AISM (2007) 59: 197–210
DOI 10.1007/s10463-006-0049-5

Tathagata Banerjee · Atanu Biswas

Testing for the absence of random effects in a two-way nested design with mixed effects model: a nonparametric approach

Received: 13 August 2004 / Revised: 11 October 2005 / Published online: 17 June 2006

Abstract Two-way nested design with mixed effects model arises in many practical situations. In the classical analysis of variance set-up, a test for the absence of the random effects is obtained under the assumption that the random effects and the errors are normally distributed. The present paper avoids this assumption and provides an asymptotically distribution-free test procedure for the above problem. The asymptotic null distribution of the test statistic is obtained. Actual implementation of the test is straight forward given the prior information on quantiles of the intra-block differences of observations. In the absence of such information, working test procedures are proposed. The performances of these tests are compared with the classical analysis of variance test through simulations. The tests are then illustrated by some real data sets.

Keywords Analysis of variance · Asymptotically distribution free · Normal error

1 Introduction

Consider a two-factor experiment with factors, say, A and B. The factor B is said to be nested within A if each of its levels is observed in conjunction with just one level of the second factor. If we denote the levels of A by $A_i, \ i = 1, \ldots, r$, then within each A_i there are s levels of B and these are denoted by $B_{ij}, \ j = 1, \ldots, s$.

T. Banerjee
Indian Institute of Management Ahmedabad,
Vastrapur, Ahmedabad, 380015, India
E-mail: btathaga@yahoo.com

On leave from University of Calcutta, India

A. Biswas (✉)
Applied Statistics Unit, Indian Statistical Institute,
203 B. T. Road, Kolkata, 700 108, India
E-mail: atanu@isical.ac.in

This is a two-factor nested design. Each of the factors is either fixed or random. In this paper, we consider a mixed model set-up with the fixed factor A and the random factor B. For a comprehensive analysis of such design under usual parametric set-up, see Scheffe (1959). For interesting applications of such design, we refer to Montgomery (1984) and Dean and Voss (1999). Morgan (1996) has given an excellent account of nested designs in his review paper.

In the classical analysis of variance, it is assumed that the underlying distributions of the random components in the model are normal. This assumption is essential for carrying out tests of hypotheses. Prevalence of nonnormality, in practice, however, restricts us to make such sweeping assumption unless supported by strong evidence. Nonparametric tests are proposed to circumvent this problem. Nonparmetric methods in design and analysis of experiments are being developed quite extensively during seventies and eighties. Good reviews of such methods are available in Brunner and Puri (1996) and Dean and Wolfe (1996).

In this paper, we propose an asymptotically distribution-free test for tesing the absence of random effects assuming a classical mixed effects model with two factors one is nested within the other. Specifically, we assume the nested factor is random. Under the same model, an asymptotically nonparametric test for testing the equality of the fixed effects is proposed by Brunner and Neumann (1982). This is discussed in Brunner and Puri (1996). However, to the best of our knowledge, testing for the absence of random effects has not been addressed in the literature so far . This problem is important from the practical point of view by its own merit.

In Sect. 2, the test procedure is described after introducing the testing problem in nonparametric set-up. Asymptotic distribution is studied in Sect. 3, and the problem of a judicious choice of the test statistic supported by simulation studies are considered and recommendation is made in Sect. 4. In Sect. 5, the proposed test procedure is applied to some real data sets. Sect. 6 ends with concluding discussion.

2 The test procedure

2.1 The problem

Suppose there are r levels of treatment A which are denoted by $A_i,\ i = 1, \dots, r$, and within each A_i there are s levels of B which are denoted by $B_{ij},\ j = 1, \dots, s$. Analysis of unbalanced models is complicated even in the parametric set-up. So we confine ourselves to the balanced case only. In the following, we consider a mixed effect model assuming that s levels of B are randomly chosen from a large number of possible levels at each level of A. The model is thus given by

$$Y_{ijk} = \mu + \alpha_i + b_{j(i)} + e_{ijk}, \quad i = 1, \dots, r, \quad j = 1, \dots, s, \quad k = 1, \dots, n, \tag{1}$$

where

Y_{ijk} kth observation corresponding to the level B_{ij},
μ main effect,
α_i effect due to A_i, assumed to be fixed with $\sum_{i=1}^{r} \alpha_i = 0$,
$b_{j(i)}$ effect due to B_{ij}, assumed to be random,
e_{ijk} random error, independent of $b_{j(i)}$.

