Tatra Mt. Math. Publ. **26** (2003), 365–379

TATRA MOUNTAINS
Mathematical Publications

A NOTE ON MULTIPLE COMPARISONS IN THE METHOD OF M-RANKING

František Rublík — Pranab Kumar Sen

ABSTRACT. A. Friedman type statistic for testing the hypothesis of no treatment effect in random block scheme is presented, and its correction for asymptotic variance is derived. Simulation results included show that a currently used method of multiple comparisons has a high risk of wrong decision substantially exceeding risk values of its competitors, considered in this paper.

1. Introduction and the main results

Suppose that the observed data fulfill the random blocks scheme assumptions, i.e., there are n random blocks and in each block k treatments are applied. Let $X_{i1}, \dots, X_{ik}$ be response variables associated with the ith block, and let $F_i(\mathbf{x})$ be their joint distribution, for $i = 1, \dots, n$. Then the null hypothesis can be posed as

$$H_0\colon \quad \text{For each } i,\ F_i(\mathbf{x}) \text{ is symmetric in its } k \text{ arguments}. \tag{1.1}$$

This allows the functional form of F_i to be quite arbitrary and possibly different for different blocks. For simplicity, we assume that all these distributions are continuous, so that ties among the observations can be neglected with probability one.

Assume that $(R_{i1}, \dots, R_{ik})$ denotes the ranking within the ith block and

$$S_j = \sum_{i=1}^{n} R_{ij} \tag{1.2}$$

2000 Mathematics Subject Classification: 62G10.
Keywords: hypotheses testing, random blocks, multiple comparisons, Iman–Davenport statistic.

The research has been supported by the Grant VEGA 1/7295/20 from the Science Grant Agency of the Slovak Republic.

denotes the sum of ranks belonging to the jth treatment. The classical F r i e d - m a n test [2] rejects H_0 if

$$T_{\mathrm{FR}} = \frac{1}{n\sigma^2} \sum_{j=1}^{k} \left(S_j - n\frac{k+1}{2} \right)^2 \tag{1.3}$$

exceeds $(1-\alpha)$ quantile $\chi^2(k-1,\alpha)$ of the chi-square distribution with $(k-1)$ degrees of freedom; here

$$\sigma^2 = \frac{1}{k-1} \sum_{j=1}^{k} \left(j - \frac{k+1}{2} \right)^2 = \frac{k(k+1)}{12} \,. \tag{1.4}$$

Against the null hypothesis of interchangeability, we consider the set of alternatives that they are not interchangeable; this entails that at least for one pair $1 \le j < r \le k$ the random variables X_{ij} and X_{ir} are not interchancheable. Although the S_j defined above are based on all the nk observations, intuitively one may consider the paired rank-differences, and formulate in terms of the $S_j - S_r$, apparently ignoring the impact of the remaining $k-2$ variables entering due to the nonlinear nature of the ranks.

In the case that the null hypothesis is rejected, the ith and the jth treatment are considered to be different if

$$\frac{|S_i - S_j|}{\sqrt{n\sigma^2}} > q_{k,\alpha} \,, \tag{1.5}$$

where $q_{k,\alpha}$ is the $1-\alpha$ quantile of the distribution

$$\mathcal{L}^* = \mathcal{L}\Big(\max_{i,j=1,\dots,k} |x_i - x_j| \,\Big|\, N_k(0, I_k) \Big) \,. \tag{1.6}$$

Validity of this rule is derived by N e m é n y i in Chapter 6 of [6]; it is used in many statistical monographs and has been included also into [3], where also tables of $q_{k,\alpha}$ can be found. In the further text this classical procedure (1.5) will be referred as *clas*.

If one can conceive of some random block effects leading to the intrablock correlation, then in the absence of treatment effects, the X_{ij} are interchangeable. Thus, the above procedure is really geared to detect treatment differences in a multiple comparison setup, eliminating the random block effects through intrablock ranking which are invariant to such a random shift. The advantage of using (1.5) is that X_{ij} need not be observable, it suffices to have their relative ranking (within each block) as is the case with the method of ranking. For such shift model, involving observable X_{ij}'s, it is shown in [11] that the aligned rank tests possess better power properties, though they might be less robust to interblock distributional differences than the intra-block ranking method based tests.

Multiple comparisons based on aligned rank statistics have also been considered, though in our present setup we would not pursue them.

