

Nonparametric Method Using Placement in One-way Layout

Taek Su Chung¹⁾ and Dongjae Kim²⁾

Abstract

Kruskal and Wallis (1952) proposed typical nonparametric method in one-way layout problem. A special feature of this procedure is use of rank in mixed samples. In this paper, the new procedure based on placement as extension of the two sample placement tests described in Orban and Wolfe (1982) was proposed. Some critical values in small sample cases and comparative results of a Monte Carlo power study are presented.

Keywords: Placement; one-way layout; rank.

1. 서론

세 군 이상의 처리효과의 차이를 검정하기 위한 일원배치모형에서의 확률변수 Y_{ij} 는 다음과 같다.

$$Y_{ij} = \mu + \tau_i + e_{ij}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, n_i. \quad (1.1)$$

여기서 τ_i 는 i 번째 처리의 효과를 나타내고, μ 는 미지의 전체 평균을 나타낸다. 만약 오차항 e_{ij} 가 서로 독립이고 정규분포 $N(0, \sigma^2)$ 을 따르는 확률변수라는 가정이 성립한다면, τ_i 가 고정효과 (fixed effects)일 때 모수적 방법인 분산분석법을 사용하여 τ_i 들이 모두 같다는 귀무가설을 검정할 수 있다. 하지만 오차항 e_{ij} 가 정규분포를 따른다는 가정을 만족하지 않을 경우에는 분산분석법의 사용시 문제점이 발생한다. 예를 들면 오차항 e_{ij} 가 꼬리가 두터운 분포를 따를 때에는 유의수준 5%에서 검정하기를 원하고 있으나 유의수준을 정확히 관리하기 힘든 문제가 발생할 수 있다. 따라서 모수적 가정을 만족하지 않는 경우에는 제 1종 오류를 제어할 수 있는 비모수적 검정법을 선택해야 한다.

일원배치모형을 위한 비모수 검정법은 대립가설의 형태에 따라 많은 방법들이 제안되었다. 각 처리의 효과가 적어도 하나는 다르다는 일반적인 가설에 관한 검정법은 Kruskal과 Wallis (1952), Bhapkar (1961), Schulzer (1969)가 제안하였고, 순서대립가설

1) Graduate Student, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea.

2) Professor, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea.

Correspondence : djkim@catholic.ac.kr

에 관해서는 Jonckheere (1954)와 Terpstra (1952), Chacko (1963), Puri (1965)가 제안한 방법들이 있다. 일반적인 가설에서의 대표적인 검정법은 Kruskal과 Wallis가 제안한 검정법을 들 수 있는데, 이 방법은 관측값 대신에 혼합표본에서의 순위(rank)를 이용하는 것이 특징이고, 분포와 무관하게 유의수준을 제어할 수 있다. 또한, 순서대립가설에서의 대표적인 검정법은 Jonckheere와 Terpstra가 제안한 검정법을 들 수 있는데, Mann과 Whitney (1947)가 제안한 U통계량을 확장하여 순서대립가설에 적용한 방법이다.

Orban과 Wolfe (1982)는 두 처리간 효과의 차이를 검정하기 위해 위치(placement)를 사용한 비모수 검정법을 제안하였다. 이 방법은 두 처리중 어느 한 처리에 대한 상대적 위치정보를 이용하여 처리효과와의 차이를 검정하는 방법으로써 대조군의 표본 크기가 처리군의 표본 크기보다 클 때 더 유용하다고 알려져 있다.

본 논문에서는 위치라는 비모수적 도구를 일원배치모형으로 확장하여 새로운 비모수적 방법론을 제안하였다. 2절에서 귀무가설에 대한 일반적인 대립가설과 그에 대한 검정방법을 제안하고, 3절에서는 검정통계량의 소표본 기각역과 Kruskal과 Wallis가 제안한 검정법과의 관계에 대하여 논의하였다. 또한 4절에서는 소표본 모의실험을 통하여 제안된 검정법을 비모수 검정법인 Kruskal과 Wallis가 제안한 검정법과 모수적 검정법인 분산분석법과의 검정력을 비교하였다.

