5. One-way Layout

Data and problems

- The data and problems in one-way layout are similar to those in the one-way analysis of variance (ANOVA), but the techniques are nonparametric instead of relying on normal distributions as in ANOVA.
- The data consist of k samples $X_{1j},\ldots,X_{n_j j}$, $j=1,\ldots,k$, from k populations. The total number of observations is $N = n_1 + n_2 + \cdots + n_k$.
- The main problem is to test differences between the k populations.

Assumption 5.1

(i) The N random variables $X_{1j},\ldots,X_{n_j j}$, $j=1,\ldots,k$, are independent.

(ii) For each $j \in \{1,\ldots,k\}$, $X_{1j},\ldots,X_{n_j j}$ are identically distributed with cdf F_j.

(iii) $F_j(t) = F(t-\tau_j)$, $t \in \mathbb{R}$, where F is a continuous cdf with unknown median θ and τ_j is the unknown *treatment effect* for population j, $j=1,\ldots,k$.

If F is normal, Assumption 5.1 is equivalent to the one-way ANOVA model:

$$X_{ij} = \theta + \tau_j + e_{ij} \text{ with i.i.d. } e_{ij} \sim N(0, \sigma^2),\ i = 1, \ldots, n_j,\ j = 1, \ldots, k.$$

Null hypothesis: $H_0 : \tau_1 = \cdots = \tau_k$ (there are no differences among the treatment effects of the k populations).

Alternative hypothesis: We consider three types of alternative hypotheses:

1. General alternatives: $H_1 : \tau_1, \ldots, \tau_k$ are not all equal.
2. Ordered alternatives: $H_1 : \tau_1 \le \cdots \le \tau_k$, with $\tau_j \ne \tau_{j'}$ for at least one pair of treatment effects $j, j' \in \{1, \ldots, k\}$.
3. Umbrella alternatives: $H_1 : \tau_1 \le \cdots \le \tau_{p-1} \le \tau_p \ge \tau_{p+1} \ge \cdots \ge \tau_k$, with $\tau_j \ne \tau_{j'}$ for at least one pair of treatment effects $j, j' \in \{1, \ldots, k\}$.

Alternatives 2 and 3 can help to identify a trend, such as increasing and turning.

Alternative 3 is also suitable when an umbrella order is natural, but a monotone order is difficult, such as environmental temperature from Jan to Dec in a year.

5.1 Tests of equal treatment effects

Kruskal-Wallis test for general alternatives

Hypotheses: $H_0 : \tau_1 = \cdots = \tau_k$ against $H_1 : \tau_1, \ldots, \tau_k$ are not all equal.

Test statistic: Order N observations $\{X_{ij}, i = 1, \ldots, n_j, j = 1, \ldots, k\}$ in ascending order. Let r_{ij} denote the rank of X_{ij}. Define

$$R_j = \sum_{i=1}^{n_j} r_{ij} \quad \text{and} \quad R_{\cdot j} = \frac{R_j}{n_j} = \frac{1}{n_j} \sum_{i=1}^{n_j} r_{ij}, \quad j = 1, \ldots, k. \tag{5.1}$$

The test statistic H of the *Kruskal-Wallis test* is defined by

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} n_j \left(R_{\cdot j} - \frac{N+1}{2} \right)^2 = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1), \tag{5.2}$$

where

$$\frac{N+1}{2} = \frac{1}{N} \sum_{i=1}^{N} i = \frac{1}{N} \sum_{j=1}^{k} \sum_{i=1}^{n_j} r_{ij} = \frac{1}{N} \sum_{j=1}^{k} n_j R_{\cdot j} = \text{Average rank of all } X_{ij}\text{'s}$$

Exact distribution of H

Assume no ties among $\{X_{ij}\}$. Then each arrangement of $\{r_{ij}\}$ is a permutation of $\{1,2,\ldots,N\}$, hence there are totally $N!$ ways to arrange $\{r_{ij}\}$. But as the order of the ranks $r_{1j}, r_{2j}, \ldots, r_{n_j,j}$ within each treatment j does not affect R_j and H, we can fix the order as $r_{1j} < r_{2j} < \cdots < r_{n_j,j}$, $j = 1,\ldots,k$, to determine the distribution of H. Under such restrictions, the total number of ways to allocate n_j of N items to treatment j, $j = 1,\ldots,k$, is the multinomial number

$$\binom{N}{n_1,\ldots,n_k} = \frac{N!}{n_1!\cdots n_k!}$$

Each way is equally likely under H_0 and determines a vector $(R_1,\ldots,R_k)$. Thus the exact distribution of H under H_0 is given by

$$\Pr(H = h) = \frac{\text{No. of } (R_1,\ldots,R_k) : H = h}{\binom{N}{n_1,\ldots,n_k}} \qquad (5.3)$$

Example 5.1 A table on page 208 of the textbook shows how to work out the distribution of H with $k=3$ and $n_1=n_2=n_3=2$, hence $N=6$. In this case,

$$\frac{N!}{n_1!n_2!n_3!}=\frac{6!}{2!2!2!}=\frac{720}{8}=90$$

and

$$H=\frac{12}{N(N+1)}\left(\frac{R_1^2}{n_1}+\frac{R_2^2}{n_2}+\frac{R_3^2}{n_3}\right)-3(N+1)=\frac{12}{6(7)}\cdot\frac{A}{2}-3(7)=\frac{A}{7}-21,$$

where $A=R_1^2+R_2^2+R_3^2$ is invariant with the order of (R_1,R_2,R_3). Hence we only need to count $90/3!=15$ of the total 90 rank assignments. For example, the next 6 assignments for treatments I, II, III give the same value of A:

I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
1	3	5	3	1	5	5	3	1	3	5	1	1	5	3	5	1	3
2	4	6	4	2	6	6	4	2	4	6	2	2	6	4	6	2	4

with $A=(1+2)^2+(3+4)^2+(5+6)^2=9+49+121=179$.

In the 15 rank assignments (a)-(o) below, A and H are calculated:

(a)

I	II	III
1	3	5
2	4	6

$A = 179$
$H = 179/7 - 21 = 4.57$

(b)

I	II	III
1	3	4
2	5	6

$A = 3^2 + 8^2 + 10^2 = 173$
$H = 173/7 - 21 = 3.71$

(c)

I	II	III
1	3	4
2	6	5

$A = 3^2 + 9^2 + 9^2 = 171$
$H = 173/7 - 21 = 3.43$

(d)

I	II	III
1	2	5
3	4	6

$A = 173,\ H = 3.71$

(e)

I	II	III
1	2	4
3	5	6

$A = 165,\ H = 2.67$

(f)

I	II	III
1	2	4
3	6	5

$A = 161,\ H = 2$

(g)

I	II	III
1	2	3
4	5	6

$A = 155,\ H = 1.14$

(h)

I	II	III
1	2	5
4	3	6

$A = 171,\ H = 3.43$

(i)

I	II	III
1	2	3
4	6	5

$A = 153,\ H = 0.86$

(j)

I	II	III
1	2	4
5	3	6

$A = 161,\ H = 2$

(k)

I	II	III
1	2	3
5	4	6

$A = 153,\ H = 0.86$

(l)

I	II	III
1	2	3
5	6	5

$A = 149,\ H = 0.29$

(m)

I	II	III
1	2	4
6	3	5

$A = 155,\ H = 1.14$

(n)

I	II	III
1	2	3
6	4	5

$A = 149,\ H = 0.29$

(o)

I	II	III
1	2	4
6	5	3

$A = 147,\ H = 0$

Then the distribution of H is obtained from the values of H in all 15 assignments together with (5.3), as tabled below:

h	0	0.29	0.86	1.14	2	2.57	3.43	3.71	4.57
$\Pr(H = h)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

Asymptotic distribution of H: Under H_0,

$$H \sim \chi^2_{k-1} \text{ approximately if } n_1, \ldots, n_k \text{ are large.}$$

Rejection rule: Reject H_0 at level α if $H \geq h_\alpha$, where h_α is a value of H such that $\Pr(H \geq h_\alpha) = \alpha$, which can be found from the exact distribution of H.

