응용통계연구 제 **12**권 1호, *1999*년, *pp. 191-202*

임의의 불완전 이원배치 순위계획법에서 효율적인 검정법

임동훈[1]

요 약

본 논문에서는 n명의 judge들이 k개의 object들에 대해 순위를 부여하는 경우 얻어질 수 있는 불완전 순위자료에서 object효과들이 같은지에 대한 검정법을 제안하고 컴퓨터 프로그램없이 쉽게 구현할 수 있는 효과적인 알고리즘을 개발하였다. 그리고 기존의 평균순위법과 Monte Carlo 연구를 통하여 검정력을 비교하였다.

1. 서론

n명의 judge들이 k개의 object들에 대해 순위를 부여하는 경우를 생각하자. judge 들이 가장 우등(best)하다고 생각되는 object에 순위 1를 부여하고 차례로 순위 2 , 3,... , 마지막으로 가장 열등(worst)하다고 생각하는 object에 순위 k를 부여한다. τ_j를 object j의 효과라 할때 모든 judge 들이 완전한 순위벡터를 갖고 있는 표준적인 이원 배치법하에서 귀무가설 $H_0 : [\tau_1 = \cdots = \tau_k]$를 일반대립가설 H_1 : [τ_j 들이 모두 같지는 않다]에 대한 검정법으로는 Friedman (1937) 검정과 Kendall 과 Babington Smith (1939) 검정이 있고 순서대립가설 $H_1^{'} : [\tau_1 \leq \cdots \leq \tau_k$, 적어도 한 부등식은 절대적]에 대한 검정으로는 Page (1963) 검정과 우산형 대립가설 $H_1^{'} : [\tau_1 \leq \cdots \leq \tau_p \geq \tau_{p+1} \geq \cdots \geq \tau_k$, 어떤 $1 \leq p \leq k$ 와 적어도 한 부등식은 절대적] 에 대한 검정으로는 Hettmansperger 와 Norton (1987)검정 등이 널리 사용되어왔다. 그러나 실제 많은 응용에서 각 judge들이 object들에 대해 불충분한 정보로 인하여 얻어지는 불완전 순위 자료를 분석하는 문제에 직면하는 경우가 허다하다.

본 논문에서는 이와같은 자료에서 Friedman (1937) 검정과 관련된 일반대립가설에 대한 검정에 관심이 있으나 제한된 순서형 혹은 우산형 대립가설에 대한 검정으로 확장이 가능하다. 지금까지 이런 불완전 자료를 분석하는데 흔히 사용되어온 한가지 방법은 평균순위법 (average ranks procedure)이다. 이는 각 judge의 결측순위를 갖고있는 object을 부여되지 않는 순위들의 평균으로 대치 (impute)한 후 Friedman 통계량을 적용시키는 방법이다 (e.g. Lordo 와 Wolfe (1994)). 그러나 이 방법은 완전한 실험계획하에서 유용한 기존의 Friedman 통계량의 귀무분포표를 아무런 검증없이 사용함으로써 잘못된 결론을 내릴 가능성이 있을 뿐 만아니라 결측율이 높은 불완전 자료에 대해서는 보수적인 검정법이다. 최근에 Lim, Lordo 와 Wolfe (1999)는 judge들이 결측 object에 대해 부여할 수 있는 모든 가능한 경우에 대해 얻어진 Friedman 통계량중에서 최대 통계량 S_{max}와 최소통계량 S_{min}에 기초한 검정법을 제안하였다. 그러나 이 방법 또한 어떤 실험계획에서 H_0 혹은 H_1에 대한 어떠한 결정도 내릴 수 없는 경우가 발생한다. 본 논문에서는 불완전 순위자료를 분석하는데 위에서 언급한 검정법들의 문제점에 대한 보안 내지는 대안의 한 방법으로 효율적인 검정

1) (660-701) 경남 진주시 가좌동 900, 경상대학교 통계학과 조교수, 정보통신연구센터 연구원

법을 제안하였고 컴퓨터 프로그램없이 구현가능한 알고리즘을 개발하였으며 모의실험을 통하여 알고리즘의 성능을 평가하였다. 또한 예제를 가지고 구체적인 적용방법을 살펴보았고 여러가지 실험환경하에서 기존의 평균순위법과 Monte Carlo 연구를 통하여 검정력을 비교하였다.