2. 방법

처리의 수가 k 개인 일원배치모형에서 $Y_{i1}, Y_{i2}, \dots, Y_{in_i}; i = 1, 2, \dots, k$ 를 i 번째 처리에서 추출한 확률표본이라 하고, 전체표본의 크기는 $N = \sum_{i=1}^k n_i$ 이라 하자.

각 처리의 효과가 모두 동일하다는 귀무가설과 일반적인 대립가설은 다음과 같다.

$$H_0 : [\tau_1 = \tau_2 = \dots = \tau_k] \quad \text{vs.} \quad H_1 : [\tau_i \text{들이 모두 같지는 않다}].$$

두 군의 처리효과 차이를 검정하기 위한 비모수적 도구인 위치(placement)의 개념을 세 군 이상으로 확장한 결합 위치(joint placement) V_{ij} 를 다음과 같이 정의한다.

$$V_{ij} = \frac{1}{N - n_i} \sum_{h=1}^k \sum_{s=1, s \neq i}^{n_h} \chi(Y_{hs}, Y_{ij}), \quad \text{여기서 } \chi(x, y) = \begin{cases} 1, & x \leq y \text{ 일 경우,} \\ 0, & \text{그 외.} \end{cases} \quad (2.1)$$

결합 위치 V_{ij} 는 i 번째 처리의 관측값을 제외한 혼합표본에서 Y_{ij} 보다 작거나 같은 관측값의 개수를 이용한 $[0, 1]$ 에서 정의된 확률변수이다. 또한 결합 위치의 처리별 평균 \bar{V}_i 과 전체평균 \bar{V} 는 다음과 같다

$$\bar{V}_i = \frac{\sum_{j=1}^{n_i} V_{ij}}{n_i}, \quad i = 1, 2, \dots, k, \quad (2.2)$$

$$\bar{V}_{..} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} V_{ij}}{N} = \frac{\sum_{i=1}^k n_i \bar{V}_i}{N} = \frac{\sum_{i=1}^k V_i}{N}. \quad (2.3)$$

따라서 각 군에서의 \bar{V}_i 와 $\bar{V}_{..}$ 의 차이를 이용한 검정통계량 JP 를 다음과 같이 정의한다.

$$JP = \sum_{i=1}^k n_i (\bar{V}_i - \bar{V}_{..})^2. \quad (2.4)$$

또는,

$$JP = \sum_{i=1}^k \frac{V_i^2}{n_i} - N \bar{V}_{..}^2. \quad (2.5)$$

귀무가설 하에서 \bar{V}_i 는 $\bar{V}_{..}$ 에 가까울 것으로 기대되기 때문에 통계량 JP 의 값이 작을수록 귀무가설을 지지하는 증거가 되며, 반대로 값이 클수록 귀무가설을 기각하는 증거가 된다. 그러므로 기각역은 $JP \geq x, (x \geq 0)$ 의 형태가 되고, 유의수준 α 에서 $JP \geq jp(\alpha, k, (n_1, \dots, n_k))$ 이면 귀무가설 H_0 를 기각한다. 여기서 $jp(\alpha, k, (n_1, \dots, n_k))$ 는 $P_0\{JP \geq jp(\alpha, k, (n_1, \dots, n_k))\} = \alpha$ 를 만족하는 상수이다.