To improve under the hypothesis the fit of the size with the nominal significance level and to obtain a test stronger under alternative, Iman and Davenport proposed in [4] to transform the Friedman statistic. Their test statistic is

$$T_{\mathrm{ID}} = \frac{(n-1)T_{\mathrm{FR}}}{n(k-1) - T_{\mathrm{FR}}}\,,$$

and H_0 is rejected whenever

$$T_{\mathrm{ID}} > F\big(k-1,\,(n-1)(k-1),\,\alpha\big)\,. \tag{1.7}$$

Since in the process of simulating power of these tests it turned out that for testing H_0 at α the modification

$$T_{\mathrm{ID}} = \begin{cases} F\big(k-1\,,\ (n-1)(k-1),\,\alpha\big)+1\,, & S^{(\cdot)} = (n, 2n, \dots, kn)\,,\\ \frac{(n-1)T_{\mathrm{FR}}}{n(k-1)-T_{\mathrm{FR}}}\,, & \text{otherwise}\,, \end{cases} \tag{1.8}$$

is necessary, because in the upper case the denominator of T_{ID} is zero, in the further text this form of the Iman–Davenport statistic and the rule (1.7) will be used.

Perhaps the only disadvantage of the Iman–Davenport approach is that it does not automatically yield a corresponding multiple comparison method. The book [1] recommends after rejecting the null hypothesis by (1.7) to declare as different the ith and the jth treatment, if

$$|S_i - S_j| > t_{1-\alpha/2}\big((n-1)(k-1)\big)\sqrt{\frac{2}{(n-1)(k-1)}\left[\frac{n^2k(k+1)(2k+1)}{6} - \sum_{m=1}^{k} S_m^2\right]}\,,$$

where $t_{1-\alpha/2}\big((n-1)(k-1)\big)$ is the quantile of the Student t distribution. Since this procedure is used in conjunction with the T_{ID} statistics, it will be referred as the CID method.

The quadratic statistic (1.3) uses for $a(R_{ij})$ the score function $J(x) = x$; the problem of optimal choice of the score function J is studied by Sen in [10], where the optimality of the Friedman test statistic is established in the case that the sampling is made from the logistic distribution. Another modification of the Friedman test statistic and corresponding multiple comparison methods are given in the following theorem.

Theorem 1.1. *Suppose that* $I = \left\langle \frac{2}{k+1}\,, \frac{2k}{k+1}\right\rangle$ *and* $g\colon I \to \mathbb{R}^1$ *is a function which has on the interior of* I *continuous the second derivative* $g^{(2)}$. *Assume* *further that the hypothesis* (1.1) *of no treatment effect holds, the distribution functions* (1.1) *are continuous and put*

$$Z_j = \frac{2S_j}{n(k+1)}. \tag{1.9}$$

(I) *Let the first derivative* $g^{(1)}(1) \neq 0$. *For the statistic*

$$T_{2,g} = n\,\frac{3(k+1)}{k\big(g^{(1)}(1)\big)^2}\sum_{j=1}^{k}\big(g(Z_j) - g(1)\big)^2 \tag{1.10}$$

the convergence in distribution

$$T_{2,g} \longrightarrow \chi^2_{k-1} \tag{1.11}$$

holds as $n \to \infty$. *Moreover, as* $n \to \infty$,

$$\sqrt{n\frac{3(k+1)}{k\big(g^{(1)}(1)\big)^2}}\max_{i,j=1,\dots,k}\big|g(Z_i) - g(Z_j)\big| \longrightarrow \mathcal{L}^* \tag{1.12}$$

in distribution. Here $\mathcal{L}^*$ *is the probability* (1.6).

(II) *The statistic* (1.10) *can be expressed by the equivalent formula*

$$T_{2,g} = \frac{k-1}{k}\,\frac{1}{V}\sum_{j=1}^{k}\Big[\sqrt{n}\big(g(Z_j) - g(1)\big)\Big]^2, \tag{1.13}$$

where

$$V = \frac{\big(g^{(1)}(1)\big)^2(k-1)}{3(k+1)} \tag{1.14}$$

is the asymptotic variance of the random variable $\Psi_j = \sqrt{n}\big(g(Z_j) - g(1)\big)$ *in the sense that*

$$\Psi_j \to N(0, V) \tag{1.15}$$

in distribution, as $n \to \infty$. *If* g *has on the interior of the interval* I *the continuous sixth derivative, then*

$$\Psi_j = \tilde{\Psi}_j + \mathcal{O}_P(n^{-5/2}), \tag{1.16}$$

with the variance

$$\operatorname{Var}(\tilde{\Psi}_j) = V + \frac{Q}{n} + \mathcal{O}(n^{-2}), \quad Q = \left[\frac{\big(g^{(2)}(1)\big)^2}{18} + \frac{g^{(1)}(1)g^{(3)}(1)}{9}\right]\frac{(k-1)^2}{(k+1)^2}. \tag{1.17}$$