2. 제안한 검정법

2.1. S_{max} / S_{min}에 기초한 검정법

$S_{max}[S_{min}]$을 주어진 불완전 실험계획에서 각 judge들의 결측 object을 부여할 수 있는 모든 가능한 순위를 부여함으로써 얻어진 완전화된 실험계획상에서 계산된 Friedman 통계량들 중에서 최대 [최소] 통계량이라 하자. 그리고 $s(\alpha, k, n)$을 Friedman 통계량의 귀무분포에 대한 상위 α번째 백분위수라 할때 Lim, Lordo 와 Wolfe (1999)은 S_{max} 와 S_{min}에 기초한 검정법으로 $S_{min} \geq s(\alpha, k, n)$ 이면 H_0를 기각하고 $S_{max} < s(\alpha, k, n)$이면 H_0를 기각하지 못하는 방법을 제안하였다. 그들은 $S_{min} < s(\alpha, k, n) \leq S_{max}$ 인 경우에 대해서는 H_0 혹은 H_1에 대한 어떠한 결정도 내릴 수 없었다. 본 논문에서는 $S_{min} < s(\alpha, k, n) \leq S_{max}$ 인 경우 S_{max} 과 S_{min}의 산술평균을 이용하여 다음의 검정법을 제안한다.

$$\begin{aligned}
&S_{min} \geq s(\alpha, k, n) \text{ 혹은} \\
&\frac{S_{max} + S_{min}}{2} \geq s(\alpha, k, n) \text{ 이면 } H_0 \text{ 기각하고} \\
&S_{max} < s(\alpha, k, n) \text{ 혹은} \\
&\frac{S_{max} + S_{min}}{2} < s(\alpha, k, n) \text{ 이면 } H_0 \text{ 기각못한다.}
\end{aligned} \tag{2.1}$$

이 검정법은 불완전 순위자료에서 유의성을 검정하는데 기존의 평균순위법을 포함한 여러가지 방법 (e.g. Lordo 와 Wolfe (1994))들이 새로운 기각치표를 필요로 하는것과는 달리 현존하는 Friedman 통계량에 대한 기각치표 (e.g. Hollander 와 Wolfe (1973))를 이용한다. 불완전자료에서 Friedman 통계량의 정확한 귀무분포표는 여러가지 object 수 k, judge 수 n와 결측치의 형태에 의존함으로 상당히 방대하다. 대안으로 Monte Carlo 시뮬레이션에 의한 근사적인 귀무분포표를 대신 사용할 수 있으나 본 논문에서는 4절에서 보듯이 제1종의 오류를 비교적 잘 통제하는 기존의 기각치표를 사용한다.

2.2. S_{max} 와 S_{min} 계산을 위한 알고리즘

완전한 자료에 대한 Friedman 통계량 S을 다음과 같이 표현할 수 있다.

$$S = [12/nk(k+1)] \sum_{j=1}^{k} R_j^2 - 3n(k+1),$$

여기서 R_j 는 object j ($j = 1, \ldots, k$)에서 순위합이다. 따라서 S를 최대 [최소]화 하는 것은 $\sum_{j=1}^{k} R_j^2$을 최대 [최소]화 하는 것이고 이는 R_j들을 가능한 멀리 떨어지도록 [가깝도록]

결측 object에 순위를 부여함으로써 얻어진다. 다음 단계들은 $S_{max}[S_{min}]$을 얻기 위한 각 judge들이 결측 object에 순위를 할당하는 시스템을 설명한다.

