Original Article

Nonparametric Closed Testing Procedures for All Pairwise Comparisons in a Randomized Block Design

Taka-aki Shiraishi and Shin-ichi Matsuda
Department of Systems and Mathematical Science, Nanzan University

e-mail: marble@nanzan-u.ac.jp, matsu@nanzan-u.ac.jp

We consider multiple comparison test procedures among treatment effects in a randomized block design. We propose closed testing procedures based on signed rank statistics and Friedman test statistics for all pairwise comparisons of treatment effects. Although anyone has been failed to discuss a distribution-free method except Bonferroni procedures as a multiple comparison test, the proposed procedures are exactly distribution-free. Next we consider the randomized block design under simple ordered restrictions of treatment effects. We propose distribution-free closed testing procedures based on one-sided signed rank statistics and rank statistics of Chacko (1963) for all pairwise comparisons. Simulation studies are performed under the null hypothesis and some alternative hypotheses. In this studies, the proposed procedures show a good performance. We also illustrate an application to death rates by using proposed procedures.

Key words: Multiple comparison procedure, Distribution-free procedures, Multi-step procedures, Simple ordered restrictions.

1. Introduction
We consider a randomized block design with $n$ blocks and $k$ treatments. The model for the randomized block design is

$$X_{ij} = \mu + \beta_i + \tau_j + e_{ij} \quad (1 \leq i \leq n, 1 \leq j \leq k),$$

(1)

where $\sum_{i=1}^n \beta_i = \sum_{j=1}^k \tau_j = 0$ and the $e_{ij}$’s are independent and identically distributed with absolutely continuous distribution function $F(x)$ such that $\int_{-\infty}^{\infty} xdF(x) = 0$. Then $\mu$, $\beta_i$ and $\tau_j$ are referred to as the overall mean, the $i$-th block effect and the $j$-th treatment effect, respectively. Many researchers considered testing the null hypothesis of no treatment effects, $H_0: \tau_1 = \cdots = \tau_k = 0$. Let $R_{ij}$ denote the rank of $X_{ij}$ in the elements of $\{X_{ij} \mid 1 \leq j \leq k\}$ for $i=1, \ldots, n$ and let
\[ T = \frac{12}{nk(k+1)} \sum_{j=1}^{k} \left[ \sum_{i=1}^{n} R_{ij} - \frac{n(k+1)^2}{2} \right]^2. \]

Friedman (1937) showed that, under \( H_0 \), as \( n \) tends to infinity,

\[ T \xrightarrow{d} Y \sim \chi^2_{k-1}. \]

where \( \xrightarrow{d} \) denotes convergence in law and \( \chi^2_{k-1} \) denotes a chi-squared distribution with \( k-1 \) degrees of freedom. As a distribution-free test for the null hypothesis \( H_0 \) vs. \( H^A : \tau_j \neq 0 \) for some \( j \), Friedman (1937) proposed to reject \( H_0 \) when \( T \) is larger than \( \chi^2_{k-1} (\alpha) \), where \( \chi^2_{k-1} (\alpha) \) is an upper 100\( \alpha \)% point of a chi-squared distribution with \( k-1 \) degrees of freedom. On the other hand, multiple tests and simultaneous confidence intervals specify the differences of treatment effects. Therefore multiple comparison procedures are used for data-analysis in the areas of medicine, pharmacy, and biology. We consider to propose multiple comparison test procedures for all pairwise comparisons of

\[ \left\{ \text{the null hypothesis } H_{(j,j)} : \tau_j = \tau_{j'} \text{ vs. the alternative } H^A_{(j,j)} : \tau_j \neq \tau_{j'} \mid (j,j) \in U \right\}, \]

where

\[ U = \left\{ (j,j') \mid 1 \leq j < j' \leq k \right\}. \]

Hsu (1996) pointed out that it is difficult to obtain any exact distribution-free multiple comparison method except Bonferroni procedures for all pairwise comparisons.

Next, under the model (1), we consider simple ordered restrictions of treatment effects, that is,

\[ \tau_1 \leq \tau_2 \leq \cdots \leq \tau_k. \]

We construct distribution-free closed testing procedures based on one-sided signed rank statistics and rank statistics of Chacko (1963) for all pairwise comparisons of

\[ \left\{ \text{the null hypothesis } H_{(j,j)} \text{ vs. the alternative } H^A_{(j,j)} : \tau_j < \tau_{j'} \mid (j,j) \in U \right\}. \]

In section 2, we propose two closed testing procedures as multi-step tests and we introduce the procedure analogous to the REGW method. In section 3, for simple ordered restrictions of treatment effects, we propose multiple comparison procedures. In section 4, by the Monte Carlo simulations, we compare the proposed procedures with the other multiple comparisons tests in the sense of all-pairs power defined in Ramsey (1978). As the result of section 4, we find that the proposed procedures are superior to the other procedures. Especially the power of the proposed closed testing procedures is fairly higher than that of the other procedures. In section 5, we illustrate an application to death rates by using proposed procedures. We conclude with a discussion in section 6.
2. Closed testing procedures

