdot})^T$, $\mathbf{R}. = (R_{1\cdot}, R_{2\cdot}, \ldots, R_{K\cdot})^T$ and $\mathbf{R}.^* = (R^*_{1\cdot}, R^*_{2\cdot}, \ldots, R^*_{K\cdot})^T$.

By definition, the sum of the components in each of the effect vectors becomes a constant. In the testing procedures, such singularities are avoided by pre-multiplication of the contrast matrix C of rank $K-1$ with the effect vectors. It should be noted that all of H_{0p}, H_{0R} and H_{0R^*} indicate the equality of the covariate-eliminated treatment effects in different directions. These three hypotheses become identical when there are two treatments under comparison. Also, the testing problems presented here are generalizations of the problem considered by Akritas et al. (2000) and Tsangari and Akritas (2004).

3 Tests using conditional distribution approach

Here, we develop tests by use of consistent estimators of $\mathbf{p}.$, $\mathbf{R}.$ and $\mathbf{R}.^*$. For each such tests, the usual χ^2-test procedure and the multiple contrast test procedure (MCTP) (Konietschke et al. 2012) are adopted.

3.1 Tests for H_{0p}

A test for H_{0p} is obtained by using a consistent estimator of $\mathbf{p}.$ in which $F_{k\mathbf{x}}(y)$ is estimated consistently by

$$\widehat{F}_{k\mathbf{x}}(y) = \sum_{j=1}^{n_k} W_{kj}(\mathbf{x}, a_{n_k}) c_{kj}(y), \tag{6}$$

where $c_{kj}(y) = I(Y_{kj} \leq y)$, a_{n_k} is a bandwidth sequence of positive constants tending to zero as n_k tends to infinite and $W_{kj}(\mathbf{x}, a_{n_k})$, $j = 1, 2, \ldots, n_k$ are the weights determining the nature of the conditional distribution function estimator, $k = 1, 2, \ldots, K$. In this study, we use the Nadaraya–Watson weights

$$W_{kj}(x, a_{n_k}) = \frac{\Psi\left(\frac{x - X_{kj}}{a_{n_k}}\right)}{\sum_{j'} \Psi\left(\frac{x - X_{kj'}}{a_{n_k}}\right)},$$

where $\Psi(x) = \psi(x_1)\psi(x_2)\cdots\psi(x_d)$ with $\psi(.)$ as a symmetric kernel on a compact support and $\int u\psi(u)du = 0$. Here, we make the following assumption on the choice of the bandwidth sequence a_n.

A1. As $n \to \infty$, $a_n \to 0$ but $na_n \to \infty$.

For one-way design with d-dimensional covariate, we have $a_n = (\text{const})n^{-\frac{1}{(d+4)}}$ as an optimum choice of bandwidth (Silverman 1986, Sect. 4.3) in which the constant depends upon the distributions of the concerned random variables and the choice of the kernel. However, in the present context, this constant is determined from simulation study by keeping parity between empirical level and nominal level.

An estimator of $p_{k\cdot}$ can be obtained by

$$\widehat{p}_{k\cdot} = \frac{1}{K}\left[\sum_{k'(\neq k)=1}^{K} \widehat{p}_{kk'} + \frac{1}{2}\right], \quad k = 1, 2, \ldots, K,$$

where

$$\begin{aligned}
\widehat{p}_{kk'} &= \int\int \widehat{p}_{kk'}(x_k, x_{k'})\, d\widehat{G}(x_k)\, d\widehat{G}(x_{k'}) \\
&= \frac{1}{N^2}\sum_{l=1}^{N}\sum_{l'=1}^{N}\int \widehat{F}^0_{kx_l}(y)\, d\widehat{F}_{k'x_{l'}}(y) \\
&= \frac{1}{n_k n_{k'}}\sum_{j=1}^{n_k}\sum_{j'=1}^{n_{k'}} e_{kj}e_{k'j'}\, U(Y_{kj}, Y_{k'j'}),
\end{aligned}$$

with

$$\begin{aligned}
U(Y_k, Y_{k'}) &= 1 \quad \text{if } Y_k > Y_{k'} \\
&= \frac{1}{2} \quad \text{if } Y_k = Y_{k'} \\
&= 0 \quad \text{if } Y_k < Y_{k'}
\end{aligned}$$

and

$$e_{kj} = \frac{n_k}{N}\sum_{l=1}^{N} W_{kj}(X_l, a_{n_k}), \quad j = 1, 2, \ldots, n_k, \quad k = 1, 2, \ldots, K. \tag{7}$$

Now, we consider some results which are used in the derivation of the proposed tests. The proofs are given in "Appendix" and are based on the following assumptions on different sample sizes and on random covariates. Similar type of assumptions is also considered by Tsangari and Akritas (2004).

A2. For each N, there are $n_k = n_K(N)$, $k = 1, 2, \ldots, K$ such that $n_k \to \infty$ as $N \to \infty$, but $\frac{n_k}{N} \to \lambda_k \in (0, 1)$, $k = 1, 2, \ldots, K$ and $\sum_{k=1}^{K} \lambda_k = 1$.

A3. (i) The range, $\mathcal{S}$, of $\mathbf{X_k}$ is same and is bounded for all k.
(ii) The density (or the probability mass function) g_k corresponding to G_k satisfies the condition $\inf\{g_k(\mathbf{x}), \mathbf{x} \in \mathcal{S}\} > 0$.

Assumption A3 ensures the finiteness of the quantity $\frac{g(\mathbf{x})}{g_k(\mathbf{x})}$ and the existence of the conditional distributions. Furthermore, if the covariate is continuous, the following additional assumption is made.

A4. The d.f. G_k has bounded second derivative for each k.

Result 1 *Under A1–A4, $\widehat{p}_{k\cdot}$ is a consistent estimator of $p_{k\cdot}$, $k = 1, 2, \ldots, K$.*

Result 2 *There exists a positive definite (p.d.) matrix Σ_{Cp} such that, under the same assumptions as in Result 1, the asymptotic distribution of $\sqrt{N}C(\widehat{\mathbf{p}}. - \mathbf{p}.)$ is $(K-1)$–variate normal with mean vector $\mathbf{0}$ and dispersion matrix Σ_{Cp}, where $\widehat{\mathbf{p}}. = (\widehat{p}_{1\cdot}, \widehat{p}_{2\cdot}, \ldots, \widehat{p}_{K\cdot})^T$.*

The elements of $\Sigma_{Cp} = (\sigma_{ij})$ are given in "Appendix." Next, we make the following tests.

χ^2-test

Let $\widehat{\Sigma}_{Cp}$ be a consistent estimator of Σ_{Cp}. Then, by Result 2, the asymptotic distribution of $T_p = N(C\widehat{\mathbf{p}}.)^T \widehat{\Sigma}_{Cp}^{-1}(C\widehat{\mathbf{p}}.)$ under H_{0p} is central χ^2 with $(K-1)$ degrees of freedom. Moreover, as $N \to \infty$,

$$\frac{T_p}{N} \to \Delta_p^2$$

in probability, where $\Delta_p^2 = (C\mathbf{p}.)^T \Sigma_{Cp}^{-1}(C\mathbf{p}.) \geq 0$. Equality holds if and only if H_{0p} is true. Hence, a right-tailed test based on T_p will be appropriate for testing H_{0p} against H_{1p}. Thus, H_{0p} is rejected asymptotically at level α if and only if

$$T_p > \chi^2_{\alpha, K-1}, \tag{8}$$

where $\chi^2_{\alpha,K-1}$ is the upper α point of χ^2 distribution with $(K-1)$ degrees of freedom.

Multiple contrast test

Here, we set $H_{0p} = \bigcap_{l=1}^{K-1} H_{0pl}$ with $H_{0pl} : \mathbf{c}_l^T \mathbf{p}. = 0$, $l = 1, 2, \ldots, K-1$ and concentrate on multiple test procedure (MTP) in which H_{0p} is rejected if and only if at least one of the component hypotheses, $\{H_{0pl}, l = 1, 2, \ldots, K-1\}$, is rejected. In the present framework, H_{0pl} is tested by using the statistic

$$T^*_{pl} = \frac{\sqrt{N}\mathbf{c}_l^T \widehat{\mathbf{p}}.}{\sqrt{\widehat{\sigma}_{ll}}},$$

where $\widehat{\sigma}_{ll}$ is the lth diagonal element of $\widehat{\Sigma}_{Cp}$, $l = 1, 2, \ldots, K-1$. By Result 2, the asymptotic distribution of T^*_{pl} is standard normal under H_{0p}, $l = 1, 2, \ldots, K-1$. Moreover, the statistics are dependent. The traditional approach in MTP is to control the probability of at least one false rejection, known as family-wise error rate (FWER). The existing methods, which control FWER, are based on the *p* values, adjusted suitably, of the component tests. See, for example, the single-step procedures in Lehmann and Romano (2005, Sect. 9.1) for the classical Bonferroni method and Simes (1986), and the stepwise procedures in Holm (1979) and Hochberg (1988). Generally, stepwise procedures are more powerful than single-step procedures. Moreover, these procedures are conservative and do not require any assumption on the dependent structure of the *p* values. However, in the present scenario, the component statistics are dependent. If the true dependent structure of the component statistics could be used in the formulation of MTP, these would provide more control over FWER and power.

Tamhane and Dunnett (1999) provide MTPs that account for the specific structure of normally distributed component test statistics (see also Bretz et al. 2010, Chap. 2–4). With this idea, we describe a nonparametric multiple contrast test procedure (MCTP), considered by Konietschke et al. (2012), which incorporates dependency among the component test statistics.

As mentioned earlier, the asymptotic null distributions of component statistics are standard normal. Next, we set $\mathbf{T}^*_{\mathbf{p}} = \left\{T^*_{pl}; l = 1, 2, \ldots, K-1\right\}$ and $\Omega = (\omega_{ij})$ with $\omega_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$ for $i, j = 1, 2, \ldots, K-1$. Then, using Result 2, it can be observed that the asymptotic distribution of $\mathbf{T}^*_{\mathbf{p}}$ under H_{0p} is $(K-1)$-variate normal with mean vector 0 and dispersion matrix Ω. Thus, $\left\{H_{0p}, \mathbf{T}^*_{\mathbf{p}}\right\}$ asymptotically constitutes joint testing family. The simultaneous test procedure (STP), described by the triple $\left\{H_{0p}, \mathbf{T}^*_{\mathbf{p}}, z_{(1-\alpha),2,\widehat{\Omega}}\right\}$, controls FWER in the strong sense (Konietschke et al. 2012), where $\widehat{\Omega}$ is a consistent estimator of Ω and $z_{(1-\alpha),2,\Omega}$ is the two-sided $(1-\alpha)$-equicoordinate quantile of a multivariate normal distribution with mean vector 0 and dispersion matrix Ω. Therefore, asymptotically level α two-sided single-step MCTP is given by

$$\text{Reject any } H_{0pl} \text{ if } |T^*_{pl}| > z_{(1-\alpha),2,\widehat{\Omega}}, l = 1, 2, \ldots, K-1. \tag{9}$$

Furthermore, associated with this MCTP, we also find the asymptotically $100(1-\alpha)\%$ simultaneous confidence intervals (SCIs)

$$\left[\mathbf{c}_l^T\widehat{\mathbf{p}}. - z_{(1-\alpha),2,\widehat{\Omega}}\sqrt{\widehat{\sigma}_{ll}/N}, \mathbf{c}_l^T\widehat{\mathbf{p}}. + z_{(1-\alpha),2,\widehat{\Omega}}\sqrt{\widehat{\sigma}_{ll}/N}\right] \tag{10}$$

on the component parameters $\mathbf{c}_l^T\mathbf{p}.$, $l = 1, 2, \ldots, K-1$. This gives an equivalent approach for applying MCTP in which H_{0pl} is rejected if 0 is not included in the interval for the lth component. In addition, this procedure is more informative than the corresponding adjusted *p* value approach where true dependent structure of the *p* values or the component test statistics is not taken into account.

3.2 Test for H_{0R}

Here, as earlier, tests are developed by using the statistics

$$\widehat{R}_{k\cdot} = \frac{1}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} e(j_1, j_2, \ldots, j_K)\, V_k(Y_{1j_1}, Y_{2j_2}, \ldots, Y_{Kj_K}),$$

$k = 1, 2, \ldots, K$, where $e(j_1, j_2, \ldots, j_K) = e_{1j_1} e_{2j_2} \cdots e_{Kj_K}$ is given by (7) and

$$\begin{aligned} V_k(Y_1, Y_2, \ldots, Y_K) &= 1 \quad \text{if } Y_k > \max(Y_m, m = 1, 2, \ldots, K; m \neq k) \\ &= \frac{1}{2} \quad \text{if } Y_k = Y_{k_1} > \max(Y_m, m = 1, 2, \ldots, K; m \neq k, k_1) \\ &= \frac{1}{3} \quad \text{if } Y_k = Y_{k_1} = Y_{k_2} \\ &\qquad > \max(Y_m, m = 1, 2, \ldots, K; m \neq k, k_1, k_2) \\ &\vdots \\ &= \frac{1}{K} \quad \text{if } Y_1 = Y_2 = \cdots = Y_K,\ k = 1, 2, \ldots, K. \end{aligned}$$

Now, similar to Results 1 and 2, we use the following results for testing H_{0R} against H_{1R}. The results are based on the same assumptions as in Result 1.