Thus the version of (1.13) *with the correction for variance is*

$$T_{2,g}^{\mathrm{mod}} = n\,\frac{k-1}{k}\,\frac{1}{V+\frac{Q}{n}}\sum_{j=1}^{k}\big(g(Z_j)-g(1)\big)^2\,, \tag{1.18}$$

and the counterpart of (1.12) *is the convergence in distribution*

$$\sqrt{n\,\frac{(k-1)}{k}\,\frac{1}{V+\frac{Q}{n}}}\;\max_{i,j=1,\dots,k}\big|g(Z_i)-g(Z_j)\big| \longrightarrow \mathcal{L}^*\,. \tag{1.19}$$

(III) *Suppose that g has on the interior of I the continuous third derivative and*

$$g(1)=0\,,\quad g^{(2)}(1)\neq 0\,. \tag{1.20}$$

The distribution of the statistic

$$T_{1,g} = n\,\frac{6(k+1)}{kg^{(2)}(1)}\sum_{j=1}^{k} g(Z_j) \tag{1.21}$$

converges weakly to χ^2_{k-1} as $n\to\infty$. Moreover, if

$$g^{(1)}(1)=0\,,\quad g^{(2)}(1)>0\,, \tag{1.22}$$

and the function g is non-negative, then

$$\sqrt{n\frac{6(k+1)}{kg^{(2)}(1)}}\;\max_{i,j=1,\dots,k}\Big|\sqrt{g(Z_i)}\,\mathrm{sign}(Z_i-1)-\sqrt{g(Z_j)}\,\mathrm{sign}(Z_j-1)\Big| \longrightarrow \mathcal{L}^* \tag{1.23}$$

in distribution as $n\to\infty$.

A similar theorem holds also in the finite values setting.

Theorem 1.2. *Suppose that $\{(X_{i1},\dots,X_{ik})\}_{i=1}^{\infty}$ are independent k-dimensional random vectors. Further suppose that for each i the variables $X_{i1},\dots,X_{ik}$ are independent and their distribution functions $F_{ij}(t)=P(X_{ij}<t)$ fulfill the hypothesis $F_{i1}=\dots=F_{ik}$ of no treatment effect. Let $(R_{i1},\dots,R_{ik})$ be the mid-ranks computed from the ordering within the ith block. Put*

$$\tilde{\sigma}_n^2 = \frac{1}{n(k-1)}\sum_{i=1}^{n}\sum_{j=1}^{k}\Big(R_{ij}-\frac{k+1}{2}\Big)^2\,,\quad Z_j=\frac{2S_j}{n(k+1)}\,, \tag{1.24}$$

where $S_j=\sum_{i=1}^{n}R_{ij}$. Assume further that there exist finite sets $\{M_i\}$ such that

$$\limsup_{i\to\infty}\,\max\{P_i(z)\colon z\in M_i\}<1\,,$$

where P_i is the probability corresponding to the distribution function F_{ij}.

Let $g\colon I = \left\langle \frac{2}{k+1}, \frac{2k}{k+1} \right\rangle \to \mathbb{R}^1$ *be a function which has on the interior of* I *the continuous second derivative* $g^{(2)}$.

(I) *If the first derivative* $g^{(1)}(1) \neq 0$, *then for the statistic*

$$\tilde{T}_{2,g} = \frac{n(k+1)^2}{4\left(g^{(1)}(1)\right)^2 \tilde{\sigma}_n^2} \sum_{j=1}^{k} \left(g(Z_j) - g(1)\right)^2 \tag{1.25}$$

the convergence in distribution

$$T_{2,g} \longrightarrow \chi^2_{k-1} \tag{1.26}$$

holds as $n \to \infty$. *Moreover, as* $n \to \infty$,

$$\sqrt{\frac{n(k+1)^2}{4\left(g^{(1)}(1)\right)^2 \tilde{\sigma}_n^2}} \max_{i,j=1,\dots,k} \left|g(Z_i) - g(Z_j)\right| \longrightarrow \mathcal{L}^* \tag{1.27}$$

in distribution.

(II) *Suppose that* g *has on the interior of* I *the continuous third derivative and* (1.20) *holds. The distribution of the statistic*

$$\tilde{T}_{1,g} = \frac{n(k+1)^2}{2\,g^{(2)}(1)\,\tilde{\sigma}_n^2} \sum_{j=1}^{k} g(Z_j) \tag{1.28}$$

converges weakly to χ^2_{k-1} *as* $n \to \infty$. *Moreover, if* (1.22) *holds and the function* g *is non-negative, then*

$$\sqrt{\frac{n(k+1)^2}{2\,g^{(2)}(1)\,\tilde{\sigma}_n^2}} \max_{i,j=1,\dots,k} \left|\sqrt{g(Z_i)}\,\mathrm{sign}(Z_i - 1) - \sqrt{g(Z_j)}\,\mathrm{sign}(Z_j - 1)\right| \longrightarrow \mathcal{L}^* . \tag{1.29}$$