We consider test procedures for all pairwise comparisons of
\[ \{ \text{the null hypothesis } H_{ij}, \text{ vs. the alternative } H_{ij}^a \mid (j, j) \subseteq U \}. \]

Let us put
\[ \mathcal{H} = \{ H_{ij} \mid (j, j) \subseteq U \}. \]

Then, the closure of $\mathcal{H}$ is given by
\[ \bar{\mathcal{H}} = \{ \bigwedge_{i \in V} H_v \mid \emptyset \subseteq V \subseteq U \}, \]

where $\bigwedge$ denotes the conjunction symbol (Refer to Enderton (2001)). Then, we get
\[ \bigwedge_{i \in V} H_v : \text{for any } (j, j) \subseteq V, \tau_j = \tau_v. \]

For an integer $J$ and disjoint sets $I_1, \ldots, I_J \subseteq \{1, \ldots, k\}$, we define the null hypothesis $H(I_1, \ldots, I_J)$ by
\[ H(I_1, \ldots, I_J) : \text{for any integer } a \text{ such that } 1 \leq a \leq J \text{ and for any } j, j' \in I_a, \tau_j = \tau_{j'}. \]

From (6) and (7), for any nonempty $V \subseteq U$, there exist an integer $J$ and disjoint sets $I_1, \ldots, I_J$ such that
\[ \bigwedge_{i \in V} H_v = H(I_1, \ldots, I_J) \]

and $\#(I_a) \geq 2$ for $a = 1, \ldots, J$, where $\#(A)$ stands for the cardinal number of set $A$. For $H(I_1, \ldots, I_J)$ of (8), we set
\[ M = M(I_1, \ldots, I_J) = \sum_{a=1}^J \ell_a, \ell_a = \#(I_a). \]

For $I_a$ satisfying $\ell_a = 2$, we define the minimum value and maximum value of the elements of $I_a$ by $j(I_a)$ and $j' (I_a)$, respectively. Let $R^+_i(I_a)$ denote the rank of $|X_{a''/I_a} - X_{a''/I_a}|$ in $|X_{a''/I_a} - X_{a''/I_a}|, |X_{a''/I_a} - X_{a''/I_a}|, \ldots, |X_{a''/I_a} - X_{a''/I_a}|$ and let
\[ S(I_a) = \sqrt{\frac{6}{n(n+1)(2n+1)}} \frac{\sum_{i=1}^n \text{sign}(X_{a''/I_a} - X_{a''/I_a}) R^+_i(I_a)}{n}, \]

where $\text{sign}(x) = 1$ if $x > 0$, $0 = 0$ if $x = 0$, $= -1$ otherwise. Then, $S(I_a)$ is Wilcoxon’s signed rank statistic. From Hettmansperger (1984), under $H_0$, we get, as $n$ tends to infinity,
\[ S(I_a) \xrightarrow{d} Z(I_a) \sim N(0, 1), \]

where $Z \sim N(0, \sigma^2)$ denotes that $Z$ is distributed according to $N(0, \sigma^2)$.

For $I_a$ satisfying $\ell_a \geq 3$, let $R(i, j) (i \in I_a)$ denote the rank of $X_{ij}$ in the elements of $(X_{ij} \mid j \in I_a)$
for $i = 1, \ldots, n$ and let

$$T(I_a) = \frac{12}{n \ell_a(\ell_a+1)} \cdot \sum_{j \in I_a} \left[ \sum_{i=1}^{n} R_{ij}(I_a) - \frac{n(\ell_a+1)^2}{2} \right]$$

(12)

Then $T(I_a)$ is Friedman’s test statistic. From Hettmansperger (1984), under $H_0$, we get, as $n$ tends to infinity,

$$T(I_a) \overset{d}{\to} Y(I_a) \sim \chi^2_{\ell_a-1}.$$  

(13)

We define rank statistic $ST(I_a)$ by

$$ST(I_a) = \begin{cases} |S(I_a)| & \text{if } \ell_a = 2 \\ T(I_a) & \text{if } \ell_a \geq 3. \end{cases}$$

(14)

For $\ell = \ell_1, \ldots, \ell_j$, we define $\alpha(M, \ell)$ by

$$\alpha(M, \ell) = 1 - (1-\alpha)^{1/M}. $$

(15)

We define the asymptotic upper 100$\alpha$% point of rank statistic $ST(I_a)$ by

$$z_c(\alpha(M, \ell_a)|\ell_a) = \begin{cases} z(\alpha(M, \ell_a)/2) & \text{if } \ell_a = 2 \\ \chi^2_{\ell_a-1, \alpha(M, \ell_a)} & \text{if } \ell_a \geq 3, \end{cases}$$

(16)

where $z(\alpha)$ is an upper 100$\alpha$% point of a standard normal distribution and $\chi^2_{\ell_a-1, \alpha(M, \ell_a)}$ is an upper 100$\alpha(M, \ell_a)$% point of a chi-squared distribution with $\ell_a-1$ degrees of freedom.

Then, we propose the closed testing procedure (I).

(I) Closed testing procedure based on rank statistics

(a) $J \geq 2$

Whenever $z_c(\alpha(M, \ell_a)|\ell_a) < ST(I_a)$ holds for an integer $a$ such that $1 \leq a \leq J$, we reject the hypothesis $\bigwedge_{v \in V} H_v$.