We assume

$$b_{j(i)} \sim F_b(.) \quad \text{and} \quad e_{ijk} \sim F_e(.),$$

independently of each other. Both the distributions are symmetric about '0'. In the parametric set-up, it is assumed that $b_{j(i)}$'s and e_{ijk}'s are independently and normally distributed with means 0 and variances σ_b^2 and σ_e^2, respectively, and the test for absence of random components $b_{j(i)}$ in model (1) is equivalent to testing

$$H : \sigma_b^2 = 0 \quad \text{against} \quad K : \sigma_b^2 > 0.$$

In the nonparametric set-up the above null hypothesis is that $F_b(.)$ is degenerate at '0', i.e., $F_b(-\epsilon) = 0,\ F_b(\epsilon) = 1$ for every $\epsilon > 0$. A nonparametric test procedure for testing H is proposed in the following sub-section.

2.2 Test procedure

We observe that

$$\begin{aligned} Y_{ijk} - Y_{ijk'} &= e_{ijk} - e_{ijk'}, \quad k \neq k', \\ Y_{ijk} - Y_{ij'k'} &= \left(b_{j(i)} - b_{j'(i)}\right) + \left(e_{ijk} - e_{ij'k'}\right), \quad j \neq j'. \end{aligned}$$

Note that each $Y_{ijk} - Y_{ij'k'}$, $j \neq j'$, is more dispersed than each of $Y_{ijk} - Y_{ijk'}$, $k \neq k'$. Now we define the indicator variables $\left\{u_{kk'}^{(j)(i)}, v_{(jk)(j'k')}^{(i)}\right\}$ as follows:

$$\begin{aligned} u_{kk'}^{(j)(i)} &= 1 \text{ or } 0 \text{ according as } |Y_{ijk} - Y_{ijk'}| > c \text{ or not,} \quad \text{and} \\ v_{(jk)(j'k')}^{(i)} &= 1 \text{ or } 0 \text{ according as } |Y_{ijk} - Y_{ij'k'}| > c \text{ or not,} \end{aligned}$$

where 'c' is a preassigned positive constant. The choice of 'c' is in the experimenter's hand. The problem of a judicious choice of c and hence the test statistic has been discussed in Sect. 4. We now define the following statistics:

$$U_i = \sum_j \sum_{k<k'} u_{kk'}^{(j)(i)}, \quad V_i = \sum_{j<j'} \sum_{k,k'} v_{(jk)(j'k')}^{(i)},$$

and

$$U = \sum_{i=1}^{r} U_i, \quad V = \sum_{i=1}^{r} V_i.$$

It is easy to find that

$$E(U_i) = s\binom{n}{2} p_1, \tag{2}$$

irrespective of whether H is true or not and

$$E_K(V_i) = n^2 \binom{s}{2} p_1' \quad \text{and} \quad E_H(V_i) = n^2 \binom{s}{2} p_1,$$

the expectations under K and H, respectively, where

$$p_1 = P_H\left\{|e_{ijk} - e_{ijk'}| > c\right\}, \quad k \neq k',$$
$$p_1' = P_K\left\{|(b_{j(i)} - b_{j'(i)}) + (e_{ijk} - e_{ij'k'})| > c\right\}, \quad j \neq j'.$$

Note that $p_1' > p_1$ under any alternative hypothesis. Then, we set

$$T_1 = \frac{V}{r\binom{s}{2}n^2} - \frac{U}{rs\binom{n}{2}}.$$

Note that T_1 is expected to be larger under any alternative than under H, and hence T_1 can be used as a test statistic. Right-tailed test is suggested which rejects H if

$$T_1 > t_1,$$

where t_1 is so chosen to have a level 'γ' test. Note that the test proposed above is not strictly distribution free as the distribution of T_1 depends on the parent distribution. However, one can get asymptotically distribution-free test based on the null distribution of T_1. Asymptotic distribution of T_1 has been derived in the next section.

3 Asymptotic distribution

Brunner and Neumann (1982), while tesing for equality of α_i's assuming the model (1), found the asymptotic approximation to the distribution of the rank test statistic with s tending to infinity. Brunner and Denker (1994), on the other hand, considered asymptotic approximation when both n and s (in a certain rate depending on n) tend to infinity for two sample location problem. In the following, we state a theorem for finding asymptotic approximation to the distribution of T_1 when both n and s (in a certain rate depending on n) tend to infinity. Let us denote

$$p_2 = P_H\left\{|Y_{ijk} - Y_{ij'k'}| > c,\ |Y_{ijk} - Y_{ij''k''}| > c\right\},$$
$$p_2' = P_K\{|Y_{ijk} - Y_{ij'k'}| > c,\ |Y_{ijk} - Y_{ij''k''}| > c\}.$$

Now we have the following two theorems whose proofs are given in the Appendix.