3. 검정통계량의 성질 및 소표본 기각역

Kruskal과 Wallis가 제안한 순위를 이용한 검정법에서 Y_{ij} 에 부여된 혼합 표본에서의 순위를 R_{ij} 라고 하면, 결합 위치의 처리별 합 V_i 와 순위의 처리별 합 R_i 은 다음과 같은 관계를 유도할 수 있다.

$$V_i = \frac{R_i - \frac{n_i(n_i + 1)}{2}}{(N - n_i)} \quad \left(\because R_i = \sum_{j=1}^{n_i} R_{ij} \right). \quad (3.1)$$

그리고, \bar{V}_i 와 $\bar{V}_{..}$ 는 다음과 같이 표현 된다.

$$\bar{V}_i = \frac{V_i}{n_i} = \frac{R_i - \frac{n_i(n_i + 1)}{2}}{(N - n_i) n_i}, \quad (3.2)$$

$$\bar{V}_{..} = \frac{\sum_{i=1}^k V_i}{N} = \frac{1}{N} \sum_{i=1}^k \left(\frac{R_i - \frac{n_i(n_i + 1)}{2}}{(N - n_i)} \right). \quad (3.3)$$

이제 식 (3.2)와 식 (3.3)을 식 (2.4)에 대입하면 통계량 JP 를 다음과 같이 순위에 관한 식으로 표현할 수 있다.

$$JP = \sum_{i=1}^k n_i \left[\frac{R_i - \frac{n_i(n_i+1)}{2}}{(N-n_i)n_i} - \frac{1}{N} \sum_{i=1}^k \left(\frac{R_i - \frac{n_i(n_i+1)}{2}}{(N-n_i)} \right) \right]^2. \quad (3.4)$$

만약 k 개 처리의 각 표본의 크기 n_i 들이 n 으로 동일하다면 전체 표본의 크기 N 은 kn 과 같고, 식 (3.4)는 다음과 같이 정리 된다.

$$JP = \frac{1}{(k-1)^2 n^2} \left(\sum_{i=1}^k \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right). \quad (3.5)$$

또한, 각 처리별 표본의 크기가 n 으로 동일한 경우에 Kruskal과 Wallis (1952)가 제안한 검정법의 검정통계량 H 는 다음과 같다.

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1). \quad (3.6)$$

따라서 결합 위치를 이용한 검정방법의 검정통계량 JP 와 Kruskal과 Wallis가 제안한 검정법의 검정통계량 H 는 다음과 같은 관계가 성립한다.

$$JP = \frac{N(N+1)}{12(k-1)^2 n^2} H. \quad (3.7)$$

그러므로 각 처리별 표본의 크기가 동일하게 n 으로 정해진 경우, 두 통계량은 서로 상수배의 관계에 놓여있기 때문에 검정통계량은 같다고 할 수 있다. 그리고 각 처리별 표본의 크기가 동일할 때 결합 위치를 이용한 검정방법과 Kruskal과 Wallis가 제안한 검정법은 같은 검정방법이 된다.

각 처리별 표본의 크기가 동일하지 않은 경우에는 관측값들의 결합 위치값에 영향을 미치는 표본의 크기가 처리별로 다르기 때문에, 전체 표본의 크기가 같더라도 각 처리별 표본의 크기의 조합에 따라 결합 위치의 전체평균이 상수가 되지 않고 n_i 들에 의해 변하는 함수가 된다. 따라서 각 처리별 표본의 크기가 동일하지 않다면 결합 위치를 이용한 검정방법과 Kruskal과 Wallis가 제안한 검정법은 서로 다른 검정법이다.

결합 위치를 이용한 검정법을 사용하기 위해 처리별 표본의 크기를 고려하여 가능한 모든 경우에 대해서 검정통계량의 값을 구한다음 정렬을 통해 확률값을 계산하여 소표본 분포를 구하였고, 처리의 수 $k = 3$ 과 $k = 4$ 에 대해 유의수준 0.05 근방의 검정통계량의 기각값을 표 3.1과 표 3.2에 정리하였다. 한 군의 최대 표본의 크기는 $k = 3$ 일때는 5, 그리고 $k = 4$ 일때는 3까지 제시하였다.

