J. Japan Statist. Soc.
Vol. 37 No. 2 2007 157–174

MULTIPLE COMPARISONS BASED ON R-ESTIMATORS IN THE ONE-WAY LAYOUT

Taka-aki Shiraishi*

In a one-way analysis of variance model, robust versions based on R-estimators are proposed for single-step multiple comparisons procedures discussed by Tukey (1953), Dunnett (1955), and Scheffé (1953). The robust procedures are two methods based on joint ranks and pairwise ranks. It is shown that the two methods are asymptotically equivalent. Although we fail to construct simultaneous tests based on linear joint ranks, we are able to propose simultaneous tests based on the R-estimators. Robustness for asymptotic properties is discussed. The accuracy of asymptotic approximation is investigated.

Key words and phrases: Asymptotic property, robust statistics, simultaneous confidence intervals, simultaneous tests, single-step procedures.

1. Introduction

Let $\mu_1, \dots, \mu_k$ be the mean responses under k treatments. Suppose that, under the i-th treatment, a random sample $X_{i1}, \dots, X_{in_i}$ is taken. Then we have the one-way model

$$X_{ij} = \mu_i + e_{ij} \qquad (j = 1, \dots, n_i, i = 1, \dots, k) \tag{1.1}$$

where e_{ij} is a random variable with $E(e_{ij}) = 0$ for all i, j's. It is further assumed that e_{ij}'s are independent and identically distributed with a continuous distribution function (d.f.) $F(x)$. Let $\mathrm{Var}(e_{ij}) = \sigma^2 > 0$. The model (1.1) is rewritten as usual by

$$X_{ij} = \nu + \tau_i + e_{ij},$$

where $\sum_{i=1}^k n_i\tau_i = 0$. Then ν and τ_i's are referred to as the grand mean and additive treatment effects, respectively. We put $N = \sum_{i=1}^k n_i$. The least squares estimator of τ_i is given by $\tilde{\tau}_i = \bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}$, where $\bar{X}_{i\cdot} = \sum_{j=1}^{n_i} X_{ij}/n_i$ and $\bar{X}_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij}/N$.

The relations of

$$\mu_i - \mu_{i'} = \tau_i - \tau_{i'} \qquad \text{and} \qquad \bar{X}_{i\cdot} - \bar{X}_{i'\cdot} = \tilde{\tau}_i - \tilde{\tau}_{i'}$$

hold. We discuss single-step procedures. Let

$$\tilde{T}_{ii'} = \frac{\tilde{\tau}_i - \tilde{\tau}_{i'} - (\tau_i - \tau_{i'})}{\sqrt{\tilde{\sigma}^2 \cdot (1/n_i + 1/n_{i'})}} \qquad \text{and} \qquad \tilde{T}^*_{ii'} = \frac{\tilde{\tau}_i - \tilde{\tau}_{i'}}{\sqrt{\tilde{\sigma}^2 \cdot (1/n_i + 1/n_{i'})}},$$

Received December 8, 2005. Revised October 31, 2006. Accepted June 25, 2007.

*Department of Mathematical Sciences, Yokohama City University, Seto 22-2, Kanazawa-ku, Yokohama 236-0027, Japan.

where $\tilde{\sigma}^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i\cdot})^2/(N-k)$. Furthermore, we put

$$G_{0\boldsymbol{n}}(t) = P_0\left(\max_{1\le i<i'\le k}|\tilde{T}_{ii'}| \le t\right),$$

where $P_0(\cdot)$ stands for a probability measure assuming that e_{ij} has a normal distribution with mean 0 and variance σ^2. We introduce studentized range distribution $D_{k,m}(t)$ with $k-1$ and m degrees of freedom which is expressed as

$$D_{k,m}(t) = k\int_0^\infty \int_{-\infty}^\infty \{\Phi(x) - \Phi(x - \sqrt{2}\cdot ts)\}^{k-1}\cdot\varphi(x)dxg(s)ds, \tag{1.2}$$

where $\Phi(x)$ and $\varphi(x)$ denotes the distribution function and density function of a standard normal distribution respectively,

$$g(s) = \frac{m^{m/2}}{\Gamma(m/2)2^{(m/2-1)}}s^{m-1}\exp(-ms^2/2), \qquad \text{and} \qquad m = N-k.$$

Hayter (1984) showed the inequality

$$D_{k,m}(t) \le G_{0\boldsymbol{n}}(t). \tag{1.3}$$

Hence using

$$|\tilde{T}_{ii'}| \le \max_{1\le i<i'\le k}|\tilde{T}_{ii'}|,$$

Tukey (1953) and Kramer (1956) proposed normal theory $100(1-\alpha)\%$ simultaneous confidence intervals for all-pairwise $\{\mu_i - \mu_{i'}; 1 \le i < i' \le k\}$ given by

$$\mu_i - \mu_{i'} \in \tilde{\tau}_i - \tilde{\tau}_{i'} \pm p(k,m;\alpha)\cdot\sqrt{\tilde{\sigma}^2\cdot(1/n_i + 1/n_{i'})} \qquad \text{for} \quad 1 \le i < i' \le k,$$

where $p(k,m;\alpha)$ satisfies $D_{k,m}(p(k,m;\alpha)) = 1-\alpha$. From (1.3), we find that the normal theory $100(1-\alpha)\%$ simultaneous confidence intervals are conservative. Normal theory simultaneous tests of level α for the null hypotheses $\{H_{(ii')} : \mu_i = \mu_{i'}\}_{\{1\le i<i'\le k\}}$ also consist in rejecting $H_{(ii')}$ for $1 \le i < i' \le k$ such that $|\tilde{T}^*_{ii'}| > p(k,m;\alpha)$.

Similarly, using the relation of the inequalities

$$|\tilde{T}_{1i}| \le \max_{2\le i\le k}|\tilde{T}_{1i}| \qquad \text{and} \qquad \tilde{T}_{1i} \le \max_{2\le i\le k}\tilde{T}_{1i},$$

Dunnett (1955) proposed normal theory multiple comparisons procedures for the differences between control and treatments $\{\mu_1 - \mu_i; i = 2,\ldots,k\}$.

Let $\mathcal{C}^k = \{\boldsymbol{c} = (c_1,\ldots,c_k) : \sum_{i=1}^k c_i = 0\}$. For any $\boldsymbol{c} \in \mathcal{C}^k$, the relations of

$$\sum_{i=1}^k c_i\mu_i = \sum_{i=1}^k c_i\tau_i \qquad \text{and} \qquad \sum_{i=1}^k c_i\bar{X}_{i\cdot} = \sum_{i=1}^k c_i\tilde{\tau}_i$$

hold. For some $\boldsymbol{c} \in \mathcal{C}^k$, we put

$$\tilde{T}_{\boldsymbol{c}} = \frac{\sum_{i=1}^k c_i(\tilde{\tau}_i - \tau_i)}{\sqrt{\tilde{\sigma}^2 \cdot \sum_{i=1}^k c_i^2/n_i}} \quad \text{and} \quad \tilde{T}^*_{\boldsymbol{c}} = \frac{\sum_{i=1}^k c_i\tilde{\tau}_i}{\sqrt{\tilde{\sigma}^2 \cdot \sum_{i=1}^k c_i^2/n_i}}.$$

Scheffé (1953) showed

$$\sup_{\boldsymbol{c}\in\mathcal{C}^k} \tilde{T}^2_{\boldsymbol{c}} = \frac{\sum_{i=1}^k n_i(\tilde{\tau}_i - \tau_i)^2}{\tilde{\sigma}^2}.$$