Theorem 3.1 *Under H,*

$$E(T_1) = 0,$$

and

$$\mathrm{Var}(V) = r\left[n^2\binom{s}{2}p_1(1-p_1) + n^2(n-1)s(s-1)\left(p_2 - p_1^2\right)\right.$$
$$\left. + n^3 s(s-1)(s-2)\left(p_2 - p_1^2\right)\right]$$

and

$$\mathrm{Var}(T_1) = p_1(1-p_1)\left[\frac{1}{r\binom{s}{2}n^2} + \frac{1}{rs\binom{n}{2}}\right] + \left(p_2 - p_1^2\right)\left[\frac{2(n-1)+2(s-2)n}{r\binom{s}{2}n^2} + \frac{2(n-2)}{rs\binom{n}{2}} - \frac{8}{rsn}\right].$$

Theorem 3.2 *Under H, as $s \to \infty$, $n = O\left(s^{1+\delta}\right)$, $\delta \geq 0$,*

$$\frac{T_1}{\sqrt{\mathrm{Var}_H\left(\left\{r\binom{s}{2}n^2\right\}^{-1} V\right)}} \xrightarrow{d} N(0,1).$$

Remark 1 As $p_1' - p_1$ is equal to or greater than 0 under H or K, a right-tailed test is appropriate.

Remark 2 From Remark 1, the proposed test is consistent against any fixed alternative.

Remark 3 Note that

$$\frac{E(V^* - \tilde{V})^2}{Var(\tilde{V})} \not\to 0 \quad \text{as } n \to \infty.$$

On the other hand, for ensuring

$$\frac{\left[r\binom{s}{2}n^2\left(p_1 - \left\{rs\binom{n}{2}\right\}^{-1} U\right)\right]}{\sqrt{\mathrm{Var}_H(V)}} \to 0,$$

as s tends to infinity does not work. We need to make n tending to infinity maintaining appropriate order relation with n.

But, it is again an interesting question: "how large s and n should be for the asymptotics to work?" We carried out a detailed simulation study for $r = 2$, $s = 10$, $n = 6$. We report our findings in Sect. 5.

4 Choice of c: some suggestions

To carry out the above test in Neyman–Pearson sense, one needs to specify the value of c before observing the data. In essence, the problem is similar to the basic problem of choosing a score in a given nonparametric testing context as mentioned in Sect. 1 of the classic book by Hajek and Sidak (1967). Note, in our context, choosing a value of c amounts to choosing a score that leads to a specific nonparametric test. For rank tests, the general problem of choosing a score was first adressed and tackled by Hajek and Sidak (1967) and later by Hajek (1969, 1970), Hogg et al. (1975). A good review is given in Sidak et al. (1999). Chatterjee and Banerjee (1986, 1991) gave a solution to this general problem in nonparmetric linear regression set-up.

Although our problem here is in essence similar to the problem noted above, the experimental set-up is, however, more complicated. In fact none of the above solutions would be directly applicable here. We offer a workable solution in the following.

A natural choice of c is a higher-order quantile of the distribution of intrablock differences. In experimental design, prior knowledge of such quantiles may usually be available from historical data from past studies. To motivate, we suppose a company produces an item in several factories located in several places of a country (several countries). The product of each factory may go to a number of locations for sales. A given location may have products from several factories too. Suppose the response is a quality characteristic (quantitative, may be lifetime of a sports shoe/television/laptop)and Y_{ijk} denotes the response of the kth unit in the jth location (where the products are being used) from the ith factory. Here α_i is the ith factory effect and $b_{j(i)}$ is the jth location effect nested within the ith factory. In such cases life testing experiments in each factory would generate enough data for having precise estimates of the quantiles of intrablock differences and hence can be effectively utilized for determining c.

Given the higher-order quantiles, the question is, which of these is to be chosen as the value of c or more specifically is there any optimum choice of c? Intuitively it is clear that larger the $r\binom{s}{2}n^2$, the power of the test against a nonlocal alternative would, in general, increase with the increase in the chosen value of c. However, for a fixed set of values of r, s and n, the value of c should be so chosen as to ensure reasonable estimates of tail probabilities both under the null and the alternatives. Thus an optimum choice of c irrespective of different configurations of r, s and n does not exist. On the other hand, for a given choice of r, s and n and a given alternative, the performance of the test would depend on the shape of the parent distribution (see Sidak et al. 1999; Hajek 1969, 1970; Hogg et al. 1975; Chatterjee and Banerjee 1986, 1991). Hajek (1969, 1970), in particular, showed that change in tail weights of the unknown parent distribution significantly affects the optimum choice of the score. Thus, in order to prescribe a workable choice of c, we conduct simulation studies to investigate the sensitivity of the optimum value of c to the tail weight, to different configurations of r, s and n and to different alternatives. The main idea is, of course, to prescribe a guideline for the choice of c given the prior information about the tail weight of the null distribution of intra-block difference.

Although we have considered different configurations of r, s and n, for brevity, we report only the computations for $r = 2$, $s = 3$, $n = 6$. Also we have chosen four distributions according to their tail weights, viz., light tailed (Double Exponential), moderate tailed (Normal distribution), heavy tailed (Cauchy) and abrupt tailed (rectangular). The pattern we observe for other configurations of r, s and n is similar. We carry out 10,000 simulations. Generating data from a distribution, we estimate c as a predetermined quantile of the null distribution of the intrablock differences (70, 75, 80, 90, 95 and 97 quantiles are considered). Consequently we obtain t_1, the cut-off point, and the powers at different alternatives keeping the level at 5%. A part of the results is presented in Table 1.