Result 3 $\widehat{R}_{k\cdot}$ *is a consistent estimator of* $R_{k\cdot}$, $k = 1, 2, \ldots, K$.

Result 4 *There exists a positive definite (p.d.) matrix* Σ_{CR} *such that the asymptotic distribution of* $\sqrt{N}C(\widehat{\mathbf{R}}_{\cdot} - \mathbf{R}_{\cdot})$ *is* $(K-1)$*-variate normal with mean vector* 0 *and dispersion matrix* Σ_{CR}*, where* $\widehat{\mathbf{R}}_{\cdot} = (\widehat{R}_{1\cdot}, \widehat{R}_{2\cdot}, \ldots, \widehat{R}_{K\cdot})^T$.

The elements of $\Sigma_{CR} = (\nu_{ij})$ are given in "Appendix."

3.3 Test for H_{0R*}

With a very similar way, as described for H_{0R}, tests can be developed by using the statistics

$$\widehat{R}^*_{k\cdot} = \frac{1}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} e(j_1, j_2, \ldots, j_K)\, V^*_k(Y_{1j_1}, Y_{2j_2}, \ldots, Y_{Kj_K}),$$

$k = 1, 2, \ldots, K$, where $e(j_1, j_2, \ldots, j_K) = e_{1j_1} e_{2j_2} \cdots e_{Kj_K}$ is given by (7) and

$$\begin{aligned} V^*_k(Y_1, Y_2, \ldots, Y_K) &= 1 \quad \text{if } Y_k < \min(Y_m, m = 1, 2, \ldots, K; m \neq k) \\ &= \frac{1}{2} \quad \text{if } Y_k = Y_{k_1} \end{aligned}$$

$$
\begin{aligned}
&\qquad < \min(Y_m, m = 1, 2, \ldots, K; m \neq k, k_1) \\
&= \frac{1}{3} \quad \text{if } Y_k = Y_{k_1} = Y_{k_2} \\
&\qquad < \min(Y_m, m = 1, 2, \ldots, K; m \neq k, k_1, k_2) \\
&\vdots \\
&= \frac{1}{K} \quad \text{if } Y_1 = Y_2 = \cdots = Y_K,\ k = 1, 2, \ldots, K.
\end{aligned}
$$

Under the same assumptions as in Result 1, we consider two results similar to Results 3 and 4.

Result 5 $\widehat{R}^*_{k.}$ *is a consistent estimator of* $R^*_{k.}$, $k = 1, 2, \ldots, K$.

Result 6 *There exists a p.d. matrix* Σ_{CR^*} *such that the asymptotic distribution of* $\sqrt{N}C(\widehat{\mathbf{R}}^*. - \mathbf{R}^*.)$ *is* $(K-1)$*-variate normal with mean vector* 0 *and dispersion matrix* Σ_{CR^*}*, where* $\widehat{\mathbf{R}}^*. = (\widehat{R}^*_{1}., \widehat{R}^*_{2}., \ldots, \widehat{R}^*_{K}.)^T$.

The elements of $\Sigma_{CR^*} = (\nu^*_{ij})$ are given in "Appendix."

The proofs of Results 3 and 5 are very similar to that of Result 1 and hence omitted. The proofs of Results 4 and 6 are given in "Appendix." Next, we apply the χ^2-test procedure and the MCTP for testing H_{0R} and H_{0R^*}. The procedures remain same as those described for H_{0p}. Thus, for the χ^2-test given by (8), we get the tests in which T_p is replaced by $T_R = N(C\widehat{\mathbf{R}}.)^T\widehat{\Sigma}^{-1}_{CR}(C\widehat{\mathbf{R}}.)$ and $T^*_R = N(C\widehat{\mathbf{R}}^*.)^T\widehat{\Sigma}^{-1}_{CR^*}(C\widehat{\mathbf{R}}^*.)$ for the two problems, respectively, where $\widehat{\Sigma}_{CR}$ and $\widehat{\Sigma}_{CR^*}$ are consistent estimators of Σ_{CR} and Σ_{CR^*}. Similarly, the tests for the MCTP are given by (9) and (10) with T^*_{pl} replaced by $T^*_{Rl} = \frac{\sqrt{N}\mathbf{c}_l^T\widehat{\mathbf{R}}.}{\sqrt{\widehat{\nu}_{ll}}}$ and $T^*_{R^*l} = \frac{\sqrt{N}\mathbf{c}_l^T\widehat{\mathbf{R}}^*.}{\sqrt{\widehat{\nu}^*_{ll}}}$ in which $\widehat{\nu}_{ll}$ and $\widehat{\nu}^*_{ll}$, $l = 1, 2, \ldots, K-1$ are, respectively, the diagonal elements of $\widehat{\Sigma}_{CR}$ and $\widehat{\Sigma}_{CR^*}$, and $\widehat{\Omega}$ is replaced by the consistent estimators, $\widehat{\Lambda}$ and $\widehat{\Lambda}^*$, of $\Lambda = \left(\left(\frac{\nu_{ij}}{\sqrt{\nu_{ii}\nu_{jj}}}\right)\right)$ and $\Lambda^* = \left(\left(\frac{\nu^*_{ij}}{\sqrt{\nu^*_{ii}\nu^*_{jj}}}\right)\right)$, respectively.

3.4 Competitor 1

We consider the test provided by Tsangari and Akritas (2004) for the testing H_{0A} : $C\theta. = 0$, where C and $\theta.$ are defined in Sect. 2. Moreover, H_{0p}, H_{0R} and H_{0R^*} are implied by H_{0A}, and hence it is more restricted than ours. Under some specific assumptions (Tsangari and Akritas 2004), it can be shown that for some positive definite (p.d.) matrix V,

$$\sqrt{N}C\widehat{\mathbf{T}} = \sqrt{N}C\int \widehat{H}d\widehat{\theta}. \to N_{K-1}\left(0, CVC^T\right) \tag{11}$$

in distribution when H_{0A} is true, where

$$\widehat{H}(y) = \frac{1}{2N} \sum_{k=1}^{K} \sum_{j=1}^{n_k} \{I(Y_{kj} \le y) + I(Y_{kj} < y)\}.$$

By (11), χ^2-test procedure and MCTP can be developed for H_{0A} in the similar way as described for H_{0p}, H_{0R} and H_{0R^*}. The elements of V, together with its consistent estimator $\widehat{V}$, are given by Tsangari and Akritas (2004). We refer the test described above as *TA test*.

4 Alternative tests

Here, we follow Bathke and Brunner (2003) to provide an alternative method by which, as in the parametric ANCOVA model, the relative effect of a treatment can be expressed in terms of a linear function of the relative treatment effects of the response variables and the covariates associated with the treatment. Moreover, the relative treatment effects are measured by ridit reliability functionals but, unlike the previous sections, they use marginal distributions of the concerned random variables. Thus, representing the relative effect due to treatment k by $\pi_k^{(0)}$ and that due to the rth component of the covariate associated with the treatment by $\pi_k^{(r)}$, $r = 1, 2, \ldots, d$, the covariate-eliminated main effect due to treatment k is defined by

$$\pi_k = \pi_k^{(0)} - \sum_{r=1}^{d} \beta_r \pi_k^{(r)},$$

where β_r is constant reflecting the dependency of the rth component of the covariate vector to the response variable corresponding to the kth treatment through the given functional, $r = 1, 2, \ldots, d$ and $k = 1, 2, \ldots, K$. Then, with the contrast matrix C defined in Sect. 2, the testing problem becomes

$$H_{0\pi} : C\boldsymbol{\pi} = 0 \;\text{ against }\; H_{1\pi} : C\boldsymbol{\pi} \neq 0,$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_K)^T$.

In particular, for BP-type functional, we obtain

$$\pi_k^{(0)} = \frac{1}{K}\left[\frac{1}{2} + \sum_{k'=1, k' \neq k}^{K} \pi_{kk'}^{(0)}\right] \text{ with}$$
$$\pi_{kk'}^{(0)} = P(Y_k > Y_{k'}) + \frac{1}{2} P(Y_k = Y_{k'}),$$

$$\pi_k^{(r)} = \frac{1}{K}\left[\frac{1}{2} + \sum_{k'=1,k'\neq k}^{K} \pi_{kk'}^{(r)}\right] \text{ with}$$

$$\pi_{kk'}^{(r)} = P(X_k^{(r)} > X_{k'}^{(r)}) + \frac{1}{2}P(X_k^{(r)} = X_{k'}^{(r)}),$$

where $r = 1, 2, \ldots, d$ and $k = 1, 2, \ldots, K$. Similarly, based on the marginal distribution of the random variables $(Y_1, Y_2, \ldots, Y_K)^T$ and $(X_1^{(r)}, X_K^{(r)}, \ldots, X_K^{(r)})^T$, we can form different sets of $\left\{\pi_k^{(0)}, \pi_k^{(r)}, r = 1, 2, \ldots, d,\ k = 1, 2, \ldots, K\right\}$ from BD-type and BC-type functionals (Bandyopadhyay and Chatterjee 2015).

Next, write $\widehat{\boldsymbol{\pi}}$ as a consistent estimator of $\boldsymbol{\pi}$ obtained by the same technique as in Sect. 3 using the sample measures of the concerned marginal d.f.'s. Then, similar to Results 4 and 6, we get the following result that provides alternative tests based on three choices of $\boldsymbol{\pi}$ as mentioned above.

Result 7 *There exists a p.d. matrix* $\Sigma_{C\pi}$ *such that, under the assumption A2, the asymptotic distribution of* $\sqrt{N}C(\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi})$ *is* $(K-1)$*-variate normal with mean vector* 0 *and dispersion matrix* $\Sigma_{C\pi}$.

The details of the dispersion matrix $\Sigma_{C\pi}$ and its consistent estimator are given in "Appendix" for three different choices of $\boldsymbol{\pi}$. Now, by Result 7 together with the associated consistent estimators, we can construct tests for $H_{0\pi}$ using both the χ^2-test procedure and the MCTP as described in Sect. 3. In this way, we get three different tests for each procedure from three different choices of $\boldsymbol{\pi}$ mentioned earlier. Such tests are referred to as *Alt. BP-type test*, *Alt. BD-type test* and *Alt. BC-type test*, respectively.

Competitor 2

Bathke and Brunner (2003), following some assumptions which include the homogeneity of the covariate distributions with respect to treatment groups, provide a test for the equality of the treatment effects by considering the testing problem

$$H_{0B} : C\mathbf{F}_0 = 0 \text{ against } H_{1B} : C\mathbf{F}_0 \neq 0,$$

where C is the contrast matrix as defined in Sect. 2 and $\mathbf{F}_0 = (F_{01}, F_{02}, \ldots, F_{0K})^T$ with F_{0k} as the marginal d.f. of Y_k, $k = 1, 2, \ldots, K$.

Now, with the definitions of $\widehat{\gamma}^{(r)}, \widehat{q}_k^{(r)}$ for $r = 0, 1, \ldots, d,\ k = 1, 2, \ldots, K$ and $\widehat{\Sigma}_N$ given in Bathke and Brunner (2003), we find

$$\widehat{q}_k^{*} = \widehat{q}_k^{(0)} - \sum_{r=1}^{d} \widehat{\gamma}^{(r)}\widehat{q}_k^{(r)},\ k = 1, 2, \ldots, K \text{ and } \widehat{\Sigma}_N^{*} = \widehat{\Gamma}^T\widehat{\Sigma}_N\widehat{\Gamma},$$

where

$$\widehat{\Gamma} = \begin{pmatrix} 1 \\ -\widehat{\boldsymbol{\gamma}} \end{pmatrix} \otimes I_K \text{ and } \widehat{\boldsymbol{\gamma}} = \left(\widehat{\gamma}^{(1)}, \widehat{\gamma}^{(2)}, \ldots, \widehat{\gamma}^{(d)}\right)^T.$$

Then, the statistics (Bathke and Brunner 2003)

$$Q_N = N(C\widehat{\mathbf{q}}^*)^T(C\widehat{\Sigma}_N^* C^T)^+ C\widehat{\mathbf{q}}^*$$

and

$$A_N = \frac{N\widehat{f_2}(\widehat{\mathbf{q}}^*)^T M\widehat{\mathbf{q}}^*}{tr(M\widehat{\Sigma}_N^*)},$$

where $M = C^T(CC^T)^{-1}C$ and $\widehat{\mathbf{q}}^* = \left(\widehat{q}_1^*, \widehat{q}_2^*, \ldots, \widehat{q}_K^*\right)^T$, follow asymptotically χ^2 distributions under H_{0B} with degrees of freedom estimated respectively by

$$\widehat{f_1} = \text{rank}(C\widehat{\Sigma}_N^*) \text{ and } \widehat{f_2} = \frac{[tr(M\widehat{\Sigma}_N^*)]^2}{tr(M\widehat{\Sigma}_N^* M\widehat{\Sigma}_N^*)}$$

We refer the tests based on the above two statistics as *Bathke* Q_N *test* and *Bathke* A_N *test*.