4. 소표본 모의실험의 계획 및 결과

처리 개수가 3이상인 일원배치모형에서의 비모수 검정방법은 일반적인 대립가설에서 Kruskal과 Wallis가 제안한 검정법을 사용하고, 모수적인 방법으로서 F 통계량을

☒ 3.1: Critical value ; $\alpha = 0.05, k = 3$

n_1	n_2	n_3	x	$P_0\{JP \geq x\}$	n_1	n_2	n_3	x	$P_0\{JP \geq x\}$
1	1	2	0.7500	0.3333	2	3	3	1.3271	0.0107
1	1	3	0.9500	0.2000				1.1438	0.0536
1	1	4	1.1000	0.0667	2	3	4	1.1874	0.0476
1	1	5	1.2143	0.0952				1.1651	0.0508
1	2	2	1.0000	0.0667	2	3	5	1.1899	0.0500
1	2	3	1.2333	0.0333	2	4	4	1.1694	0.0483
			1.0875	0.0667				1.1313	0.0502
1	2	4	1.3238	0.0381	2	4	5	0.1808	0.0488
			1.3143	0.0571				1.1752	0.0505
1	2	5	1.5893	0.0119	2	5	5	1.1776	0.0499
			1.1438	0.0595				1.1600	0.0523
1	3	3	1.2976	0.0429	3	3	3	1.1667	0.5000
			1.1905	0.0714	3	3	4	1.1476	0.0486
1	3	4	1.2667	0.5000				1.1357	0.0533
1	3	5	1.3250	0.0476	3	3	5	1.1564	0.0498
			1.3056	0.0516				1.1523	0.0511
1	4	4	1.3050	0.0476	3	4	4	1.1420	0.0490
			1.3000	0.0508				1.1404	0.0507
1	4	5	1.2962	0.0476	3	4	5	1.1359	0.0499
			1.2757	0.0508				1.1339	0.0507
1	5	5	1.3121	0.0469	3	5	5	1.1439	0.0498
			1.3091	0.0512				1.1338	0.0513
2	2	2	1.0000	0.0667	4	4	4	1.1563	0.0487
2	2	3	1.1371	0.0381				1.1484	0.0546
			1.0971	0.0571	4	4	5	1.1669	0.0490
2	2	4	1.1259	0.0476				1.1343	0.0501
			1.0703	0.0571	4	5	5	1.1446	0.0500
2	2	5	1.1265	0.0476	5	5	5	1.1560	0.0490
			1.0653	0.0582				1.1320	0.0510

☒ 3.2: Critical value ; $\alpha = 0.05, k = 4$

n_1	n_2	n_3	n_4	x	$P_0\{JP \geq x\}$	n_1	n_2	n_3	n_4	x	$P_0\{JP \geq x\}$
1	1	1	2	0.8000	0.2000	1	3	3	3	1.1347	0.0500
1	1	1	3	1.0400	0.1000	2	2	2	2	1.0278	0.0381
1	1	2	2	1.0200	0.0222					1.0000	0.0667
				0.8833	0.0667	2	2	2	3	1.0580	0.0492
1	1	2	3	1.0810	0.0476					1.0527	0.0524
				1.0675	0.0571	2	2	3	3	1.0939	0.0495
1	1	3	3	1.1490	0.0500					1.0905	0.0502
1	2	2	2	0.9505	0.0476	2	3	3	3	1.0966	0.0499
				0.9486	0.0667					1.0947	0.0506
1	2	2	3	1.0496	0.0500	3	3	3	3	1.1235	0.0435
1	2	3	3	1.1191	0.0500					1.1070	0.0502