Approximate rejection rule: Reject H_0 at level α if $H \geq \chi^2_{k-1,\alpha}$.

Ties: If there are ties among X_{ij}'s, assign the average rank to each tied group. In this case, the test statistic H is adjusted to

$$H' = \frac{H}{1-A} \quad \text{with} \quad A = \frac{1}{N^3 - N}\sum_{j=1}^{g}\left(t_j^3 - t_j\right), \tag{5.4}$$

where g is the number of tied groups and t_j is the size of group j, $j = 1, \ldots, k$.

After the adjustment in (5.4), the above rejection rules are valid approximately for large samples. The exact rule can also be worked out by (5.3) in the same way as in the case with no ties, and the adjustment in (5.4) is not needed.

Example 5.2 Example 6.1 of the textbook (page 205) present the data (rank):

X_{i1} : 2.9 (8) 3.0 (9) 2.5 (4) 2.6 (5) 3.2 (10) $\quad R_1 = 8+9+4+5+10 = 36$

X_{i2} : 3.8 (13) 2.7 (6) 4.0 (14) 2.4 (3) $\quad R_2 = 13+6+14+3 = 36$

X_{i3} : 2.8 (7) 3.4 (11) 3.7 (12) 2.2 (2) 2.0 (1) $\quad R_3 = 7+11+12+2+1 = 33$

In this example, $k=3$, $n_1=5$, $n_2=4$, $n_3=5$ and so $N=5+4+5=14$.

By (5.2), we can calculate

$$H = \frac{12}{14(14+1)}\left(\frac{36^2}{5}+\frac{36^2}{4}+\frac{33^2}{5}\right) - 3(14+1) = 0.771$$

By R, $h_{0.6871} = 0.783$. Hence $H = 0.771 < 0.783 \Rightarrow H_0$ can be accepted at a very high level over 68%. The approximate p-value is

$$\Pr(H \geq 0.771) \approx \Pr\left(\chi^2_{3-1} \geq 0.771\right) = \Pr\left(\chi^2_2 \geq 0.771\right) = 0.68$$

Both show little evidence against H_0 in favor of H_1. As a result, the test concludes no significant difference between the three treatment effects.

Jonckheere-Terpstra test for ordered alternatives

Hypotheses: $H_0 : \tau_1 = \cdots = \tau_k$ against $H_1 : \tau_1 \le \cdots \le \tau_k$ and $\tau_i < \tau_j$ for at least one pair $i < j$ ($\tau_1, \ldots, \tau_k$ are not all equal).

Test statistic: The *Jonckheere-Terpstra test* statistic is defined by

$$J = \sum_{u<v} U_{uv} = \sum_{v=2}^{k} \sum_{u=1}^{v-1} U_{uv} \quad \text{with} \quad U_{uv} = \sum_{i=1}^{n_u} \sum_{j=1}^{n_v} I_{\{X_{iu} < X_{jv}\}}, \; 1 \le u < v \le k, \qquad (5.5)$$

where U_{uv} represents the number of pairs (X_{iu}, X_{jv}) such that $X_{iu} < X_{jv}$. If there are no ties, then U_{uv} and J can be found from the joint ranks of X_{ij}'s.

Example 5.3 Consider the case $k = 3$, $n_1 = 2$, $n_2 = 1$, $n_3 = 3$, $N = 6$.

Suppose that the joint ranks of the 6 data points under treatments I, II and III are $(r_{11}, r_{21}) = (2, 4)$, $r_{12} = 5$ and $(r_{13}, r_{23}, r_{33}) = (1, 3, 6)$, respectively.

Then $U_{12} = 2$ $(2 < 5, 4 < 5)$, $U_{13} = 3$ $(2 < 3, 2 < 6, 4 < 6)$ and $U_{23} = 1$ $(5 < 6)$.

Thus $J = U_{12} + U_{13} + U_{23} = 2 + 3 + 1 = 6$.

Mean and variance of J: It follows from significant algebraic manipulations (see Appendix) that under H_0, the mean and variance of J are given by

$$\mathrm{E}_0[J] = \frac{1}{4}\left(N^2 - \sum_{u=1}^{k} n_u^2\right) \tag{5.6}$$

and

$$\mathrm{Var}_0(J) = \frac{1}{72}\left[N^2(2N+3) - \sum_{u=1}^{k} n_u^2(2n_u+3)\right] \tag{5.7}$$

Rejection rule: Reject H_0 at level α and conclude $\tau_1 \le \cdots \le \tau_k$ if $J \ge j_\alpha$, where j_α is a value of J such that $\Pr(J \ge j_\alpha) = \alpha$.

The exact distribution of J can be obtained by counting the number of ways to allocate $n_1, \ldots, n_k$ of N items to treatments $1, \ldots, k$, respectively, such that $J = j$, similar to that of the Kruskal-Wallis test statistic H. The critical point j_α is then determined by the exact distribution of J.

An example is shown on page 220 of the textbook.

Large-sample approximation

Under H_0, when the sample sizes are large,

$$J^* = \frac{J - \mathrm{E}_0[J]}{\sqrt{\mathrm{Var}_0(J)}} \sim N(0,1) \text{ approximately}$$

where $\mathrm{E}_0[J]$ and $\mathrm{Var}_0(J)$ are given by (5.6) and (5.7), respectively.

Approximate rejection rule

Reject H_0 at level α to conclude $\tau_1 \le \cdots \le \tau_k$ if $J^* \ge z_\alpha$.

Ties: If there are ties, modify the U_{uv} in (5.5) to

$$U_{uv} = \sum_{i=1}^{n_u} \sum_{j=1}^{n_v} \left[I_{\{X_{iu} < X_{jv}\}} + \frac{1}{2} I_{\{X_{iu} = X_{jv}\}} \right]$$

That is, assign $1/2$ to each tied pair $X_{iu} = X_{jv}$ in the definitions of U_{uv} and J. In such a case, $\mathrm{Var}_0(J)$ is adjusted to (6.19) of the textbook (page 217), and the above rejection rules remain valid approximately.