(b) $J = 1$ ($M = \ell_1$)

Whenever $z_c(\alpha|\ell_1) < ST(I_1)$, we reject the hypothesis $\bigwedge_{v \in V} H_v$.

By using the methods of (a) and (b), when $\bigwedge_{v \in V} H_v$ is rejected for any $V$ such that $(j, j') \in V \subset U$, the null hypothesis $H_{(j, j')}$ is rejected as a multiple comparison test.

**Theorem 1.** The test procedure (I) is an asymptotic multiple comparison test of level $\alpha$ as $n$ tends to infinity.

The proof of Theorem 1 is similar to the proof of Theorem 2.1 of Shiraishi and Matsuda (2018). The values of $z_c(\alpha(M, \ell)|\ell)$ for $\alpha = 0.05$, $0.01$, $2 \leq \ell \leq M$ and $2 \leq M \leq 10$ are provided in Tables 1 and 2.
3. Simple ordered restrictions

We assume that the simple ordered restrictions (4) is satisfied. We consider the null hypothesis $H_0$ vs. the alternative $H^{OA}_0$: $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_k$ with at least one strict inequality, which is equivalent to $H_0^\dagger$: $\tau_1 = \tau_k$ vs. $H^{OA}_0$: $\tau_1 < \tau_k$. We define $\{\hat{\tau}_j(R)\} j=1,\ldots,k$ by $\{u_j\} j=1,\ldots,k$ which minimize $\sum_{j=1}^{k} (u_j-R_j)^2$ under simple ordered restrictions $u_1 \leq u_2 \leq \cdots \leq u_k$, i.e.,

$$\sum_{j=1}^{k} (\hat{\tau}_j(R) - R_j)^2 = \min_{u_1 \leq \cdots \leq u_k} \sum_{j=1}^{k} (u_j-R_j)^2,$$

where $R_j = \frac{\sum_{i=1}^{n} R_{ij}}{n}$. $\hat{\tau}_1(R), \ldots, \hat{\tau}_k(R)$ are computed by using the pool-adjacent-violators algorithm stated in Robertson et al. (1988) and the following equation holds.

$$\hat{\tau}_j(R) = \max_{1 \leq p < j} \min_{j < q \leq k} \frac{\sum_{i=1}^{n} R_{ij}}{q - p + 1}.$$

We put

$$\hat{\tau}_k^2(R) = \frac{12n}{k(k+1)} \sum_{j=1}^{k} \left[ \hat{\tau}_j(R) - \frac{k+1}{2} \right]^2.$$

Let $P(L, k)$ be the probability that $\hat{\tau}_1(R), \ldots, \hat{\tau}_k(R)$ takes exactly $L$ distinct values under $H_0$. Then by using Theorem 2.3.1 of Robertson et al. (1988), we get

| $M \setminus \ell$ | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|-------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 10                | 2.569 | 8.364 | 9.804 | 11.113 | 12.349 | 13.536 | 14.689 | *   | 16.919 |
| 9                 | 2.532 | 8.155 | 9.576 | 10.868 | 12.087 | 13.259 | *   | 15.507 |
| 8                 | 2.491 | 7.921 | 9.320 | 10.592 | 11.793 | *   | 14.067 |
| 7                 | 2.443 | 7.657 | 9.031 | 10.279 | *   | 12.592 |
| 6                 | 2.388 | 7.352 | 8.696 | *   | 11.070 |
| 5                 | 2.321 | 6.993 | *   | 9.488 |
| 4                 | 2.236 | *   | 7.815 |
| 3                 | *   | 5.991 |
| 2                 | 1.960 |

| $M \setminus \ell$ | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|-------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 10                | 3.089 | 11.611 | 13.310 | 14.855 | 16.310 | 17.706 | 19.058 | *   | 21.666 |
| 9                 | 3.058 | 11.401 | 13.085 | 14.616 | 16.059 | 17.443 | *   | 20.090 |
| 8                 | 3.022 | 11.166 | 12.833 | 14.349 | 15.777 | *   | 18.475 |
| 7                 | 2.982 | 10.899 | 12.547 | 14.045 | *   | 16.812 |
| 6                 | 2.934 | 10.592 | 12.216 | *   | 15.086 |
| 5                 | 2.877 | 10.228 | *   | 13.277 |
| 4                 | 2.806 | *   | 11.345 |
| 3                 | *   | 9.210 |
| 2                 | 2.576 |

Table 1. Critical values $ze(a(M, \ell) | \ell)$ for the closed testing procedure (I) with $a=0.05$

Table 2. Critical values $ze(a(M, \ell) | \ell)$ for the closed testing procedure (I) with $a=0.01$
\[
\lim_{n \to \infty} P_0 (\chi_k^2 (R) \geq t) = \sum_{L=2}^{k-1} P (L, k) P (\chi_{L-1}^2 \geq t) \quad (t > 0),
\]
(17)

where \( P_0 () \) denotes the probability measure under \( H_0 \) and \( \chi_{L-1}^2 \) is a chi-square variable with \( L - 1 \) degrees of freedom. Barlow et al. (1972) gives the following recurrence formula.

\[
P (1, k) = \frac{1}{k},
\]

\[
P (L, k) = \frac{1}{k} [(k-1)P(L, k-1) + P(L-1, k-1)], \quad (2 \leq L \leq k - 1)
\]

\[
P (k, k) = \frac{1}{k!}.
\]

We consider test procedures for all pairwise comparisons of

\[
\left\{ \text{the null hypothesis } H_{(j,j')} vs. \text{ the alternative } H_{(j,j')}^{0A}; \quad \tau_j < \tau_{j'} \quad \left| (j,j') \in \mathcal{U} \right. \right\}
\]

Let \( I_1^*, \ldots, I_J^* \) be disjoint sets satisfying the following property (C1).