Theorem 1.1 provides a possibility to construct some new version of the Friedman test with the advantage that such a modification will automatically yield the corresponding new multiple comparison method (MCM) which should be applied after H_0 is rejected with the mentioned new statistic. The test rule rejects H_0 if $T > \chi^2(k-1,\alpha)$. The MCM based on (1.19) declares as different the ith and the jth treatment if

$$\sqrt{n\,\frac{(k-1)}{k}\,\frac{1}{V + \frac{Q}{n}}}\,\left|g(Z_i) - g(Z_j)\right| > q_{k,\alpha}\,, \tag{1.30}$$

and when (1.27) is employed, if

$$\sqrt{\frac{n(k+1)^2}{4\left(g^{(1)}(1)\right)^2 \tilde{\sigma}_n^2}}\,\left|g(Z_i) - g(Z_j)\right| > q_{k,\alpha}\,, \tag{1.31}$$

analogous rules can be based also on (1.23) and (1.29).

The quality of tests is as usual evaluated by means of its power; since the most frequently used rule is the Friedman test or the Iman–Davenport test, the power comparisons should be related to these tests. In what follows the focus is on the simplest situation, when $k = 3$ or $k = 4$. If $k = 3$, then the simulations show that as far as testing at $\alpha = 0.05$ or at $\alpha = 0.1$ is concerned, then both the power and the actual level of significance of T_{FR} and T_{ID} are the same for $n \leq 50$. A different situation occurs for $k = 4$, where also the structure of the alternatives is more complicated. As for the fit with the nominal level of significance for $k = 4$, when $9 \leq n \leq 50$ the fit of T_{ID} size with nominal α appears to be good and in many cases this size fits better than the size of T_{FR}. What is more important, simulations showed that T_{ID} usually (for various alternatives and underlying probabilities including the logistic and discrete distributions) yields for $n \leq 50$ a better power than T_{FR} and the identical power occurs only rarely. This is the reason that in the simulation results on performance of multiple comparisons mentioned in this paper, only the statistic T_{ID} is considered.

Although the formulas of the previous theorems have been applied to several types of functions, no decisive results yielding in some way uniformly best rule have been achieved. It turns out that if for some g the resulting test is stronger than the Iman–Davenport test (1.7), this occurs only for some type of ordering of the means of the treatments, while for some other alternative such a dominance does not hold. To demonstrate this, put $g(x) = x\log(x)$ and the statistic resulting from (1.18) of Theorem 1.1 label by T_{sqe} (it could be called square-entropic). In the following three tables the simulation results are obtained from $N = 7000$ trials for $k = 4$.

TABLE 1. Simulation estimates of the actual significance level when the functions (1.1) are continuous.

n	11		16		22		28		35	
α	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1
T_{sqe}	0.048	0.102	0.042	0.092	0.048	0.097	0.047	0.1	0.046	0.102
T_{FR}	0.047	0.097	0.044	0.093	0.049	0.104	0.045	0.1	0.045	0.103
T_{ID}	0.058	0.107	0.049	0.098	0.051	0.105	0.049	0.104	0.048	0.105

This table suggests that except for some rare exceptions the magnitude of the size of the mentioned tests fulfills the usual precision requirements not to differ more than 10 % from the nominal value. Now suppose that

$$X_{ij} \sim N(\mu_j, 1), \qquad j = 1, \ldots, k. \tag{1.32}$$

The following estimates of the probability of rejection the null hypothesis for $k = 4$, the significance level $\alpha = 0.05$, $\alpha = 0.1$ and particular alternatives show the complexity of the situation.

TABLE 2. Simulation estimates of the power when $\mu_1 = 0$, $\mu_2 = 0$, $\mu_3 = 0.9$, $\mu_4 = 0.9$.

n	11		16		22		28		35	
α	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1
T_{sqe}	0.514	0.647	0.714	0.820	0.858	0.922	0.935	0.969	0.980	0.991
T_{FR}	0.516	0.656	0.717	0.825	0.859	0.927	0.934	0.970	0.981	0.991
T_{ID}	**0.553**	**0.672**	**0.738**	**0.829**	**0.864**	**0.928**	**0.939**	**0.972**	**0.981**	**0.991**

TABLE 3. Simulation estimates of the power when $\mu_1 = 0$, $\mu_2 = 0$, $\mu_3 = 0$, $\mu_4 = 0.9$.

n	11		16		22		28		35	
α	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1
T_{sqe}	**0.449**	**0.581**	**0.611**	**0.727**	**0.765**	**0.853**	**0.866**	**0.924**	**0.946**	**0.971**
T_{FR}	0.391	0.543	0.561	0.695	0.727	0.837	0.845	0.909	0.933	0.966
T_{ID}	0.425	0.556	0.585	0.708	0.737	0.838	0.850	0.912	0.934	0.967

It can be seen from Table 2 and Table 3 that for the first alternative the statistic T_{ID} yields better results than T_{sqe}, while for the second alternative the statistic T_{sqe} is better than T_{ID}. Since after inspecting also some other not included simulation results, the statistic T_{ID} does not appear to be dominated for $\alpha = 0.05$ by T_{sqe} in the sense of the test power, in the further tables concerning multiple comparisons only T_{ID} is included.