Example 5.4 Refer to Example 6.2 of the textbook (from page 217). Table 6.6 on page 218 presents the following data (ranks) of 3 treatments:

1. Control	2. Group B	3. Group C
40 (5.5)	38 (2.5)	48 (18)
35 (1)	40 (5.5)	40 (5.5)
38 (2.5)	47 (17)	45 (15)
43 (10.5)	44 (13)	43 (10.5)
44 (13)	40 (5.5)	46 (16)
41 (8)	42 (9)	44 (13)

For 38 in Group 2: Group 1 has one value $35 < 38$ and one value $38 = 38$. Hence this 38 contributes score 1.5 to U_{12} (1 from 35 and 0.5 from 38 in Group 1). Similarly, the scores contributed to U_{12} by other values in Group 2 are:

2.5 for 40 (by 35, 38, 40 in Group 1), 6 for 47, 5.5 for 44, 2.5 for 40, 4 for 42

Hence $U_{12} = 1.5 + 2.5 + 6 + 5.5 + 2.5 + 4 = 22$. Similarly,

$U_{13} = 6 + 2.5 + 6 + 4.5 + 6 + 5.5 = 30.5$ and $U_{23} = 6 + 2 + 5 + 4 + 5 + 4.5 = 26.5$

It follows that $J = 22 + 30.5 + 26.5 = 79$. By (5.6) with $n_1 = n_2 = n_3 = 6$, $N = 18$,

$$\mathrm{E}_0[J] = \frac{1}{4}\left(18^2 - 3\times 6^2\right) = \frac{324-108}{4} = \frac{216}{4} = 54$$

By formula (6.19) of the textbook with $g = 4$, $t_1 = t_2 = 2$ (38 and 43), $t_3 = 3$ (44), $t_4 = 4$ (40) (ignore groups with $t_j = 1$),

$$\mathrm{Var}_0(J) = \frac{18(17)(41) - 3(6)(5)(17) - 2(2)(1)(9) - 3(2)(11) - 4(3)(13)}{72}$$
$$+ \frac{3(6)(5)(4)[3(2)(1) + 4(3)(2)]}{36(18)(17)(16)} + \frac{3(6)(5)[2(2)(1) + 3(2) + 4(3)]}{8(18)(17)}$$
$$= 150.29$$

Thus

$$J^* = \frac{79-54}{\sqrt{150.29}} = 2.04 \;\Rightarrow\; p\text{-value} \approx \Pr(Z \geq 2.04) = 0.0207$$

This shows strong evidence for $\tau_1 \leq \tau_2 \leq \tau_3$, with at least one strict inequality.

Mack-Wolfe tests for umbrella alternatives

We first consider umbrella alternatives with known peak.

Hypotheses: $H_0: \tau_1 = \cdots = \tau_k$ against $H_1: \tau_1 \le \cdots \le \tau_{p-1} \le \tau_p \ge \tau_{p+1} \ge \cdots \ge \tau_k$ with at least one strict inequality, where p is a known integer.

Test statistic: Let U_{uv} be defined in (5.5). The *Mack-Wolfe test* statistic is

$$A_p = \sum_{v=2}^{p} \sum_{u=1}^{v-1} U_{uv} + \sum_{v=p+1}^{k} \sum_{u=p}^{v-1} U_{vu} = \sum_{u<v\le p} U_{uv} + \sum_{p\le u<v} U_{vu}, \tag{5.8}$$

where $U_{vu} = n_u n_v - U_{uv}$ is the number of pairs $X_{iu} > X_{jv}$, for $u < v$.

Rejection rule: Reject H_0 at an achievable level α if $A_p \ge a_{p,\alpha}$, where $a_{p,\alpha}$ is a value of A_p such that $\Pr(A_p \ge a_{p,\alpha}) = \alpha$; or the p-value $\Pr(A_p \ge a_p) \le \alpha$ with the observed value a_p of A_p for any level α.

The values of $a_{p,\alpha}$ and $\Pr(A_p \ge a_p)$ can be found by enumerating allocations of n_j observations to treatment j, $j = 1, \ldots, k$. An example is given in Comment 25 on page 231 of the textbook with $k = 4$, $n_1 = n_2 = n_4 = 1$, $n_3 = 2$ and $p = 3$.

Mean and variance of A_p: Under H_0 (see Appendix for derivations),

$$\mathrm{E}_0[A_p] = \frac{1}{4}\left(N_1^2 + N_2^2 - \sum_{i=1}^{k} n_i^2 - n_p^2 \right), \tag{5.9}$$

where $N_1 = n_1 + \cdots + n_p$ and $N_2 = n_p + \cdots + n_k = N - N_1 + n_p$, and

$$\mathrm{Var}_0(A_p) = \frac{1}{72}\left[2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^{k} n_i^2(2n_i + 3) - n_p^2(2n_p + 3) \right] + \frac{1}{6}\left(n_p N_1 N_2 - n_p^2 N \right) \tag{5.10}$$

Approximate rejection rule:

Reject H_0 at level α to conclude $H_1 : \tau_1 \le \cdots \le \tau_{p-1} \le \tau_p \ge \tau_{p+1} \ge \cdots \ge \tau_k$ if

$$A_p^* = \frac{A_p - \mathrm{E}_0[A_p]}{\sqrt{\mathrm{Var}_0(A_p)}} \ge z_\alpha$$

where $\mathrm{E}_0[A_p]$ and $\mathrm{Var}_0(A_p)$ are given by (5.9) and (5.10), respectively.

Example 5.5 Refer to Example 6.3 of the textbook (from page 228): Table 6.8 presents the following data of the Fasting Metabolic Rate (FMR) for white-tailed deer in 6 two-month periods (1 – 6) over a year:

1. Jan-Feb	2. Mar-Apr	3. May-Jun	4. Jul-Aug	5. Sep-Oct	6. Nov-Dec
36.0	39.9	44.6	53.8	44.3	31.7
33.6	29.1	54.4	53.9	34.1	22.1
26.9	43.4	48.2	62.5	35.7	30.7
35.8		55.7	46.6	35.6	
30.1		50.0			
31.2					
35.3					

The problem of interest is to test whether the FMR is increasing in temperature. As Jul-Aug is the hottest period, it is reasonable to take $p = 4$.

Since all 7 FMR values in period 1 are smaller than 39.9 and 43.4, and one is less than 29.1, in period 2, $U_{12} = 7+1+7 = 15$. On the other hand, as all 4 FMR values in period 5 are greater than the 3 values in period 6, $U_{65} = 4\times 3 = 12$.

Similarly, $U_{13} = 5 \times 7 = 35$, $U_{14} = 4 \times 7 = 28$, $U_{23} = 5 \times 3 = 15$, $U_{24} = 4 \times 3 = 12$, $U_{34} = 3 + 3 + 5 + 1 = 12$, $U_{54} = 4 \times 4 = 16$ and $U_{64} = U_{65} = 4 \times 3 = 12$.