Remark 1 Note that for each of the choices of $\boldsymbol{\pi}$, the hypothesis, $H_{0\pi}$, is more general than that provided by Bathke and Brunner (2003) and it differs from the hypotheses defined in Sect. 2. Hence, the tests that are developed in this section are not directly comparable with those developed in Sect. 3. Now, to compare the tests, we need to modify those by estimating the dispersion matrices under H_{0B} with the assumptions considered by Bathke and Brunner (2003). The dispersion matrix estimators are given in “Appendix.” In addition, we can construct a *Combined test* for H_{0B} based on *Alt. BD-type test* and *Alt. BC-type test* using the approach that is described in Bandyopadhyay and Chatterjee (2015). However, *Combined test* based on *BD-type test* and *BC-type test* is not possible, since the choice of the bandwidth constant for the two tests is not same.

Remark 2 Due to some specific assumptions on bandwidth (Tsangari and Akritas 2004), *TA test* cannot be applied for more than three covariates. On the other hand, there is no restriction on the number of covariates for our tests as well as the tests based on the alternative method.

5 Small sample approximation and simulation results

In the simulation study, we use the Gaussian kernel function

$$\psi(x) = \frac{1}{\sqrt{2\pi}(2\Phi(10) - 1)} e^{-\frac{x^2}{2}} I(|x| \leq 10)$$

on the compact support $[-10, 10]$ for the tests developed in Sect. 3, where $\Phi(x)$ is the standard normal d.f. Clearly, the final expressions of the asymptotic distributions of $\sqrt{N}(\widehat{\mathbf{q}}.-\mathbf{q}.)$, $\sqrt{N}(\widehat{\mathbf{R}}.-\mathbf{R}.)$ and $\sqrt{N}(\widehat{\mathbf{R}}^*.-\mathbf{R}^*.)$ are free from $\Psi(\cdot)$. Therefore, the given consistent estimators and the developed test procedures do not depend on particular choice of the kernel function, provided it satisfies the conditions mentioned in Sect. 3.1. Note that the test procedures described in Sects. 3 and 4, including *competitor 1*, are valid for large sample sizes. The simulation study shows that the rates of convergence of $\sqrt{N}C(\widehat{\mathbf{p}}.-\mathbf{p}.)$, $\sqrt{N}C(\widehat{\mathbf{R}}.-\mathbf{R}.)$, $\sqrt{N}C(\widehat{\mathbf{R}}^*.-\mathbf{R}^*.)$, $\sqrt{N}C\int\widehat{H}d\widehat{\boldsymbol{\theta}}.$ and $\sqrt{N}C(\widehat{\boldsymbol{\pi}}-\boldsymbol{\pi})$ toward their asymptotic distributions are rather slow. Hence, the tests, conducted with the help of such asymptotic distributions, are not appropriate when the sample sizes are small. So, in that situation, we adopt Box-type approximation (Brunner et al. 1997 and Gao et al. 2008) by which the distributions of $\sqrt{N}C(\widehat{\mathbf{p}}.-\mathbf{p}.)$, $\sqrt{N}C(\widehat{\mathbf{R}}.-\mathbf{R}.)$, $\sqrt{N}C(\widehat{\mathbf{R}}^*.-\mathbf{R}^*.)$ and $\sqrt{N}C(\widehat{\boldsymbol{\pi}}-\boldsymbol{\pi})$ are approximated by multivariate t-distributions with ν_p, ν_R, ν_{R^*} and ν_π degrees of freedom having common expectation 0 and dispersion matrices $\widehat{\Sigma}_{Cp}$, $\widehat{\Sigma}_{CR}$, $\widehat{\Sigma}_{CR^*}$ and $\widehat{\Sigma}_{C\pi}$, respectively. Hence, instead of the χ^2-test, F-test is performed by using the approximation

$$N(C\widehat{\boldsymbol{\phi}})^T\widehat{\Sigma}_{C\phi}^{-1}(C\widehat{\boldsymbol{\phi}})/(K-1)\sim F_{(K-1),\nu_\phi}$$

under $H_{0\pi}$ with $\phi=p, R, R^*$ and π, where $F_{a,b}$ denotes the F-distribution with a and b degrees of freedom. For the MCTP, we perform the tests in which the two-sided $(1-\alpha)$-equicoordinate quantiles of multivariate normal distribution are replaced by that of the corresponding multivariate t-distribution with suitable degrees of freedom. The degrees of freedom can be computed in the same way as that described by Konietschke et al. (2012).

On the other hand, setting D_M as the diagonal matrix with the diagonal elements of $M=C^T[CC^T]^{-1}C$ (Brunner et al. 1997), the χ^2-test, provided by Tsangari and Akritas, can be replaced by an F-test in which the distribution of the test statistic

$$\frac{N}{tr(D_M\widehat{V})}\widehat{\mathbf{T}}^TM\widehat{\mathbf{T}}$$

under H_{0A} is approximated by $F_{\widehat{f}_3,\widehat{f}_4}$, where $\widehat{f}_3=\frac{[tr(D_M\widehat{V})]^2}{tr(M\widehat{V}M\widehat{V})}$, $\widehat{f}_4=\frac{[tr(D_M\widehat{V})]^2}{tr(D_M^2\widehat{V}^2\Lambda_N)}$ with $\Lambda_N=\text{diag}\left\{(n_l-1)^{-1}, l=1,2,\ldots,K\right\}$.

Simulation study is performed considering both continuous and discrete cases. For the tests that are developed in Sect. 3, including *competitor 1*, our first target is to find the best choice of the bandwidth by comparing the empirical levels with the corresponding nominal levels and also verifying the QQ-plots. Next, with that particular choice of the bandwidth, the tests are compared with respect to their empirical powers. Note that all the tests, that appear in this study, correspond to different testing problems. Therefore, we must restrict ourselves to a single testing problem to get a meaningful comparison of the performance of the tests. Thus, for simulation purpose, we consider the testing of equality of the covariate eliminating distribution functions when the covariate distributions are different for different treatment groups. The same problem is considered by Tsangari and Akritas (2004). For the given testing problem, we slightly modify the test statistics of *BP-type test*, *BD-type test* and *BC-type test* by estimating the covariance matrices under H_{0A}. The detail expressions of such modified test statistics are given in "Appendix" (*Note 3*). However, the asymptotic distributions of the test statistics remain unchanged. On the other hand, the alternative procedure deals with the marginal distributions of the associated random variables and obtain tests for the equality of the covariate-eliminated treatment effects by considering a linear model of the relative treatment effects on the effects of the respective covariates.

A challenging task for testing the equality of the treatment effects under the given setup is to construct test that attains nominal levels even when the sample sizes are not very large. After detailed investigation, it is observed that there exist situations where *TA test* fails to attain the nominal levels even when the sample sizes are large enough. However, *BP-type*, *BD-type* and *BC-type tests* reach the target quite appropriately. For illustration, let us consider a simple model with single continuous covariate as $Y_{kj} = 1 + 0.5X_{kj} + E_{kj}$, $j = 1, 2, \ldots, n_k$, $k = 1, 2, 3$ in which E_{kj} are iid (independently and identically distributed) according to $N(0, 1)$; X_{1j} are iid according to $|B_1|$ with $B_1 \sim N(0, 1)$ restricted on $(-4, 4)$; X_{2j} are iid according to $|B_2|$ with $B_2 \sim DE(0, 1)$ restricted on $(-4, 4)$; X_{3j} are iid according to $|B_3|$ with $B_3 \sim C(0, 3)$ restricted on $(-4, 4)$, where $N(\mu, \sigma^2)$, $DE(\mu, \sigma)$ and $C(\mu, \sigma)$ represent, respectively, normal, double exponential and Cauchy distributions with location μ and scale σ. The model clearly shows that the main effects due to treatments are equal for all the groups. For our tests, we consider the form of the bandwidth as $a_n = (\text{const})n^{-0.2}$ because this is the optimum form as shown by Silverman (1986, Sect. 4.3). For the competitor, we take the bandwidth $a_n = (\text{const})n^{-0.26}$ because of the assumption for continuous covariate considered by Akritas et al. (2000). Tables 1, 2, 3 and 4 present the empirical levels of the tests using both large sample distributions and also small sample approximate distributions corresponding to different choices of bandwidth constants. Here, we consider four different sample sizes described by *sample size 1* : $N = 37$ with $(n_1, n_2, n_3) = (11, 12, 14)$; *sample size 2* : $N = 75$ with $(n_1, n_2, n_3) = (24, 25, 26)$; *sample size 3* : $N = 120$ with $(n_1, n_2, n_3) = (35, 40, 45)$; *sample size 4* : $N = 195$ with $(n_1, n_2, n_3) = (60, 65, 70)$. The nominal level, α, is taken as 0.05, and consider that bandwidth constant to be the appropriate one for which the empirical levels of the tests attain the nominal level at each of the four different sample sizes. On the other hand, performances of the tests based on alternative procedure with respect to empirical levels are shown in Table 5 for different sample sizes. A total of 10,000 simulated samples, for each N, are generated to compute empirical levels and powers using the software *R 3.2.1*, and equicoordinate quantiles of multivariate normal and t distributions are calculated by using R-package *mvtnorm* (R Development Core Team 2013).

From Tables 1, 2, 3, 4 and 5, we observe 1.2 to be an appropriate bandwidth constant for *BP-type test*, whereas for *BD-* and *BC-type tests* these are 1.1 and 1.55, respectively. However, for *sample size 1* and *sample size 2*, the large sample tests become little liberal and thus we can perform the tests using the small sample distributions. On the other hand, we find no satisfactory result from *TA test* toward finding an appropriate bandwidth even when the sample sizes are large enough. Moreover, we observe that, in this setup, the empirical levels of *BP-type*, *BD-type* and *BC-type tests* increase with the increase in the bandwidth constant. Thus, it is not difficult to find the band-

Table 1 Empirical levels for BP-type test: χ^2-test procedure along with MCTP in the bracket

Test	Constant	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.0762	0.0468	0.0428	0.0378
	0.75	(0.0717)	(0.0469)	(0.0422)	(0.0371)
		0.0836	0.0543	0.0482	0.0466
	1.0	(0.0764)	(0.0544)	(0.0503)	(0.0458)
		0.0889	0.0641	0.0553	0.0512
	1.2	(0.0841)	(0.0620)	(0.0563)	(0.0502)
		0.1037	0.0762	0.0776	0.0710
	1.5	(0.0928)	(0.0775)	(0.0760)	(0.0677)
		0.1268	0.1065	0.1149	0.1225
	2.0	(0.1128)	(0.0972)	(0.1066)	(0.1165)
		0.1418	0.1323	0.1540	0.1748
	2.5	(0.1246)	(0.1202)	(0.1399)	(0.1618)
Small sample		0.0499	0.0386	0.0364	0.0350
	0.75	(0.0496)	(0.0405)	(0.0379)	(0.0344)
		0.0584	0.0463	0.0425	0.0425
	1.0	(0.0550)	(0.0447)	(0.0445)	(0.0429)
		0.0595	0.0530	0.0494	0.0515
	1.2	(0.0562)	(0.0523)	(0.0506)	(0.0514)
		0.0721	0.0653	0.0697	0.0672
	1.5	(0.0672)	(0.0668)	(0.0688)	(0.0642)
		0.0925	0.0929	0.1033	0.1158
	2.0	(0.0842)	(0.0858)	(0.0975)	(0.1109)
		0.1085	0.1156	0.1411	0.1672
	2.5	(0.0926)	(0.1043)	(0.1277)	(0.1557)

width constants appropriately using such tests. Furthermore, the empirical levels of the alternative tests are found inappropriate.