As for the multiple comparison method used after the test procedure rejects the null hypothesis, it was proposed in [9] that its quality should be described by the conditional probabilities of right and wrong decision. In this random block setting the criterial quantities are the conditional probabilities

$$rd = P(\text{method correctly detects at least one pair of different treatments} \mid A),$$
$$wd = P(\text{method wrongly declares at least two identical treatments as different} \mid A),$$

where A is the event that the underlying test procedure rejects the null hypothesis of no treatment effect (everything at the same α). Larger values of rd can be interpreted as a better performance and smaller values of wd as a better reliability of the comparison rule.

In addition to the already mentioned procedures also another two rules will be considered. The rule (1.30) with $g(x) = x\log(x)$ will be denoted by *entr* (entropic) and (1.30) with $g(x) = \log(x)$ will be denoted by *logar* (logarithmic). The following simulation results, where F stands for the critical value $F\big(k-1, (n-1)(k-1), \alpha\big)$ of the F distribution, are obtained from $N = 16000$ trials.

TABLE 4. Simulation estimates of the performance when $\mu_1 = 0$, $\mu_2 = 0$, $\mu_3 = 0.5$.

n	17				20			
α	0.05		0.1		0.05		0.1	
$P(T_{\mathrm{ID}} > F)$	0.215		0.344		0.265		0.380	
	rd	wd	rd	wd	rd	wd	rd	wd
CID	0.983	0.204	0.978	0.236	0.985	0.177	0.990	0.236
class	**0.957**	0.087	**0.961**	0.137	**0.886**	0.070	**0.934**	0.117
entr	0.904	**0.053**	0.909	**0.072**	0.856	**0.037**	0.097	**0.064**
logar	0.820	0.123	0.862	0.156	0.792	0.100	0.870	0.132

n	30				40			
α	0.05		0.1		0.05		0.1	
$P(T_{\mathrm{ID}} > F)$	0.384		0.489		0.476		0.603	
	rd	wd	rd	wd	rd	wd	rd	wd
CID	0.995	0.128	0.995	0.176	0.998	0.102	0.998	0.154
class	0.870	0.037	**0.963**	0.081	**0.960**	0.039	0.949	0.057
entr	**0.921**	**0.026**	0.954	**0.050**	0.955	**0.022**	**0.984**	**0.041**
logar	0.881	0.064	0.932	0.103	0.923	0.055	0.955	0.086

TABLE 5. Simulation estimates of the performance when $\mu_1 = 0$, $\mu_2 = 0.3$, $\mu_3 = 0.6$.

n	17		20		30		40	
α	0.05	0.1	0.05	0.1	0.5	0.1	0.5	0.1
$P(T_{\mathrm{ID}} > F)$	0.223	0.360	0.280	0.405	0.415	0.520	0.511	0.638
	rd	wd	rd	wd	rd	wd	rd	wd
CID	1	1	1	1	1	1	1	1
class	1	1	0.908	0.961	0.885	0.983	0.976	0.965
entr	0.896	0.917	0.848	0.920	0.902	0.963	0.956	0.983
logar	0.905	0.921	0.850	0.918	0.919	0.963	0.957	0.982

It can be seen from Table 4 and Table 5 that for $k = 3$ the *CID* rule has a high performance but a low reliability; such a situation has been detected by means of simulation results in [8] for an analogous rule, used in conjunction with the Kruskal–Wallis test. The simulation results suggest that if one wants to have a low risk of false labelling identical treatment as different, then for the sampling from the normal distribution the entropic rule can be used. Since the classical rule seems to have relatively balanced quality for various underlying distributions and types of alternatives, it can be considered as an universal rule for $k = 3$.

For $k = 4$ the citeria rd and wd did not detect an ordering between the mentioned rules, because their performance and reliability vary with circumstancies. Similarly as in the previous case, also for $k = 4$ the CID rule has a low reliability. Since the classical rule seems to have relatively balanced quality under various circumstances, it can be considered as an universal rule also for $k = 4$.

2. Appendix: Proofs

Proof of Theorem 1.1. (I). The proof is the same as in the classical case, except for a use of the delta method.