$(\sup_{\boldsymbol{c}\in\mathcal{C}^k} \tilde{T}^2_{\boldsymbol{c}})/(k-1)$ has a F distribution with $k-1$ and $N-k$ degrees of freedom. Hence he was able to propose a normal theory $100(1-\alpha)\%$ simultaneous confidence intervals for contrasts $\sum_{i=1}^k c_i\mu_i : \boldsymbol{c} \in \mathcal{C}^k$ given by

$$\sum_{i=1}^k c_i\mu_i \in \sum_{i=1}^k c_i\tilde{\tau}_i \pm \sqrt{(k-1)\cdot F^{k-1}_{N-k,\alpha}\cdot \tilde{\sigma}^2 \cdot \sum_{i=1}^k c_i^2/n_i},$$

where $F^{k-1}_{N-k,\alpha}$ denotes the upper α point of the F distribution with $k-1$ and $N-k$ degrees of freedom. Normal theory simultaneous tests of level α for the null hypotheses $\{H_{\boldsymbol{c}} : \sum_{i=1}^k c_i\mu_i = 0\}_{\boldsymbol{c}\in\mathcal{C}^k}$ also consist in rejecting $H_{\boldsymbol{c}}$ for $\boldsymbol{c} \in \mathcal{C}$ satisfying

$$\tilde{T}^{*2}_{\boldsymbol{c}} > (k-1)F^{k-1}_{N-k,\alpha}.$$

As nonparametric tests based on pairwise ranks, Steel (1960) and Dwass (1960) discussed simultaneous tests for the null hypotheses of all-pairwise $\{H_{(ii')}\}_{\{1\le i<i'\le k\}}$. Steel (1959) discussed simultaneous tests for the null hypotheses of control vs. treatments $\{H_{(1i)}\}_{\{2\le i\le k\}}$. As a nonparametric test based on joint ranks, Dunn (1964) proposed simultaneous tests for the null hypotheses of all-pairwise comparison. Her rank test procedure is (asymptotically) distribution-free only under the overall null hypothesis

$$H_0;\ \tau_1 = \cdots = \tau_k = 0,$$

that is, the (asymptotic) distribution of her rank test statistic does not depend on $F(x)$ under H_0. However her rank procedure testing the null hypothesis H_{12} is not (asymptotically) distribution-free under H_{12} when H_{12} is true and H_{13} is not true. Therefore Oude Voshaar (1980) and Hsu (1996) pointed out that the test procedures based on joint ranks are not recommended as simultaneous tests for the null hypotheses of $\{H_{(ii')}\}_{\{1\le i<i'\le k\}}$ and for the null hypotheses of $\{H_{(1i)}\}_{\{2\le i\le k\}}$.

Sen (1966) and Sen (1980) stated simultaneous confidence intervals of $\{\mu_i - \mu_{i'}; 1 \le i < i' \le k\}$ as a nonparametric T-method based on pairwise ranks in detail for $n_1 = \cdots = n_k$. He also discussed simultaneous confidence intervals and tests for contrasts. However his procedures are laborious and it is hard to make the algorithms. The versions based on R-estimators are proposed for single-step multiple comparisons procedures discussed by Tukey (1953), Dunnett

(1955), and Scheffé (1953). The proposed procedures are the two methods based on joint ranks and based on pairwise ranks, and they are more simple than Sen's procedures. It is shown that the two methods are asymptotically equivalent. Although the exact distributions for the normal theory pocedures are given by double integrals, the asymptotic distributions for the proposed procedures and the normal theory pocedures are expressed as single integrals. Although we fail to construct simultaneous tests based on linear joint ranks, we are able to propose simultaneous tests based on joint rank estimators. Robustness for asymptotic properties is discussed. The accuracy of asymptotic approximation is investigated.

2. R-estimators based on joint ranks

For the k-dimensional row vector $\boldsymbol{s} = (s_1, \ldots, s_k)$, we put $X_{ij}(\boldsymbol{s}) = X_{ij} - s_i$. Setting $N = \sum_{i=1}^k n_i$, let $R_{ij}(\boldsymbol{s})$ be the rank of $X_{ij}(\boldsymbol{s})$ among the N observations $X_{11}(\boldsymbol{s}), \ldots, X_{kn_k}(\boldsymbol{s})$. Using these ranks and the score functions $a_N(\cdot)$ which is a map from $\{1, \ldots, N\}$ to real values, for $\boldsymbol{s}$, let us put

$$S_i(\boldsymbol{s}) = \sum_{j=1}^{n_i} \{a_N(R_{ij}(\boldsymbol{s})) - \bar{a}_N\}/\sqrt{N}, \tag{2.1}$$

where $\bar{a}_N = \sum_{\ell=1}^N a_N(\ell)/N$.

Let

$$\mathcal{A}_N(R) = \left\{ \boldsymbol{\theta} : \sum_{i=1}^k |S_i(\boldsymbol{\theta})| = \text{minimum subject to } \sum_{i=1}^k n_i\theta_i = 0 \right\},$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$. Then Shiraishi (1990) proposed one point $\hat{\boldsymbol{\tau}} = (\hat{\tau}_1, \ldots, \hat{\tau}_k) \in \mathcal{A}_N(R)$ as an R-estimator of the row vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_k)$. It is simple to verify

$$\left\{ \boldsymbol{\theta} : \sum_{i=1}^k |S_i(\boldsymbol{\theta} + \boldsymbol{\tau})| = \text{minimum subject to } \sum_{i=1}^k n_i\theta_i = 0 \right\}$$
$$= \{\boldsymbol{\theta} + \boldsymbol{\tau} : \boldsymbol{\theta} \in \mathcal{A}_N(R)\}.$$

The follwing are the minimum assumptions needed to discuss the asymptotic theory.

Assumption 1. The score function $a_N(\cdot)$ is taken as

$$a_N(\ell) = E\{\psi(U_N(\ell))\} \quad \text{or} \quad \psi(\ell/(N+1)) \quad \text{for} \quad \ell = 1, \ldots, N,$$

where $U_N(\ell)$ is the ℓ-th order statistic in a sample of size N from uniform $(0, 1)$ distribution. The score generating function $\psi(u)$ is non-constant, nondecreasing and square integrable.

Assumption 2. $\lim_{N\to\infty}(n_i/N) = \lambda_i > 0$ for $i = 1, \ldots, k$.

ASSUMPTION 3. $f(x)$ has finite Fisher's information

$$0 < \int_{-\infty}^{\infty} \{-f'(x)/f(x)\}^2 f(x) dx < \infty.$$

As in the proof of Lemma 2.1 of Shiraishi (1990), we can derive the following asymptotic linearity for the rank statistic $S_i(\boldsymbol{s})$.

PROPOSITION 1. *Let* $\|\boldsymbol{z}\| = \sqrt{\boldsymbol{z}\boldsymbol{z}'}$ *for the k-dimensional row vector* $\boldsymbol{z}$. *Then under Assumptions* 1–3, *for any positive* ε, C_1 *and* C_2,

$$\lim_{N\to\infty} P\left\{ \sup_{\|\boldsymbol{\Delta}\|<C_1 \quad |\Delta^*|<C_2} |S_i(\boldsymbol{\tau} + \boldsymbol{\Delta}/\sqrt{N} + \Delta^* \boldsymbol{\omega}_i/\sqrt{N}) - S_i(\boldsymbol{\tau}) \right.$$
$$\left. + \lambda_i d(\Delta_i - \bar{\Delta}.) + d\Delta^* \lambda_i (1-\lambda_i)| > \varepsilon \right\} = 0,$$

where $\boldsymbol{\Delta} = (\Delta_1, \ldots, \Delta_k)$, $\bar{\Delta}. = \sum_{i=1}^k \lambda_i \Delta_i$, $\boldsymbol{\omega}_i$ *is a k-dimensional row vector with* 1 *at the i-th element and* 0 *elsewhere, and* $d = -\int_0^1 \{\psi(u) \cdot f'(F^{-1}(u))/ f(F^{-1}(u))\} du$.