사용한 분산분석법 (ANOVA Test)을 사용한다. 따라서 결합 위치를 이용한 검정방법을 Kruskal과 Wallis가 제안한 검정법, 분산분석법과 검정력을 비교하였고, 순서형 대립가설에서의 검정력도 살펴보기 위해 Jonckheere와 Terpstra가 제안한 방법과도 비교해 보았다. 실행 프로그램은 SAS를 사용하였다. 모집단의 분포로는 정규분포, 이중지수분포, Cauchy분포를 채택하였으며, 정규분포와 Cauchy분포의 난수생성은 SAS의 RANNOR 함수와 RANCAU 함수를 이용하였다. 이중지수분포의 난수생성은 균일분포의 난수생성을 위한 SAS의 RANUNI 함수와 역변환기법을 이용하여 생성하였고, 척도모수 β 는 3으로 하였다. 생성된 난수를 실제 표본으로 간주하고, 계산된 검정통계량이 기각역에 포

표 4.1: One-way layout Monte Carlo power estimates ; $\alpha = 0.05, k = 3$

Distribution	n_1	n_2	n_3	τ_1	τ_2	τ_3	JP	$K-W$	$J-T$	F
Normal	3	4	5	0	0	0	0.0486	0.0467	0.0503	0.0483
				0	0	1	0.2665	0.2204	0.3847	0.2310
				0	1	0	0.2221	0.2158	0.0147	0.2187
				1	0	0	0.1267	0.1652	0.0037	0.1887
				1	0	2	0.6546	0.5839	0.3622	0.6100
				2	0	1	0.3718	0.4535	0.0027	0.5011
	4	4	5	0	0	0	0.0495	0.0483	0.0491	0.0488
				0	0	1	0.2727	0.2385	0.3960	0.2568
				0	1	0	0.2056	0.2198	0.0249	0.2332
				0	1	2	0.6179	0.5866	0.8426	0.6256
	4	5	5	0	2	1	0.4787	0.5436	0.1467	0.5804
				0	0	0	0.0485	0.0468	0.0537	0.0450
				0	0	1	0.2722	0.2528	0.4016	0.2676
				1	0	0	0.2009	0.2250	0.0010	0.2415
	Double	3	4	5	1	0	2	0.7049	0.6567	0.2903
2					0	1	0.5864	0.6070	0.0005	0.6444
0					0	0	0.0503	0.0496	0.0530	0.0310
0					0	1	0.2113	0.1801	0.3111	0.1197
0					1	0	0.1917	0.1868	0.0166	0.1181
1					0	0	0.1116	0.1446	0.0070	0.1000
Exponential	4	4	5	1	0	2	0.4104	0.3599	0.2609	0.2980
				2	0	1	0.2697	0.3185	0.0118	0.2453
				0	0	0	0.0480	0.0484	0.0472	0.0318
				0	0	1	0.2196	0.2004	0.3225	0.1354
				0	1	0	0.1769	0.1867	0.0281	0.1190
	4	5	5	0	1	2	0.3910	0.3694	0.6387	0.2981
				0	2	1	0.3338	0.3704	0.1257	0.2826
				0	0	0	0.0499	0.0483	0.0490	0.0304
				0	0	1	0.2215	0.2093	0.3228	0.1272
4	5	5	1	0	0	0.1786	0.1979	0.0035	0.1257	
			1	0	2	0.4405	0.4062	0.2133	0.3299	
			2	0	1	0.3815	0.3994	0.0037	0.3062	

(continued)

Distribution	n_1	n_2	n_3	τ_1	τ_2	τ_3	JP	K-W	J-T	F
Cauchy	3	4	5	0	0	0	0.0502	0.0480	0.0509	0.0212
				0	0	1	0.1135	0.1020	0.1779	0.0427
				0	1	0	0.1038	0.0990	0.0266	0.0397
				1	0	0	0.0756	0.0870	0.0135	0.0407
				1	0	2	0.2259	0.1977	0.1841	0.0928
				2	0	1	0.1426	0.1682	0.0200	0.0794
	4	4	5	0	0	0	0.0449	0.0443	0.0491	0.0164
				0	0	1	0.1162	0.1050	0.1909	0.0421
				0	1	0	0.0963	0.0972	0.0353	0.0364
				0	1	2	0.2118	0.1995	0.4003	0.0897
				0	2	1	0.1630	0.1866	0.1221	0.0760
	4	5	5	0	0	0	0.0509	0.0498	0.0523	0.0191
				0	0	1	0.1183	0.1135	0.1923	0.0415
				1	0	0	0.0986	0.1077	0.0085	0.0399
				1	0	2	0.2418	0.2202	0.1647	0.0947
2				0	1	0.2001	0.2096	0.0117	0.0907	