단계 1. 가장 적은 결측순위를 갖고 있는 judge들로 구성된 집합 A를 만든다.

단계 2. 전체 k개의 완전한 순위를 갖고 있는 judge들로 구성된 집합 B를 만든다. 만약 집합 B에 둘 이상의 judge들이 있으면 각각에 object에 대해 순위합을 구한다. 완전 순위를 갖는 judge가 없는 경우는 가장 적은 결측순위를 갖고 있는 judge들을 선택하여 가능한 모든 경우에 대해 순위를 부여 한후 집합 B에 포함시킨다.

단계 3. $S_{max}[S_{min}]$를 얻기위해 집합 A의 judge들이 결측 object에 대해 할당할 순위는 집합 B에 있는 judge들에 대해 순위를 오름차순[내림차순]으로 부여한 다음 이 순서에 기초하여 R_j들을 가능한 멀리 떨어지도록 [가깝도록] 부여안된 순위를 할당한다.

단계 4. 먼저 집합 A에 있는 하나의 judge들에 대해 순위할당이 끝나면 이를 집합 B에 포함시키고 다른 judge를 선택하여 단계 2로 간다. 집합 A에 있는 모든 judge들에 대해 가능한 모든 경우의 할당순서에 의해 위 과정을 반복한다.

단계 1 -4는 집합 B에 할당받은 모든 judge들이 포함될 때까지 계속 반복수행한다. 이렇게 하여 얻어진 완전화된 순위자료에서 $S_{max}[S_{min}]$ 에 상응하는 근사통계량 $S^*_{max}[S^*_{min}]$ 이 얻어진다. 참고로, Lim, Lordo 와 Wolfe (1999)는 단계 4에서 집합 A의 judge들에 대해 서로 독립적으로 순위할당이 이루어짐으로써 $S^*_{max}[S^*_{min}]$를 적은 반복수로서 얻을 수 있으나 정밀도(precision)면에서 떨어진다.
다음은 간단한 불완전 순위자료를 가지고 개발한 알고리즘을 적용시켜보자.

		objects			
		1	2	3	4
judges	1	–	2	–	1
	2	3	1	2	4
	3	1	2	–	–
	4	–	–	–	1

여기서 –는 결측순위를 나타낸다.

먼저 최대 통계량 계산을 위해 위의 알고리즘 단계 1 과 2에서 A={1, 3 }와 B={ 2 }가 얻어진다. 단계 3에서 judge 1과 3의 결측 object들에 대한 순위부여는 judge 1를 먼저 부여하고나서 judge 3를 부여하는 방법과 judge 3를 먼저 부여한후 judge 1를 부여하는 두가지 경우가 있다. 여기서는 최대 통계량값을 얻는 후자인 경우만을 살펴보겠다. B={ 2 }에 있는 judge 2의 object 4에 대한 순위가 object 3의 순위보다 높기 때문에 judge 3의 object 3 과 4에 순위 3 과 4를 할당한다. 단계 4에서는 집합 B={2, 3}가 얻어지고 A={1}에 있는 judge 1에 대한 순위부여를 위해 단계 2로 되돌아간다. judge 1의 결측 object 1, 3에 순위부여는 집합 B에 있는 judge들의 이들 object에 대응하는 순위합이 4 와 5 이므로 각각 순위

3 과 4를 부여한다. 그리고 judge 4의 결측 object 1, 2, 3에 대한 순위부여는 B={1, 2, 3}의 순위합이 7, 5, 9이므로 각각 3, 2, 4가 부여된다. 최소 통계량값 계산도 위와 똑같은 방법에 의해 얻어진다. 따라서 S^*_{max}와 S^*_{min}를 낳는 완전한 순위자료는 각각 다음과 같다.

		objects 1	2	3	4
judges	1	**3**	2	**4**	1
	2	3	1	2	4
	3	1	2	**3**	**4**
	4	**3**	**2**	**4**	1

		objects 1	2	3	4
judges	1	**3**	2	**4**	1
	2	3	1	2	4
	3	1	2	**3**	**4**
	4	**3**	**4**	**2**	1

위 자료에서 S^*_{max} =2.7이고 S^*_{min}= 0.3이다.