ASSUMPTION 4. $d > 0$.

In many cases, using integration by parts yields $d = \int_0^1 \psi'(F(x)) \cdot \{f(x)\}^2 dx$. Thus Assumption 4 is feasible. Under Assumptions 1–4, by applying $\sqrt{N}(\hat{\tau}_i - \tau_i)$ and 0 to Δ_i and Δ^* respectively in Proposition 1, the proof for Theorem 3.1 of Shiraishi (1990) implies

$$\sqrt{N}(\hat{\tau}_i - \tau_i) - S_i(\boldsymbol{\tau})/(d \cdot \lambda_i) \xrightarrow{\text{P}} 0, \tag{2.2}$$

where $\xrightarrow{\text{P}}$ denotes convergence in probability. From the proof of VI.1.5 Theorem 1 of Hájek *et al.* (1999), we get

$$S_i(\boldsymbol{\tau}) - V_i \xrightarrow{\text{P}} 0, \tag{2.3}$$

where

$$V_i = \sum_{j=1}^{n_i} \{\psi(F(e_{ij})) - \bar{\psi}(F(e..))\}/\sqrt{N},$$

and $\bar{\psi}(F(e..)) = \sum_{i=1}^k \sum_{j=1}^{n_i} \psi(F(e_{ij}))/N$. From (2.2) and (2.3), we get

$$\sqrt{N}(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})' - (1/d)\text{diag}(1/\lambda_1, \ldots, 1/\lambda_k)\, \boldsymbol{V}' \xrightarrow{\text{P}} \boldsymbol{0}, \tag{2.4}$$

where $\boldsymbol{V} = (V_1, \ldots, V_k)$. Hence using the Cramér-Wold technique to $\boldsymbol{V}'$, it follows that

$$\sqrt{N}(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})' \xrightarrow{\text{L}} N_k(\boldsymbol{0}, (\gamma^2/d^2)\Lambda) \tag{2.5}$$

where $\xrightarrow{\text{L}}$ denotes convergence in law, $N_k(\boldsymbol{\theta}, \Sigma)$ stands for the k-dimensional normal variable with mean $\boldsymbol{\theta}$ and variance-covariance matrix Σ, $\gamma^2 = \int_0^1 \{\psi(u) - \bar{\psi}\}^2 du$, $\bar{\psi} = \int_0^1 \psi(u)du$, and $\Lambda = (\delta_{ii'}/\lambda_i - 1)_{ii'=1,\ldots,k}$ with $\delta_{ii'}$ denoting Kronecker's delta.

(2.5) is also expressed as

$$\text{(2.6)} \qquad \sqrt{N}(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})' \xrightarrow{\text{L}} (\gamma/d) \left(Y_1 - \sum_{j=1}^{k} \lambda_j Y_j, \ldots, Y_k - \sum_{j=1}^{k} \lambda_j Y_j \right)',$$

where $Y_1, \ldots, Y_k$ are independent and Y_i has a normal distribution with mean 0 and variance $1/\lambda_i$ $(i = 1, \ldots, k)$.

3. R-estimators based on pairwise ranks

For the scalar t, let $X_{ij}(t) = X_{ij} - t$. Setting $N_{ii'} = n_i + n_{i'}$, we define $R_{ij}^{i'}(t)$ by the rank of $X_{ij}(t)$ among the $N_{ii'}$ observations $X_{i1}(t), \ldots, X_{in_i}(t)$, $X_{i'1}, \ldots, X_{i'n_{i'}}$. Using these ranks and the score functions $a_{N_{ii'}}(\cdot)$, for t, let us put

$$\text{(3.1)} \qquad S_{ii'}(t) = \sum_{j=1}^{n_i} \{a_{N_{ii'}}(R_{ij}^{i'}(t)) - \bar{a}_{N_{ii'}}\}/\sqrt{N}.$$

Then $S_{ii'}(t)$ is nonincreasing in t. Using a method similar to Hodges and Lehmann (1963), we propose an estimator of $\eta_{ii'} = \mu_i - \mu_{i'}$,

$$\breve{\eta}_{ii'} = \frac{1}{2}[\inf\{t : S_{ii'}(t) < 0\} + \sup\{t : S_{ii'}(t) > 0\}].$$

When $a_{N_{ii'}}(\cdot)$ is of Wilcoxon's type, that is, $a_{N_{ii'}}(\ell) = 2\ell/(N_{ii'} + 1) - 1$, this R-estimator is expressed as

$$\breve{\eta}_{ii'} = \text{the sample median of } \{X_{ij} - X_{i'j'} : j = 1, \ldots, n_i,\ j' = 1, \ldots, n_{i'}\}.$$

Since $\tau_i = (1/N) \sum_{i'=1}^{k} n_{i'} \eta_{ii'}$, we may propose as an R-estimator of τ_i

$$\breve{\tau}_i = (1/N) \sum_{i'=1}^{k} n_{i'} \breve{\eta}_{ii'},$$

where we set $\breve{\eta}_{ii} = 0$ for convenience. Then we can derive Proposition 2 similar to Proposition 1.

Proposition 2. *Under Assumptions* 1–3, *for any positive* ε *and* C,

$$\lim_{N \to \infty} P \left\{ \sup_{|\Delta| < C} |S_{ii'}(\eta_{ii'} + \Delta/\sqrt{N}) - S_{ii'}(\eta_{ii'}) + \lambda_i d(n_{i'}/N_{ii'})\Delta| > \varepsilon \right\} = 0.$$

Under Assumptions 1–4, by using Proposition 2, we get (3.2) and (3.3) similar to (2.2) and (2.3).

$$\text{(3.2)} \qquad \sqrt{N}(\breve{\eta}_{ii'} - \eta_{ii'}) - N_{ii'} \cdot S_{ii'}(\eta_{ii'})/(n_{i'} \cdot d \cdot \lambda_i) \xrightarrow{\text{P}} 0$$

and

$$S_{ii'}(\eta_{ii'}) - W_{ii'} \xrightarrow{\mathrm{P}} 0, \tag{3.3}$$

where

$$W_{ii'} = \sum_{j=1}^{n_i} \left\{ \psi(F(e_{ij})) - \left[\sum_{j=1}^{n_i} \psi(F(e_{ij})) + \sum_{j'=1}^{n_{i'}} \psi(F(e_{i'j'})) \right] / N_{ii'} \right\} / \sqrt{N}.$$

From (3.2) and (3.3), we get

$$\sqrt{N}(\breve{\boldsymbol{\tau}} - \boldsymbol{\tau})' - (1/d)\mathrm{diag}(1/\lambda_1, \ldots, 1/\lambda_k)\, \boldsymbol{V}' \xrightarrow{\mathrm{P}} \mathbf{0}, \tag{3.4}$$

where $\breve{\boldsymbol{\tau}} = (\breve{\tau}_1, \ldots, \breve{\tau}_k)$ and $\boldsymbol{V}$ is defined in (2.4). Hence, from (2.4) and (3.4), we get