Next, we compare the performance of the four tests in terms of empirical power. For this, we consider the model: $Y_{kj} = 1 + \tau_k + 0.5X_{kj} + E_{kj}$, $j = 1, 2, \ldots, n_k$, $k = 1, 2, 3$, where E_{kj} are iid according to $N(0, 1)$; X_{1j} are iid according to $N(0, 1)$ restricted on $(-2, 2)$; X_{2j} are iid according to Uniform $(-2, 2)$; X_{3j} are iid according to $C(1, 1)$ restricted on $(-2, 2)$. Clearly, the model represents a simple linear structure with standard normal distribution of the error components and continuous covariates. Also, note that the covariate distributions are different in shapes, scales and even in locations for the three groups. The primary task is to find the bandwidth constants for which the tests, developed in Sect. 3, perform most accurately taking the same forms of the bandwidths and the different sample sizes as mentioned earlier. From the simulation study, we find the best choice of the bandwidth constant for each test. The constants are $1.45, 1.45, 1.55$ and 1.2 for *BP*-, *BD*- and *BC-type tests* and *TA tests*, respectively. Table 6 shows the empirical levels of the four tests corresponding

Table 2 Empirical levels for BD-type test: χ^2-test procedure along with MCTP in the bracket

Test	Constant	Sample size 1	Sample size 2	Sample size 3	sample size 4
Large sample		0.0754	0.0483	0.0474	0.0379
	0.75	(0.0703)	(0.0493)	(0.0494)	(0.0408)
		0.0818	0.0535	0.0511	0.0427
	1.0	(0.0756)	(0.0543)	(0.0529)	(0.0472)
		0.0837	0.0608	0.0544	0.0516
	1.1	(0.0779)	(0.0611)	(0.0518)	(0.0532)
		0.1041	0.0828	0.0783	0.0748
	1.5	(0.0944)	(0.0743)	(0.0779)	(0.0772)
		0.1289	0.1076	0.1170	0.1344
	2.0	(0.1173)	(0.1043)	(0.1149)	(0.1307)
		0.1507	0.1430	0.1768	0.1968
	2.5	(0.1262)	(0.1306)	(0.1595)	(0.1826)
Small sample		0.0497	0.0399	0.0425	0.0354
	0.75	(0.0485)	(0.0409)	(0.0449)	(0.0389)
		0.0548	0.0449	0.0463	0.0398
	1.0	(0.0539)	(0.0473)	(0.0472)	(0.0448)
		0.0548	0.0510	0.0472	0.0484
	1.1	(0.0530)	(0.0519)	(0.0470)	(0.0506)
		0.0724	0.0692	0.0713	0.0703
	1.5	(0.0684)	(0.0644)	(0.0712)	(0.0732)
		0.0926	0.0927	0.1078	0.1274
	2.0	(0.0861)	(0.0896)	(0.1058)	(0.1249)
		0.1138	0.1230	0.1642	0.1873
	2.5	(0.0964)	(0.1137)	(0.1485)	(0.1763)

to the given choices of the bandwidth constants using large sample and small sample procedures. Empirical levels of the alternative tests (Table 7) for different sample sizes are also obtained setting $\tau_1 = \tau_2 = \tau_3 = 0$.

From Tables 6 and 7, it is observed that empirical levels of *BP*-, *BD*- and *BC-type tests*, based on large sample procedures, are slightly larger than the nominal level, $\alpha = 0.05$, for *sample size 1*, whereas in case of *TA test* such levels are significantly larger than the chosen nominal level for the sample sizes considered here. In addition, we can use small sample approximate test procedure here with the given choice of the bandwidth constant. Finally, we compare the empirical powers of the tests for *sample size 3* corresponding to different types of alternatives.

Clearly, for *BP*-, *BD*- and *BC-type tests* we can use large sample procedures and for *TA test* we use small sample approximate test. Unfortunately, as before, empirical levels of the tests correspond to the alternative procedure are significantly larger than the chosen nominal level. Hence, these tests are not suitable for comparison. Table 8 gives the empirical powers of the four tests. For *BP-, BD-* and *BC-type tests*, we

Table 3 Empirical levels for BC-type test: χ^2-test procedure along with MCTP in the bracket

Test	Constant	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.0634	0.0435	0.0350	0.0311
	1.0	(0.0606)	(0.0421)	(0.0366)	(0.0304)
		0.0675	0.0536	0.0402	0.0392
	1.25	(0.0632)	(0.0511)	(0.0422)	(0.0392)
		0.0750	0.0639	0.0543	0.0476
	1.45	(0.0666)	(0.0606)	(0.0480)	(0.0453)
		0.0795	0.0613	0.0534	0.0497
	1.55	(0.0754)	(0.0571)	(0.0551)	(0.0501)
		0.0918	0.0782	0.0640	0.0654
	1.75	(0.0821)	(0.0701)	(0.0621)	(0.0634)
		0.0983	0.0882	0.0853	0.0871
	2.0	(0.0891)	(0.0813)	(0.0767)	(0.0793)
Small sample		0.0395	0.0337	0.0286	0.0290
	1.0	(0.0418)	(0.0348)	(0.0326)	(0.0277)
		0.0428	0.0441	0.0344	0.0359
	1.25	(0.0432)	(0.0440)	(0.0371)	(0.0361)
		0.0485	0.0537	0.0476	0.0444
	1.45	(0.0473)	(0.0519)	(0.0435)	(0.0432)
		0.0526	0.0507	0.0474	0.0474
	1.55	(0.0535)	(0.0487)	(0.0482)	(0.0477)
		0.0616	0.0650	0.0574	0.0606
	1.75	(0.0575)	(0.0588)	(0.0566)	(0.0594)
		0.0652	0.0737	0.0761	0.0819
	2.0	(0.0619)	(0.0695)	(0.0693)	(0.0745)

provide empirical powers corresponding to the χ^2-test procedure along with that of the MCTP in the bracket. Empirical power comparison among the different tests under the given model shows that none of the tests can be taken as the best considering all types of alternatives at a time. Specifically, it is observed that *BP-type test* for the χ^2-test procedure has the highest empirical power under *convex* and *linear* alternatives, whereas for *concave* alternative, *BP*- and *BC-type tests* provide almost same empirical power and the corresponding MCTP has significantly larger empirical power than the others. However, under *umbrella* alternative, *TA test* gives the maximum power. On the other hand, *BC-type test* under *U-shaped* alternative produces larger empirical power than that of its competitors.

We also compare the empirical powers of the tests under discrete response model. Here, we take single covariate and consider the response model: $Y_{kj} \sim \mathrm{Bernoulli}(\mu_k)$ with $\log\left(\frac{\mu_k}{1-\mu_k}\right) = 0.25 + \tau_k + 0.25X_{kj},\ j = 1, 2, \ldots, n_k, k = 1, 2, 3$, where X_{kj} are iid according to $N(0, 1)$ or $C(0, 1)$ restricted on $(-2, 2)$. As earlier, we first compute the bandwidth constants for the four tests by setting $\tau_1 = \tau_2 = \tau_3 = 0$. The

Table 4 Empirical levels for TA test: χ^2-test procedure along with MCTP in the bracket for large sample and approximate F test for small sample

Test	Constant	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.1759	0.1292	0.1174	0.1055
	0.5	(0.1554)	(0.1215)	(0.1164)	(0.1056)
		0.1348	0.1017	0.0936	0.0850
	0.75	(0.1210)	(0.0955)	(0.0924)	(0.0860)
		0.1252	0.0953	0.0844	0.0725
	1.0	(0.1111)	(0.0889)	(0.0811)	(0.0712)
		0.1178	0.0852	0.0798	0.0704
	1.2	(0.1046)	(0.0788)	(0.0786)	(0.0684)
		0.1140	0.0864	0.0848	0.0784
	1.5	(0.1005)	(0.0834)	(0.0818)	(0.0752)
		0.1232	0.1064	0.1058	0.0994
	2.0	(0.1088)	(0.0965)	(0.0930)	(0.0950)
		0.1431	0.1282	0.1363	0.1377
	2.5	(0.1279)	(0.1151)	(0.1175)	(0.1258)
Small sample	0.5	0.1205	0.0985	0.0989	0.0892
	0.75	0.0891	0.0791	0.0761	0.0724
	1.0	0.0817	0.0709	0.0703	0.0625
	1.2	0.0790	0.0671	0.0669	0.0615
	1.5	0.0745	0.0697	0.0694	0.0679
	2.0	0.0861	0.0869	0.0894	0.0902
	2.5	0.1004	0.1084	0.1177	0.1279

Table 5 Empirical levels for alternative tests: χ^2-test procedure along with MCTP in the bracket

	Test	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.1404	0.1119	0.1048	0.1055
	Alt. BP type	(0.1193)	(0.0986)	(0.0918)	(0.0976)
		0.1293	0.0980	0.0894	0.0798
	Alt. BD type	(0.0905)	(0.0778)	(0.0818)	(0.0782)
		0.1416	0.1012	0.0911	0.0886
	Alt. BC type	(0.0911)	(0.0687)	(0.0688)	(0.0671)
	Bathke Q_N	0.1113	0.0890	0.0895	0.0908
Small sample		0.1093	0.0958	0.0941	0.0991
	Alt. BP type	(0.0935)	(0.0867)	(0.0851)	(0.0937)
		0.0997	0.0838	0.0808	0.0763
	Alt. BD type	(0.0664)	(0.0672)	(0.0752)	(0.0744)
		0.1073	0.0906	0.0828	0.0834
	Alt. BC type	(0.0648)	(0.0602)	(0.0623)	(0.0620)
	Bathke A_N	0.0925	0.0810	0.0822	0.0862

Table 6 Empirical levels of the tests: χ^2-test procedure along with MCTP in the bracket

	Test	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.0826	0.0618	0.0532	0.0508
	BP type	(0.0742)	(0.0585)	(0.0508)	(0.0501)
		0.0892	0.0585	0.0503	0.0523
	BD type	(0.0803)	(0.0575)	(0.0514)	(0.0520)
		0.0795	0.0586	0.0534	0.0497
	BC type	(0.0754)	(0.0584)	(0.0551)	(0.0501)
		0.1145	0.0809	0.0687	0.0643
	TA test	(0.1002)	(0.0727)	(0.0635)	(0.0622)
Small sample		0.0576	0.0539	0.0479	0.0502
	BP type	(0.0540)	(0.0509)	(0.0488)	(0.0503)
		0.0574	0.0499	0.0464	0.0488
	BD type	(0.0561)	(0.0494)	(0.0461)	(0.0489)
		0.0526	0.0507	0.0474	0.0459
	BC type	(0.0535)	(0.0487)	(0.0482)	(0.0477)
	TA test	0.0755	0.0583	0.0522	0.0508

Table 7 Empirical levels for alternative tests: χ^2-test procedure along with MCTP in the bracket

	Test	Sample size 1	Sample size 2	Sample size 3	Sample size 4
Large sample		0.1174	0.0894	0.0840	0.0779
	BP type	(0.1008)	(0.0807)	(0.0779)	(0.0699)
		0.1175	0.0908	0.0860	0.0798
	BD type	(0.0805)	(0.0740)	(0.0757)	(0.0730)
		0.1451	0.0983	0.0924	0.0851
	BC type	(0.0919)	(0.0719)	(0.0642)	(0.0656)
	Bathke Q_N	0.0988	0.0816	0.0704	0.0688
Small sample		0.0863	0.0759	0.0751	0.0744
	BP type	(0.0774)	(0.0701)	(0.0700)	(0.0662)
		0.0869	0.0789	0.0777	0.0740
	BD type	(0.0574)	(0.0656)	(0.0687)	(0.0695)
		0.1107	0.0859	0.0838	0.0793
	BC type	(0.0652)	(0.0626)	(0.0579)	(0.0618)
	Bathke A_N	0.0911	0.0728	0.0653	0.0657

constants come out as $2.0(2.0)$, $2.0(3.25)$, $2.75(2.5)$ and $4.5(3.0)$ for *BP-type*, *BD-type*, *BC-type* and *TA tests*, respectively, when the covariates are iid $N(0, 1)(C(0, 1))$. We also perform the simulation for the alternative tests. In both the situations, it is observed that the large sample procedures have inflated empirical type I error rates when the sample sizes are small. On the other hand, small sample procedures seem to be

Table 8 Empirical powers of the four tests

Alternative	(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	TA test
Convex	(0,0,0.5)	0.6820 (0.5766)	0.6687 (0.5987)	0.6215 (0.4836)	0.6218
Concave	(0,0.5,0.5)	0.6087 (0.6599)	0.5759 (0.6127)	0.5974 (0.6546)	0.5789
Umbrella	(0,0.5,0)	0.4517 (0.4208)	0.4653 (0.4314)	0.3727 (0.3413)	0.5134
U-shaped	(0.5,0,0.5)	0.5471 (0.3548)	0.4876 (0.2883)	0.5689 (0.3963)	0.5539
Linear	(0,0.2,0.5)	0.5694 (0.5788)	0.5455 (0.5642)	0.5171 (0.5241)	0.5079

Table 9 Empirical powers of the four tests considering normal distribution for the covariate

Alternative	(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	TA test
Under H_0	(0,0,0)	0.0515 (0.0479)	0.0573 (0.0568)	0.0577 (0.0557)	0.0558
Convex	(0,0,1.2)	0.5094 (0.3846)	0.5469 (0.4436)	0.4987 (0.3648)	0.4773
Concave	(0,1.2,1.2)	0.4508 (0.5193)	0.4367 (0.5071)	0.4827 (0.5481)	0.5018
Umbrella	(0,1.2,0)	0.5081 (0.3889)	0.5311 (0.4253)	0.4956 (0.3639)	0.4846
U-shaped	(1.2,0,1.2)	0.4635 (0.3700)	0.4460 (0.3401)	0.5040 (0.4210)	0.5065
Linear	(0,0.6,1.2)	0.3793 (0.4014)	0.3850 (0.4034)	0.3902 (0.4149)	0.3865

Table 10 Empirical powers of the four alternative tests considering normal distribution for the covariate

Alternative	(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	Bathke A_N test
Under H_0	(0,0,0)	0.0547 (0.0555)	0.0544 (0.0520)	0.0606 (0.0529)	0.0635
Convex	(0,0,1.2)	0.5189 (0.4053)	0.4774 (0.3888)	0.5519 (0.4025)	0.4943
Concave	(0,1.2,1.2)	0.4632 (0.5326)	0.4720 (0.5213)	0.4653 (0.5373)	0.5241
Umbrella	(0,1.2,0)	0.5156 (0.4000)	0.4746 (0.3786)	0.5487 (0.3864)	0.4958
U-shaped	(1.2,0,1.2)	0.4825 (0.3854)	0.4756 (0.3797)	0.4761 (0.3766)	0.5039
Linear	(0,0.6,1.2)	0.3870 (0.4214)	0.3752 (0.4029)	0.4077 (0.4191)	0.3742

appropriate even for small sample sizes. Thus, considering small sample procedures, we compute empirical powers, shown in Tables 9 and 10, for *sample size 2* taking covariate distribution as normal.