Then $A_p = A_4$ is calculated by

$$A_4 = \sum_{u<v\le 4} U_{uv} + \sum_{4\le u<v} U_{vu} = U_{12} + U_{13} + U_{14} + U_{23} + U_{24} + U_{34} + U_{54} + U_{64} + U_{65}$$

$$= 15 + 35 + 28 + 15 + 12 + 12 + 16 + 12 + 12 = 157$$

By R, $a_{4,0.0008} = 137$. Hence the exact p-value $= \Pr(A_4 \ge 157) < 0.0008$.

Since $(n_1, \ldots, n_6) = (7,3,5,4,4,3) \;\Rightarrow\; N = 26$, $N_1 = 19$, $N_2 = 11$ and $n_p = n_4 = 4$, by (5.9) – (5.10) we calculate $\mathrm{E}_0[A_4] = 85.5$ and $\mathrm{Var}_0(A_4) = 291.92$. Thus

$$A_4^* = \frac{A_4 - \mathrm{E}_0[A_4]}{\sqrt{\mathrm{Var}_0(A_4)}} = \frac{157 - 85.5}{\sqrt{291.92}} = 4.18 > 3.09 = z_{0.001}$$

Both results show overwhelming evidence for $H_1 : \tau_1 \le \tau_2 \le \tau_3 \le \tau_4 \ge \tau_5 \ge \tau_6$, which support the claim that FMR is increasing with environmental temperature.

Mack-Wolfe test with unknown peak

In umbrella alternatives $H_1 : \tau_1 \le \cdots \le \tau_p \ge \cdots \ge \tau_k$, if the peak p is unknown, we first estimate p and then apply the Mack-Wolfe test based on the estimated p.

Estimation of the peak: Let

$$U_{\bullet q} = \sum_{i \ne q} U_{iq} = \sum_{i<q} U_{iq} + \sum_{i>q} \left(n_i n_q - U_{qi} \right), \quad q = 1, \ldots, k.$$

It is proved in Appendix that

$$\mathrm{E}_0[U_{\bullet q}] = \frac{n_q (N - n_q)}{2} \quad \text{and} \quad \mathrm{Var}_0(U_{\bullet q}) = \frac{n_q (N - n_q)(N+1)}{12} \tag{5.11}$$

Define

$$U_{\bullet q}^{*} = \frac{U_{\bullet q} - \mathrm{E}_0[U_{\bullet q}]}{\sqrt{\mathrm{Var}_0(U_{\bullet q})}}, \quad q = 1, \ldots, k. \tag{5.12}$$

The estimate $\hat{p}$ of the peak p is then determined by the maximum value among $\{U_{\bullet 1}^{*}, \ldots, U_{\bullet k}^{*}\}$ defined in (5.12).

Test statistic: If $U^*_{\cdot\hat{p}}$ is the unique maximum among $\{U^*_{\cdot 1},\ldots,U^*_{\cdot k}\}$, i.e., $U^*_{\cdot\hat{p}} > U^*_{\cdot p}$ for all $p \neq \hat{p}$, then the (standardized) Mack-Wolfe test statistic based on $\hat{p}$ is defined by

$$A^*_{\hat{p}} = \frac{A_p - \mathrm{E}_0[A_p]}{\sqrt{\mathrm{Var}_0(A_p)}} \quad \text{with } p = \hat{p} \tag{5.13}$$

If $\{U^*_{\cdot 1},\ldots,U^*_{\cdot k}\}$ have $r \geq 2$ maximum points $U^*_{\cdot j} \in \{U^*_{\cdot 1},\ldots,U^*_{\cdot k}\}$, $j \in B \subset \{1,\ldots,k\}$, then we take the average of A^*_j in B as the test statistic:

$$A^*_{\hat{p}} = \frac{1}{r}\sum_{j \in B} A^*_j = \frac{1}{r}\sum_{j \in B} \frac{A_j - \mathrm{E}_0[A_j]}{\sqrt{\mathrm{Var}_0(A_j)}} \tag{5.14}$$

In most practical problems, $r = 1$ and hence (5.13) is sufficient.

Rejection rule: Reject H_0 at level α if $A^*_{\hat{p}} \geq a^*_{\hat{p},\alpha}$, where $a^*_{\hat{p},\alpha}$ is a value of $A^*_{\hat{p}}$ such that $\Pr\left(A^*_{\hat{p}} \geq a^*_{\hat{p},\alpha}\right) = \alpha$.

The distribution of $A_{\hat{p}}^*$: The distribution of $A_{\hat{p}}^*$ under H_0, the critical point $a_{\hat{p},\alpha}^*$ and the *p*-value of the test can be obtained by enumeration in a similar way to those of other statistics introduced earlier.

An example for $k=3$, $n_1=1$, $n_2=2$ and $n_3=1$ is provided in Comment 36 from page 246 of the textbook, with $4!/(1!2!1!)=4\cdot 3=12$ possible rank allocations. Some details of calculations are explained in pages 29 – 32 of Appendix.

The result for the distribution of $A_{\hat{p}}^*$ is shown below:

$a_{\hat{p}}^*$	0.361	0.775	1.084	1.549	1.806
$\Pr\left(A_{\hat{p}}^* = a_{\hat{p}}^*\right)$	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$

Note that $\hat{p}$ varies in different rank allocations. Hence if $\hat{p}=3$, say, in a problem, $A_{\hat{p}}^*$ does not have the same distribution as A_3^* with known $p=3$, and $A_{\hat{p}}^*$ is not approximately $N(0,1)$ for the same reason.

Example 5.6 Example 6.4 of the textbook (from page 242) presents the data of Wechsler Adult Intelligence Score (WAIS) in five age groups (G1 – G5):

G1 (16-19)	G2 (20-34)	G3 (35-54)	G4 (55-69)	G5 (≥ 70)
8.62	9.85	9.98	9.12	4.80
9.94	10.43	10.69	9.89	9.18
10.06	11.31	11.40	10.57	9.27

In this example, $k=5$, $n_1=n_2=n_3=n_4=n_5=3$ and $N=15$.

Similar to Example 5.5, it is easy to obtain $U_{12}=7$, $U_{13}=8$, $U_{14}=5$, $U_{15}=2$, $U_{23}=6$, $U_{24}=3$, $U_{25}=0$, $U_{34}=1$, $U_{35}=0$, $U_{45}=2$. Consequently,

$$U_{\bullet 1}=U_{21}+U_{31}+U_{41}+U_{51}=4(3\times 3)-U_{12}-U_{13}-U_{14}-U_{15}=36-22=14$$

$$U_{\bullet 2}=U_{12}+U_{32}+U_{42}+U_{52}=7+3(9)-U_{23}-U_{24}-U_{25}=34-9=25$$

$$U_{\bullet 3}=U_{13}+U_{23}+U_{43}+U_{53}=8+6+2(9)-U_{34}-U_{35}=32-1=31$$

$$U_{\bullet 4}=U_{14}+U_{24}+U_{34}+U_{54}=5+3+1+9-U_{45}=18-2=16$$

$$U_{\bullet 5}=U_{15}+U_{25}+U_{35}+U_{45}=2+0+0+2=4 \quad (\text{as } U_{uv}=n_u n_v-U_{vu} \text{ for } u>v)$$