PROPOSITION 3. *Under Assumptions* 1–4, $\sqrt{N}(\breve{\boldsymbol{\tau}} - \boldsymbol{\tau})'$ *is asymptotically equivalent to* $\sqrt{N}(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})'$ *as* N *tends to infinity.*

4. Tukey-type procedures

Let us put

$$\hat{\breve{\boldsymbol{\tau}}} = \hat{\boldsymbol{\tau}} \quad \text{or} \quad \breve{\boldsymbol{\tau}}, \qquad \text{that is,} \quad \hat{\breve{\tau}}_i = \hat{\tau}_i \quad \text{or} \quad \breve{\tau}_i \ \ (i = 1, \ldots, k).$$

Hájek *et al.* (1999) showed

$$\hat{\gamma}^2 = \sum_{\ell=1}^{N} \{a_N(\ell) - \bar{a}_N\}^2/(N-1) \to \gamma^2 \qquad (\text{as} \quad N \to \infty). \tag{4.1}$$

Let us define, for a constant b,

$$\hat{d}_{N(i)} = \{S_i(\hat{\breve{\boldsymbol{\tau}}} - b\boldsymbol{\omega}_i/\sqrt{N}) - S_i(\hat{\breve{\boldsymbol{\tau}}} + b\boldsymbol{\omega}_i/\sqrt{N})\}/\{2b(n_i/N)(1 - n_i/N)\},$$

where $\boldsymbol{\omega}_i$ is defined in Proposition 1. Then we put

$$\hat{d} = \sum_{i=1}^{k} n_i \hat{d}_{N(i)}/N.$$

LEMMA 1. *Suppose that Assumptions* 1–4 *are satisfied. Then as* N *tends to infinity*, $\hat{d}$ *is a consistent estimator of* d.

PROOF. Applying $\Delta_i = \sqrt{N}(\hat{\breve{\tau}}_i - \tau_i)$ for $i = 1, \ldots, k$ and $\Delta^* = \mp b$ in Proposition 1, we get

$$S_i(\hat{\breve{\boldsymbol{\tau}}} - b\boldsymbol{\omega}_i/\sqrt{N}) - S_i(\boldsymbol{\tau}) + \lambda_i d\sqrt{N}(\hat{\breve{\tau}}_i - \tau_i) - bd\lambda_i(1 - \lambda_i) \xrightarrow{\mathrm{P}} 0, \tag{4.2}$$

and

$$S_i(\hat{\breve{\boldsymbol{\tau}}} + b\boldsymbol{\omega}_i/\sqrt{N}) - S_i(\boldsymbol{\tau}) + \lambda_i d\sqrt{N}(\hat{\breve{\tau}}_i - \tau_i) + bd\lambda_i(1 - \lambda_i) \xrightarrow{\mathrm{P}} 0. \tag{4.3}$$

Then from (4.2) and (4.3), we find

$$\hat{d}_{N(i)} \xrightarrow{\mathrm{P}} d,$$

which implies the conclusion. By using $S_{ii'}(\hat{\mathring{\boldsymbol{\tau}}} \pm b\boldsymbol{\omega}_i/\sqrt{N})$, we may also construct a consistent estimator of d, which is similar to $\hat{d}$.

Let

$$\hat{T}_{ii'} = \frac{\hat{\mathring{\tau}}_i - \hat{\mathring{\tau}}_{i'} - (\tau_i - \tau_{i'})}{\sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot (1/n_i + 1/n_{i'})}}$$

and

$$\hat{T}^*_{ii'} = \frac{\hat{\mathring{\tau}}_i - \hat{\mathring{\tau}}_{i'}}{\sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot (1/n_i + 1/n_{i'})}}.$$

Then we get

THEOREM 1. *Under Assumptions* 1–4, *for any positive* t,

$$A(t) \leq \lim_{N\to\infty} P\left(\max_{1\leq i<i'\leq k} |\hat{T}_{ii'}| \leq t\right) \leq B(t) \tag{4.4}$$

holds, and when $\lambda_i = \frac{1}{k}$ $(1 \leq i \leq k)$, *both the equalities in* (4.4) *hold, where*

$$A(t) = k\int_{-\infty}^{\infty} \{\Phi(x) - \Phi(x - \sqrt{2}\cdot t)\}^{k-1}\varphi(x)dx, \tag{4.5}$$

and

$$B(t) = \int_{-\infty}^{\infty} \sum_{j=1}^{k} \prod_{i=1 \ i\neq j}^{k} \left\{\Phi\left(\sqrt{\frac{\lambda_i}{\lambda_j}}\cdot x\right) - \Phi\left(\sqrt{\frac{\lambda_i}{\lambda_j}}\cdot x - \sqrt{\frac{\lambda_i+\lambda_j}{\lambda_j}}\cdot t\right)\right\}\varphi(x)dx.$$

PROOF. From (2.6), (4.1) and Lemma 1, we have

$$\lim_{N\to\infty} P\left(\max_{1\leq i<i'\leq k} |\hat{T}_{ii'}| \leq t\right) = P\left(\max_{1\leq i\leq j\leq k} \frac{|Y_i - Y_j|}{\sqrt{\dfrac{1}{\lambda_i} + \dfrac{1}{\lambda_j}}} \leq t\right),$$

where $Y_1, \ldots, Y_k$ are defined in (2.6). Let us put the events

$$C(t) = \left\{\max_{1\leq i\leq j\leq k} \frac{|Y_i - Y_j|}{\sqrt{\dfrac{1}{\lambda_i} + \dfrac{1}{\lambda_j}}} \leq t\right\}, \qquad D_j = \{Y_j = \max\{Y_1, \ldots, Y_k\}\},$$

and

$$E_j(t) = \left\{ Y_j - t\sqrt{\frac{1}{\lambda_i} + \frac{1}{\lambda_j}} \le Y_i \le Y_j,\ i = 1, 2, \ldots, k,\ i \ne j \right\}.$$

Then we get

$$C(t) \cap D_j \subset E_j(t). \tag{4.6}$$

Also we have

$$\begin{aligned} P(E_j(t)) &= \int_{-\infty}^{\infty} P\left(y_j - t\sqrt{\frac{1}{\lambda_i} + \frac{1}{\lambda_j}} \le Y_i \le y_j,\ i = 1, 2, \ldots, k,\ i \ne j \right) \\ &\quad \times \sqrt{\lambda_j}\varphi\left(\sqrt{\lambda_j} y_j\right) dy_j \\ &= \int_{-\infty}^{\infty} \prod_{i=1\ i\ne j}^{k} \left\{ \Phi\left(\sqrt{\lambda_i} y_j\right) - \Phi\left(\sqrt{\lambda_i} y_j - t \cdot \sqrt{1 + \frac{\lambda_i}{\lambda_j}} \right) \right\} \\ &\quad \times \sqrt{\lambda_j}\varphi(\sqrt{\lambda_j} y_j) dy_j. \end{aligned}$$

Furthermore, by using the change of variable $x = \sqrt{\lambda_j} y_j$, we may derive

$$\begin{aligned} P(E_j(t)) = \int_{-\infty}^{\infty} \prod_{i=1\ i\ne j}^{k} &\left\{ \Phi\left(\sqrt{\frac{\lambda_i}{\lambda_j}} x \right) \right. \\ &\left. - \Phi\left(\sqrt{\frac{\lambda_i}{\lambda_j}} x - \sqrt{\frac{\lambda_i + \lambda_j}{\lambda_j}} t \right) \right\} \varphi(x) dx. \end{aligned} \tag{4.7}$$

Hence we get the right hand side of the inequalities in (4.4) from (4.6) and (4.7). The left hand side of the inequalities in (4.4) is the main result of Hayter (1984).