참고로, Lim, Lordo 와 Wolfe (1999)의 최대 [최소] 통계량값을 구하는 방법에 대해 살펴보자. 먼저 최대 통계량 값을 얻기위해 알고리즘 단계 1, 2에서 얻어진 A={1, 3 }에 있는 judge 1, 3의 순위 할당은 위의 제안된 방법에서는 judge 1과 3의 두가지 할당순서 모두를 고려한 반면 이 논문에서는 할당순서와 무관하게 B={ 1 } 에 기초해 judge 1의 object 1과 3에 순위 4와 3, judge 3의 object 3 과 4에 순위 3과 4를 부여한다. judge 4에 순위할당은 B={1, 2, 3}의 순위합에 기초해 행해진다. 최소 통계량 값도 위와 같은 방법에 의해 얻어진다. 이렇게 하여 얻어진 완전한 순위자료는 각각 다음과 같다.

		objects 1	2	3	4
judges	1	**4**	2	**3**	1
	2	3	1	2	4
	3	1	2	**3**	**4**
	4	**3** or **4**	**2**	**4** or **3**	1

		objects 1	2	3	4
judges	1	**3**	2	**4**	1
	2	3	1	2	4
	3	1	2	**4**	**3**
	4	**3**	**4**	**2**	1

위 자료에서 S^*_{max} =2.1와 S^*_{min}= 0.9가 얻어진다. 지금까지 살펴보았듯이 Lim, Lordo 와 Wolfe의 방법은 결측 순위을 할당하는데 있어서 단순한 반면에 제안된 알고리즘에 의해 얻어진 통계량값들(이 경우 S^*_{max} =2.7이고 S^*_{min}= 0.3)보다 덜 정확함을 알 수 있다.

2.3. 제안된 알고리즘 성능평가

2.2절에서 제안된 알고리즘에 의해 계산된 S^*_{max} 와 S^*_{min} 은 결측 object들에 대해 부여할 수 있는 모든 순위를 고려함으로써 얻어진 참값 S_{max} 와 S_{min} 와는 언제나 일치하지는 않는다. $j = 1, \ldots, k$에 대하여 $n_{(j)}$를 object 1에서 j까지 순위를 부여한 judge들의 수라 할때 불완전 계획법 (k, n)=(4,4) 와 (5,5) 그리고 여러가지 $(n_{(1)}, \ldots, n_{(k-2)}, n_{(k)})$하에서 1000번 반복하여 $S_{max} = S^*_{max}$ 과 $S_{min} = S^*_{min}$의 되는 횟수를 비교하고자한다. 그리고 식 (2.1)에서 기초한 검정법과 식 (2.1)에서 S_{max}와 S_{min} 대신 S^*_{max}와 S^*_{min}를 사용한 검정법 즉,

표 2.1: 알고리즘 평가를 위한 모의실험비교 (1000 번 반복실험)

실험환경	같은값을 갖는 횟수		기각수	
$(n_{(1)},\cdots,n_{(k-2)},n_{(k)})$	$\sharp[S_{max}=S^*_{max}]$	$\sharp[S_{min}=S^*_{min}]$	♯[모든가능한경우]	♯[알고리즘]
	(n,k)=(4,4), α=0.052, $s(.052,4,4)$=7.5			
(0, 2, 2)	974	992	77	77
(2, 0, 2)	957	834	202	202
(2, 2, 0)	921	618	457	457
(2, 1, 1)	736	708	304	312
(1, 1, 2)	930	942	134	134
(1, 2, 1)	866	899	203	206
	(n,k)=(5,5), α=0.049, $s(.049,5,5)$=8.96			
(1,1,1,2)	755	773	222	229
(1,2,1,1)	587	674	318	346
(2,0,3,0)	929	691	427	427
(0,3,0,2)	822	912	213	213
(2,1,0,2)	678	557	432	443
(0,2,2,1)	742	821	167	174