Since $n_1 = \cdots = n_5$, we have $\mathrm{E}_0[U_{.1}] = \cdots = \mathrm{E}_0[U_{.5}]$ and $\mathrm{Var}_0(U_{.1}) = \cdots = \mathrm{Var}_0(U_{.5})$ by (5.11). Hence $U_{.p}^* = \max\{U_{.1}^*, \ldots, U_{.5}^*\}$ if and only if

$$U_{.p} = \max\{U_{.1}, \ldots, U_{.5}\} = \max\{14, 25, 31, 16, 4\} = 31 = U_{.3} \;\Rightarrow\; \hat{p} = 3 \;\Rightarrow$$

$$A_{\hat{p}} = A_3 = U_{12} + U_{13} + U_{23} + U_{43} + U_{53} + U_{54} = 7 + 8 + 6 + 3(9) - 1 - 0 - 2 = 45$$

By (5.9) – (5.10), $\mathrm{E}_0[A_3] = 27$ and $\mathrm{Var}_0(A_3) = 58.5$. Then (5.12) leads to

$$A_{\hat{p}}^* = \frac{A_3 - \mathrm{E}_0[A_3]}{\sqrt{\mathrm{Var}_0(A_3)}} = \frac{45 - 27}{\sqrt{58.5}} = 2.353 \quad \text{(not approximately } N(0,1))$$

By R, $\Pr(A_{\hat{p}}^* \geq 2.353) = 0.0342$ (note that $\Pr(A_3^* \geq 2.353) = \Pr(A_3 \geq 45) = 0.0086$ and $\Pr(Z \geq 2.353) = 0.0093$ for $Z \sim N(0,1)$).

Hence H_0 is rejected at the 5% level in favor of $H_1: \tau_1 \leq \tau_2 \leq \tau_3 \geq \tau_4 \geq \tau_5$.

This result points to sufficient evidence (at the 5% level) to support the claim that WAIS (a measure of the ability to comprehend ideas and learn) is rising with age up to age group 35-54, then turn to decline with further aging.

Fligner-Wolfe test for treatments versus a control

If there is a *control* (or *baseline*) in a one-way layout, a specific test is designed to assess the difference between other treatments and the control.

Null hypothesis: Let treatment 1 represent the control, and treatments $2,\ldots,k$ the other treatments. Then the null hypothesis is $H_0: \tau_i = \tau_1,\ i = 2,\ldots,k$. This is obviously equivalent to $H_0: \tau_1 = \cdots = \tau_k$.

Alternative hypotheses: We consider two types of one-sided alternatives:

$H_1: \tau_i \geq \tau_1,\ i = 2,\ldots,k,\ \tau_i > \tau_1$ for some $i \in \{2,\ldots,k\}$ (shortened to $H_1: \tau_i > \tau_1$);

$H_2: \tau_i \leq \tau_1,\ i = 2,\ldots,k,\ \tau_i < \tau_1$ for some $i \in \{2,\ldots,k\}$ (shortened to $H_2: \tau_i < \tau_1$).

Test statistic: Let r_{ij} be the rank of X_{ij} in combined samples and $N^* = N - n_1$. The *Fligner-Wolfe* test statistic, its mean and variance under H_0 are given by

$$FW = \sum_{j=2}^{k} \sum_{i=1}^{n_j} r_{ij}, \quad \mathrm{E}_0[FW] = \frac{N^*(N+1)}{2}, \quad \mathrm{Var}_0(FW) = \frac{n_1 N^*(N+1)}{12}$$

Rejection rule: Reject H_0 at level α if

$$FW \geq f_\alpha \text{ against } H_1: \tau_i > \tau_1; \quad \text{or} \quad FW \leq N^*(N+1) - f_\alpha \text{ against } H_2: \tau_i < \tau_1,$$

where f_α is a value of FW such that $\Pr(FW \geq f_\alpha) = \alpha$.

The distribution of FW: Under H_0, FW has the same distribution as that of the two-sample Wilcoxon rank sum statistic W, with $X = \{X_{i1}\}$ of size $m = n_1$ and $Y = \{X_{ij}, j \geq 2\}$ of size $n = N^*$.

Approximate rejection rule: Reject H_0 at level α if

$$FW^* \geq z_\alpha \text{ against } H_1: \tau_i > \tau_1; \quad \text{or} \quad FW^* \leq -z_\alpha \text{ against } H_2: \tau_i < \tau_1,$$

where

$$FW^* = \frac{FW - \mathrm{E}_0[FW]}{\sqrt{\mathrm{Var}_0(FW)}} = \frac{FW - N^*(N+1)/2}{[n_1 N^*(N+1)/12]^{1/2}}$$

Ties: If there are ties among X_{ij}'s, we can assign average ranks to tied values. In that case, $\mathrm{Var}_0(FW)$ is reduced to (6.58) of the textbook (page 251).

5.2 Multiple comparisons

In a one-way layout problem, if $H_0: \tau_1 = \cdots = \tau_k$ is rejected, so we conclude that the k treatments do not have the same effect, then the next question is to identify the differences among $\tau_1, \ldots, \tau_k$, such as $\tau_1 = \tau_2$ or $\tau_1 > \tau_2$?

The idea of multiple comparison is to test pairwise differences between τ_i and τ_j jointly for all $k(k-1)/2$ pairs (i, j) with $i < j \in \{1, \ldots, k\}$.

Two-sided multiple comparisons

For each pair (i, j), $i < j \in \{1, \ldots, k\}$, let $R_{1j}, \ldots, R_{n_j j}$ be the ranks of $X_{1j}, \ldots, X_{n_j j}$ among combined data $X_{1i}, \ldots, X_{n_i i}, X_{1j}, \ldots, X_{n_j j}$ in samples i and j. Define

$$W_{ij} = \sum_{b=1}^{n_j} R_{bj} \quad \text{and} \quad W_{ij}^* = \frac{W_{ij} - \mathrm{E}_0[W_{ij}]}{\sqrt{\mathrm{Var}_0(W_{ij})/2}} = \frac{W_{ij} - n_j(n_i + n_j + 1)/2}{\sqrt{n_i n_j (n_i + n_j + 1)/24}} \tag{5.15}$$

Let w_α^* satisfy

$$\Pr\left(\left|W_{uv}^*\right| < w_\alpha^*, 1 \le u < v \le k\right) = 1 - \alpha \quad \text{under } H_0: \tau_1 = \cdots = \tau_k. \tag{5.16}$$

The Steel-Dwass-Critchlow-Fligner (SDCF) two-sided all-treatment multiple comparison procedure is defined as follows. For each pair (u, v) with $u < v$:

$$\text{Decide } \tau_u \neq \tau_v \text{ if } \left|W_{uv}^*\right| \geq w_\alpha^*; \quad \text{otherwise accept } \tau_u = \tau_v. \tag{5.17}$$

It ensures that when H_0 is true, the probability of any errors is no more than α.

Comment 55 on page 261 of the textbook discusses how to find w_α^* and presents an example with $k = 3$ and $n_1 = n_2 = n_3 = 2$. It can also be found by R program.