Put

$$\eta_j = \frac{1}{\sqrt{n}}\left(S_j - n\frac{k+1}{2}\right). \tag{2.1}$$

It is well known (e.g., p. 270 of [7]), that

$$\eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_k \end{pmatrix} \longrightarrow N_k(0, \sigma^2 W), \quad W = I_k - \frac{1}{k}\mathbf{1}\mathbf{1}', \ \mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^k, \tag{2.2}$$

where σ^2 is defined in (1.4). Since

$$Z_j - 1 = \frac{2}{\sqrt{n}(k+1)}\,\eta_j\,, \tag{2.3}$$

making use of the Taylor formula, (2.2) and the fact that a sequence converging in distribution is bounded in probability, one obtains that

$$\begin{aligned} g(Z_j) - g(1) &= g^{(1)}(1)\,(Z_j - 1) + \frac{1}{2}\,g^{(2)}\big(\alpha Z_j + (1-\alpha)\big)\,(Z_j - 1)^2 \\ &= \frac{2g^{(1)}(1)}{\sqrt{n}(k+1)}\,\eta_j + \mathcal{O}_P\left(\frac{1}{n}\right). \end{aligned} \tag{2.4}$$

This together with (2.2), (1.4) means that

$$\sqrt{n\frac{3(k+1)}{k}}\,\frac{1}{g^{(1)}(1)}\begin{pmatrix} g(Z_1) - g(1) \\ \vdots \\ g(Z_k) - g(1) \end{pmatrix} \longrightarrow N_k(0, W). \tag{2.5}$$

Similarly as in the case of the Friedman test statistic, the convergence (1.11) follows from (2.5), because $\mathcal{L}\big(x'x \,|\, N_k(0,W)\big) = \chi^2_{k-1}$. Further, (2.5) implies that

$$\sqrt{n\frac{3(k+1)}{k(g^{(1)})^2}}\ \max_{i,j=1,\dots,k}\ \big|g(Z_i) - g(Z_j)\big| \longrightarrow \mathcal{L}\Big(\max_{i,j=1,\dots,k}|x_i - x_j|\,\big|\,N_k(0,W)\Big) \tag{2.6}$$

in distribution, as $n \to \infty$. As it is explained in chapter 6 of [6],

$$\mathcal{L}\Big(\max_{i,j=1,\ldots,k} |x_i - x_j| \,\big|\, N_k(0,W)\Big) = \mathcal{L}^* \tag{2.7}$$

and therefore (1.12) holds. □

Lemma 2.1. *Suppose that $k \geq 2$ is an integer and $(R_{i1},\ldots,R_{ik})$, $i=1,\ldots,n$, are independent random vectors such that*

$$P\big((R_{i1},\ldots,R_{ik}) = r\big) = \frac{1}{k!}$$

for every permutation r of the set $\{1,\ldots,k\}$. Then for $j=1,2,\ldots,k$ (cf. (1.2), (1.9))

$$E\big((Z_j - 1)^2\big) = \frac{(k-1)}{3(k+1)}\,\frac{1}{n}\,, \tag{2.8}$$

$$E\big((Z_j - 1)^p\big) = 0\,, \qquad p = 3,\,5,\,7,\,9\,, \tag{2.9}$$

$$E\big((Z_j - 1)^4\big) = \frac{(k-1)^2}{3(k+1)^2}\,\frac{1}{n^2} + \mathcal{O}\Big(\frac{1}{n^3}\Big)\,, \tag{2.10}$$

$$E\big((Z_j - 1)^p\big) = \mathcal{O}\Big(\frac{1}{n^3}\Big)\,, \qquad p = 6,\,8,\,10\,. \tag{2.11}$$

P r o o f. Validity of (2.9) follows from the fact that the random variable S_j is distributed symmetrically about its mean $n(k+1)/2$. We are going to prove only (2.10), because the rest of the lemma can be proved similarly.

Obviously

$$E\left(\left(S_j - n\frac{k+1}{2}\right)^4\right) = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4\,, \tag{2.12}$$

where

$$\Sigma_t = \sum_{\mathbf{i}\in A_t} E\left(\left(R_{i_1 j} - \frac{k+1}{2}\right)\left(R_{i_2 j} - \frac{k+1}{2}\right)\left(R_{i_3 j} - \frac{k+1}{2}\right)\left(R_{i_4 j} - \frac{k+1}{2}\right)\right),$$

$$A_t = \big\{\mathbf{i} = (i_1,i_2,i_3,i_4) \in \{1,2,\ldots,n\}^4 :\ \#\{i_1,i_2,i_3,i_4\} = t\big\}\,.$$