(C1) There exist integers \( \ell_1^*, \ldots, \ell_J^* \geq 2 \) and integers \( 0 \leq s_1 < \cdots < s_J < k \) such that

\[
I_a^* = (s_a + 1, s_a + 2, \ldots, s_a + \ell_a^*) \quad (a = 1, \ldots, J)
\]

and \( s_a + \ell_a^* \leq s_{a+1} \quad (a = 1, \ldots, J - 1) \).

We define the null hypothesis \( H^*(I_1^*, \ldots, I_J^*) \) by

\[
H^*(I_1^*, \ldots, I_J^*): \text{for any } a \text{ such that } 1 \leq a \leq J \text{ and } \forall j, j' \in I_a^*, \quad \tau_j = \tau_{j'} \quad \text{holds}.
\]

(18)

The elements of \( I_a^* \) are consecutive integers and \( \# (I_a^*) \geq 2 \). From (18), for any nonempty \( V \subset \mathcal{U} \), there exist an integer \( J \) and some subsets \( I_1^*, \ldots, I_J^* \subset \{ 1, \ldots, k \} \) satisfying (C1) such that

\[
\bigwedge_{a=1}^{J} H_V = H^*(I_1^*, \ldots, I_J^*).
\]

(19)

Furthermore \( H^*(I_1^*, \ldots, I_J^*) \) is expressed as

\[
H^*(I_1^*, \ldots, I_J^*): \tau_{s_a+1} = \tau_{s_a+2} = \cdots = \tau_{s_a+\ell_a^*} \quad (a = 1, \ldots, J).
\]

(20)

For \( a = 1, \ldots, J \), we define \( (\tilde{\ell}^*_a (I_a^*)) \) by \( (u_{s_a+1}, \ldots, u_{s_a+\ell_a^*}) \) which minimize \( \sum_{j \in I_a} (u_j - \overline{R}_a)^2 \) under simple ordered restrictions \( u_{s_a+1} \leq u_{s_a+2} \leq \cdots \leq u_{s_a+\ell_a^*} \), i.e.,

\[
\sum_{j \in I_a} (\tilde{\ell}^*_a (I_a^*) - \overline{R}_a)^2 = \min_{u_{s_a+1} \leq u_{s_a+2} \leq \cdots \leq u_{s_a+\ell_a^*}} \sum_{j \in I_a} (u_j - \overline{R}_a)^2,
\]

where \( I_a^*, s_a \) and \( \ell_a^* \) are defined in (C1). For \( H^*(I_1^*, \ldots, I_J^*) \) of (19), we set

\[
M = M(I_1^*, \ldots, I_J^*) = \sum_{a=1}^{J} \ell_a^* = \# (I_a^*).
\]

For \( \ell = \ell_1^*, \ldots, \ell_J^* \), we define \( \alpha (M, \ell) \) by (15). Let us put, for \( \ell_a^* \geq 3 \),
Then from (17), we have, for \( \ell^*_a \geq 3, \)
\[
\lim_{n \to \infty} \mathbb{P}(T^*(I^*_a) \geq t) = \sum_{\ell_a = 2}^{\ell^*_a} P(L, \ell_a^*) P(\chi^2_{L-1} \geq t) \quad (t > 0).
\]

We denote a solution of \( t \) satisfying the following equation by \( \tilde{\chi}^{L}_{a}^2(\alpha(M, \ell^*_a)). \)
\[
\sum_{\ell_a = 2}^{\ell^*_a} P(L, \ell_a^*) P(\chi^2_{L-1} \geq t) = \alpha(M, \ell^*_a).
\]

Since we consider simple ordered restrictions (4) of treatment effects, we reject a null hypothesis when \( S(I^*_a) \) is too large for \( \ell^*_a = 2. \) We define rank statistic \( ST^*(I^*_a) \) by
\[
ST^*(I^*_a) = \begin{cases} 
S(I^*_a) & \text{if } \ell^*_a = 2 \\
T^*(I^*_a) & \text{if } \ell^*_a \geq 3.
\end{cases}
\]
(21)

We define the upper 100\( \alpha \) percent point of rank statistic \( ST^*(I^*_a) \) by
\[
zc^*(\alpha(M, \ell^*_a)) = \begin{cases} 
z(\alpha(M, \ell^*_a)) & \text{if } \ell^*_a = 2 \\
\tilde{\chi}^{L}_{a}^2(\alpha(M, \ell^*_a)) & \text{if } \ell^*_a \geq 3.
\end{cases}
\]
(22)