Large-sample approximation

If $\min\{n_1, \ldots, n_k\}$ is large, w_α^* can be approximated by q_α that satisfies

$$\Pr\left(\max\{Z_1, \ldots, Z_k\} - \min\{Z_1, \ldots, Z_k\} \geq q_\alpha\right) = \alpha \tag{5.18}$$

with independent $Z_1, \ldots, Z_k \sim N(0,1)$. q_α can be found by R program.

Ties: If there are ties among X_{ij}'s, assign average ranks to tied values among two samples, and adjust $\text{Var}_0(W_{ij})$ to (6.65) of the textbook (page 257).

Example 5.7 Table 6.3 on page 213 of the textbook lists the following data:

Site I	Site II	Site III	Site IV
46 (15.5)	42 (9.5)	38	31
28 (2)	60 (20)	33	30
46 (15.5)	32 (3.5)	26	27
37 (5)	42 (9.5)	25	29
32 (3.5)	45 (13.5)	28	30
41 (7)	58 (19)	28	25
42 (9.5)	27 (1)	26	25
45 (13.5)	51 (17)	27	24
38 (6)	42 (9.5)	27	27
44 (12)	52 (18)	27	30

To explain the calculations in Example 6.6 of the textbook (from page 257), the ranks for combined samples 1 and 2 are provided in brackets. By (5.15),

$$W_{12} = 9.5 + 20 + 3.5 + 9.5 + 13.5 + 19 + 1 + 17 + 9.5 + 18 = 120.5$$

(The sum of ranks of sample 2 in combined samples 1 and 2).

Similarly, $W_{13} = 61.5$, $W_{14} = 60$, $W_{23} = 62.5$, $W_{24} = 61$, $W_{34} = 105$.

Next, $n_1 = n_2 = n_3 = n_4 = 10 \Rightarrow \mathrm{E}_0[W_{12}] = 10(10+10+1)/2 = 105$ and by (6.65) of the textbook with $g_{12} = 14$, $t_b = 2$ for $b = 3,9,10$, $t_7 = 4$ and $t_b = 1$ for other b,

$$\frac{1}{2}\mathrm{Var}_0(W_{12}) = \frac{10(10)}{24}\left[21 - \frac{3(1)(2)(3)+(3)(4)(5)}{20(20-1)}\right] = 86.64 \Rightarrow$$

$$W_{12}^* = \frac{W_{12} - \mathrm{E}_0[W_{12}]}{\sqrt{\mathrm{Var}_0(W_{12})/2}} = \frac{120.5 - 105}{\sqrt{86.64}} = 1.67$$

Similarly, $W_{13}^* = -4.67$, $W_{14}^* = -4.82$, $W_{23}^* = -4.57$, $W_{24}^* = -4.73$, $W_{34}^* = 0$.

By R, $w_{0.0076}^* = 4.276$. $|W_{12}^*| = 1.67 < 4.276 \Rightarrow \tau_1 = \tau_2$, $|W_{34}^*| = 0 \Rightarrow \tau_3 = \tau_4$,

$|W_{13}^*|$, $|W_{14}^*|$, $|W_{23}^*|$, $|W_{24}^*| > 4.276 \Rightarrow \tau_1 \neq \tau_3$, $\tau_1 \neq \tau_4$, $\tau_2 \neq \tau_3$, $\tau_2 \neq \tau_4$.

Thus we can conclude $(\tau_1 = \tau_2) \neq (\tau_3 = \tau_4)$ at $\alpha = 0.0076$.

The large-sample approximation of $w_{0.01}^*$ is $q_{0.01} = 4.404$ by R, which produces the same conclusions $(\tau_1 = \tau_2) \neq (\tau_3 = \tau_4)$ at $\alpha = 0.01$.

More details can be found in Example 6.6 of the textbook (from page 257).

Interpretation of multiple comparisons

In multiple comparisons, the choice is between the null hypothesis $H_0 : \tau_1 = \cdots = \tau_k$ and the alternatives H_1 specified by the decision rules in (5.17), assuming H_1 includes $\tau_u \neq \tau_v$ for at least one pair (u, v). By (5.16), if H_0 is true, then

$$\Pr\left(\left|W_{uv}^*\right| \geq w_\alpha^* \text{ for at least one pair } (u,v)\right) = \alpha$$

Thus the probability to conclude H_1 in (5.17) under H_0 is α. It is in this sense we may interpret the results in (5.17) as being supported by evidence of level α. This does not mean that the probability of any error in H_1 is α in general because other scenarios of $\tau_1, \ldots, \tau_k$ than H_0 and H_1 are ignored.

For example, in the case of Example 5.7, the decision is between $H_0 : \tau_1 = \cdots = \tau_4$ and $H_1 : (\tau_1 = \tau_2) \neq (\tau_3 = \tau_4)$. At $\alpha = 0.0076$, we can claim that "the probability of error to decide $(\tau_1 = \tau_2) \neq (\tau_3 = \tau_4)$ is below 1%" in the sense that

$$\Pr\left(\left|W_{uv}^*\right| \geq 4.276 \text{ for some } 1 \leq u < v \leq 4\right) = 0.0076 \quad \text{under } H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4$$

(not under any other scenario than H_0, such as $\tau_1 = \tau_2 = \tau_3 \neq \tau_4$).

One-sided multiple comparisons

With W_{ij}^* defined in (5.15), the Hayter-Stone one-sided all-treatments multiple comparison procedure sets the following rules: For each pair (u, v) with $u < v$:

$$\text{Decide } \tau_v > \tau_u \text{ if } W_{uv}^* \geq c_\alpha^*; \quad \text{otherwise accept } \tau_u = \tau_v,$$

where c_α^* is determined by $\Pr\left(W_{uv}^* < c_\alpha^*, 1 \leq u < v \leq k\right) = 1 - \alpha$ under H_0.

Large-sample approximation

Let $Z_i \sim N(0, 1/n_i)$, $i = 1, \ldots, k$, be independent normal random variables. Define

$$D = \max_{1 \leq i < j \leq k} \frac{Z_j - Z_i}{\sqrt{(n_i + n_j)/(2n_i n_j)}}$$

Then c_α^* can be approximated by d_α such that $\Pr(D \geq d_\alpha) = \alpha$ for large samples. Thus the approximate rules decide $\tau_v > \tau_u$ if $W_{uv}^* \geq d_\alpha$; otherwise accept $\tau_u = \tau_v$.

Example 6.7 of the textbook (page 267) demonstrates this procedure.