From Tables 9 and 10, it is observed that, except *Alt. BC-type test* for the χ^2 procedure and *Bathke* A_N *test*, empirical levels of all the other tests attain the nominal level. On the other hand, it is seen that the empirical powers of *BD-type test* and *Alt. BC-type test* for the χ^2 procedure are larger than the others under *Convex* alternative. The MCTP corresponding to *BC-type test* shows the highest power for *Concave* alternative. Under *Umbrella* alternative, *Alt.BC-type test* for the χ^2 procedure gives the highest power, but the empirical level of this test is slightly higher than the chosen nominal level. It is further observed that under *linear* alternative, *Alt. BP-type test* for the MCTP has the maximum empirical power, whereas *BC-type test* for the χ^2 procedure, *TA test* and *Bathke* A_N *test* give almost same empirical power under *U-shaped* alternative.

Table 11 Empirical powers of the four tests

Alternative	(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	TA test
Under H_0	(0,0,0)	0.0544 (0.0555)	0.0552 (0.0563)	0.0507 (0.0493)	0.0582
Convex	(0,0,0.6)	0.5648 (0.5080)	0.4944 (0.4787)	0.4856 (0.4014)	0.5645
Concave	(0,0.6,0.6)	0.5447 (0.5846)	0.4862 (0.5108)	0.4963 (0.5504)	0.5433
Umbrella	(0,0.6,0)	0.3208 (0.3170)	0.3502 (0.3472)	0.2441 (0.2427)	0.3793
U-shaped	(0.6,0,0.6)	0.3736 (0.2390)	0.2681 (0.1642)	0.4042 (0.2847)	0.4284
Linear	(0,0.3,0.6)	0.4827 (0.5105)	0.4331 (0.4647)	0.4320 (0.4553)	0.4814

The corresponding values under Cauchy covariate are also computed, and we get, not shown here, almost similar interpretation.

Furthermore, we consider a simple linear model with two covariates having three groups as: $Y_{kj} = 1 + \tau_k + 0.5X_{kj} + 0.5Z_{kj} + E_{kj}, j = 1, 2, \ldots, n_k, k = 1, 2, 3$, where E_{kj} are iid according to $N(0, 1)$; X_{1j} are iid according to $N(0, 1)$ restricted on $(-2, 2)$; X_{2j} are iid according to Uniform $(-2, 2)$; X_{3j} are iid according to $C(0, 1)$ restricted on $(-2, 2)$; Z_{1j} are iid according to $N(0, 1)$ restricted on $(-3, 3)$; Z_{2j} are iid according to Uniform $(-3, 3)$; Z_{3j} are iid according to $C(1, 1)$ restricted on $(-3, 3)$. Like univariate cases, here also our basic task is to determine the bandwidth constants, C_1 and C_2, for the two covariates. For our tests, we consider the optimum form of bandwidths (Silverman 1986, Sect. 4.3) as $C_1 n^{-1/6}$ and $C_2 n^{-1/6}$, respectively, for the first and second covariates. However, the form of the bandwidth remains same for *competitor 1* as suggested under univariate model due to some restrictions (Tsangari and Akritas 2004). The constants are chosen by the same technique as used earlier. For *BP-*, *BD-* and *BC-type tests*, the choices of (C_1, C_2) are respectively given by $(2.5, 1.7)$, $(2.8, 1.6)$ and $(2.4, 1.8)$. The choice becomes $(2.1, 1.9)$ for *TA test*. Simulation study shows that the empirical levels of *TA test* based on large sample procedure are significantly greater than the nominal level for all the chosen sample sizes, whereas small sample procedure of this test seems to be appropriate for *sample size 2*, *sample size 3* and *sample size 4*. On the other hand, large sample procedures for *BP-*, *BD-* and *BC-type tests* attain the nominal level quite appropriately corresponding to *sample size 2*, *sample size 3* and *sample size 4*. However, corresponding to *sample size 1*, small sample procedures seem to be appropriate for these tests. The tests based on alternative procedure, including *competitor 2*, show very high empirical levels even under large sample size. Table 11 gives the empirical powers for the four tests considering *sample size 2*. For *TA test*, the small sample procedure is adopted and for the other tests the empirical powers are given for the χ^2-test procedure along with MCTP in the bracket using large sample procedure.

The empirical levels of all the four tests attain the nominal level. Like single covariate case, under this model, none of the above tests can be considered as the best for all the situations. From Table 11, we observe that *BP-type test* for the χ^2 procedure and *TA test* give the same empirical power which is larger than the others under *Convex* alternative and under *Concave* and *linear* alternatives; *BP-type test* provides the high-

Table 12 Empirical powers of the four tests for testing H_{0B}

(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	TA test
(0,0,0)	0.0485 (0.0487)	0.0491 (0.0471)	0.0528 (0.0512)	0.0607
(0,0,0.8)	0.4062 (0.2292)	0.4173 (0.3286)	0.3386 (0.2430)	0.4078
(0,0.8,0.8)	0.3618 (0.4131)	0.3043 (0.3562)	0.3741 (0.4273)	0.4055
(0,0.8,0)	0.3773 (0.2843)	0.4001 (0.3181)	0.3143 (0.2237)	0.4097
(0.8,0,0.8)	0.3717 (0.2798)	0.3116 (0.2146)	0.3912 (0.3082)	0.4062
(0,0.4,0.8)	0.2977 (0.3207)	0.2788 (0.2943)	0.2759 (0.2910)	0.3245

Table 13 Empirical powers of the alternative tests for testing H_{0B}

(τ_1, τ_2, τ_3)	BP-type test	BD-type test	BC-type test	Combined test	Bathke A_N test
(0,0,0)	0.0478 (0.0484)	0.0462 (0.0495)	0.0509 (0.0483)	0.0522 (0.0512)	0.0613
(0,0,0.8)	0.4193 (0.3186)	0.4273 (0.3316)	0.3474 (0.2420)	0.4269 (0.3270)	0.4194
(0,0.8,0.8)	0.3755 (0.4285)	0.3123 (0.3644)	0.3860 (0.4364)	0.3842 (0.4380)	0.4066
(0,0.8,0)	0.4029 (0.3002)	0.4114 (0.3237)	0.3345 (0.2364)	0.4145 (0.3081)	0.4118
(0.8,0,0.8)	0.3930 (0.3028)	0.3302 (0.2364)	0.4127 (0.3256)	0.4131 (0.3142)	0.4156
(0,0.4,0.8)	0.3029 (0.3227)	0.2874 (0.3026)	0.2977 (0.3173)	0.3158 (0.3363)	0.3396

est power for the MCTP. On the other hand, *TA test* provides the highest empirical powers for *Umbrella* and *U-shaped* alternatives.

Finally, we perform simulation study under the assumptions given by Bathke and Brunner (2003) and test for H_{0B} against H_{1B}. The model for simulation is given by: $Y_{kj} = 1 + \tau_k + 0.5X_{kj} + 0.5Z_{kj} + E_{kj}$, $j = 1, 2, \ldots, n_k$, $k = 1, 2, 3$, where E_{kj} are iid according to $C(0, 1)$; X_{1j}, X_{2j} and X_{3j} are iid according to $N(0, 1)$ restricted on $(-2, 2)$ and Z_{1j}, Z_{2j} and Z_{3j} are iid according to $DE(0, 1)$ restricted on $(-3, 3)$. Clearly, the model represents equal distribution of each covariate for all the treatment groups. Along with the marginal distributions, we further assume that all the treatment groups have the same joint distribution for covariates. Here, we compare the three tests developed in Sect. 3 and the tests based on alternative procedure together with *TA test*, *Bathke test* and *Combined test* (*Remark 1*). For *Combined test*, we calculate the cutoff points through simulation by considering different types of linear and nonlinear models with different distributions. The simulated cutoff point of *Combined test* for the χ^2 procedure (MCTP) is given by 11.26711 (11.22373). With the help of simulation study, it is observed that our developed tests, including *Combined test*, attain the nominal level when $n_k \geq 4$, whereas *TA test* and *Bathke test* based on small sample procedures attain the level when $n_k \geq 35$. Tables 12 and 13 give the empirical levels and powers for all the tests for *sample size 3*. The performances of the tests with respect to empirical powers are nearly similar to that of univariate case. The empirical powers of *Combined test* is generally larger than the other tests. The empirical levels of *TA test* and *Bathke* A_N test are slightly larger than the chosen nominal level, whereas the other tests attain nominal levels quite accurately.

Table 14 Table showing the nonparametric estimates of the relative treatment effects and the standard errors in the bracket

Treatment	Conditional approach			Alternative approach		
	BP	*BD*	*BC*	*BP*	*BD*	*BC*
SF	0.4687	0.2008	0.2948	0.4550	0.1655	0.3168
	(0.0406)	(0.0630)	(0.0642)	(0.0408)	(0.0596)	(0.0726)
APF	0.3420	0.1384	0.6123	0.3409	0.1260	0.6236
	(0.0432)	(0.0650)	(0.0703)	(0.0419)	(0.0525)	(0.0796)
W	0.6892	0.6608	0.0930	0.6800	0.6150	0.1050
	(0.0421)	(0.0742)	(0.0583)	(0.0389)	(0.0773)	(0.0418)

6 Data study

Consider the fluoride data obtained by Cartwright et al. (1968) from an experiment to reduce dental caries. The data correspond to two treatments, *stannous fluoride* (SF) and *acid phosphate fluoride* (APF) and a placebo treatment *distilled water* (W). The treatments were applied over 69 female children to observe the number of decayed, missing or filled teeth (DMFT) before (B) and after (A) the study. The responses to be analyzed are the differences on $\mathrm{Y} = \mathrm{A} - \mathrm{B}$. Besides the response, the age for each child is considered as covariate in this analysis. Firstly, we find the bandwidth constant taking the form of the bandwidth as $a_n = (\text{const})n^{-0.2}$ when there is no treatment effect. For this, we adopt resampling technique by drawing with replacement samples from the combined data set of 69 observations and then distributing the resampled data at random among the three groups. This strategy justifies no treatment effect. Here, we use three different sample sizes as: $N = 49, (n_1, n_2, n_3) = (15, 16, 18)$; $N = 75, (n_1, n_2, n_3) = (24, 25, 26)$; $N = 120, (n_1, n_2, n_3) = (35, 40, 45)$ to generate data from the combined data set and observe 5.0 as the bandwidth constant for the *BP*-, *BD*-, *BC-type tests*. Table 14 demonstrates nonparametric estimates of the covariate-eliminated relative treatment effects (with the estimates of asymptotic standard errors in the bracket) derived from the three functionals using conditional and alternative approaches.

The estimates clearly show the same ordering in the treatment effects for the functionals under both the approaches. Further, it is observed that the placebo treatment *distilled water* (W) has an influential effect on the increase in DMFT count. Next, we test for the significance of the treatment effects using the covariate (age) and without using covariate. Table 15 shows the observed values of the test statistics and the corresponding *p* values based on the small sample test procedures because the tests based on large sample procedures show little liberal for the given sample sizes. We also adopt the same technique for the alternative approach.

We observe a strong evidence for rejection of the null hypothesis of absence of treatment effects by all the tests considered here. That is, there exists a significant difference among the three treatments applied. Moreover, the *p* values of the tests using covariate tend to be smaller compared to those without using covariate. Hence,

Table 15 Values of the test statistics and the corresponding p values

Test	Conditional approach		Alternative approach		Without covariate	
	Statistic	p value	Statistic	p value	Statistic	p value
BP	11.1108	1.517×10^{-4}	12.6895	6.361×10^{-5}	16.6883	2.378×10^{-4}
	(3.1280)	(6.422×10^{-3})	(3.3200)	(3.900×10^{-3})	(2.5645)	(1.957×10^{-2})
BD	9.8276	3.774×10^{-4}	8.6861	8.970×10^{-4}	18.1713	1.133×10^{-4}
	(3.7912)	(1.057×10^{-3})	(3.5195)	(2.443×10^{-3})	(3.4360)	(1.159×10^{-3})
BC	10.7959	1.856×10^{-4}	17.5544	2.315×10^{-6}	15.7325	3.835×10^{-4}
	(2.6156)	(2.359×10^{-2})	(2.4144)	(3.396×10^{-2})	(2.3378)	(3.620×10^{-2})

Table 16 SCI for the treatment differences

Treatment	Conditional approach		Alternative approach	
	SF-APF	*SF-W*	*SF-APF*	*SF-W*
BP	[−0.0402, 0.2936]	[−0.3828, −0.0583]	[−0.0535, 0.2818]	[−0.3806, −0.0692]
BD	[−0.1785, 0.3034]	[−0.7402, −0.1796]	[−0.1490, 0.2280]	[−0.7452, −0.1538]
BC	[−0.5961, −0.0390]	[−0.0290, 0.4323]	[−0.6358, 0.0221]	[0.0148, 0.4087]

we can say that the use of covariate is worthwhile in this study. Now, we focus on pairwise comparison based on SCI (Eq. 10) based on both the approaches using age as covariate.