Since $R_{i,s}$, R_{i^*,s^*} are independent if $i \neq i^*$,

$$\Sigma_4 = 0\,, \quad \Sigma_3 = 0\,. \tag{2.13}$$

But the set $A_2^* = \{\mathbf{i} \in A_2;\ \textit{no number appears in } \mathbf{i} \textit{ three times}\}$ consists of $3n(n-1)$ elements, and by means of

$$E\left(\left(R_{1j} - \frac{k+1}{2}\right)^2\right) = \frac{1}{k}\sum_{j=1}^{k}\left(j - \frac{k+1}{2}\right)^2 = \frac{k^2-1}{12}$$

one obtains that

$$\Sigma_2 = \frac{n(n-1)(k^2-1)^2}{48}\,. \tag{2.14}$$

Finally, since

$$\sum_{s=1}^{k}\Big(s-\frac{k+1}{2}\Big)^4 = \frac{k(k^2-1)(3k^2-7)}{240}\,,$$

obviously

$$\Sigma_1 = \sum_{i=1}^{n}\sum_{s=1}^{k}\Big(s-\frac{k+1}{2}\Big)^4\,\frac{1}{k} = \frac{n(k^2-1)(3k^2-7)}{240}\,, \tag{2.15}$$

and (2.12)–(2.15) yield (2.10). □

Proof of Theorem 1.1. (II) and (III).

(II) Since (2.2), (1.4), (2.4) hold, validity of (1.13)–(1.15) is obvious. An application of the Taylor theorem gives the formulas

$$g(Z_j) - g(1) = \kappa_j + R_6\,, \tag{2.16}$$

$$\kappa_j = \sum_{t=1}^{5}\frac{g^{(t)}(1)}{t!}(Z_j-1)^t\,, \quad R_6 = \frac{g^{(6)}\big(\alpha Z_j + (1-\alpha)\big)}{6!}\,(Z_j-1)^6\,.$$

By means of the Markov inequality and (2.11) for any positive ε

$$P\big(|Z_j-1|^6 > \frac{\varepsilon}{n^3}\big) \le \frac{n^3}{\varepsilon}\,E\big(|Z_j-1|^6\big) = \frac{\mathcal{O}(1)}{\varepsilon}\,,$$

$$(Z_j-1)^6 = \mathcal{O}_P\Big(\frac{1}{n^3}\Big)\,,$$

which together with the continuity of $g^{(6)}$ means that $g^{(6)}\big(\alpha Z_j + (1-\alpha)\big) \longrightarrow g^{(6)}(1)$ and

$$R_6 = \mathcal{O}_P\Big(\frac{1}{n^3}\Big)\,. \tag{2.17}$$

Since the Lemma 2.1 holds,

$$\begin{aligned} E(\kappa_j) &= \frac{g^{(2)}(1)}{2}E\big((Z_j-1)^2\big) + \frac{g^{(4)}(1)}{24}E\big((Z_j-1)^4\big) \\ &= g^{(2)}(1)\frac{k-1}{6(k+1)}\,\frac{1}{n} + g^{(4)}(1)\,\frac{(k-1)^2}{72(k+1)^2}\,\frac{1}{n^2} + \mathcal{O}\Big(\frac{1}{n^3}\Big)\,, \end{aligned} \tag{2.18}$$

$$E(\kappa_j^2) = \sum_{t=1}^{5}\Big(\frac{g^{(t)}(1)}{t!}\Big)^2 E\big((Z_j-1)^{2t}\big) + 2\sum_{t=1}^{4}\sum_{s=t+1}^{5}\frac{g^{(t)}(1)}{t!}\,\frac{g^{(s)}(1)}{s!}E\big((Z_j-1)^{t+s}\big)$$

$$= \big(g^{(1)}(1)\big)^2 E\big((Z_j-1)^2\big) + \Big(\frac{g^{(2)}(1)}{2}\Big)^2 E\big((Z_j-1)^4\big)$$
$$+ 2g^{(1)}(1)\,\frac{g^{(3)}(1)}{6}\,E\big((Z_j-1)^4\big) + \mathcal{O}\Big(\frac{1}{n^3}\Big)$$
$$= \frac{V}{n} + \left[\frac{\big(g^{(2)}(1)\big)^2}{12} + \frac{g^{(1)}(1)g^{(3)}(1)}{9}\right]\frac{(k-1)^2}{(k+1)^2}\,\frac{1}{n^2} + \mathcal{O}\Big(\frac{1}{n^3}\Big), \tag{2.19}$$

where V is defined in (1.14). Validity of (2.18) and (2.19) means that

$$\operatorname{Var}(\kappa_j) = E(\kappa_j^2) - E(\kappa_j)^2 = \frac{V}{n} + \frac{Q}{n^2} + \mathcal{O}\Big(\frac{1}{n^3}\Big). \tag{2.20}$$

Hence for

$$\tilde{\Psi}_j = \sqrt{n}\,\kappa_j$$

from (2.16), (2.17) one obtains (1.16), and obviously (2.20) implies (1.17).

Validity of (1.19) follows from the equality

$$\frac{k}{k-1}\,V = \frac{k\big(g^{(1)}(1)\big)^2}{3(k+1)}$$

and from (1.12).