Next assume that $\lambda_i = \frac{1}{k}$ $(1 \le i \le k)$. Then we get $C(t) \cap D_j = E_j(t)$, which implies

$$P(C(t) \cap D_j) = P(E_j(t)) = \int_{-\infty}^{\infty} \{\Phi(x) - \Phi(x - \sqrt{2} \cdot t)\}^{k-1} \varphi(x) dx.$$

Therefore we get all the conclusions.

Example 1. Suppose

$$k = 3, \qquad \lambda_1 = 1/6, \qquad \lambda_2 = 1/3, \qquad \lambda_3 = 1/2. \tag{4.8}$$

Without any loss of generality, we put $t = 1$. If we set $Y_1 = 1.2$, $Y_2 = -1.7$, and $Y_3 = 0.9$, we have $(Y_1, Y_2, Y_3) \in E_1(1) \cap D_1$ and $(Y_1, Y_2, Y_3) \notin C(1)$. Hence under (4.8), we get $C(1) \cap D_1 \ne E_1(1)$, which implies

$$\lim_{N\to\infty} P\left(\max_{1\le i<i'\le k} |\hat{T}_{ii'}| \le t \right) < B(t).$$

We define $p_1(k; \alpha)$ and $q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$ by the upper $100\alpha\%$ points of $A(t)$ and $B(t)$ respectively, that is, $1 - A(p_1(k; \alpha)) = \alpha$ and $1 - B(q_1(k, \lambda_1, \ldots,$

$\lambda_k;\alpha)) = \alpha$. Let $t(k, \lambda_1, \ldots, \lambda_k; \alpha)$ be a unique solution of t satisfying $\lim_{N\to\infty} P(\max_{1\le i<i'\le k} |\hat{T}_{ii'}| \le t) = 1 - \alpha$. Then from Theorem 1, we find

$$p_1(k;\alpha) \ge t(k, \lambda_1, \ldots, \lambda_k; \alpha) \ge q_1(k, \lambda_1, \ldots, \lambda_k; \alpha). \tag{4.9}$$

The values of $q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$ for $\alpha = 0.05$, 0.01 and $k = 3(1)10$ are provided in Table 1. Since $B(t)$ depends on $\lambda_i' s$, computations were restricted to $\lambda_i = (1/k)[1 + 2u\{i - (k+1)/2\}/\{3(k-1)\}]$; $u = 0.0(0.1)1.0$. When $u = 0$, we find $p_1(k;\alpha) = q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$. From Table 1, it can be seen that (i) the value of $q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$ decreases in u, and that (ii) the value of $q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$ is nearly equal to $p_1(k;\alpha)$ when $1 < \max\{\lambda_i : i = 1, \ldots, k\}/\min\{\lambda_i : i = 1, \ldots, k\} \le 2$.

Table 1. The values of $q_1(k, \lambda_1, \ldots, \lambda_k; \alpha)$ for $\alpha = 0.05$, 0.01 and $k = 3(1)10$ where $\lambda_i = (1/k)[1 + 2u\{i - (k+1)/2\}/\{3(k-1)\}]$ $(i = 1, \ldots, k)$.

(i) $\alpha = 0.05$

	k							
u	3	4	5	6	7	8	9	10
0.0	2.344	2.569	2.728	2.850	2.948	3.031	3.102	3.164
0.1	2.344	2.569	2.728	2.850	2.948	3.031	3.102	3.164
0.2	2.343	2.569	2.727	2.849	2.948	3.030	3.101	3.163
0.3	2.343	2.568	2.727	2.849	2.947	3.030	3.101	3.162
0.4	2.343	2.568	2.726	2.848	2.946	3.029	3.100	3.161
0.5	2.342	2.567	2.725	2.847	2.945	3.028	3.098	3.160
0.6	2.341	2.566	2.724	2.846	2.944	3.026	3.097	3.159
0.7	2.340	2.564	2.723	2.844	2.942	3.025	3.095	3.157
0.8	2.339	2.563	2.721	2.842	2.940	3.023	3.093	3.155
0.9	2.338	2.561	2.719	2.840	2.938	3.020	3.091	3.152
1.0	2.336	2.559	2.717	2.838	2.936	3.018	3.088	3.149

(ii) $\alpha = 0.01$

	k							
u	3	4	5	6	7	8	9	10
0.0	2.913	3.113	3.255	3.364	3.452	3.526	3.590	3.646
0.1	2.913	3.113	3.255	3.364	3.452	3.526	3.590	3.646
0.2	2.913	3.113	3.254	3.363	3.452	3.526	3.590	3.646
0.3	2.913	3.113	3.254	3.363	3.451	3.526	3.589	3.645
0.4	2.913	3.112	3.253	3.362	3.451	3.525	3.589	3.645
0.5	2.912	3.111	3.252	3.361	3.450	3.524	3.588	3.644
0.6	2.911	3.111	3.252	3.360	3.449	3.523	3.587	3.642
0.7	2.910	3.109	3.251	3.359	3.447	3.522	3.585	3.641
0.8	2.909	3.108	3.249	3.358	3.446	3.520	3.583	3.639
0.9	2.908	3.107	3.248	3.356	3.444	3.518	3.582	3.637
1.0	2.907	3.105	3.246	3.354	3.442	3.516	3.579	3.635

Hence, from (4.9), the value of $t(k, \lambda_1, \ldots, \lambda_k; \alpha)$ is approximately equal to $p_1(k; \alpha)$. Furthermore, we may not compute the value of $t(k, \lambda_1, \ldots, \lambda_k; \alpha)$. Therefore, using $p_1(k; \alpha)$, from Theorem 1, we have

$$\lim_{N\to\infty} P\left(\max_{1\le i<i'\le k} |\hat{T}_{ii'}| \le p_1(k;\alpha)\right) \ge A(p_1(k;\alpha)) = 1-\alpha,$$

which implies

$$\lim_{N\to\infty} P\left(|\hat{T}_{ii'}| \le p_1(k;\alpha), 1 \le i < i' \le k\right) \ge 1-\alpha. \tag{4.10}$$

As a conclusion, by using $\mu_i - \mu_{i'} = \tau_i - \tau_{i'}$, from (4.10), we find that

$$\mu_i - \mu_{i'} \in \hat{\tilde{\tau}}_i - \hat{\tilde{\tau}}_{i'} \pm p_1(k;\alpha) \cdot \sqrt{(\hat{\gamma}^2/\hat{d}^2)\cdot(1/n_i + 1/n_{i'})} \quad \text{for} \quad 1 \le i < i' \le k \tag{4.11}$$

forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals for $\{\mu_i - \mu_{i'}; 1 \le i < i' \le k\}$. Similarly asymptotic simultaneous tests of level α for the null hypotheses $\{H_{(ii')} : \mu_i = \mu_{i'}\}_{\{1\le i<i'\le k\}}$ consist in rejecting $H_{(ii')}$ for $1 \le i < i' \le k$ such that $|\hat{T}^*_{ii'}| > p_1(k;\alpha)$. As a non-robust procedure,

$$\mu_i - \mu_{i'} \in \tilde{\tau}_i - \tilde{\tau}_{i'} \pm p_1(k;\alpha) \cdot \sqrt{\tilde{\sigma}^2 \cdot (1/n_i + 1/n_{i'})} \quad \text{for} \quad 1 \le i < i' \le k \tag{4.12}$$

also forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals.