Example 5.8 Consider the data in Example 5.7. Since W_{ij}^* is symmetric about 0 under H_0, we can set the "other side" rule of the one-sided multiple comparisons as follows: For each pair (u,v) with $u<v$:

$$\text{Decide } \tau_v < \tau_u \text{ if } W_{uv}^* \le -c_\alpha^*; \quad \text{otherwise accept } \tau_u = \tau_v.$$

Then the probability of error under H_0 is

$$\Pr\left(W_{uv}^* \le -c_\alpha^* \text{ for some } u<v\right) = 1 - \Pr\left(W_{uv}^* > -c_\alpha^*, 1 \le u < v \le k\right) = \alpha$$

By the W_{ij}^* values calculated in Example 5.7 and $c_{0.0091}^* = 4.062$ from R,

$$W_{12}^* = 1.67 > -4.062 \;\Rightarrow\; \tau_1 = \tau_2, \qquad W_{13}^* = -4.67 < -4.062 \;\Rightarrow\; \tau_1 < \tau_3,$$

$$W_{14}^* = -4.82 < -4.062 \;\Rightarrow\; \tau_1 < \tau_4, \qquad W_{23}^* = -4.57 < -4.062 \;\Rightarrow\; \tau_2 < \tau_3,$$

$$W_{24}^* = -4.73 < -4.062 \;\Rightarrow\; \tau_2 < \tau_4, \qquad W_{34}^* = 0 > -4.062 \;\Rightarrow\; \tau_3 = \tau_4.$$

In summary, $(\tau_1 = \tau_2) < (\tau_3 = \tau_4)$ at $\alpha = 0.0091$.

The approximate $d_{0.01} = 4.098$ of $c_{0.01}^*$ leads to the same results at $\alpha = 0.01$.

One-sided treatments-versus-control multiple comparisons

First, place the baseline control as treatment 1 as before.

Let N^* be the least common multiple of $n_1, \ldots, n_k$, and $R_{\bullet j}$ the average rank for treatment j, $j = 2, \ldots, k$, as defined in (5.1).

Then the Nemenyi-Damico-Wolfe one-sided treatments-versus-control multiple comparison procedure is stated as follows. For $u = 2, \ldots, k$,

$$\text{Decide } \tau_u > \tau_1 \text{ if } N^*(R_{\bullet u} - R_{\bullet 1}) \geq y_\alpha^*; \quad \text{otherwise accept } \tau_u = \tau_1,$$

where y_α^* satisfies

$$\Pr\left(N^*(R_{\bullet u} - R_{\bullet 1}) < y_\alpha^*, u = 2, \ldots, k\right) = 1 - \alpha \text{ under } H_0.$$

Comment 68 on page 275 of the textbook explains how to obtain the distribution of $N^*(R_{\bullet u} - R_{\bullet 1})$ and the critical point y_α^*, and demonstrates the approach by an example with $k = 3$, $n_1 = n_2 = 1$ and $n_3 = 2$. We can also find y_α^* by R.

Large-sample approximation

First consider the case $n_1 = b$, $n_2 = \cdots = n_k = n$, $N = b + (k-1)n$. Let $(Z_2, \ldots, Z_k)$ be a $(k-1)$-variate normal random vector with $\mathrm{E}[Z_i] = 0$, $\mathrm{Var}(Z_i) = 1$, $i = 2, \ldots, k$, and $\mathrm{Corr}(Z_i, Z_j) = \rho = n/(b+n)$, $2 \le i \ne j \le k$. Define $m^*_{\alpha,\rho}$ by

$$\Pr\left(Z_{\max} \ge m^*_{\alpha,\rho}\right) = \alpha \text{ for } Z_{\max} = \max\{Z_2, \ldots, Z_k\} \tag{5.19}$$

The following rules are approximately valid with large b and n: For $u = 2, \ldots, k$,

$$\text{Decide } \tau_u > \tau_1 \text{ if } R_{\cdot u} - R_{\cdot 1} \ge m^*_{\alpha,\rho}\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{b} + \frac{1}{n}\right)}; \quad \text{otherwise accept } \tau_u = \tau_1.$$

In the general case with unequal $n_2, \ldots, n_k$, a conservative approximate procedure is given as follows: Let $\alpha^* = \alpha/(k-1)$. For $u = 2, \ldots, k$,

$$\text{Decide } \tau_u > \tau_1 \text{ if } R_{\cdot u} - R_{\cdot 1} \ge z_{\alpha^*}\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_1} + \frac{1}{n_u}\right)}; \quad \text{otherwise accept } \tau_u = \tau_1,$$

See Example 6.8 of the textbook (page 273) for an illustration of the procedure.

5.3 Estimation of treatment effects

It is of interest to estimate a linear combination of treatment effects:

$$\theta = \sum_{i=1}^{k} a_i \tau_i \text{, where } a_1, \ldots, a_k \in \mathbb{R} \text{ satisfy } \sum_{i=1}^{k} a_i = 0 \tag{5.20}$$

(referred to as *contrast*). In particular, (5.20) covers the difference $\tau_i - \tau_j$ for any $i \neq j \in \{1, \ldots, k\}$ (with $a_i = 1$, $a_j = -1$ and $a_l = 0$ for $l \neq i, j$).

Let $\Delta_{ij} = \tau_i - \tau_j$, which is termed *simple contrast*, and $d_{ij} = a_i / k$. Since

$$\sum_{i=1}^{k} a_i = 0 \quad \text{and} \quad k\tau_i = \sum_{j=1}^{k} \tau_i = \sum_{j=1}^{k} \left(\tau_i - \tau_j + \tau_j \right) = \sum_{j=1}^{k} \Delta_{ij} + \sum_{j=1}^{k} \tau_j ,$$

the contrast θ in (5.20) can be equivalently expressed by

$$\theta = \sum_{i=1}^{k} \frac{a_i}{k} (k\tau_i) = \sum_{i=1}^{k} \frac{a_i}{k} \left(\sum_{j=1}^{k} \Delta_{ij} + \sum_{j=1}^{k} \tau_j \right) = \sum_{i=1}^{k} \sum_{j=1}^{k} d_{ij} \Delta_{ij} \tag{5.21}$$

Estimators of contrasts: For $u, v \in \{1, \ldots, k\}$, let

$$Z_{uv} = \text{median}\{X_{iu} - X_{jv}, i = 1, \ldots, n_u, j = 1, \ldots, n_v\} \tag{5.22}$$

It is obvious that $Z_{uv} = -Z_{vu}$ and $Z_{uu} = 0$. For $u, v \in \{1, \ldots, k\}$, define

$$\bar{\Delta}_u = \frac{1}{N}\sum_{v=1}^{k} n_v Z_{uv} = \frac{1}{N}\sum_{v \neq u} n_v Z_{uv} \quad \text{and} \quad \hat{\Delta}_{uv} = \bar{\Delta}_u - \bar{\Delta}_v \tag{5.23}$$

Then $\hat{\Delta}_{uv}$ is an estimator of the simple contrast Δ_{uv}. An estimator of the contrast θ is given by

$$\hat{\theta} = \sum_{u=1}^{k} a_u \bar{\Delta}_u = \sum_{u=1}^{k}\sum_{v=1}^{k} d_{uv}(\bar{\Delta}_u - \bar{\Delta}_v) = \sum_{u=1}^{k}\sum_{v=1}^{k} d_{uv}\hat{\Delta}_{uv} \tag{5.24}$$

In the special case $n_1 = \cdots = n_k = n$, $n_v/N = n/N = 1/k$ for $j = 1, \ldots, k$. Hence

$$\bar{\Delta}_u = \frac{1}{k}\sum_{v=1}^{k} Z_{uv} = Z_{u\cdot} = \frac{1}{k}\sum_{v \neq u} Z_{uv} \quad \text{and} \quad \hat{\Delta}_{uv} = Z_{u\cdot} - Z_{v\cdot} \tag{5.25}$$