**JP : 결합 위치를 이용한 검정방법

K-W : Kruskal과 Wallis 검정법

**J-T : Jonckheere와 Terpstra 검정법

F : 분산분석법

합되는지를 판단하는 과정을 10000번 반복하는 Monte Carlo Study를 사용하였다.

처리의 수는 3개와 4개일 경우를 선택하였다. 처리의 수가 3개일 때는 각 처리별로 세 군의 표본의 크기가 모두 다를 경우와 두 표본의 크기는 같고 한 표본의 크기는 다른 경우를 고려하였는데, 같은 두 표본의 크기와 비교하여 클 때와 작을 때로 구분하였다. 처리의 수가 4개일 때는 세 표본의 크기는 같고 한 표본의 크기는 다른 경우와 두 표본의 크기는 같고 다른 두 표본의 크기가 똑같이 다른 경우를 고려하였다. 각 처리의 표본의 크기가 모두 같은 경우는 결합 위치를 사용한 검정법과 Kruskal과 Wallis 검정법이 같은 방법이기 때문에 모의실험에서 제외하였다. 각 처리별 효과를 나타내는 τ_i 들은 처리의 수가 3개이고 표본의 크기가 모두 다른 경우에 6가지 조합을, 처리의 수와 상관없이 나머지 경우에는 5가지 조합의 실험을 실시하였다. 그리고 유의수준 α 를 0.05로 보정하기 위해 확률화 검정을 사용하였다. 각 분포에서의 방법별 검정력의 비교결과를 처리의 개수가 3일 때는 표 4.1에, 처리의 개수가 4일 때는 표 4.2에 정리 하였다.

각 처리의 효과가 모두 동일할 때에 유의수준이 0.05를 만족하는지를 살펴보면, 분산분석법의 유의수준은 처리의 개수와는 상관없이 이중지수분포에서는 0.03, Cauchy분포에서는 0.02에 근사한 값들을 얻었다. 이는 정규성을 만족하지 않을 경우에 제 1종 오류를 제어하기 힘들다는 앞의 내용과 일치하는 결과이다. 결합 위치를 사용한 검정법과 Kruskal과 Wallis 검정법, 그리고 Jonckheere와 Terpstra가 제안한 검정법은 세 가지 분포에서 유의수준의 값이 모두 0.05 근방의 값들을 얻었는데, 이는 제 1종 오류를 제어하는데 문제가 없음을 보여주고 있다.

전체적으로 분산분석법의 검정력은 정규분포에서 일반적인 대립가설의 검정을 위한 다른 두 방법보다 검정력이 높았고, 이중지수와 Cauchy분포에서는 다른 두 방법보다 낮았다. 처리효과 τ_i 들이 순서형일 경우에는 Jonckheere와 Terpstra 검정법의 검정력