5. Dunnett-type procedures

In order to derive robust Dunnett-type procedures, we get Theorem 2 corresponding to Theorem 1.

THEOREM 2. *Under Assumptions* 1–4, *for any positive* t,

$$\lim_{N\to\infty} P\left(\max_{2\le i\le k} |\hat{T}_{1i}| \le t\right) = \int_{-\infty}^{\infty} \prod_{i=2}^{k} \left\{\Phi\left(\sqrt{\frac{\lambda_i}{\lambda_1}}\cdot x + \sqrt{\frac{\lambda_i+\lambda_1}{\lambda_1}}\cdot t\right) - \Phi\left(\sqrt{\frac{\lambda_i}{\lambda_1}}\cdot x - \sqrt{\frac{\lambda_i+\lambda_1}{\lambda_1}}\cdot t\right)\right\}\varphi(x)dx \tag{5.1}$$

and

$$\lim_{N\to\infty} P\left(\max_{2\le i\le k} \hat{T}_{1i} \le t\right) = \int_{-\infty}^{\infty} \prod_{i=2}^{k} \left\{\Phi\left(\sqrt{\frac{\lambda_i}{\lambda_1}}\cdot x + \sqrt{\frac{\lambda_i+\lambda_1}{\lambda_1}}\cdot t\right)\right\}\varphi(x)dx \tag{5.2}$$

hold.

PROOF. From (2.6), (4.1) and Lemma 1, we have

$$\lim_{N\to\infty} P\left(\max_{2\le i\le k} |\hat{T}_{1i}| \le t\right) = P\left(\max_{2\le i\le k} \frac{|Y_1 - Y_i|}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le t\right),$$

where $Y_1, \ldots, Y_k$ are defined in (2.6). Since

$$\left\{\max_{2\le i\le k} \frac{|Y_1 - Y_i|}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le t\right\} = \left\{Y_1 - t\sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_i}} \le Y_i \le Y_1 + t\sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_i}}, i = 2, \ldots, k\right\}$$

and $Y_1, \ldots, Y_k$ are independent, we get

$$P\left(\max_{2\le i\le k} \frac{|Y_1 - Y_i|}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le t\right) = \int_{-\infty}^{\infty} P\left(y_1 - t\sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_i}} \le Y_i \le y_1 + t\sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_i}}, i = 2, \ldots, k\right) \times \sqrt{\lambda_1}\varphi(\sqrt{\lambda_1}y_1)dy_1.$$

By using the change of variable $x = \sqrt{\lambda_1}y_1$, we get (5.1). Similarly from the equality

$$P\left(\max_{2\le i\le k} \frac{Y_1 - Y_i}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le t\right) = \int_{-\infty}^{\infty} P\left(y_1 - t\sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_i}} \le Y_i, i = 2, \ldots, k\right) \sqrt{\lambda_1}\varphi(\sqrt{\lambda_1}y_1)dy_1,$$

we have (5.2).

We put $G_2(t)$ = (the right hand side of (5.1)) and $G_3(t)$ = (the right hand side of (5.2)). Let $p_2(k, \lambda_1, \ldots, \lambda_k; \alpha)$ and $p_3(k, \lambda_1, \ldots, \lambda_k; \alpha)$ be the upper $100\alpha\%$ points of $G_2(t)$ and $G_3(t)$, respectively. Then from Theorem 2, we have

$$\lim_{N\to\infty} P\left(\max_{2\le i\le k} |\hat{T}_{1i}| \le p_2(k, \lambda_1, \ldots, \lambda_k; \alpha)\right) = 1 - \alpha,$$

which implies

$$\lim_{N\to\infty} P(|\hat{T}_{1i}| \le p_2(k, \lambda_1, \ldots, \lambda_k; \alpha), 2 \le i \le k) = 1 - \alpha. \tag{5.3}$$

As a conclusion, by using $\mu_1 - \mu_i = \tau_1 - \tau_i$, from (5.3), we find that

$$\mu_1 - \mu_i \in \hat{\hat{\tau}}_1 - \hat{\hat{\tau}}_i \pm p_2(k, \lambda_1, \ldots, \lambda_k; \alpha) \cdot \sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot (1/n_1 + 1/n_i)} \quad \text{for} \quad 2 \le i \le k$$

forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals for $\{\mu_1 - \mu_i; 2 \le i \le k\}$. Asymptotic simultaneous tests of level α for the null hypotheses $\{H_{(1i)} : \mu_1 = \mu_i\}_{\{2\le i\le k\}}$ consist in rejecting $H_{(1i)}$ for $2 \le i \le k$ such that $|\hat{T}^*_{1i}| > p_2(k, \lambda_1, \ldots, \lambda_k; \alpha)$.

$$\mu_1 - \mu_i \in \tilde{\tau}_1 - \tilde{\tau}_i \pm p_2(k, \lambda_1, \ldots, \lambda_k; \alpha) \cdot \sqrt{\tilde{\sigma}^2 \cdot (1/n_1 + 1/n_i)} \quad \text{for} \quad 2 \le i \le k$$

also forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals.

Next for the simplicity of notation, we put $p_3 = p_3(k, \lambda_1, \ldots, \lambda_k; \alpha)$. Then we have

$$\lim_{N\to\infty} P(\hat{T}_{1i} \ge -p_3, 2 \le i \le k) = \lim_{N\to\infty} P\left(\min_{2\le i\le k} \hat{T}_{1i} \ge -p_3\right). \tag{5.4}$$

For $Y_1, \ldots, Y_k$ defined in (2.6), $(Y_1, \ldots, Y_k)$ and $(-Y_1, \ldots, -Y_k)$ have the same normal distribution. Hence we get

$$\begin{aligned}
\lim_{N\to\infty} P\left(\min_{2\le i\le k} \hat{T}_{1i} \ge -p_3\right) &= P\left(\min_{2\le i\le k} \frac{Y_1 - Y_i}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \ge -p_3\right) \\
&= P\left(\max_{2\le i\le k} \frac{-Y_1 + Y_i}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le p_3\right) \\
&= P\left(\max_{2\le i\le k} \frac{Y_1 - Y_i}{\sqrt{\dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_i}}} \le p_3\right) \\
&= \lim_{N\to\infty} P\left(\max_{2\le i\le k} \hat{T}_{1i} \le p_3\right) \\
&= 1 - \alpha.
\end{aligned}$$

Combining this fact with (5.4), we find that, under the one-sided restriction $\{\mu_1 \le \mu_2, \ldots, \mu_k\}$,

$$\mu_1 - \mu_i < \hat{\hat{\tau}}_1 - \hat{\hat{\tau}}_i + p_3(k, \lambda_1, \ldots, \lambda_k; \alpha) \cdot \sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot (1/n_1 + 1/n_i)} \quad \text{for} \quad 2 \le i \le k$$

forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals for $\{\mu_1-\mu_i; 2 \le i \le k\}$. Asymptotic simultaneous tests of level α for the null hypotheses $\{H_{(1i)} : \mu_1 = \mu_i\}_{\{2\le i\le k\}}$ consist in rejecting $H_{(1i)}$ for $2 \le i \le k$ such that $\hat{T}^*_{1i} < -p_3(k, \lambda_1, \ldots, \lambda_k; \alpha)$.

$$\mu_1 - \mu_i < \tilde{\tau}_1 - \tilde{\tau}_i + p_3(k, \lambda_1, \ldots, \lambda_k; \alpha) \cdot \sqrt{\tilde{\sigma}^2 \cdot (1/n_1 + 1/n_i)} \quad \text{for} \quad 2 \le i \le k$$

also forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals.