From the SCIs of the treatment differences (Table 16), it is identified that, except for the *BC-type test* based on conditional distribution approach, the difference between SF and W is significant.

7 Conclusion

This work develops nonparametric tests for equality of treatment effects in the presence of multiple covariates through ridit reliability functionals based on two different approaches. In the first approach, we develop tests by considering a dependent structure of the response variables and the corresponding covariates through conditional distributions. The alternative approach considers the marginal distributions and ignores the direct dependency between variables. Thus, the tests based on alternative approach assume lesser restrictions and the use of such tests is more easier than that developed under conditional distribution approach. But the simulation study demonstrates that the performance of the conditional approach in terms of type I error rate is more accurate than that of the marginal approach. The convergence of the proposed test statistics using conditional distribution approach to their asymptotic distributions are more rapid than that of the test statistics provided by Tsangari and Akritas (2004) and Bathke and Brunner (2003). The simulation study shows that the tests are asymptotically distribution free and are applicable to the situation when the response and covariate are not necessarily continuous. An important fact is that the choice of the bandwidth constants depends on the choice of the response model. Thus, it is necessary to determine the bandwidth constants before applying the tests based on conditional distribution approach. Empirical power study shows that none of the tests can be taken as superior to the others uniformly for all types of alternatives. There is no such restrictions on the number of covariates for the tests developed under both the approaches considered here. However, in practice, the conditional approach becomes intractable when there is a large number of covariates. In this case, the optimum bandwidth converges to zero in an extremely slow rate. Hence, it requires larger sample sizes than the alternative approach. Moreover, based on the alternative approach, it is also possible to develop tests for the significance of covariates as described by Bathke and Brunner (2003).

Acknowledgements The authors wish to thank the Editor, the corresponding Associate Editor and the anonymous referees for their helpful comments and suggestions that have improved the paper.

Appendix

Proof of Result 1 It is enough to show that for any $k, k'(\neq k)$, $\widehat{p}_{kk'}$ is a consistent estimator of $p_{kk'}$. On this matter, we set

$$\widehat{p}_{kk'}^{(1)} = \frac{1}{n_k n_{k'}} \sum_{j=1}^{n_k} \sum_{j'=1}^{n_{k'}} \frac{g(\mathbf{X}_{kj}) g(\mathbf{X}_{k'j'})}{g_k(\mathbf{X}_{kj}) g_{k'}(\mathbf{X}_{k'j'})} U(Y_{kj}, Y_{k'j'})$$

and

$$\widehat{p}_{kk'}^{(2)} = \frac{1}{n_k n_{k'}} \sum_{j=1}^{n_k} \sum_{j'=1}^{n_{k'}} \left(e_{kj} e_{k'j'} - \frac{g(\mathbf{X}_{kj}) g(\mathbf{X}_{k'j'})}{g_k(\mathbf{X}_{kj}) g_{k'}(\mathbf{X}_{k'j'})} \right) U(Y_{kj}, Y_{k'j'}),$$

where $g(.)$ is the density (or probability mass function) corresponding to $G(.)$. Then, we can rewrite $\widehat{p}_{kk'}$ as

$$\widehat{p}_{kk'} = \widehat{p}_{kk'}^{(1)} + \widehat{p}_{kk'}^{(2)}. \tag{12}$$

Now,

$$\begin{aligned} E\left(\widehat{p}_{kk'}^{(1)}\right) &= \frac{1}{n_k n_{k'}} \sum_{j=1}^{n_k} \sum_{j'=1}^{n_{k'}} E\left[\frac{g(\mathbf{X}_{kj}) g(\mathbf{X}_{k'j'})}{g_k(\mathbf{X}_{kj}) g_{k'}(\mathbf{X}_{k'j'})} p_{kk'}(\mathbf{X}_{kj}, \mathbf{X}_{k'j'}) \right] \\ &= \int\int p_{kk'}(\mathbf{x}_k, \mathbf{x}_{k'})\, dG(\mathbf{x}_k)\, dG(\mathbf{x}_{k'}) \\ &= p_{kk'} \end{aligned}$$

and

$$
\begin{aligned}
Var\left(\widehat{p}_{kk'}^{(1)}\right) &= \frac{1}{n_k^2 n_{k'}^2} Var\left[\sum_{j=1}^{n_k}\sum_{j'=1}^{n_{k'}}\left\{\frac{g(\mathbf{X_{kj}})g(\mathbf{X_{k'j'}})}{g_k(\mathbf{X_{kj}})g_{k'}(\mathbf{X_{k'j'}})}U(Y_{kj},Y_{k'j'})\right\}\right] \\
&= \frac{1}{n_k^2 n_{k'}^2}\left[\sum_{j=1}^{n_k}\sum_{j'=1}^{n_{k'}} Va_p(j,j')\right. \\
&\quad \left. + \sum_{j(\neq j_1)=1}^{n_k}\sum_{j'=1}^{n_{k'}} C_p(j,j_1,j') + \sum_{j=1}^{n_k}\sum_{j'(\neq j'_1)=1}^{n_{k'}} C_p(j,j',j'_1)\right].
\end{aligned}
$$

In the above expression, $Va_p(.,.)$ and $C_p(.,.,.)$ denote, respectively, the variance and covariance terms and these terms are bounded under A3. Therefore, $Var\left(\widehat{p}_{kk'}^{(1)}\right) \to 0$ under A2. Thus, we get

$$
\widehat{p}_{kk'}^{(1)} \to p_{kk'} \tag{13}
$$

in probability. Furthermore, the assumptions A1–A4 imply that

$$
\max_j \left| e_{kj} - \frac{g(\mathbf{X_{kj}})}{g_k(\mathbf{X_{kj}})}\right| \to 0
$$

almost surely (Tsangari and Akritas 2004), and hence, we get

$$
\begin{aligned}
0 \leq \left|\widehat{p}_{kk'}^{(2)}\right| &\leq \max_{j,\,j'} \left| e_{kj}e_{k'j'} \right. \\
&\quad \left. - \frac{g(\mathbf{X_{kj}})g(\mathbf{X_{k'j'}})}{g_k(\mathbf{X_{kj}})g_{k'}(\mathbf{X_{k'j'}})}\right|\left\{\frac{1}{n_k n_{k'}}\sum_j\sum_{j'} U(Y_{kj},Y_{k'j'})\right\} = o_p(1)O_p(1).
\end{aligned} \tag{14}
$$

Now, combining (13) and (14), the required result follows from (12). □

Proof of Result 2 To find the asymptotic distribution of $\sqrt{N}(\widehat{\mathbf{q}}. - \mathbf{q}.)$, we write, for any $k = 1, 2, \ldots, K$,

$$
\begin{aligned}
T_k^p &= \sqrt{N}(\widehat{p}_{k\cdot} - p_{k\cdot}) \\
&= \frac{\sqrt{N}}{K}\left[\sum_{\substack{k_1=1\\k_1\neq k}}^{K} (\widehat{p}_{kk_1} - p_{kk_1})\right] \\
&= T_{K1}^p + T_{K2}^p,
\end{aligned} \tag{15}
$$

where

$$T^p_{k1} = \frac{\sqrt{N}}{K} \sum_{\substack{k_1=1 \\ k_1 \neq k}}^{K} \frac{1}{n_k n_{k_1}} \sum_{j=1}^{n_k} \sum_{j_1=1}^{n_{k_1}} \left\{ \frac{g(\mathbf{X_{kj}})g(\mathbf{X_{k_1 j_1}})}{g_k(\mathbf{X_{kj}})g_{k_1}(\mathbf{X_{k_1 j_1}})} \left(U(Y_{kj}, Y_{k_1 j_1}) - p^0_{kk_1} \right) - \left(p_{kk_1} - p^0_{kk_1} \right) \right\},$$

$$T^p_{k2} = \frac{1}{n_k n_{k_1}} \sum_{j=1}^{n_k} \sum_{j_1=1}^{n_{k_1}} \left(e_{kj} e_{k_1 j_1} - \frac{g(\mathbf{X_{kj}})g(\mathbf{X_{k_1 j_1}})}{g_k(\mathbf{X_{kj}})g_{k_1}(\mathbf{X_{k_1 j_1}})} \right) \left(U(Y_{kj}, Y_{k_1 j_1}) - p^0_{kk_1} \right)$$

and $p^0_{kk_1} = E(U(Y_{kj}, Y_{k_1 j_1}))$ for any (j, j_1). Now, since

$$0 \leq \left| T^p_{k2} \right| \leq \frac{1}{K} \sum_{\substack{k_1=1 \\ k_1 \neq k}}^{K} \max_{j,\, j_1} \left| e_{kj} e_{k_1 j_1} - \frac{g(\mathbf{X_{kj}})g(\mathbf{X_{k_1 j_1}})}{g_k(\mathbf{X_{kj}})g_{k_1}(\mathbf{X_{k_1 j_1}})} \right| \times \left| \frac{\sqrt{N}}{n_k n_{k_1}} \sum_{j=1}^{n_k} \sum_{j_1=1}^{n_{k_1}} \left(U(Y_{kj}, Y_{k_1 j_1}) - p^0_{kk_1} \right) \right| = o_p(1) O_p(1),$$

we get

$$T^p_{k2} \to 0$$

in probability. Hence, the asymptotic distribution of $\sqrt{N}(\widehat{\mathbf{q}}. - \mathbf{q}.)$ is same as that of $\mathbf{T^p} = \left\{ T^p_{k1}, k = 1, 2, \ldots, K \right\}$. Now, using Hajek's projection theorem in multivariate setup, we get that the asymptotic distribution of $\mathbf{T^p}$ is same as that of $\mathbf{Z^p} = \left\{ Z^p_k, k = 1, 2, \ldots, K \right\}$, with

$$Z^p_k = \sum_{k_1=1}^{K} \frac{\sqrt{N}}{n_{k_1}} \sum_{j=1}^{n_{k_1}} \left\{ \frac{g(\mathbf{X_{k_1 j}})}{g_{k_1}(\mathbf{X_{k_1 j}})} Z^p_{kk_1}(Y_{k_1 j}) - (p_{k\cdot} - p_k) \right\},$$

where

$$Z^p_{kk}(Y_{kj}) = \frac{1}{K} \sum_{\substack{k_1=1 \\ k_1 \neq k}}^{K} \left(1 - F^0_{k_1}.(Y_{kj}) - p^0_{kk_1} \right) \text{ and}$$

$$Z^p_{kk_1}(Y_{k_1 j}) = \frac{1}{K} \left(F^0_k.(Y_{k_1 j}) - p^0_{kk_1} \right)$$

for $k, k_1 (\neq k) = 1, 2, \ldots, K$.