The results show that both tail weights and the design in terms of r, s and n have significant effect on the choice of c. It also depends on the chosen alternative. For abrupt tailed (rectangular) higher quantile seems to be a better choice while for heavy tailed (Cauchy) not so high quantile will be a good choice. We keep the

Table 1 Power at different alternatives for different parent distributions where c is considered as different quantiles (here $\sigma_e^2 = 1$)

σ_b^2	c as different quantiles						T_1^*
	70%	75%	80%	90%	95%	97%	
Normal parent							
1.0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
1.6	0.1884	0.2864	0.3842	0.7134	0.8061	0.7977	0.6440
2.2	0.3968	0.6317	0.8280	0.9942	0.9978	0.9952	0.8440
2.8	0.5949	0.8624	0.9757	0.9999	0.9999	0.9999	0.8860
Double exponential parent							
1.0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
1.6	0.2058	0.2860	0.4074	0.6868	0.7596	0.7502	0.7423
2.2	0.4290	0.6294	0.8477	0.9840	0.9837	0.9741	0.9232
2.8	0.6312	0.8589	0.9793	0.9990	0.9982	0.9961	0.9750
Cauchy parent							
1.0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
1.6	0.2234	0.3102	0.4187	0.5625	0.5889	0.6579	0.7424
2.2	0.4540	0.6682	0.8310	0.8670	0.8497	0.8519	0.8322
2.8	0.6682	0.8818	0.9634	0.9379	0.9306	0.9010	0.8783
Rectangular parent							
1.0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
1.6	0.1811	0.2512	0.3204	0.6601	0.8557	0.8962	0.3700
2.2	0.3648	0.5406	0.7189	0.9925	0.9991	0.9994	0.4560
2.8	0.5455	0.7822	0.9294	0.9997	1.0000	1.0000	0.8333

choice at the discretion of the experimenter subject to the availability of the prior information about the quantiles. See Banerjee (1984) for a discussion on utilization of additional information in nonparametric tests.

In case such information is not at all available, one could as well perform the test procedure by choosing several possible values of c (which will correspond to different quantiles). If c is sufficiently small or sufficiently large, then the detection power of the test on the basis of a finite sample will be almost nil (as T_1 will be zero in both the cases). Thus, one can think of $T_1(c)$ for $c \in \mathcal{R}$, defining a class of tests and it is rational to choose c to maximize the power of the test. But, in this case, since the parent distribution is unknown, choosing 'c' by maximizing power is not feasible. Instead, we can maximize a sort of empirical power, that is we can minimize the P value. Thus, one can take a decision on the basis of the observed P value.

As one referee has suggested, one sensible approach might be to consider

$$T_1^* = \max_{0<c<\infty} T_1(c)$$

as a test statistic. We carried out a detailed simulation study to investigate the null and nonnull distributions of T_1^*. The power of the test is compared with that of the tests based on $T_1(c)$ for different c, in Table 1. The P-values are compared with the proposed test and that of F-test or permutation test for some real data sets in Table 3. We observe that the performance of the test with T_1^* as the test statistic is reasonably good in terms of power and P-values. So this can be recommended for application. But, certainly the power is much more if one can identify the appropriate percentile point. Since that is not possible, T_1^* is a very good working solution.

5 Example and simulation studies

In the following example, we illustrate our procedure with a real life data set in the absence of any prior information. First, we choose 'c'minimizing the P value and then apply the asymptotic test based on T_1.

5.1 Example

We use an unpublished data set (S. Sarkar, unpublished data) obtained from an experimental study of the effects of different hormones on the ovarian weight of the soft-shelled turtles *Lissemys punctata punctata*, which is an endangered species. The data are collected at two different time points, i.e., $r = 2$. At each time point there are three blocks (i.e., $s = 3$) corresponding to control and two different hormones. Two blocks at each time point record responses of six turtles ($n = 6$) treated with control and Leutinizing hormone (LH). The other two blocks one each at a given time point record responses of the turtles treated with Estradiol-17β (E-17β) anf Follicle stimulating hormone (FSH). Dose applied was 15-μg per 100-g body weight for 15 days. The data are provided in Table 2

We take different percentile points as determined from the intra-block differences of the data as the choices of c and the P values of permutation test based on T_1 are reported in Table 3. We take $c = 5.7$ as it minimizes the P value of the permutation test based on T_1. It is to be noted that the P value corresponding to the asymptotic test based on T_1 broadly agrees with the P values corresponding to the permutation tests based on T_1 and T^*. On the otherhand, it should be noted that the P value of the paprametric F test differs from others by a factor 1,000.