Then, we propose the closed testing procedure (II).
Table 5. Values of familywise error rate for [C1]-[H4] with $n=16$, $a=0.05$ and $H_0$

| $k$ | procedure | [C1] | [H3] | [C2] | [H4] |
|-----|------------|------|------|------|------|
| 3   |            | 0.0472 | 0.0363 | 0.0479 | 0.0398 |
| 4   |            | 0.0333 | 0.0275 | 0.0500 | 0.0352 |
| 5   |            | 0.0239 | 0.0241 | 0.0496 | 0.0334 |

Table 6. Values of familywise error rate for [C1]-[H4] with $n=16$, $a=0.01$ and $H_0$

| $k$ | procedure | [C1] | [H3] | [C2] | [H4] |
|-----|------------|------|------|------|------|
| 3   |            | 0.0056 | 0.0052 | 0.0091 | 0.0059 |
| 4   |            | 0.0035 | 0.0036 | 0.0085 | 0.0050 |
| 5   |            | 0.0020 | 0.0020 | 0.0089 | 0.0036 |

(II) Proposed closed testing procedure based on $\bar{\chi}^2$-statistics

(a) $J \geq 2$

Whenever $z\hat{c}^*(\alpha(M, \ell^*_a)|\ell_v^a) < ST^*(J^*_a)$ holds for an integer $a$ such that $1 \leq a \leq J$, we reject the hypothesis $\bigwedge_{v \in V} H_v$.

(b) $J = 1 \quad (M = \ell^*_1)$

Whenever $z\hat{c}^*(\alpha|\ell_1^*) < ST^*(J_1^*)$, we reject the hypothesis $\bigwedge_{v \in V} H_v$.

By using the methods of (a) and (b), when $\bigwedge_{v \in V} H_v$ is rejected for any $V$ such that $(j,j') \in V \subset U$, the null hypothesis $H_{(0,j)}$ is rejected as a multiple comparison test. We get Theorem 2.

Theorem 2. The test procedure (II) is an asymptotic multiple comparison test of level $\alpha$ as $n$ tends to infinity.

The proof of Theorem 2 is similar to the proof of Theorem 2.1 of Shiraishi and Matsuda (2018).

The values of $z\hat{c}^*(\alpha(M, \ell^*)|\ell^*)$ for $\alpha = 0.05, 0.01, 2 \leq \ell^* \leq M$ and $2 \leq M \leq 10$ are provided in Tables 3 and 4.

4. Simulation studies

We give results of several simulations to evaluate the performance of the procedures, each based on 100,000 Monte Carlo replicates. We deal two types of procedures, i.e. all pairwise comparisons with no restriction in sections 2, and all pairwise comparisons with simple ordered restrictions in section 3. The former type has (I) closed testing procedure based on rank statistics, denoted as [C1].

The latter type has (II) proposed closed testing procedure based on $\bar{\chi}^2$-statistics, as [C2].

For $S(I)$ defined by (10), let

$$S_{ij} = S((j,j')) \quad ((j,j') \in U).$$

Then we consider (III) two-sided Bonferroni-Holm method based on $|S_{ij}|$'s for another former type procedure, as [H3] and (IV) one-sided Bonferroni-Holm method based on $S_{ij}$'s for another
latter type procedure, as [H4].

On simulation studies, we use the standard normal distribution for \( e_{ij} \)’s. We also check the performance on other distributions, i.e., \( t \)-distribution with D.F. 5, logistic distribution and Laplace distribution. The other distributions have similar results, then we do not show other results.

At first, we check familywise error rate for those procedures for null hypothesis, \( H_0 \).

Tables 5 and 6 shows the results of familywise error rate for all pairwise comparisons for \( n=16 \), \( \mu_1 =0 \) and \( k =3, 4, 5 \).

In this situation, we can use Tables 1, 2, 3 and 4 in simulations.

We find that all procedures control the familywise error at level \( \alpha \) and the others except for [C2] are conservative in Table 5 and 6. Moreover, [C1] and [H3] are strongly conservative tending to increasing \( k \). These trends are kept even if we increase \( n \).

At second, we investigate the all-pairs power, which is shown on Ramsey (1978), on alternative hypotheses. All-pairs power is defined by the probability that all hypotheses which is false are rejected. For \( k =4 \), alternative hypotheses are the following one:

\[
H^{A1}, \mu_i=i\Delta/5 \quad (i=1, 2, 3, 4),
\]

\[
H^{A2}, \mu_1=\mu_2=0, \mu_3=\mu_4=\Delta/5.
\]

Tables 7, 8, 9 and 10 show the results of all-pairs power for all pairwise comparisons for \( n=16 \) and \( \Delta=4, 5, 6 \).

From Tables 7 and 9, the order of the power is the following:

\[ [C2] >[H4] >[C1] >[H3]. \]

Moreover, from Tables 8 and 10, the order of the power is the following:

\[ [C2] >[C1] >[H4] >[H3]. \]

[C2] has high power because they assume simple order restrictions. Even though assuming them, [H4] is lower than [C1] for \( H^{A2} \). We find that these results of order for procedures may be same for \( n\geq9 \) by investigating for \( n=9, 30 \).