$$\begin{aligned}&S^*_{min}\geq s(\alpha,k,n)\ \text{혹은}\\&\frac{S^*_{max}+S^*_{min}}{2}\geq s(\alpha,k,n)\ \text{이면}\ H_0\ \text{기각하고}\\&S^*_{max}< s(\alpha,k,n)\ \text{혹은}\\&\frac{S^*_{max}+S^*_{min}}{2}< s(\alpha,k,n)\ \text{이면}\ H_0\ \text{기각못한다}\end{aligned} \tag{2.2}$$

이 H_0를 기각하는 횟수를 비교하였다. 그 결과는 표 2.1에 나타나 있다. 표 2.1는 제안된 알고리즘이 참값을 얻는데 매우 효과적임을 보여주고 있다. 비록 (k,n)=(5,5)에서 $S_{max}=S^*_{max}$와 $S_{min}=S^*_{min}$이 되는 횟수가 60 %미만인 경우에도 H_0를 기각하는 횟수에서는 거의 차이가 없음을 알 수 있다.

그리고 우리 알고리즘에 의해 S^*_{max}와 S^*_{min}이 S_{max}와 S_{min}에 비하여 어느정도 쉽게 얻어지는가를 보기 위해 각각의 결과값들을 얻는데 요구되는 경우의 수(실험계획 수)를 평균적으로 비교하였다. 그 결과는 표 2.2에 수록되어있다. (k,n)=(4,4)에서 결측율이 낮은 $(n_{(1)},n_{(2)},n_{(4)})$=(0,2,2)를 제외한 모든 경우에 대하여 우리의 알고리즘은 상대적으로 작은 경우의 수를 필요함을 알 수 있었다. 특히, (k,n)=(5,5)에서 $(n_{(1)},n_{(2)},n_{(3)},n_{(5)})=$ (1,2,1,1)와 (2,1,0,2)인 경우 평균비율이 1 % 미만으로 경우의 수에서 현저한 차이가 있음을 알 수 있다.

3. 예제

A, B, C, D, E 다섯종류의 수면제를 15명의 불면증 환자에게 임상실험하려고 한다. 이때에 15명의 환자는 연령, 건강상태, 불면증의 정도 등이 서로 심각하게 다르기 때문에 이

표 2.2: $S_{max}[S_{min}]$와 $S^*_{max}[S^*_{min}]$를 얻는데 요구되는 평균수에 대한 비율 (1000 번 반복실험)

실험환경 $(n_{(1)}, n_{(2)}, \cdots, n_{(k-1)}, n_{(k)})$	평균비율 $\sharp[S^*_{max}]/\sharp[S_{max}]$	$\sharp[S^*_{min}]/\sharp[S_{min}]$
(n,k)=(4,4)		
(0, 2, 2)	1.281	1.326
(2, 0, 2)	0.207	0.271
(2, 2, 0)	0.092	0.151
(2, 1, 1)	0.045	0.075
(1, 1, 2)	0.249	0.276
(1, 2, 1)	0.136	0.144
(n,k)=(5,5)		
(1,1,1,2)	0.013	0.021
(1,2,1,1)	0.003	0.006
(2,0,3,0)	0.015	0.038
(0,3,0,2)	0.117	0.171
(2,1,0,2)	0.003	0.009
(0,2,2,1)	0.047	0.064

들을 비슷한 그룹끼리 묶어 다섯그룹으로 만들고 각 그룹내에서 랜덤하게 A, B, C, D, E를 적용하여 약효를 비교하였다. 표 3.1은 약효를 나타낸 순위자료이다. 각 그룹내에서 가장 효능이 좋다고 생각되는 약에 1, 가장 효능이 떨어진다고 생각되는 약에 5의 순위를 표시하고 관측이 불가능하다든지 혹은 불충분한 정보로 인해 순위 부여가 어려운 경우 −로 나타내었다. 제안된 검정법을 사용하여 수면제의 종류에 따라 약효에 심각한 차이가 있는지를 알아보고자 한다.