Simultaneous confidence intervals for all simple contrasts

Let $D_{ij}^{uv} = X_{iu} - X_{jv}$, $i = 1,\ldots,n_u$, $j = 1,\ldots,n_v$, and $D_{(1)}^{uv} \le \cdots \le D_{(n_u n_v)}^{uv}$ the ordered values of $\{D_{ij}^{uv}\} = \{D_{ij}^{uv}, i = 1,\ldots,n_u, j = 1,\ldots,n_v\}$. Then (5.22) is equivalent to

$$Z_{uv} = \text{median}\{D_{ij}^{uv}\} = \text{median}\{D_{(1)}^{uv},\ldots,D_{(n_u n_v)}^{uv}\} \tag{5.26}$$

For $1 \le u < v \le k$, define

$$a_{uv} = \frac{n_u n_v}{2} - w_\alpha^* \sqrt{\frac{n_u n_v (n_u + n_v + 1)}{24}} + 1 \quad \text{and} \quad b_{uv} = a_{uv} - 1, \tag{5.27}$$

where w_α^* is defined in (5.16). A set of simultaneous $100(1-\alpha)\%$ confidence intervals for all $k(k-1)/2$ simple contrasts are given by

$$\left[D_{([a_{uv}])}^{uv}, D_{(n_u n_v - [b_{uv}])}^{uv}\right), \quad 1 \le u < v \le k, \tag{5.28}$$

where $[a]$ denote the integer part of a. These intervals satisfy

$$\Pr\left(D_{([a_{uv}])}^{uv} \le \tau_u - \tau_v < D_{(n_u n_v - [b_{uv}])}^{uv}, 1 \le u < v \le k\right) = 1 - \alpha \tag{5.29}$$

Example 5.9 For the data in Example 5.4 (Example 6.2 of the textbook),

$$\left\{D_{ij}^{12}, \begin{matrix} i=1,\ldots,6 \\ j=1,\ldots,6 \end{matrix}\right\} = \left\{\begin{matrix} 2,0,-7,-4,0-2,-3,-5,-12,-9,-5,-7,0,-2,-9,-6, \\ -2,-4,5,3,-4,-1,3,1,6,4,-3,0,4,2,3,1,-6,-3,1,-1 \end{matrix}\right\} \Rightarrow$$

$$\left\{D_{(1)}^{12},\ldots,D_{(36)}^{12}\right\} = \left\{\begin{matrix} -12,-9,-9,-7,-7,-6,-6,-5,-5,-4,-4,-4,-3,-3, \\ -3,-2,-2,-2,-1,-1,0,0,0,0,1,1,1,2,2,3,3,3,4,4,5,6 \end{matrix}\right\} \qquad (5.30)$$

It follows from (5.26) and (5.30) that

$$Z_{12} = \left(D_{(18)}^{12} + D_{(19)}^{12}\right)/2 = (-2-1)/2 = -3/2 = -1.5$$

Similarly, $Z_{13} = -4$, $Z_{23} = -3$, and so $Z_{21} = -Z_{12} = 3/2 = 1.5$, $Z_{31} = -Z_{13} = 4$ and $Z_{32} = -Z_{23} = 3$. Thus by (5.25),

$$\bar{\Delta}_1 = \frac{1}{3}\sum_{v\neq 1} Z_{1v} = \frac{Z_{12}+Z_{13}}{3} = \frac{-1.5-4}{3} = \frac{-5.5}{3} = -\frac{11}{6},$$

$$\bar{\Delta}_2 = \frac{Z_{21}+Z_{31}}{3} = \frac{1.5-3}{3} = -\frac{1}{2} \quad \text{and} \quad \bar{\Delta}_3 = \frac{Z_{31}+Z_{32}}{3} = \frac{4+3}{3} = \frac{7}{3}$$

Then the simple contrasts are estimated by

$$\widehat{\Delta}_{12} = \overline{\Delta}_1 - \overline{\Delta}_2 = -\frac{11}{6} + \frac{1}{2} = \frac{1}{6}, \quad \widehat{\Delta}_{13} = -\frac{11}{6} - \frac{7}{3} = -\frac{25}{6}, \quad \widehat{\Delta}_{23} = -\frac{1}{2} - \frac{7}{3} = -\frac{17}{6}$$

By (5.24), the contrast $\theta = \tau_3 - (\tau_1 + \tau_2)/2$ (treatment 3 effect versus the average effect of treatments 1 and 2) is estimated by

$$\hat{\theta} = \overline{\Delta}_3 - \frac{\overline{\Delta}_1 + \overline{\Delta}_2}{2} = \frac{7}{3} - \frac{1}{2}\left(-\frac{11}{6} - \frac{1}{2}\right) = \frac{7}{3} + \frac{7}{6} = \frac{7}{2} = 3.5$$

To obtain simultaneous (approximate) 90% confidence intervals for the simple contrasts $\Delta_{12} = \tau_1 - \tau_2$, $\Delta_{13} = \tau_1 - \tau_3$ and $\Delta_{23} = \tau_2 - \tau_3$, use $k = 3$, $n_1 = n_2 = n_3 = 6$, $w^*_{0.0997} = 2.9439$ (by R) and (5.27) to calculate

$$a_{12} = a_{13} = a_{23} = \frac{6(6)}{2} - 2.9439\sqrt{\frac{6(6)(6+6+1)}{24}} + 1 = 6.00$$

and $b_{12} = b_{13} = b_{23} = 6.00 - 1 = 5.00$. Then by (5.30),

$$\left[D^{12}_{([a_{12}])}, D^{12}_{(36-[b_{12}])}\right) = \left[D^{12}_{(6)}, D^{12}_{(36-5)}\right) = \left[D^{12}_{(6)}, D^{12}_{(31)}\right) = [-6, 3)$$

Similarly we can find $\left[D^{13}_{(6)}, D^{13}_{(31)}\right) = [-8,0)$ and $\left[D^{23}_{(6)}, D^{23}_{(31)}\right) = [-6,2)$. Hence by (5.28), the simultaneous 90.03% confidence intervals for Δ_{12}, Δ_{13} and Δ_{23} are given by $[-6,3)$, $[-8,0)$ and $[-6,2)$ respectively.

Large-sample approximation: For large samples, the w^*_α used to define a_{uv} in (5.27) can be approximately by q_α in (5.18).

Example 5.10 For the data in Example 5.4, $q_{0.10} = 2.903$ (by R). Hence

$$a_{12} = a_{13} = a_{23} \approx \frac{6(6)}{2} - 2.903\sqrt{\frac{6(6)(6+6+1)}{24}} + 1 = 6.18$$

and $b_{12} = b_{13} = b_{23} \approx 6.18 - 1 = 5.18$. Thus

$$[a_{uv}] = [6.18] = 6 \quad \text{and} \quad [b_{uv}] = [5.18] = 5 \quad \text{for } uv = 12, 13, 23,$$

which produce the same simultaneous 90% confidence intervals for the simple contrasts as in Example 5.9.