표 4.2: One-way layout Monte Carlo power estimates ; $\alpha = 0.05, k = 4$

Distribution	n_1	n_2	n_3	n_4	τ_1	τ_2	τ_3	τ_4	JP	K-W	J-T	F
Normal	2	2	2	3	0	0	0	0	0.0529	0.0502	0.0514	0.0475
					0	0	0	2	0.4550	0.3042	0.5094	0.3428
					0	0	2	0	0.1270	0.1653	0.0301	0.2807
					0	0	1	2	0.3794	0.2927	0.6630	0.2971
					0	0	2	1	0.1292	0.2061	0.2280	0.2460
	2	2	3	3	0	0	0	0	0.0542	0.0540	0.0515	0.0512
					0	2	0	0	0.1197	0.1560	0.0023	0.3128
					0	0	0	2	0.3952	0.3154	0.5135	0.3931
					0	1	0	2	0.4207	0.3632	0.3510	0.3653
	2	3	3	3	0	2	0	1	0.2076	0.2457	0.0347	0.3039
					0	0	0	0	0.0508	0.0500	0.0537	0.0520
					2	0	0	0	0.1071	0.1637	0.0005	0.3404
0					0	0	2	0.3719	0.3246	0.5030	0.4505	
1					0	0	2	0.4346	0.3980	0.1765	0.4199	
Double	2	2	2	3	2	0	0	1	0.2477	0.2966	0.0042	0.3553
					0	0	0	0	0.0492	0.0495	0.0458	0.0349
					0	0	0	2	0.3207	0.2242	0.3621	0.1888
					0	0	2	0	0.0965	0.1244	0.0412	0.1505
					0	0	1	2	0.2840	0.2296	0.4818	0.1645
	2	2	3	3	0	0	2	1	0.1352	0.1756	0.1848	0.1359
					0	0	0	0	0.0471	0.0489	0.0471	0.0300
					0	2	0	0	0.0921	0.1157	0.0072	0.1567
					0	0	0	2	0.2670	0.2207	0.3572	0.1993
	2	3	3	3	0	1	0	2	0.2957	0.2626	0.2685	0.1816
					0	2	0	1	0.1836	0.1989	0.0422	0.1491
					0	0	0	0	0.0476	0.0489	0.0508	0.0297
2					0	0	0	0.0845	0.1194	0.0046	0.1681	
0					0	0	2	0.2541	0.2270	0.3612	0.2234	
Exponential	2	2	2	3	1	0	0	2	0.2824	0.2606	0.1378	0.1994
					2	0	0	1	0.1855	0.2135	0.0190	0.1713
					0	0	0	0	0.0484	0.0499	0.0523	0.0243
					0	0	0	2	0.1968	0.1465	0.2461	0.0703
					0	0	2	0	0.0808	0.0947	0.0497	0.0590
	2	2	3	3	0	0	1	2	0.1568	0.1341	0.2974	0.0648
					0	0	2	1	0.0772	0.1005	0.1392	0.0535
					0	0	0	0	0.0493	0.0501	0.0492	0.0220
					0	2	0	0	0.0725	0.0868	0.0146	0.0568
	2	2	3	3	0	0	0	2	0.1665	0.1448	0.2396	0.0669
					0	1	0	2	0.1681	0.1513	0.1903	0.0642
					0	2	0	1	0.0937	0.1042	0.0428	0.0553
0					0	0	0	0.0500	0.0509	0.0454	0.0204	
2					0	0	0	0.0709	0.0885	0.0111	0.0587	
2	3	3	3	0	0	0	2	0.1578	0.1447	0.2484	0.0712	
				1	0	0	2	0.1558	0.1452	0.1216	0.0641	
				2	0	0	1	0.0952	0.1074	0.0254	0.0541	