We check the rate of reject of each pair. Tables 11 and 12 are the result of \( n=16 \) and \( \alpha=0.05 \).

The performance of a pair is similar to one of all-pairs power. The proposed procedure [C2] shows the best performance among all procedures on simple order restrictions. If we cannot assume simple order restrictions, we should use the procedure [C1].

In addition, as we use the normal distribution, we can compare the result above with the result in Shiraishi and Matsuda (2018). If the normal distribution is true, the power of nonparametric procedures in this paper are less than one of parametric procedures in Shiraishi and Matsuda (2018) with the maximum difference as 0.1815, which occur on [C2] and [B6] in Shiraishi and Matsuda.
Table 7. Values of all-pairs power for \([C1]-[H4]\) with \(n=16, \alpha=0.05\) and \(H^{41}\)

| \(\Delta\) | procedure | [C1]   | [H3]   | [C2]   | [H4]   |
|---------|-----------|--------|--------|--------|--------|
| 4       |           | 0.0824 | 0.0619 | 0.2373 | 0.2046 |
| 5       |           | 0.3258 | 0.2921 | 0.5606 | 0.5421 |
| 6       |           | 0.6396 | 0.6224 | 0.8157 | 0.8121 |

Table 8. Values of all-pairs power for \([C1]-[H4]\) with \(n=16, \alpha=0.05\) and \(H^{42}\)

| \(\Delta\) | procedure | [C1]   | [H3]   | [C2]   | [H4]   |
|---------|-----------|--------|--------|--------|--------|
| 4       |           | 0.1200 | 0.0426 | 0.4086 | 0.1072 |
| 5       |           | 0.2982 | 0.1443 | 0.6341 | 0.2795 |
| 6       |           | 0.5360 | 0.3343 | 0.8201 | 0.5194 |

Table 9. Values of all-pairs power for \([C1]-[H4]\) with \(n=16, \alpha=0.01\) and \(H^{41}\)

| \(\Delta\) | procedure | [C1]   | [H3]   | [C2]   | [H4]   |
|---------|-----------|--------|--------|--------|--------|
| 4       |           | 0.0017 | 0.0009 | 0.0129 | 0.0081 |
| 5       |           | 0.0274 | 0.0181 | 0.1040 | 0.0814 |
| 6       |           | 0.1479 | 0.1189 | 0.3477 | 0.3161 |

Table 10. Values of all-pairs power for \([C1]-[H4]\) with \(n=16, \alpha=0.01\) and \(H^{42}\)

| \(\Delta\) | procedure | [C1]   | [H3]   | [C2]   | [H4]   |
|---------|-----------|--------|--------|--------|--------|
| 4       |           | 0.0136 | 0.0021 | 0.1519 | 0.0090 |
| 5       |           | 0.0585 | 0.0124 | 0.3214 | 0.0418 |
| 6       |           | 0.1740 | 0.0510 | 0.5353 | 0.1370 |

Table 11. Rate of reject of a pair for \([C1]-[H4]\) with \(\Delta=4, \alpha=0.05\) and \(H^{41}\)

| Pair | procedure | [C1] | [H3] | [C2] | [H4] |
|------|-----------|------|------|------|------|
| (1,2)|           | 0.4429 | 0.3582 | 0.6194 | 0.5511 |
| (1,3)|           | 0.9188 | 0.9188 | 0.9725 | 0.9684 |
| (1,4)|           | 0.9985 | 0.9988 | 1.0000 | 0.9999 |
| (2,3)|           | 0.5196 | 0.3592 | 0.6725 | 0.5529 |
| (2,4)|           | 0.9177 | 0.9183 | 0.9733 | 0.9683 |
| (3,4)|           | 0.4426 | 0.3584 | 0.6203 | 0.5527 |

Table 12. Rate of reject of a pair for \([C1]-[H4]\) with \(\Delta=4, \alpha=0.05\) and \(H^{42}\)

| Pair | procedure | [C1] | [H3] | [C2] | [H4] |
|------|-----------|------|------|------|------|
| (1,2)|           | 0.0176 | 0.0083 | 0.0231 | 0.0108 |
| (1,3)|           | 0.2945 | 0.2297 | 0.5875 | 0.3650 |
| (1,4)|           | 0.2939 | 0.2274 | 0.8022 | 0.3618 |
| (2,3)|           | 0.2957 | 0.2292 | 0.4086 | 0.3655 |
| (2,4)|           | 0.2948 | 0.2310 | 0.5877 | 0.3656 |
| (3,4)|           | 0.0170 | 0.0082 | 0.0234 | 0.0108 |
Matsuda (2018) with the case of $H^{12}$ and $\Delta = 4$.

5. Application to death rates

In this section we show results for the application of the proposed procedures. We use the same notation of procedures in section 4.

In Shiraishi and Matsuda (2018), data of age-adjusted death rates in the Southern United States was used. This data have 4 major death rates, i.e. Heart Disease, Cancer, Chronic Lower Respiratory Diseases (CLRD) and Accident for 16 states from Centers for Disease Control and Prevention (2015). In point of view for homoscedasticity, we may use logarithmic transformed data for calculating estimates in this section. We apply proposed procedures in this paper to the data in which states are randomized block.