Next, using central limit theorem, the asymptotic distribution of $\mathbf{Z^p}$ is K-variate normal with mean vector 0 and dispersion matrix $S_p = ((s^p_{kk'}))$, where

$$s^p_{kk} = \sum_{k_1=1}^{K} \frac{1}{\lambda_{k_1}} \left[Var_{X_{k_1}} \left\{ \frac{g(\mathbf{X_{k_1 j}})}{g_{k_1}(\mathbf{X_{k_1 j}})} E_{Y_{k_1}|X_{k_1}} Z^p_{kk_1} \right\} + E_{X_{k_1}} \left\{ \left(\frac{g(\mathbf{X_{k_1 j}})}{g_{k_1}(\mathbf{X_{k_1 j}})} \right)^2 Var_{Y_{k_1}|X_{k_1}} Z^p_{kk_1} \right\} \right],$$

$k = 1, 2, \ldots, K$ and for $k_1 \neq k_2 = 1, 2, \ldots, K$, we have

$$s^p_{k_1 k_2} = \sum_{k_3=1}^{K} \frac{1}{\lambda_{k_3}} Cov_{X_{k_3}} \times \left\{ \frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} E_{Y_{k_3}|X_{k_3}} Z^p_{k_1 k_3}, \frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} E_{Y_{k_3}|X_{k_3}} Z^p_{k_2 k_3} \right\} + E_{X_{k_3}} \left\{ \left(\frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} \right)^2 Cov_{Y_{k_3}|X_{k_3}} \left(Z^p_{k_1 k_3}, Z^p_{k_2 k_3} \right) \right\},$$

in which

$$p^0_k = \frac{1}{K} \left[\frac{1}{2} + \sum_{\substack{k_1=1 \\ k_1 \neq k}}^{K} p^0_{kk_1} \right].$$

Therefore, the asymptotic distribution of $\sqrt{N} C(\widehat{\mathbf{q}}. - \mathbf{q}.)$ is $(K-1)$-variate normal with mean vector 0 and dispersion matrix $\Sigma_{Cp} = C S_p C^T$. Hence, the result follows. □

Proof of Result 4 We can prove the result with very similar approach as suggested for Result 2. To find the asymptotic distribution of $\sqrt{N}(\widehat{\mathbf{R}}. - \mathbf{R}.)$, we set

$$V_k = V_k(Y_{1j_1}, Y_{2j_2}, \ldots, Y_{Kj_K})$$

and write

$$\begin{aligned} T^R_k &= \sqrt{N}(\widehat{R}_{k\cdot} - R_{k\cdot}) \\ &= \sqrt{N} \left[\frac{1}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} e(j_1, j_2, \ldots, j_K)\, V_k - R_{k\cdot} \right] \\ &= T^R_{K1} + T^R_{K2}, \end{aligned} \tag{16}$$

where

$$T_{k1}^R = \frac{\sqrt{N}}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} \left\{ \prod_{l=1}^{K} \frac{g(\mathbf{X}_{lj_l})}{g_l(\mathbf{X}_{lj_l})} (V_k - R_k) - (R_{k.} - R_k) \right\},$$

$$T_{k2}^R = \frac{\sqrt{N}}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} \left(e(j_1, j_2, \ldots, j_K) - \prod_{l=1}^{K} \frac{g(\mathbf{X}_{lj_l})}{g_l(\mathbf{X}_{lj_l})} \right) \times (V_k - R_k)$$

and $R_k = E(V_k)$ for any $k = 1, 2, \ldots, K$. Now, since

$$0 \leq \left| T_{k2}^p \right| \leq \max_{j_1 \cdots j_K} \left| e(j_1, j_2, \ldots, j_K) - \prod_{l=1}^{K} \frac{g(\mathbf{X}_{lj_l})}{g_l(\mathbf{X}_{lj_l})} \right| \times \left| \frac{\sqrt{N}}{n_1 n_2 \cdots n_K} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_K=1}^{n_K} (V_k - R_k) \right| = o_p(1) O_p(1),$$

we get

$$T_{k2}^R \to 0$$

in probability, and hence the asymptotic distribution of $\sqrt{N}(\widehat{\mathbf{R}}. - \mathbf{R}.)$ is same as that of $\mathbf{T^R} = \left\{ T_{k1}^R, k = 1, 2, \ldots, K \right\}$. Applying Hajek's projection theorem in multivariate setup, we get that the asymptotic distribution of $\mathbf{T^R}$ is same as that of $\mathbf{Z^R} = \left\{ Z_k^R, k = 1, 2, \ldots, K \right\}$ with

$$Z_k^R = \sum_{k_1=1}^{K} \frac{\sqrt{N}}{n_{k_1}} \sum_{j=1}^{n_{k_1}} \left\{ \frac{g(\mathbf{X}_{k_1 j})}{g_{k_1}(\mathbf{X}_{k_1 j})} Z_{kk_1}^R (Y_{k_1 j}) - (R_{k.} - R_k) \right\},$$

where

$$Z_{kk}^R(Y_{kj}) = \prod_{\substack{l=1 \\ l \neq k}}^{K} F_{l.}^{-}(Y_{kj}) + \frac{1}{2} \sum_{\substack{1 \leq l_1 \leq K \\ l_1 \neq k}} f_{l_1 .}(Y_{kj}) \prod_{\substack{l=1 \\ l \neq l_1,\, k}}^{K} F_{l.}^{-}(Y_{kj}) + \frac{1}{3} \sum_{\substack{1 \leq l_1 \\ l_1,\, l_2}} \sum_{\substack{< l_2 \leq K \\ \neq k}} f_{l_1 .}(Y_{kj}) f_{l_2 .}(Y_{kj}) \prod_{\substack{l=1 \\ l \neq l_1,\, k}}^{K} F_{l.}^{-}(Y_{kj}) + \cdots + \frac{1}{K} \prod_{\substack{l=1 \\ l \neq k}}^{K} f_{l.}(Y_{kj}) - R_k$$

for $k = 1, 2, \ldots, K$ and for $k_1 \neq k$, we have

$$\begin{aligned}
Z^R_{kk_1}(Y_{k_1 j}) =& \{1 - F_{k\cdot}(Y_{k_1 j})\} \prod_{\substack{l=1 \\ l \neq k, k_1}}^{K} F_{l\cdot}(Y_{k_1 j}) + \frac{1}{2} f_{k\cdot}(Y_{k_1 j}) \prod_{\substack{l=1 \\ l \neq k, k_1}}^{K} F_{l\cdot}^{-}(Y_{k_1 j}) \\
&+ \frac{1}{3} \sum_{\substack{1 \le k_2 \le K \\ k_2 \neq k, k_1}} f_{k\cdot}(Y_{k_1 j}) f_{k_2\cdot}(Y_{k_1 j}) \\
&\prod_{\substack{l=1 \\ l \neq k, k_1, k_2}}^{K} F_{l\cdot}^{-}(Y_{k_1 j}) + \cdots + \frac{1}{K} \prod_{\substack{l=1 \\ l \neq k_1}}^{K} f_{l\cdot}(Y_{k_1 j}) \\
&+ \sum_{\substack{1 \le k_2 \le K \\ k_2 \neq k, k_1}} \prod_{\substack{l=1 \\ l \neq k, k_1, k_2}}^{K} F_{l\cdot}(Y_{k_1 j}) \, E\{R_{Y_k}(Y_{k_2}) \, I(Y_{k_2} > Y_{k_1 j})\} \\
&+ \sum_{\substack{1 \le k_2 < k_3 \le K \\ k_2, k_3 \neq k, k_1}} \prod_{\substack{l=1 \\ l \neq k, k_1, k_2, k_3}}^{K} F_{l\cdot}(Y_{k_1 j}) \, E\{R_{Y_k}(Y_{k_2}, Y_{k_3}) \\
&\times I(\min(Y_{k_2}, Y_{k_3}) > Y_{k_1 j})\} + \cdots \\
&+ E\{R_{Y_k}(Y_1, Y_2, \ldots, Y_{k_1 - 1}, Y_{k_1 + 1}, \ldots Y_K) \\
&\times I(\min(Y_1, Y_2, \ldots, Y_{k_1 - 1}, Y_{k_1 + 1}, \ldots Y_K) > Y_{k_1 j})\} - R_k
\end{aligned}$$

with $I(.)$ representing the indicator of the corresponding set,

$$F_{k\cdot}^{-}(y) = \int P(Y_k < y|x) \, dG(x), \quad f_{k\cdot}(y) = \int P(Y_k = y|x) \, dG(x)$$

and

$$\begin{aligned}
R_Y(Y_{s_1}, \ldots, Y_{s_{r-1}}) =& P(Y_{s_1} < Y, Y_{s_2} < Y, \ldots, Y_{s_{r-1}} < Y) \\
&+ \frac{1}{2} \sum_{1 \le q_1 \le p} P(Y_{q_1} = Y, Y_{s_1} < Y, Y_{s_2} < Y, \ldots, Y_{s_{r-1}} < Y) \\
&+ \cdots + \frac{1}{r} P(Y_{s_1} = Y, Y_{s_2} = Y, \ldots, Y_{s_{r-1}} = Y). \qquad (17)
\end{aligned}$$

Thus, using central limit theorem, the asymptotic distribution of $\mathbf{Z}^{\mathbf{R}}$ is K-variate normal with mean vector $\mathbf{0}$ and dispersion matrix $S_R = ((s^R_{kk'}))$, where

$$\begin{aligned}
s^R_{kk} = \sum_{k_1=1}^{K} \frac{1}{\lambda_{k_1}} \Bigg[& Var_{X_{k_1}} \left\{ \frac{g(\mathbf{X}_{k_1 j})}{g_{k_1}(\mathbf{X}_{k_1 j})} E_{Y_{k_1}|X_{k_1}} Z^R_{kk_1} \right\} \\
&+ E_{X_{k_1}} \left\{ \left(\frac{g(\mathbf{X}_{k_1 j})}{g_{k_1}(\mathbf{X}_{k_1 j})} \right)^2 Var_{Y_{k_1}|X_{k_1}} Z^R_{kk_1} \right\} \Bigg],
\end{aligned}$$

$k = 1, 2, \ldots, K$ and for $k_1 \neq k_2 = 1, 2, \ldots, K$, we have

$$s^R_{k_1 k_2} = \sum_{k_3=1}^{K} \frac{1}{\lambda_{k_3}} Cov_{X_{k_3}} \times \left\{ \frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} E_{Y_{k_3}|X_{k_3}} Z^R_{k_1 k_3}, \frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} E_{Y_{k_3}|X_{k_3}} Z^R_{k_2 k_3} \right\} + E_{X_{k_3}} \left\{ \left(\frac{g(\mathbf{X_{k_3 j}})}{g_{k_3}(\mathbf{X_{k_3 j}})} \right)^2 Cov_{Y_{k_3}|X_{k_3}} \left(Z^R_{k_1 k_3}, Z^R_{k_2 k_3} \right) \right\}.$$

Therefore, the asymptotic distribution of $\sqrt{N} C(\widehat{\mathbf{R}}. - \mathbf{R}.)$ is $(K-1)$-variate normal with mean vector 0 and dispersion matrix $\Sigma_{CR} = CS_R C^T$. Hence, the result follows. □

Note 1. The proof of Result 6 is exactly same as that of Result 4. Hence, the proof is not given here. The elements of the dispersion matrix $\Sigma_{CR^*} = CS_{R^*}C^T$ can be obtained by replacing the terms $F_k.(y)$ and $F_k.^-(y)$ by $\bar{F}_k.(y)$ and $F_k.^+(y)$ and finally

$$E\{R_{Y_k}(Y_1, Y_2, \ldots, Y_{k_1-1}, Y_{k_1+1}, \ldots Y_K) \; I(\min(Y_1, Y_2, \ldots, Y_{k_1-1}, Y_{k_1+1}, \ldots Y_K) > Y_{k_1 j})\}$$

by

$$E\left\{R^*_{Y_k}(Y_1, Y_2, \ldots, Y_{k_1-1}, Y_{k_1+1}, \ldots Y_K) \; I(\max(Y_1, Y_2, \ldots, Y_{k_1-1}, Y_{k_1+1}, \ldots Y_K) < Y_{k_1 j})\right\},$$

where

$$F_k.^+(y) = \int P(Y_k > y|x)\, dG(x), \quad \bar{F}_k.(y) = \int P(Y_k \geq y|x)\, dG(x)$$

and R^*_Y can be defined similarly as (17) corresponding to *BC-type functional*.

Note 2. The consistent estimators of the dispersion matrices can be provided under the assumptions A1–A4 by replacing the quantities $\lambda_k, F_k.(y), F^0_k.(y), F_k.^-(y), F_k.^+(y), \bar{F}_k.(y), f_k.(y)$ and $\frac{g(\mathbf{X_{kj}})}{g_k(\mathbf{X_{kj}})}$ by their sample versions $\frac{n_k}{N}, \widehat{F}_k.(y), \widehat{F}^0_k.(y), \widehat{F}_k.^-(y), \widehat{F}_k.^+(y), \widehat{\bar{F}}_k.(y), \widehat{f}_k.(y)$ and e_{kj}, respectively, where $\widehat{F}^0_k.(y) = \frac{1}{n_k} \sum_{j=1}^{n_k} e_{kj} \left\{ I(Y_{kj} > y) + \frac{1}{2} I(Y_{kj} = y) \right\}$ and similarly the others. Note that, as we do not assume any specific nature of the distribution, the estimators under A1–A4 are consistent even for tie cases.

Note 3. To compare the tests with the test provided by Tsangari and Akritas (2004), we modify the tests by using the estimator of the variance components under H_{0A}. Here, we use the combined estimator of $F^0_k.(y)$ as

$$\sum_{k=1}^{K} \frac{n_k}{N} \widehat{F}^0_{k\cdot}(y)$$

for any $k = 1, 2, \ldots, K$. Similarly, we can also estimate the other quantities.