**JP : 결합 위치를 이용한 검정방법

K-W : Kruskal과 Wallis 검정법

**J-T : Jonckheere와 Terpstra 검정법

F : 분산분석법

이 매우 높게 나왔고, 그렇지 않은 경우에는 검정력이 상당히 낮았다. 결합 위치를 사용한 검정방법은 τ_i 들이 순서형일 경우에 Jonckheere와 Terpstra 검정법보다 낮았지만 Kruskal과 Wallis 검정법보다 높았다. 각 처리들의 표본의 크기가 모두 다른 경우와 한 처리의 표본의 크기만이 클 경우에는 표본의 크기가 큰 처리의 τ 값이 표본의 크기가 작은 처리의 τ 값보다 클 때, 결합 위치를 사용한 검정방법의 검정력이 Kruskal과 Wallis 검정법보다 높았다. 한 처리의 표본의 크기만 작을 경우에는 표본의 크기가 큰 처리의 τ 값이 더 클 때, Kruskal과 Wallis 검정법보다 검정력이 높았다. 처리의 개수가 4개일 경우에 두 처리는 같고 다른 두 처리가 똑같은 값으로 다를 경우에도 표본의 크기가 큰 처리의 τ 값이 클 때 검정력이 높았다. 따라서 표본의 크기가 큰 처리의 τ 값이 표본의 크기가 작은 처리의 τ 값보다 클 때, 결합 위치를 사용한 검정법은 Kruskal과 Wallis 검정법보다 검정력이 높다는 것을 알 수 있다.

5. 토의 및 결론

각 처리별 표본의 크기와 τ 값의 조합에 따른 검정력의 경향을 검정 통계량과 연관지어 살펴보면, 표본의 크기가 작은 처리의 τ 값이 클 때에 비해 표본의 크기가 큰 처리의 τ 값이 클 때의 \bar{V}_i 와 $\bar{V}_{..}$ 의 차이가 커지는 경향을 보였다는 점을 알 수 있었다. 이러한 경향이 나오게 된 원인은 \bar{V}_i 와 $\bar{V}_{..}$ 의 차이가 표본의 크기에 따라 달라진다는 점에서 찾을 수 있다. 즉 각 처리별 표본의 크기의 차이가 관측값의 결합 위치에 영향을 미치기 때문이다. 따라서 $\bar{V}_{..}$ 가 일정한 값을 보이지 않고 검정에도 영향을 미치게 된다.

모의실험의 결과로 나타나게 된 검정력의 경향을 생각해 본다면 본 논문에서 제안한 검정방법은 작은 표본의 크기를 가지는 처리의 τ 값이 작고 가장 큰 표본의 크기를 가진 처리의 τ 값이 가장 클 때 Kruskal과 Wallis가 제안한 검정법보다 더 효율적이라 말할 수 있다. 따라서 실험 설계에서 이와 같은 상황이 기대될 경우에는 Kruskal과 Wallis가 제안한 검정방법보다 본 논문에서 제시한 결합 위치를 이용한 검정방법을 사용하는 것이 더 효율적인 실험이 될 것이다. 하지만 처리의 수가 3개일 경우에 3절에서 제시한 소표본 기각역은 한 처리의 최대 표본의 크기가 5까지 이므로 그 이상의 표본의 크기를 가진 경우에는 검정을 할 수 없는 한계점을 안고 있다. 따라서 검정통계량의 분포와 성질을 더 밝혀내고, 점근분포를 찾아낸다면 본 논문에서 제안한 검정방법은 특정한 실험설계를 위한 유용한 비모수 검정방법이 될 것으로 기대한다.

참고문헌

- Bhappkar, V. P. (1961). A nonparametric test for the problem of several samples. *The Annals of Mathematical Statistics*, **32**, 1108–1117.
- Chacko, V. J. (1963). Testing homogeneity against ordered alternatives. *The Annals of Mathematical Statistics*, **34**, 945–956.
- Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, **41**, 133–145.

- Kruskal, W. H. and Wallis, W. A (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, **47**, 583–621.
- Mann, H. B. and Whitney, D. R.(1947). On a test of whether one of the two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, **18**, 50–60.
- Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placement. *Journal of the American Statistical Association*, **77**, 666–672.
- Puri, M. L. (1965). On the combination of independent two sample tests of a general class. *Review of the International Statistical Institute*, **33**, 229–241.
- Schulzer, M. (1969). Contributions to the k-sample problem : A symmetric statistic. *The Annals of Mathematical statistics*, **40**, 1933–1949.
- Terpstra, T. J. (1952). The asymptotic normality and consistency of kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae*, **14**, 327–333.
- SAS Release 8.1 (1998). Cary SAS Institute.

[Received July 2007, Accepted September 2007]