6. Scheffé-type procedures

It is not simple to construct multiple comparisons procedures based on rank statistics of (2.1) and (3.1). We shall propose robust procedures based on the R-estimators $\hat{\boldsymbol{\tau}}$ and $\breve{\boldsymbol{\tau}}$ for the family of $\{\sum_{i=1}^k c_i\mu_i\}_{\boldsymbol{c}\in\mathcal{C}}$.

For $\boldsymbol{c} \in \mathcal{C}^k$, let

$$\hat{T}_{\boldsymbol{c}} = \frac{\sum_{i=1}^k c_i(\hat{\mathring{\tau}}_i - \tau_i)}{\sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot \sum_{i=1}^k c_i^2/n_i}}, \qquad \hat{T}^*_{\boldsymbol{c}} = \frac{\sum_{i=1}^k c_i\hat{\mathring{\tau}}_i}{\sqrt{(\hat{\gamma}^2/\hat{d}^2) \cdot \sum_{i=1}^k c_i^2/n_i}},$$

and

$$KW = \frac{\sum_{i=1}^k n_i(\hat{\mathring{\tau}}_i - \tau_i)^2}{\hat{\gamma}^2/\hat{d}^2}.$$

From the Cauchy-Schwarz inequality, we find

$$\hat{T}^2_{\boldsymbol{c}} \le \sup_{\boldsymbol{c}\in\mathcal{C}^k} \hat{T}^2_{\boldsymbol{c}} = KW. \tag{6.1}$$

Hence we get

THEOREM 3. *Suppose that Assumptions* 1–4 *are satisfied. Then as N tends to infinity, KW has asymptotically a* χ^2*-distribution with* $(k-1)$ *degrees of freedom.*

PROOF. From (2.6), (4.1) and Lemma 1, we have

$$KW \xrightarrow{\text{L}} T = \sum_{i=1}^k \lambda_i \left(Y_i - \sum_{j=1}^k \lambda_j Y_j \right)^2.$$

From Theorem 2.4.1.1 of Hájek *et al.* (1999), T has a χ^2-distribution with $(k-1)$ degrees of freedom. Hence we get the conclusion.

Let $\chi^2_{k-1,\alpha}$ be the upper $100\alpha\%$ point of a χ^2-distribution with $(k-1)$ degrees of freedom. Then from (6.1) and Theorem 3, we have

$$\lim_{N\to\infty} P(\hat{T}^2_{\boldsymbol{c}} \le \chi^2_{k-1,\alpha}, \boldsymbol{c} \in \mathcal{C}^k) = \lim_{N\to\infty} P(KW \le \chi^2_{k-1,\alpha}) = 1-\alpha. \tag{6.2}$$

Hence, by using $\sum_{i=1}^k c_i\mu_i = \sum_{i=1}^k c_i\tau_i$, (6.2) implies that

$$\sum_{i=1}^k c_i\mu_i \in \sum_{i=1}^k c_i\hat{\mathring{\tau}}_i \pm \sqrt{\chi^2_{k-1,\alpha} \cdot (\hat{\gamma}^2/\hat{d}^2) \cdot \sum_{i=1}^k c_i^2/n_i}$$

forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals for $\sum_{i=1}^{k} c_i\mu_i$; $\boldsymbol{c} \in \mathcal{C}^k$. Asymptotic simultaneous tests of level α for the null hypotheses $\{H_{\boldsymbol{c}} : \sum_{i=1}^{k} c_i\mu_i = 0 \text{ for } \boldsymbol{c} \in \mathcal{C}^k\}$ consist in rejecting $H_{\boldsymbol{c}}$ for $\boldsymbol{c} \in \mathcal{C}^k$ such that $|\hat{T}_{\boldsymbol{c}}^{*2}| > \chi^2_{k-1,\alpha}$. As Scheffé's method,

$$\sum_{i=1}^{k} c_i\mu_i \in \sum_{i=1}^{k} c_i\tilde{\tau}_i \pm \sqrt{\chi^2_{k-1,\alpha} \cdot \tilde{\sigma}^2 \cdot \sum_{i=1}^{k} c_i^2/n_i}$$

also forms a set of asymptotic $100(1-\alpha)\%$ simultaneous confidence intervals.

7. Efficiency and simulation

Proposition 3 implies that the asymptotic relative efficiency (ARE) of the joint R-estimator $\hat{\boldsymbol{\tau}}$ with respect to the pairwise R-estimator $\breve{\boldsymbol{\tau}}$ is 1. Let $MS(\hat{\boldsymbol{\tau}})$ be the mean squared error of the estimator $\hat{\boldsymbol{\tau}}$, that is, $MS(\hat{\boldsymbol{\tau}}) = E\{(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})'(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau})\}$. Then we define the relative risk efficiency (RRE) of $\hat{\boldsymbol{\tau}}$ with respect to $\breve{\boldsymbol{\tau}}$ by $RRE(\hat{\boldsymbol{\tau}}, \breve{\boldsymbol{\tau}}) = MS(\breve{\boldsymbol{\tau}})/MS(\hat{\boldsymbol{\tau}})$. Under certain regularity conditions,

$$\lim_{n\to\infty} RRE(\hat{\boldsymbol{\tau}}, \breve{\boldsymbol{\tau}}) = ARE(\hat{\boldsymbol{\tau}}, \breve{\boldsymbol{\tau}})$$

holds. We simulate $RRE(\hat{\boldsymbol{\tau}}, \breve{\boldsymbol{\tau}})$. We limited attention to $k = 3$, $n_1 = n_2 = n_3 = 10, 20, 30$ and $F(x) = N(0,1)$, logistic $LG(0, \sqrt{3}/\pi)$ with density function $\exp(-\pi x/\sqrt{3})/\{1+\exp(-\pi x/\sqrt{3})\}^2$, and double exponential $DE(0, 1/\sqrt{2})$ with density function $(1/\sqrt{2})\exp(-\sqrt{2}|x|)$. The Wilcoxon-type score functions were taken, that is, $a_N(\ell) = 2\ell/(N+1) - 1$ and $a_{N_{ii'}}(\ell) = 2\ell/(N_{ii'}+1) - 1$. The

Table 2. The ARE when $F(x)$ =normal $N(0,1)$, logistic $LG(0, \sqrt{3}/\pi)$, and double exponential $DE(0, 1/\sqrt{2})$.

$F(x)$	$N(0,1)$	$LG(0, \sqrt{3}/\pi)$	$DE(0, 1/\sqrt{2})$
ARE	0.955	1.097	1.5

Table 3. The ARE when the underlying distribution is an ϵ-contaminated distribution.

ϵ	0.01	0.02	0.03	0.04	0.05
ARE	1.009	1.060	1.108	1.153	1.196

ϵ	0.06	0.07	0.08	0.09	0.10
ARE	1.236	1.274	1.309	1.342	1.373

Table 4. The ARE when the underlying distribution is a t-distribution with m degrees of freedom.

m	3	4	5	6	7	8	9	10	11
ARE	1.900	1.401	1.241	1.164	1.119	1.089	1.069	1.054	1.042

m	12	13	14	15	16	17	18	19	20
ARE	1.033	1.025	1.019	1.014	1.009	1.006	1.002	0.999	0.997

values of the RRE were estimated by Monte-Carlo simulation of 5,000 samples. Then rounding the simulated RRE off to the two decimal places, it becomes $RRE(\hat{\boldsymbol{\tau}}, \breve{\boldsymbol{\tau}}) = 1$.