Proof of Result 7 For *BP*-, *BD*- and *BC-type functionals*, it is not difficult to show that

$$\sqrt{N}\{(\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}) - \mathbf{B}\} \to 0 \tag{18}$$

in probability, where $\mathbf{B} = (B_1, B_2, \ldots, B_K)^T$ with

$$B_k = B_k^{(0)} - \sum_{r=1}^{d} \beta_r B_k^{(r)} \text{ and } B_k^{(r)} = \sum_{k_1=1}^{K} \frac{1}{n_{k_1}} \sum_{j=1}^{n_{k_1}} B_{kk_1}^{(r)}(X_{k_1 j}^{(r)}),$$

$k = 1, 2, \ldots, K$ and $r = 0, 1, \ldots, d$. For convenience, without loss of generality, we write $X_{kj}^{(0)} = Y_{kj}$. The mathematical expression of $B_{kk_1}^{(r)}$ depends on the choices of the functionals. For *BP-type functionals*, we have,

$$B_{kk}^{(r)}(X_{kj}^{(r)}) = \frac{1}{K} \sum_{\substack{k_1=1 \\ k_1 \neq k}}^{K} \left(1 - F^0_{(r)k_1}(X_{kj}^{(r)}) - p_{kk_1}^{(r)}\right) \text{ and}$$

$$B_{kk_1}^{(r)}(X_{k_1 j}^{(r)}) = \frac{1}{K} \left(F^0_{(r)k}(X_{k_1 j}^{(r)}) - p_{kk_1}^{(r)}\right),$$

where $F^0_{(r)k}(x) = P(X_k^{(r)} > x) + \frac{1}{2} P(X_k^{(r)} > x)$ for $k, k_1 (\neq k) = 1, 2, \ldots, K$ and $r = 0, 1, \ldots, d$.

Again, for *BD-type functionals* the expression will be

$$\begin{aligned} B_{kk}^{(r)}(X_{kj}^{(r)}) = &\prod_{\substack{l=1 \\ l \neq k}}^{K} F^-_{(r)l}(X_{kj}^{(r)}) + \frac{1}{2} \sum_{\substack{1 \leq l_1 \leq K \\ l_1 \neq k}} f_{(r)l_1}(X_{kj}^{(r)}) \prod_{\substack{l=1 \\ l \neq l_1,\, k}}^{K} F^-_{(r)l}(X_{kj}^{(r)}) \\ &+ \frac{1}{3} \sum_{\substack{1 \leq l_1 \\ l_1,\, l_2}} \sum_{\substack{< l_2 \leq K \\ \neq k}} f_{(r)l_1}(X_{kj}^{(r)}) f_{(r)l_2}(X_{kj}^{(r)}) \\ &\prod_{\substack{l=1 \\ l \neq l_1,\, k}}^{K} F^-_{(r)l}(X_{kj}^{(r)}) + \cdots + \frac{1}{K} \prod_{\substack{l=1 \\ l \neq k}}^{K} f_{(r)l}(X_{kj}^{(r)}) - R_k^{(r)} \end{aligned}$$

for $k = 1, 2, \ldots, K$ and for $k_1 \neq k$, we have

$$\begin{aligned}
B^{(r)}_{kk_1}(X^{(r)}_{k_1 j}) &= \left\{1 - F_{(r)k}\left(X^{(r)}_{k_1 j}\right)\right\} \prod_{\substack{l=1\\ l\neq k,k_1}}^{K} F_{(r)l}\left(X^{(r)}_{k_1 j}\right)\\
&\quad + \frac{1}{2} f_{(r)k}\left(X^{(r)}_{k_1 j}\right) \prod_{\substack{l=1\\ l\neq k,k_1}}^{K} F^{-}_{(r)l}\left(X^{(r)}_{k_1 j}\right)\\
&\quad + \frac{1}{3} \sum_{\substack{1\leq k_2\leq K\\ k_2\neq k,k_1}} f_{(r)k}\left(X^{(r)}_{k_1 j}\right) f_{(r)k_2}\left(X^{(r)}_{k_1 j}\right)\\
&\quad \prod_{\substack{l=1\\ l\neq k,k_1,k_2}}^{K} F^{-}_{(r)l}\left(X^{(r)}_{k_1 j}\right) + \cdots + \frac{1}{K} \prod_{\substack{l=1\\ l\neq k_1}}^{K} f_{(r)l}\left(X^{(r)}_{k_1 j}\right)\\
&\quad + \sum_{\substack{1\leq k_2\leq K\\ k_2\neq k,k_1}} \prod_{\substack{l=1\\ l\neq k,k_1,k_2}}^{K} F_{(r)l}\left(X^{(r)}_{k_1 j}\right)\\
&\quad \times E\left\{R_{X^{(r)}_k}\left(X^{(r)}_{k_2}\right) I\left(X^{(r)}_{k_2} > X^{(r)}_{k_1}\right)\right\}\\
&\quad + \sum_{\substack{1\leq k_2<k_3\leq K\\ k_2,k_3\neq k,k_1}} \prod_{\substack{l=1\\ l\neq k,k_1,k_2,k_3}}^{K} F_{(r)l}\left(X^{(r)}_{k_1 j}\right) E\left\{R_{X^{(r)}_k}\left(X^{(r)}_{k_2}, X^{(r)}_{k_3}\right)\right.\\
&\quad \left.\times\, I\left(\min\left(X^{(r)}_{k_2}, X^{(r)}_{k_3}\right) > X^{(r)}_{k_1 j}\right)\right\} + \cdots\\
&\quad + E\left\{R_{X^{(r)}_k}\left(X^{(r)}_1, X^{(r)}_2, \ldots, X^{(r)}_{k_1-1}, X^{(r)}_{k_1+1}, \ldots X^{(r)}_K\right)\right.\\
&\quad \times\, I\left(\min\left(X^{(r)}_1, X^{(r)}_2, \ldots, X^{(r)}_{k_1-1}, X^{(r)}_{k_1+1}, \ldots X^{(r)}_K\right)\right.\\
&\quad \left.\left. > X^{(r)}_{k_1 j}\right)\right\} - R^{(r)}_k,
\end{aligned}$$

where $F_{(r)k}(x) = P(X^{(r)}_k \geq x)$, $F^{-}_{(r)k}(x) = P(X^{(r)}_k < x)$ and $f_{(r)k}(x) = P(X^{(r)}_k = x)$, $r = 0, 1, \ldots, d$. In case of *BC-type functional*, we get the expressions by altering the $>$ (or $\geq$) sign by $<$ (or $\leq$) and "min" by "max" in the above expressions.

Therefore, from (18) we can say that $\sqrt{N}(\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi})$ and $\sqrt{N}\mathbf{B}$ have the same asymptotic distribution. Using central limit theorem, $\sqrt{N}\mathbf{B}$ asymptotically follows K-variate normal distribution with mean vector 0 and dispersion matrix Σ_B. Thus, $\sqrt{N}C\mathbf{B}$ and hence $\sqrt{N}C(\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi})$ asymptotically follow $(K-1)$-variate normal distribution with mean vector 0 and dispersion matrix $\Sigma_{C\pi}$. □

Note 4. Consistent estimator of the dispersion matrix $\Sigma_{C\pi}$ can be provided by estimating the elements with the usual sample versions. Note that $\Sigma_{C\pi}$ involves unknown parameter $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_d)^T$. We estimate $\boldsymbol{\beta}$ consistently by minimizing the variance of $\frac{1}{N}\sum_{k=1}^{K} n_k B_k$ (Bathke and Brunner 2003) and solving the set of d equations

$$\sum_{r_1=1}^{d} \beta_{r_1} \left\{ \sum_{k=1}^{K} n_k^2 \sum_{k_1=1}^{K} \frac{1}{n_{k_1}} Cov\left(B_{kk_1}^{(r)}, B_{kk_1}^{(r_1)}\right) \right.$$
$$\left. + \sum_{k} \sum_{\neq k'} n_k n_{k'} \sum_{k_1=1}^{K} \frac{1}{n_{k_1}} Cov\left(B_{kk_1}^{(r)}, B_{k'k_1}^{(r_1)}\right) \right\}$$
$$= \sum_{k=1}^{K} n_k^2 \sum_{k_1=1}^{K} \frac{1}{n_{k_1}} Cov\left(B_{kk_1}^{(0)}, B_{kk_1}^{(r)}\right)$$
$$+ \sum_{k} \sum_{<k'} n_k n_{k'} \sum_{k_1=1}^{K} \frac{1}{n_{k_1}} \left\{ Cov\left(B_{kk_1}^{(0)}, B_{k'k_1}^{(r)}\right) + Cov\left(B_{kk_1}^{(r)}, B_{k'k_1}^{(0)}\right) \right\},$$

$r = 1, 2, \ldots, d.$

References

Akritas, M. G., Keilegom, I. V. (2001). Non-parametric estimation of the residual distribution. *Scandinavian Journal of Statistics*, *28*(3), 549–567.

Akritas, M. G., Arnold, S. F., Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association*, *92*(437), 258–265.

Akritas, M. G., Arnold, S. F., Du, Y. (2000). Nonparametric models and methods for nonlinear analysis of covariance. *Biometrika*, *87*(3), 507–526.

Bandyopadhyay, U., Chatterjee, D. (2015). Nonparametric homogeneity test based on ridit reliability functional. *Journal of the Korean Statistical Society*, *44*(4), 577–591.

Bandyopadhyay, U., De, S. (2011). On multi-treatment adaptive allocation design for dichotomous response. *Communications in Statistics Theory and Methods*, *40*(22), 4104–4124.

Bathke, A., Brunner, E. (2003). A nonparametric alternative to analysis of covariance. In M. G. Akritas, D. N. Politis (Eds.), *Recent advances and trends in nonparametric statistics* (pp. 109–120). Amsterdam: Elsevier.

Bretz, F., Hothorn, T., Westfall, P. (2010). *Multiple comparisons using R*. London: Chapman and Hall.

Bross, I. D. J. (1958). How to use ridit analysis. *Biometrics*, *14*(1), 18–38.

Brunner, E., Munzel, U. (2000). The nonparametric behrens-fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal*, *42*(1), 17–25.

Brunner, E., Puri, M. L. (2001). Nonparametric methods in factorial designs. *Statistical Papers*, *42*(1), 1–52.

Brunner, E., Dette, H., Munk, A. (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association*, *92*(440), 1494–1502.

Brunner, E., Konietschke, F., Pauly, M., Puri, M. L. (2017). Rank-based procedures in factorial designs: Hypotheses about non-parametric treatment effects. *Journal of Royal Statistical Society Series B*, *79*(5), 1463–1485.

Cartwright, H. V., Lindahl, R. L., Bawden, J. W. (1968). Clinical findings on the effectiveness of stannous fluoride and acid phosphate fluoride as caries reducing agents in children. *Journal of Dentistry for Children*, *35*(1), 36–40.

Dette, H., Neumeyer, N. (2001). Nonparametric analysis of covariance. *The Annals of Statistics*, *29*(5), 1361–1400.

Fischer, D., Oja, H., Schleutker, J., Sen, P. K., Wahlfors, T. (2014). Generalized Mann-Whitney type tests for microarray experiments. *Scandinavian Journal of Statistics*, *41*(3), 672–692.

Friedrich, S., Konietschke, F., Pauly, M. (2017). A wild bootstrap approach for nonparametric repeated measurements. *Computational Statistics and Data Analysis*, *113*, 38–52.

Gao, X., Alvo, M., Chen, J., Li, G. (2008). Nonparametric multiple comparison procedures for unbalanced one-way factorial designs. *Journal of Statistical Planning and Inference*, *138*(8), 2574–2591.

Grigoletto, M., Akritas, M. G. (1999). Analysis of covariance with incomplete data via semiparametric transformations. *Biometrics*, *55*(4), 1177–1187.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–802.

Lehmann, E. L., Romano, J. P. (2005). *Testing statistical hypotheses* (3rd ed.). New York: Springer.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Konietschke, F., Hothorn, L. A., Brunner, E. (2012). Rank-based multiple test procedures and simultaneous confidence intervals. *Electronic Journal of Statistics*, *6*, 738–759.

Munk, A., Neumeyer, N., Scholz, A. (2007). Non-parametric analysis of covariance: The case of inhomogeneous and heteroscedastic noise. *Scandinavian Journal of Statistics*, *34*(3), 511–534.

Neve, J. D., Thas, O. (2015). A regression framework for rank tests based on the probabilistic index model. *Journal of the American Statistical Association*, *110*(511), 1276–1283.

R Development Core Team. (2013). R package mvtnorm: Multivariate Normal and t Distributions. In: A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, B. Bornkamp, T. Hothorn (Eds.), *License: GPL-2*. Maintainer: Hothorn T. < Torsten.Hothorn@R − project.org >.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. New York: Chapman and Hall/CRC.

Simes, R. J. (1986). An improved bonferroni procedure for multiple tests of significance. *Biometrika*, *73*(3), 655–660.

Tamhane, A. C., Dunnett, C. W. (1999). Stepwise multiple test procedures with biometric applications. *Journal of Statistical Planning and Inference*, *82*(1–2), 55–68.

Terpstra, J. T., Magel, R. C. (2003). A new nonparametric test for the ordered alternative problem. *Journal of Nonparametric Statistics*, *15*(3), 289–301.

Thas, O., Neve, J. D., Clement, L., Ottoy, J. P. (2012). Probabilistic index models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *74*(4), 623–671.

Tsangari, H., Akritas, M. G. (2004). Nonparametric ANCOVA with two and three covariates. *Journal of Multivariate Analysis*, *88*(2), 298–319.

Wang, L., Akritas, M. G. (2006). Testing for covariate effects in the fully nonparametric analysis of covariance model. *Journal of the American Statistical Association*, *101*(474), 722–736.