표 3.1: 수면제의 약효자료

그룹	수면제 A	B	C	D	E
1	1	-	2	-	-
2	-	3	2	1	-
3	4	2	1	3	5
4	-	1	-	-	-
5	-	2	-	1	-

2.2절에 소개된 알고리즘에 의해 S^*_{max} 와 S^*_{min} 를 낳는 완전화된 순위자료는 각각 다음과 같다.

	수면제				
그룹	A	B	C	D	E
1	1	**4**	2	**3**	**5**
2	4	3	2	1	**5**
3	4	2	1	3	5
4	4	1	**3**	**2**	**5**
5	**4**	2	**3**	1	**5**

	수면제				
그룹	A	B	C	D	E
1	1	**5**	2	**4**	**3**
2	**5**	3	2	1	**4**
3	4	2	1	3	5
4	**3**	1	**4**	**5**	**2**
5	**4**	2	**5**	1	**3**

위 완전화된 자료에서 S^*_{max} =12.32 와 S^*_{min}= 1.12 이 얻어진다. 따라서 식 (2.2)에서 $(S^*_{max}+S^*_{min})/2$= 6.72 이고 α =0.049 에서 기각치 $s(0.049,5,5)$=8.96 이므로 H_0를 기각할 수 없다. 즉, 다섯 수면제의 약효에는 심각한 차이가 있다는 근거가 없다는 결론을 내린다.

4. 모의실험

우리는 식 (2.2)에 주어진 검정법과 평균순위법과 검정력을 비교하고자 한다. 이를 위해 judge수 와 object수가 같은 3개의 불완전 계획법 (k,n)=(4,4), (5,5), (6,6)를 생각하였고 object들의 유사(similarity) 측도로서 평균 $\theta_1,\ldots,\theta_k$ 이고 분산 1인 정규분포로 부터 $\Delta=[\sum_{j=1}^{k}(\theta_j-\bar{\theta})^2]^{1/2}$ 를 생각하였다. 여기서 $\bar{\theta}=\sum_{j=1}^{k}\theta_j/k$ 이다.

(k,n)=(4,4)인 경우 Δ=0,2,4,6에 대응하는 $(\theta_1,\ldots,\theta_4)$ = (0,0,0,0), (-1,-1,1,1), (-2,-2,2,2), (-3,-3,3,3)와 (k,n) = (5,5)인 경우 $(\theta_1,\ldots,\theta_5)$ = (0,0,0,0,0), (-1,-1,0,1,1), (-2,-2,0,2,2),(-3,-3,0,3,3) 그리고 (k,n) = (6,6)인 경우 $(\theta_1,\ldots,\theta_6)$ = (0,0,0,0,0,0), (-1,-1,0,0,1,1), (-2,-2,0,0,2,2), (-3,-3,0,0,3,3)를 생각하였다.

(k,n) = (4,4)인 경우에 고려한 유의수준과 이에 대한 기각치는 $\alpha=0.052$ 와 0.094에서 $s(0.052,4,4)=7.5$와 $s(0.094,4,4)=6.3$이고 (k,n) = (5,5)인 경우는 $\alpha=0.049$ 와 0.094에서 $s(0.049,5,5)=8.96$와 $s(0.094,5,5)=7.68$이다. (k,n) = (6,6)인 경우에는 대표본에서 Friedman 통계량의 귀무분포가 카이제곱분포에 근사한다는 성질을 이용하여 근사적인 기각치 $\alpha=0.05$와 0.10에서 각각 $s(0.05,6,6)=11.07$ 와 $s(0.10,6,6)=9.24$를 이용하였다.