We get the values of $ST(I_2)$ of equation (14) and $-ST(I_2)$ of equation (21) as in Table 13, where we use $-ST(I_2)$ instead of $ST(I_2)$ because the rate of data has the decreasing order, which discuss in section 6. Moreover, statistics with tie is treated as the least favorable statistics, which also discuss in section 6.

In [C1], pair (1, 2), i.e. Heart Disease vs. Cancer, is rejected for $\alpha = 0.05$ but retained for $\alpha = 0.01$. Moreover, pair (3, 4), i.e. CLRD vs. Accident, is retained for $\alpha = 0.05$. The other pairs are rejected for $\alpha = 0.01$. In detail, we can use Tables 1 and 2 for this situation. If we want to reject for $H ((1, 2))$ in [C1], we should reject all $H ((1, 2)), H ((1, 2, 3)), H ((1, 2, 4)), H ((1, 2), (3, 4))$ and $H ((1, 2, 3, 4))$. For $\alpha = 0.01$, those evaluation are $2.741 > 2.576, 26 > 9.210, 26 > 9.210, 2.741 < 2.806$ and $39.675 > 11.345$. (We notice that the rejection of $H ((1, 2), (3, 4))$ may be satisfied by either $ST ((1, 2)) > 2.806$ or $ST ((3, 4)) > 2.806$.) Thus we retain $H ((1, 2))$ in [C1] for $\alpha = 0.01$.

In [C2], we assume the simple order restriction by using the order of result of United States simply. If we have a reasonable external evidence with respect to the order of the cause of death, it is desirable to use it. In the Southern United States, however, order of CLRD vs. Accident is exchanged. Therefore, the value of $-ST(I_2)$ of $H ((3, 4))$ becomes the minus. Regardless of this, the test result of [C2] is same as [C1] except for $H ((1, 2))$, i.e. pair (3, 4) is retained for $\alpha = 0.05$ and the other pairs is rejected for $\alpha = 0.01$. In detail, we can use Tables 3 and 4 for this situation. If we want to reject for $H ((1, 2))$ in [C2], we should reject all $H ((1, 2)), H ((1, 2, 3)), H ((1, 2), (3, 4))$ and $H ((1, 2, 3, 4))$. For $\alpha = 0.01$, those evaluation are $2.741 > 2.326, 26 > 6.823, 2.741 > 2.575$ and $39.675 > 7.709$. Thus we reject $H ((1, 2))$ in [C2] for $\alpha = 0.01$.

The result of [C2] is different from one of [B6] in Shiraishi and Matsuda (2018). [C2] is more powerful than [B6]. The reason is that the logarithmic transformed data could not be enough for homoscedasticity. For non homoscedasticity data the nonparametric method is sometimes useful like this.
Table 13. Statistics for 4 major death rates

| Cause numbers | $ST(I_a)$ | $-ST^*(I^*_a)$ |
|---------------|-----------|----------------|
| (1, 2)        | 2.741     | 2.741          |
| (1, 3)        | 3.516     | -              |
| (1, 4)        | 3.516     | -              |
| (2, 3)        | 3.516     | 3.516          |
| (2, 4)        | 3.516     | -              |
| (3, 4)        | 0.517     | -0.672         |
| (1, 2, 3)     | 26        | 26             |
| (1, 2, 4)     | 26        | -              |
| (1, 3, 4)     | 24.125    | -              |
| (2, 3, 4)     | 24.125    | 24             |
| (1, 2, 3, 4)  | 39.675    | 39.6           |

6. Discussion

We considered the randomized block design of (1). Hsu (1996) pointed out that it is difficult to obtain any exact distribution-free multiple comparison method except Bonferroni procedures for all pairwise comparisons in a randomized block design. We proposed closed testing procedures (I) and (II) based on signed rank statistics and Friedman test statistics for all pairwise comparisons of treatment effects. The two proposed procedures were distribution-free. We found that the proposed procedures were superior to the other procedures. Especially the power of the proposed closed testing procedures was fairly higher than that of the other procedures.

We proposed asymptotic procedures as the number of blocks tends to infinity. However we may discuss exact distribution-free procedures. Suppose that $Y_1, \ldots, Y_n$ are random sample observations from a standard normal distribution $N(0, 1)$. Let $R^+_i$ denote the rank of $|Y_1|, \ldots, |Y_n|$ and let $S = \sqrt{6/[n(n+1)(2n+1)]} \sum_{i=1}^{n} \text{sign}(Y_i) \cdot R^+_i$. Since the random variable $S$ has a discrete distribution, for $\alpha(M, \ell)$ defined by (15), we may give the following two inequalities of parameter $t$.

$$P(S \geq t) > \frac{\alpha(M, 2)}{2} \quad \text{and} \quad P(S > t) \leq \frac{\alpha(M, 2)}{2}.$$ 