If we take the ratio of the squares of the width of the confidence intervals as the asymptotic efficiency, the asymptotic relative efficiency of the proposed Tukey-type simultaneous confidence intervals of (4.11) with respect to the Tukey-Kramer method of (4.12) is given by

$$\frac{\sigma^2[\int_0^1\{\psi(u)\cdot f'(F^{-1}(u))/f(F^{-1}(u))\}du]^2}{\int_0^1\{\psi(u)-\bar{\psi}\}^2du}, \tag{7.1}$$

which is equivalent to the well-known ARE result of the two-sample rank test with respect to the t-test. The ARE of the proposed Tukey-type simultaneous tests based on $\{|\hat{T}^*_{ii'}| : 1 \le i < i' \le k\}$ with respect to the Tukey-Kramer method based on $\{|\tilde{T}^*_{ii'}| : 1 \le i < i' \le k\}$ is also given by (7.1). The ARE results of the other proposed multiple comparisons procedures relative to the normal theory multiple comparisons procedures remain the same in this case too. The values of the ARE, when the Wilcoxon-type score functions are taken, appear in Tables 2–4.

Lemma 1 implies that the asymptotic procedures do not depend on b. However we must decide the value of b. Hence a simulation study for the goodness of $\hat{d}$ estimating d is done when $\hat{\hat{\boldsymbol{\tau}}} = \hat{\boldsymbol{\tau}}$ and $a_{N_{ii'}}(\ell) = 2\ell/(N_{ii'}+1) - 1$.

The underlying distributions $F(x)$ chosen here are normal ($N(0,1)$), logistic distribution, and double exponential. We simulate the mean squared error of $\hat{d}$ (MSE) given by $E\{(\hat{d}-d)^2\}$ in Table 5 for $k=3$, $n_1 = n_2 = n_3 = 15, 30$ and $b = 1(1)10$. The values of the MSE are estimated by Monte-Carlo simulation from 10,000 samples. From Table 5, we may decide $b = 6$ as the best choice.

Table 5. The simulated mean squared error of $\hat{d}$.

(i) $F(x) = N(0,1)$

$n_1 \setminus b$	1	2	3	4	5	6	7	8	9	10
15	.0111	.0070	.0056	.0046	.0041	.0040	.0042	.0052	.0068	.0089
30	.0044	.0031	.0026	.0024	.0022	.0022	.0022	.0023	.0027	.0033

(ii) $F(x) = LG(0, \sqrt{3}/\pi)$

$n_1 \setminus b$	1	2	3	4	5	6	7	8	9	10
15	.0133	.0090	.0070	.0063	.0056	.0057	.0068	.0086	.0114	.0150
30	.0053	.0043	.0038	.0033	.0032	.0031	.0032	.0036	.0044	.0054

(iii) $F(x) = DE(0, 1/\sqrt{2})$

$n_1 \setminus b$	1	2	3	4	5	6	7	8	9	10
15	.0212	.0152	.0123	.0119	.0130	.0155	.0206	.0270	.0350	.0442
30	.0095	.0077	.0069	.0066	.0068	.0077	.0092	.0118	.0152	.0190

Table 6. The simulated values of $P(\max_{1\le i<i'\le k}|\hat{T}_{ii'}|\le p_1(k,\alpha))$.

(i) $\alpha = 0.05$

$n_1 \setminus F(x)$	$N(0,1)$	$LG(0,\sqrt{3}/\pi)$	$DE(0,1/\sqrt{2})$
15	.9576	.9615	.9746
30	.9530	.9523	.9629

(ii) $\alpha = 0.01$

$n_1 \setminus F(x)$	$N(0,1)$	$LG(0,\sqrt{3}/\pi)$	$DE(0,1/\sqrt{2})$
15	.9897	.9925	.9956
30	.9895	.9909	.9934

Table 7. The simulated values of $P(\max_{2\le i\le k}|\hat{T}_{1i}|\le p_2(k,\lambda_1,\ldots,\lambda_k;\alpha))$.

(i) $\alpha = 0.05$

$n_1 \setminus F(x)$	$N(0,1)$	$LG(0,\sqrt{3}/\pi)$	$DE(0,1/\sqrt{2})$
15	.9595	.9610	.9737
30	.9525	.9532	.9633

(ii) $\alpha = 0.01$

$n_1 \setminus F(x)$	$N(0,1)$	$LG(0,\sqrt{3}/\pi)$	$DE(0,1/\sqrt{2})$
15	.9907	.9923	.9954
30	.9893	.9907	.9933

Hence we set $b = 6$ and, under the same settings, we investigate the accuracy of asymptotic approximation for the coverage probabilities

$$P\left(\max_{1\le i<i'\le k}|\hat{T}_{ii'}|\le p_1(k,\alpha)\right) \quad \text{and} \quad P\left(\max_{2\le i\le k}|\hat{T}_{1i}|\le p_2(k,\lambda_1,\ldots,\lambda_k;\alpha)\right)$$

for $\alpha = 0.05$, 0.01 in Tables 6 and 7. The values of the coverage probabilities are estimated by Monte-Carlo simulation from 10,000 samples. The values are nearly equal to or larger than $1-\alpha$. Therefore when $b = 6$, the proposed procedures are approximately conservative.

Acknowledgements

The author would like to thank the referees for their valuable comments.

References

Dunn, O. J. (1964). Multiple comparisons using rank sums, *Technometrics*, **6**, 241–252.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control, *J. Amer. Statist. Assoc.*, **50**, 1096–1121.

Dwass, M. (1960). *Some k-sample Rank Order Tests*, *Contributions to Probability and Statistics*, Stanford University Press, 198–202.

Hájek, J., Šidák, Z. and Sen, P. K. (1999). *Theory of Rank Tests*, *2nd Edition*, Academic Press.

Hayter, A. J. (1984). A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative, *Ann. Statist.*, **12**, 61–75.

Hodges, J. L., Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests, *Ann. Math. Statist.*, **34**, 598–611.

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

Kramer, C. Y. (1956). Extension of multiple range tests to goup means with unequal numbers of replications, *Biometrics*, **8**, 75–86.

Oude Voshaar, J. H. (1980). $(k-1)$-mean significance levels of nonparametric multiple comparisons procedures, *Ann. Statist.*, **8**, 75–86.

Scheffé, H. (1953). A method for judging all contrasts in analysis of variance, *Biometrika*, **40**, 87–104.

Sen, P. K. (1966). On nonparametric simultaneous confidence regions and tests for the one criterion analysis of variance problem, *Ann. Inst. Statist. Math.*, **18**, 319–336.

Sen, P. K. (1980). *Nonparametric Simultaneous Inference for Some MANOVA Models. Handbook of Statistics*, North-Holland, 673–702.

Shiraishi, T. (1990). R-estimators and confidence regions in one-way MANOVA, *J. Statist. Plan. Infer.*, **24**, 203–214.

Steel, R. G. D. (1959). A multiple comparison rank sum test: Treatments versus control, *Biometrics*, **15**, 560–572.

Steel, R. G. D. (1960). A rank sum test for comparing all pairs of treatments, *Technometrics*, **2**, 197–207.

Tukey, J. W. (1953). *The Problem of Multiple Comparisons*, Mimeographed monograph.