Next we present a limited simlulation study to investigate (i) how the asymptotic approximation to the distribution of T_1 works even for moderate values of s and n? (ii) when would we expect our test to outperform the parametric F test completely? We investigate both these questions by comparing the attained level of significance with the stipulated level of significance.

Table 2 Ovarian weight of soft-shelled turtle at two different times using different treatments and control

A_1 (Time point 1)			A_2 (Time point 2)		
B_1: Control	B_2: LH-treated	B_3: E-17β treated	B_1: Control	B_2: LH-treated	B_3: FSH-treated
22.8	38.0	30.0	24.0	28.4	30.2
27.4	37.3	26.7	27.8	35.0	29.6
32.4	40.3	27.2	28.0	35.7	31.2
25.4	35.4	28.0	25.3	38.9	28.4
27.5	36.3	32.2	31.0	36.4	30.3
30.0	41.3	33.1	25.3	37.0	31.0

Table 3 P values of the turtle data using the permutation test and the asymptotic test for different percentile points taken as the choice of c

Percentile	Exact permutation test				
points taken	c	U	V	T_1	P value
70	3.2	32	163	0.3991	1.67×10^{-4}
75	3.8	29	160	0.4185	1.00×10^{-4}
80	4.2	24	154	0.4463	1.00×10^{-4}
90	5.7	12	116	0.4037	0.67×10^{-4}
95	7.0	6	96	0.3778	0.67×10^{-4}
97	8.0	3	75	0.3139	1.67×10^{-4}

P value of the asymptotic test (T_1-statistic) = 1.44×10^{-4}
P value of the parametric test (F-statistic) = 4.29×10^{-8}
P value of the test based on T_1^* = 0.75×10^{-4}

5.2 Simulation studies

We design our simulation study with samples from normal distribution assuming $\mu = 0, \alpha_{(i)} = 0$, $b_{j(i)} = 0$ and $e_{ijk} \sim \text{Normal}(0, 1)$. The configuration chosen is $r = 2, s = 10$ and $n = 6$. For applying our test, we take c as 80% quantile of the null distribution of intrablock difference. We observe that the level attained by the asymptotic test based on T_1 is almost same for $\alpha = 0.05$ and above. For $\alpha = 0.01$ there is a small discrepancy. The asymptotic procedure, thus, seems to work well even for a value of s as small as 10 and n equal to 6. In addition, we have carried out a simulation study to understand the robustness of our proposed nonparametric test. For this we contaminate 15% of e_{ijk}'s by observations from $N(0, 5^2)$. Assuming absence of contamination we carry out usual papramentric F test and the asymptotic test based on T_1 at 5% level of significance repetetively for 10,000 samples. The attained levels of significance of the asymptotic test based on T_1 and the the parametric test based on F are found to be 0.038 and 0.56, respectively. Thus the test based on T_1 outperforms the usual F test. It shows that the test based on T_1 is highly robust especially in the presence of outliers.

6 Concluding remarks

In this paper, we have proposed a nonparametric test for testing the presence of a variance component in the nested two-way mixed model. We now discuss two possible generalizations of the proposed test. Suppose, instead of c, we have a sequence of constants $\{c_1, c_2, \ldots, c_t\}$ such that $0 < c_1 < c_2 < \cdots < c_t$. Setting the intervals

$$I_j = [c_{j-1}, c_j), \quad j = 1(1)t + 1,$$

with $c_0 = 0$ and $c_{t+1} = \infty$, we can consider the following general scores:

$$u_{kk'}^{(j)(i)} = d \quad \text{if} |Y_{ijk} - Y_{ijk'}| \in I_d,$$

and

$$v_{(jk)(j'k')}^{(i)} = d \quad \text{if} |Y_{ijk} - Y_{ij'k'}| \in I_d,$$

to propose a generalization of our proposed procedure. For each such interval, we can find a statistic of type T_1 defining an indicator. First one needs to take a convex combination of these two types of quantities separately. The weights are so chosen that it increases (or decreases) with d, i.e., the subscript of c so that the observations away from central part get increasing weights with the increase of the distance. Then the difference of U and V after adjusting for number of such differences will serve as a test statistic. Choice of c_j's are also to be made using some quantiles as before. The study will be taken up in future, and we intend to communicate it in a separate issue. A random effect model like (1) is also true in different bio-medical problems. For example, in teratological experiments some dose levels are considered, and there may be many other such possible dose levels. The same is true in dose response studied in toxicity (phase I trial). The asymptotic results in terms of $s \to \infty$ is conceptually alright and has some physical interpretation. For example, if we are interested in the production rate by a machine we can take a block consisting the production in one hour. Instead of increasing the time, we can take many such one hour slots for our purpose. Such types of works where asymptotics are done in terms of s are available in literature. [see, for example, the asymptotics of Friedman statistic as the number of blocks tends to infinity in Hajek's (1967) book.] But, instead of s, if we want the asymptotic results in terms of n, there is some problem with the test statistic. Here $n \times \mathrm{Var}_H(T_1)$ converges to zero as n goes to infinity. To interpret the results in terms of n we do the following. Define