We denote a solution of $t$ satisfying these inequalities by $sr(\alpha(M,2)/2|n)$. Let $Z_{1n}, \ldots, Z_{n\ell}$ be random sample observations from $i$-th standard normal distribution ($i=1, \ldots, n$). Let $R_{i\ell}$ denote the rank of $Z_i$ in $Z_{1n}, \ldots, Z_{n\ell}$ ($i=1, \ldots, n$) and let $T(\ell) = 12/[n\ell(\ell+1)] - \sum_{i=1}^{n} (\sum_{j=1}^{i} R_{ij\ell} - n(\ell+1)/2)^2$. We denote a solution of $t$ satisfying the following two inequalities by $fr(\alpha(M,\ell)|\ell,n)$.

$$P(T(\ell) \geq t) > \alpha(M, \ell) \quad \text{and} \quad P(T(\ell) > t) \leq \alpha(M, \ell).$$

Hence we put
Nonparametric Closed Testing Procedures for All Pairwise Comparisons in a Randomized Block Design

\[ zc(\alpha(M, \ell) \mid \ell, n) = \begin{cases} 
  sr(\alpha(M, 2) \mid n) & \text{if } \ell = 2 \\
  fr(\alpha(M, \ell) \mid \ell, n) & \text{if } \ell \geq 3, 
\end{cases} \]

In the procedure (I), by replacing \( zc(\alpha(M, \ell_a) \mid \ell_a) < ST(I_a) \) and \( zc(\alpha(\ell_i) < ST(I_i) \) with \( zc(\alpha(M, \ell_a) \mid \ell_a, n) < ST(I_a) \) and \( zc(\alpha(\ell_i, n) < ST(I_i) \), respectively, we get a closed testing procedure. The closed testing procedure is exactly distribution-free. \( \lim_{n \to \infty} zc(\alpha(M, \ell_a) \mid \ell_a, n) = zc(\alpha(M, \ell_a) \mid \ell_a) \) holds. Similarly, we may construct a distribution-free version of the procedure (II) under simple order restrictions.

We suppose the reverse order restrictions:

\[ \tau_1 \geq \tau_2 \geq \cdots \geq \tau_k. \]  

For \( i = 1, \ldots, n \) and \( j = 1, \ldots, k \), let us put \( X_{ij} = -X_{ij}, \mu'_i = -\mu, \beta' i = -\beta_i, \tau'_j = -\tau_j, \) and \( e'_j = -e_j \). Then we have a randomized block design \( X'_0 = \mu' + \beta'_i + \tau'_j + e'_j \) \((1 \leq i \leq n, 1 \leq j \leq k)\), where \( \sum_{i=1}^{n} \beta'_i = \sum_{j=1}^{k} \tau'_j = 0 \) and the \( e'_j \)'s are independent and identically distributed with absolutely continuous distribution function. (23) is equivalent to the simple ordered restrictions of \( \tau'_1 \)'s: \( \tau'_1 \leq \tau'_2 \leq \cdots \leq \tau'_k \). By replacing \( X_{ij} \) with \( X'_{ij} \) in all statistics of sections 3, we can discuss the multiple comparison procedures under the restrictions (23).

If there are sample observations tied in value, the multiple comparison test procedures are solved by using least favorable statistics stated in Gibbons and Chakraborti (2003). The least favorable statistics are given by choosing the different ranks among the tied observations which minimizes the probability of rejection.

Acknowledgements

The authors would like to thank the referees for some valuable comments. This research was supported in part by Nanzan University Pache Research Subsidy I-A-2 for 2018 academic year and a Grant-in-Aid for Co-operative Research (C) 18K11204 and 19K11870 from the Japanese Ministry of Education.

REFERENCES

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference under Order Restrictions, Wiley, London.

Centers for Disease Control and Prevention (2015). National Center for Health Statistics. https://www.cdc.gov/nchs/pressroom/stats_of_the_states.htm.

Chacko, V. J. (1963). Testing homogeneity against ordered alternatives. Ann. Math. Statist. 34, 945–956.

Enderton, H. B. (2001). A Mathematical Introduction to Logic: Second Edition, Academic Press.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance.

Jpn J Biomet Vol. 40, No. 1, 2019
Gibbons, J. D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference; Fourth Edition*, Marcel Dekker.

Hettmansperger, T. P. (1984). *Statistical Inference based on Ranks*, Wiley & Sons.

Hsu, J. C. (1996). *Multiple Comparisons; Theory and Methods*, Chapman & Hall.

Ramsey, P. H. (1978). Power differences between pairwise multiple comparisons. *J. Am. Stat. Assoc.* **73**, 479–485.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, New York.

Shiraishi, T. (2011). *Multiple Comparison Procedures in Multi-Sample Models: Continuous Case*. Tokyo, Japan: Kyoritsu Shuppan (in Japanese).

Shiraishi, T. and Matsuda, S. (2018). Closed testing procedures for all pairwise comparisons in a randomized block design. *Communications in Statistics - Theory and Methods* **47**, 3571–3587.

Shiraishi, T. and Sugiuira, H. (2015). The Upper \(100\alpha\) \(\star\)th Percentiles of the Distributions Used in Multiple Comparison Procedures Under a Simple Order Restriction. *Journal of the Japan Statistical Society* **44**, 271–314 (in Japanese).

Tamhane, A. C. (2009). *Statistical Analysis of Designed Experiments*, John Wiley, New York.