k, n, 그리고 $(\theta_1,\ldots,\theta_k)$에 대한 불완전 순위자료를 얻기위해 먼저 IMSL 루틴 RNNOR으로부터 평균 $\theta_1,\ldots,\theta_k$ 이고 분산이 1인 kn개의 정규난수를 생성시키고 지정된 결측율을 만족하도록 IMSL루틴 RNSRI로부터 자료를 제거한다. 여기서 결측율은 10, 20, 30, 40, 50 % 를 고려하였다. 표 4.1, 4.2, 4.3에 각각 (k,n) = (4,4), (5,5), (6,6)에 대한 모의실험 결과가 수록되어있다. 우리는 위 실험으로부터 몇가지 결론을 얻을 수 있다. 첫째, 평균순위법은 제안된 검정법에 비하여 매우 보수적이다. 특히, 결측율이 증가할수록 더욱 그렇다. 둘째, 제안된 검정법은 평균순위법보다 좋은 검정력을 갖고 있다. 특히, 결측율이 증가할수록 차이가 커짐을 알 수 있다. 몇몇 실험환경하에서는 약 2배정도의 차이가 있음을 보여준다.

표 4.1: (k,n)=(4,4)인 경우 추정된 검정력 (10,000 번 반복실험)

Δ	결측율(%)	α	제안된 검정법	평균순위법
0	10	0.052 (0.094)	0.0506 (0.0924)	0.0506 (0.0924)
	20		0.0450 (0.0908)	0.0424 (0.0905)
	30		0.0394 (0.0829)	0.0349 (0.0813)
	40		0.0344 (0.0685)	0.0250 (0.0580)
	50		0.0301 (0.0630)	0.0150 (0.0321)
2	10	0.052 (0.094)	0.4274 (0.5699)	0.4274 (0.5699)
	20		0.1889 (0.3256)	0.1809 (0.3254)
	30		0.1375 (0.2405)	0.1240 (0.2401)
	40		0.0884 (0.1338)	0.0621 (0.1164)
	50		0.0418 (0.0742)	0.0255 (0.0456)
4	10	0.052 (0.094)	0.5823 (0.7736)	0.5823 (0.7736)
	20		0.2558 (0.4602)	0.2450 (0.4602)
	30		0.1874 (0.3242)	0.1693 (0.3242)
	40		0.1094 (0.1611)	0.0765 (0.1395)
	50		0.0473 (0.0815)	0.0296 (0.0509)
6	10	0.052 (0.094)	0.5847 (0.7797)	0.5847 (0.7797)
	20		0.2598 (0.4647)	0.2466 (0.4647)
	30		0.1889 (0.3278)	0.1707 (0.3278)
	40		0.1105 (0.1625)	0.0771 (0.1406)
	50		0.0476 (0.0820)	0.0297 (0.0513)

표 4.2: (k, n)=(5;5)인 경우 추정된 검정력 (10,000 번 반복실험)

Δ	결측율(%)	α	제안된 검정법	평균순위법
0	10	0.049(0.094)	0.0462 (0.0924)	0.0458 (0.0923)
	20		0.0438 (0.0936)	0.0415 (0.0882)
	30		0.0391 (0.0883)	0.0342 (0.0769)
	40		0.0410 (0.0905)	0.0262 (0.0598)
	50		0.0429 (0.0930)	0.0191 (0.0444)
2	10	0.049 (0.094)	0.5409 (0.6860)	0.5397 (0.6852)
	20		0.2699 (0.3940)	0.2608 (0.3833)
	30		0.1646 (0.2686)	0.1505 (0.2459)
	40		0.0932 (0.1644)	0.0676 (0.1241)
	50		0.0699 (0.1333)	0.0360 (0.0728)
4	10	0.049 (0.094)	0.8927 (0.9614)	0.8915 (0.9614)
	20		0.4854 (0.6102)	0.4734 (0.6009)
	30		0.2775 (0.4033)	0.2580 (0.3799)
	40		0.1346 (0.2161)	0.1005 (0.1718)
	50		0.0874 (0.1582)	0.0486 (0.0914)
6	10	0.049 (0.094)	0.9248 (0.9925)	0.9229 (0.9925)
	20		0.5306 (0.6590)	0.5210 (0.6513)
	30		0.3040 (0.4391)	0.2840 (0.4158)
	40		0.1460 (0.2276)	0.1102 (0.1843)
	50		0.0926 (0.1624)	0.0517 (0.0971)