$$V_i^* = \sum_{\substack{j<j' \\ (j,j')\neq(s-1,s)}} \sum_{k\neq k'} v^{(i)}_{(jk)(j'k')},$$

and consequently setting $V^* = \sum_{i=1}^{s} V_i^*$, we note that

$$E_H(V_i^*) = n^2 \left[\binom{n}{2} - 1\right] p_1.$$

Then setting

$$T_2 = \frac{V^*}{r\left[\binom{s}{2} - 1\right] n^2} - \frac{U}{rs\binom{n}{2}},$$

a right-tailed test based on T_2 can be suggested. As in (4) and (5), we can obtain

$$\begin{aligned}\mathrm{Var}_H(V^*) = r\Bigg\{ & n^2 \times \left[\binom{s}{2} - 1\right] p_1(1-p_1) + 2n^2(n-1) \\ & \times \left[\binom{s}{2} - 1\right] \left(p_2 - p_1^2\right) + n^3[s(s-1) - 4](s-2)\left(p_2 - p_1^2\right)\Bigg\},\end{aligned}$$

$$\mathrm{cov}_H(U, V^*) = rn^2(n-1)(s+1)(s-2)\left(p_2 - p_1^2\right).$$

Then, as in Theorem 3.2, we have as $n \to \infty$, under H,

$$\sqrt{n}T_2 \xrightarrow{d} N(0, \sigma_2^2),$$

where

$$\sigma_2^2 = \frac{4(s^3 - 4s^2 - s + 13)}{r(s^2 - s - 2)} \left(p_2 - p_1^2\right) > 0.$$

The test is also consistent. There is one problem with the test provided by T_2. The behavior and performance of the test depends on the pair excluded from the possible $\binom{s}{2}$ possible pairs for each level of A considered in defining T_1. Here we have excluded the $(s-1, s)$th pair, but if we could exclude any other pair, T_2 would have any other realization with a possibility of other decision. It is to be noted that, as we are considering n tending to infinity and s being kept fixed, probably dropping $\binom{s}{2}$ observations do not have serious effect on inference.

Appendix

Proof of Theorem 3.1 The expectation part is immediate from (2) and (3). In variance of U, there are $\binom{n}{2}$ components of $\mathrm{Var}_H\left(u_{kk'}^{(j)(i)}\right)$ $[= p_1(1-p_1)]$ and $n(n-1)(n-2)$ components of $\mathrm{cov}_H\left(u_{kk'}^{(j)(i)}, u_{kk''}^{(j)(i)}\right)\left[= p_2 - p_1^2\right]$ in each block of each level. Hence

$$\mathrm{Var}_H(U) = rs\left[\binom{n}{2} p_1(1-p_1) + n(n-1)(n-2)\left(p_2 - p_1^2\right)\right]. \tag{3}$$

In the variance of V within each level of A, there are $\binom{s}{2}$ between block comparisons, yielding $n^2\binom{s}{2}$ variance components. We call these $\binom{s}{2}$ as reduced blocks. Within each block there are $n^2(n-1)s(s-1)$ covariance parts. [There are n subblocks within each reduced block, $\binom{n}{2}$ covariances for each subblock and there are $\binom{s}{2}$ reduced blocks.] Between the reduced blocks, there are some covariance factors. There are n subblocks within each reduced block and one subblock has n^2 covariance components with one other block. There is a term $s(s-1)(s-2)$ which comes in the same way as $n(n-1)(n-2)$ comes in $\mathrm{Var}_H(U)$. Thus

$$\begin{aligned}\mathrm{Var}_H(V) = r\Bigg[&n^2\binom{s}{2} p_1(1-p_1) + n^2(n-1)s(s-1)(p_2 - p_1^2) \\ &+ n^3 s(s-1)(s-2)\left(p_2 - p_1^2\right)\Bigg].\end{aligned} \tag{4}$$

Similarly, to find the $\mathrm{cov}_H(U, V)$ we note that, for U there are s blocks, in V there are n subblocks within each reduced block, there are $n(n-1)$ covariance terms between one block of U and one subblock of a reduced block V, and components (Y_{ijk}) of one block of U are spread into $(s-1)$ reduced blocks of V. Thus

$$\mathrm{cov}_H(U, V) = an^2(n-1)s(s-1)\left(p_2 - p_1^2\right). \tag{5}$$

Using (3)–(5), the theorem is immediate. □

Proof of Theorem 3.2 Note that

$$\frac{T_1}{\sqrt{\mathrm{Var}_H\left(\left\{r\binom{s}{2}n^2\right\}^{-1} V\right)}} = \frac{V - E_H(V)}{\sqrt{\mathrm{Var}_H(V)}} + \frac{\left[r\binom{s}{2}n^2\left(p_1 - \left\{rs\binom{n}{2}\right\}^{-1} U\right)\right]}{\sqrt{\mathrm{Var}_H(V)}}. \tag{6}$$