(III) By means of the Taylor formula, (2.3) and (2.2)

$$g(Z_j) = g^{(1)}(1)(Z_j-1) + \frac{g^{(2)}(1)}{2}(Z_j-1)^2 + \mathcal{O}\Big(\frac{1}{n^{3/2}}\Big). \tag{2.21}$$

This equality, (1.21) and the identity $\sum\limits_{j=1}^{k}(Z_j-1)=0$ yield that

$$T_{1,g} = T_{\mathrm{FR}} + o_P(1),$$

and the convergence of T_{1g} in distribution follows from the properties of the Friedman test statistics.

Now suppose that also (1.22) holds and the function g is non-negative. Then according to (2.21)

$$\begin{aligned} g(Z_j) &= \frac{g^{(2)}(1)}{2}(Z_j-1)^2 + \mathcal{O}_P\Big(\frac{1}{n^{3/2}}\Big),\\ \Big|ng(Z_j) - n\frac{g^{(2)}(1)}{2}\,(Z_j-1)^2\Big| &= o_P(1). \end{aligned} \tag{2.22}$$

But by (2.3), (2.2) the random variables $\sqrt{n}(Z_j-1)$ are asymptotically normal. This property, the equality

$$\frac{g^{(2)}(1)}{2}n(Z_j-1)^2 = \mathcal{O}_P(1)$$

and (2.22) imply that

$$\left| \sqrt{n}\,\sqrt{g(Z_j)} - \sqrt{\frac{g^{(2)}(1)}{2}}\sqrt{n}\,|Z_j - 1| \right| = o_P(1)\,.$$

Employ (2.3), (1.4) to show that

$$\sqrt{n\frac{6(k+1)}{kg^{(2)}(1)}}\sqrt{g(Z_j)}\ \mathrm{sign}(Z_j - 1) = \frac{1}{\sqrt{\sigma^2}}\eta_j + o_P(1)\,. \tag{2.23}$$

Since (2.7) holds, (1.23) follows from (2.23) and (2.2). □

Proof of Theorem 1.2. Suppose that

$$\gamma_j = S_j - n\frac{k+1}{2}$$

and put (cf. (1.24))

$$\eta = \frac{1}{\sqrt{\sigma^2}}(\gamma_1, \dots, \gamma_k)'\,.$$

It can be shown that under the assumption of the theorem on $\{(X_{i1}, \dots, X_{ik})\}_{i=1}^{\infty}$ the convergence in distribution

$$\eta \longrightarrow N_k(0, W) \tag{2.24}$$

holds, where W is defined in (2.2); a proof of such an assertion can be found on pp. 391–392 of [5] or on pp. 150–151 of [8]. Since this theorem can be proved by means of (2.24) similarly as Theorem 1.1 by means of (2.2), the rest of the proof is omitted. □

REFERENCES

[1] CONOVER, W. J.: *Practical Nonparametric Statistics.* Third edition. J. Wiley, New York, 1999.

[2] FRIEDMAN, M.: *The use of ranks to avoid the assumption of normality implicit in the analysis of variance*, J. Amer. Statist. Assoc. **32** (1937), 675–701.

[3] HOLLANDER, M.—WOLFE, D. A.: *Nonparametric Statistical Methods.* Second edition. J. Wiley, New York, 1999.

[4] IMAN, R. L.—DAVENPORT, J. M.: *Approximation of the critical region of the Friedman statistic*, Comm. Statist. Theory Methods **A9** (1980), 571–595.

[5] LEHMANN, E. L.: *Nonparametrics: Statistical Methods Based on Ranks*, in: Probability and Statistic, Holden–Day, San Francisco, 1975.

[6] MILLER, G. R.: *Simultaneous Statistical Inference*, Springer-Verlag, Berlin, 1985.

[7] PURI, M. L.—SEN, P. K.: *Nonparametric Methods in Multivariate Analysis*, J. Wiley, New York, 1971.

[8] RUBLÍK, F.: *Nonparametric Methods*, Lecture Notes, Comenius University in Bratislava, 1993. (In Slovak)

[9] RUBLÍK, F.: *On performance of multiple comparison methods used in conjunction with the Kruskal-Wallis test*, in: MEASUREMENT 2001, Proceedings, Smolenice, 2001, pp. 36–40.

[10] SEN, P. K.: *Asymptotically efficient tests by the method of n rankings*, J. Roy. Statist. Soc. B **30** (1968), 312–317.

[11] SEN, P. K.: *On a class of aligned rank order tests in two-way layouts*, Ann. Math. Stat. **39** (1968), 1115–1124.

Received April 25, 2002

František Rublík
Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9
SK–841 04 Bratislava
SLOVAKIA
E-mail: umerrubl@savba.sk

Pranab Kumar Sen
Department of Biostatistics and Statistics
University of North Carolina
Chapel Hill, NC 27599-7420
U. S. A.
E-mail: pksen@bios.unc.edu