표 4.3: (k,n)=(6,6)인 경우 추정된 검정력 (10,000 번 반복실험)

Δ	결측율(%)	α	제안된 검정법	평균순위법
0	10	0.05 (0.10)	0.0382 (0.0886)	0.0382 (0.0886)
	20		0.0369 (0.0877)	0.0352 (0.0851)
	30		0.0339 (0.0808)	0.0297 (0.0705)
	40		0.0353 (0.0869)	0.0212 (0.0560)
	50		0.0454 (0.1167)	0.0104 (0.0378)
2	10	0.05 (0.10)	0.6076 (0.7616)	0.6075 (0.7614)
	20		0.3115 (0.4654)	0.3035 (0.4585)
	30		0.1870 (0.3155)	0.1690 (0.2890)
	40		0.0999 (0.1919)	0.0692 (0.1402)
	50		0.0747 (0.1628)	0.0256 (0.0672)
4	10	0.05 (0.10)	0.9736 (0.9948)	0.9736 (0.9948)
	20		0.6317 (0.7750)	0.6237 (0.7690)
	30		0.3636 (0.5256)	0.3408 (0.5003)
	40		0.1622 (0.2972)	0.1179 (0.2194)
	50		0.1003 (0.2126)	0.0402 (0.0928)
6	10	0.05 (0.10)	0.9979 (1.0000)	0.9979 (1.0000)
	20		0.7195 (0.8462)	0.7130 (0.8413)
	30		0.4287 (0.5926)	0.4039 (0.5664)
	40		0.1868 (0.3333)	0.1403 (0.2538)
	50		0.1088 (0.2257)	0.0445 (0.1000)

참고문헌

[1] Friedman,M.(1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 32*, 675-701.

[2] Hettmansperger,T.P. and Norton,R.M.(1987). Tests for patterned alternatives in k-sample problem. *Journal of the American Statistical Association 82*, 292-99.

[3] Hollander,M. and Wolfe,D.A.(1973). *Nonparametric Statistical Methods.* New York: John Wiley & Sons.

[4] Kendall,M.G. and Babington Smith,B.(1939). The problem of m rankings. *Annals of Mathematical Statistics 10*, 275-87.

[5] Lim,D.H., Lordo,R.A., and Wolfe,D.A.(1999). A screening analysis for incomplete ranking data. to appear in *Journal of Royal Statistical Society Series D, The Statistician.*

[6] Lordo,R.A. and Wolfe,D.A.(1994). Statistical methodology for incomplete ranked data. *Parisankhyan Samikkha 1*, 7-15.

[7] Page,E.B.(1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks. *Journal of the American Statistical Association 58*, 216-30.

[1998년 7월 접수, 1999년 1월 최종수정]

An Efficient Test in a Randomly Incomplete Two-way Rank Design

Dong Hoon Lim [1]

ABSTRACT

Consider the setting where a group of n judges are to independently rank a series of k objects, but where complete rankings are not received from each of the judges. In this paper we propose an efficient procedure to test whether the treatment effects are the same, versus a general alternative that the effects differ with incomplete rankings. We describe an algorithm for implementing our procedure without the aid of a computer program and present the results of a Monte Carlo simulation comparing power performances of the average ranks procedure with our proposed procedure.

1) Assistant Professor, Department of Statistics, Information & Telecommunication Research Center, Gyeongsang National University, Chinju, 660-701, Korea