Let $\tilde{V}$ be the projection of V^* (the standardized version of V, i.e., $V - E(V)$). For any fixed $i,\ j < j'$, consider

$$\begin{aligned} V_i^{(j,j')} &= \sum_{k \neq k'} v^{(i)}_{(jk)(j'k')} - n^2 p'_1 \\ &= V\left(Y_{ij1}, \ldots, Y_{ijn}, Y_{ij'1}, \ldots, Y_{ij'n}\right) \quad \text{(say).} \end{aligned}$$

Define

$$\begin{aligned} \tilde{V}(y_{ij1}) &= EV\left(y_{ij1}, Y_{ij2}, \ldots, Y_{ijn}, Y_{ij'1}, \ldots, Y_{ij'n}\right) \\ &= nP\left\{|y_{ij1} - Y| > c\right\} + n(n-1)P\left\{|Y_{ij1} - Y_{ij'1}| > c\right\} - n^2 p'_1 \\ &= n\left[P\left\{|y_{ij1} - Y| > c\right\} - p'_1\right], \end{aligned}$$

where Y has same distribution as any Y_{ijk}. Clearly,

$$\begin{aligned} &n^{-2}\mathrm{Var}\left(\tilde{V}(Y_{ij1})\right) \\ &= E\left[P\left\{|Y_{ij1} - Y| > c|Y_{ij1}\right\}\right]^2 - \left[EP\left\{|Y_{ij1} - Y| > c|Y_{ij1}\right\}\right]^2 \\ &= E\left[P\left\{|Y_{ij1} - Y_{ij'k'}| > c|Y_{ij1}\right\} P\left\{|Y_{ij1} - Y_{ij''''}| > c|Y_{ij1}\right\}\right] - p_1'^2 \\ &= E\left[P\left\{|Y_{ij1} - Y_{ij'k'}| > c, |Y_{ij1} - Y_{ij''k''}| > c|Y_{ij1}\right\}\right] - p_1'^2 \\ &= P\left\{|Y_{ij1} - Y_{ij'k'}| > c, |Y_{ij1} - Y_{ij''k''}| > c\right\} - p_1'^2 \\ &= p'_2 - p_1'^2. \end{aligned}$$

Now, define

$$\tilde{V}_i^{(j,j')} = \tilde{V}(Y_{ij1}) + \cdots + \tilde{V}(Y_{ijn}) + \tilde{V}(Y_{ij'1}) + \cdots + \tilde{V}(Y_{ij'n}),$$

and

$$\tilde{V}_i = \sum_{1 \le j < j' \le s} \tilde{V}_i^{(j,j')} = (s-1)\sum_{j=1}^{s}\left\{\tilde{V}(Y_{ij1}) + \cdots + \tilde{V}(Y_{ijn})\right\}.$$

Hence

$$\tilde{V} = \sum_{i=1}^{r} \tilde{V}_i.$$

Consequently,

$$\mathrm{Var}(\tilde{V}) = (s-1)^2 rsn\ \mathrm{Var}\left(\tilde{V}(Y_{ij1})\right) = (s-1)^2 rsn^3\left(p'_2 - p_1'^2\right).$$

Since $\tilde{V}$ is the projection of V^*, it can be easily shown that

$$\frac{E(V^*-\tilde{V})^2}{\text{Var}(\tilde{V})}=\frac{\text{Var}(V^*)}{\text{Var}(\tilde{V})}-1$$
$$=\frac{r\left[n^2\binom{s}{2}p_1'(1-p_1')+n^2(n-1)s(s-1)\left(p_2'-p_1'^2\right)+n^3s(s-1)(s-2)\left(p_2'-p_1'^2\right)\right]}{(s-1)^2srn^3\left(p_2'-p_1'^2\right)}-1$$
$$\to 0 \quad \text{as } s\to\infty.$$

Hence $\left(\tilde{V}/\sqrt{\text{Var}(\tilde{V})}\right)$ and $\left(V^*/\sqrt{\text{Var}(V)}\right)$ have the same asymptotic distribution. Since $\tilde{V}$ is the sum of bounded independent random variables

$$\frac{\tilde{V}}{\sqrt{\text{Var}(\tilde{V})}}\xrightarrow{d} N(0,1) \quad \text{as } s\to\infty.$$

Hence

$$\frac{V-E_H(V)}{\sqrt{\text{Var}_H(V)}}=\frac{V^*}{\sqrt{\text{Var}_H(V)}}\xrightarrow{d} N(0,1) \quad \text{as } s\to\infty.$$

So it is enough to prove that the second term in (6) converges in probability to zero. Now,

$$\frac{\left[r\binom{s}{2}n^2\left(p_1-\left\{rs\binom{n}{2}\right\}^{-1}U\right)\right]}{\sqrt{\text{Var}_H(V)}}$$
$$=\frac{r\binom{s}{2}n^2(rs)^{-1}\sum_{i=1}^r\sum_{j=1}^s\sqrt{\text{Var}(U_{ij})}(p_1-U_{ij})/\sqrt{\text{Var}\left(U_{ij}\right)}}{O\left(n^{3/2}s^{3/2}\right)}$$
$$=\frac{s^{-1}\binom{s}{2}n^2\sum_{i=1}^r\sum_{j=1}^s\sqrt{\text{Var}(U_{ij})}Z_{ij}}{O\left(n^{3/2}s^{3/2}\right)},$$

where $U_{ij}=\binom{n}{2}^{-1}\sum_{k<k'}u_{k,k'}^{(ij)}$, and $Z_{ij}=(p_1-U_{ij})/\sqrt{\text{Var}(U_{ij})}$ are iid random variables with mean 0 and variance 1. Noting that $\sqrt{\text{Var}(U_{ij})}=\lambda n^{-1}+O(n^{-\gamma})$, $\gamma>1$, the leading term of the above becomes

$$\frac{\binom{s}{2}n\lambda\sum_{i=1}^r\left[s^{-1}\sum_{j=1}^s Z_{ij}\right]}{O(n^{3/2}s^{3/2})}.$$

Now, $s^{-1}\sum_{j=1}^s Z_{ij}\xrightarrow{P} 0$ (by WLLN), and $n=O\left(s^{1+\delta}\right)$, $\delta\geq 0$, entail the proof of the Theorem. □

Acknowledgements The authors thank two anonymous referees for their helpful comments which led to considerable improvement over the earlier version of the manuscript.

References

Banerjee, T. (1984). Utilization of additional information in non-parametric tests. *Calcutta Statistical Association Bulletin 33*, 69–84.
Brunner, E., Denker, M. (1994). Rank statistics under dependent observations and applications to factorial designs. *Journal of Statistical planning and Inference 42*, 353–378.
Brunner, E., Neumann, N. (1982). Rank tests for correlated random variables. *Biometrical Journal 24*, 373–389.
Brunner, E., Puri, M. L. (1996). Nonparametric methods in design and analysis of experiments. In S. Ghosh, C. R. Rao (Eds.), *Handbook of statistics* (Vol. 13 pp. 631–703). Amsterdam: Elsevier Science Publishers.
Chatterjee, S. K., Banerjee, T. (1986). Combining alternative rank tests for the multiple regression problem. *Calcutta Statistical Association Bulletin 35*, 169–188.
Chatterjee, S. K., Banerjee, T. (1991). Combination of multiple scores for nonparametric testing in multivariate linear regression set-up. *Communications in Statistics Theory and Methods 19*, 2967–2999.
Dean, A. M., Wolfe, D. A. (1996). Nonparametric analysis of experiments. In S. Ghosh, C. R. Rao (Eds.), *Handbook of statistics* (Vol. 13 pp. 705–758). Amsterdam: Elsevier Science Publishers.
Dean, A. M., Voss, D. (1999). *Design and analysis of experiments*. New York: Springer.
Friedmann, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association 32*, 675–699.
Hajek, J. (1969). *Non-parametric statistics.* San Francisco: Holden Day.
Hajek, J. (1970). *Miscellaneous problems of rank test theory*. In M. L. Puri (Ed.), *Non-parametric techniques in statistical inference.* Cambridge: Cambridege university Press.
Hajek, J., Sidak, Z. (1967). *Theory of rank tests*. New York: Academic.
Hogg, R. V., Fisher, D. M., Randles, R. H.(1975). A two-sample adaptive distribution-free test. *Journal of the American Statistical Association 70*, 656–661.
Kruskal, W. H., Wallis, W. A. (1952). The use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association 47*, 583–621.
Lemmer, H. H., Stoker, D. J. (1967). A distribution-free analysis of variance for the two-way classification. *South African Statistical Journal 1*, 67–74.
Montgomery, D. C. (1984). *Design and analysis of experiments* (2nd ed.). New York: Wiley.
Morgan, J. P. (1996). Nested designs. In S. Ghosh, C. R. Rao (Eds.), *Handbook of statistics* (Vol. 13. pp. 939–976). Amsterdam: Elsevier Science Publishers.
Scheffe, H. (1959). *The analysis of variance*. New York: John Wiley.
Sidak, Z., Sen, P. K., Hajek, J. (1999). *Theory of rank tests* (2nd ed.). San Diego: Academic.