

Nonparametric procedures using aligned method and linear placement statistics in randomized block design

Jinjoo Han^a · Dongjae Kim^{a,1}

^aDepartment of Biomedicine · Health Science, The Catholic University of Korea

(Received September 30, 2016; Revised November 14, 2016; Accepted November 23, 2016)

Abstract

Nonparametric procedures in randomized block design was proposed by Friedman (1937) as a general alternative. This method is used to find out the difference in treatment effect. It can cause a loss of inter-block information using the ranking in each block. This paper proposed nonparametric procedures using an aligned method proposed by Hodges and Lehmann (1962) to reduce block information based on joint placement suggest by Jo and Kim (2013) in a randomized block design. We also compared the power of the test of the proposed procedures and established method through a Monte Carlo simulation.

Keywords: nonparametric, randomized block design, aligned method, joint placement

1. 서론

치료 효과를 비교하는 임상시험에서 환자의 연령, 성별 또는 병의 중증도는 치료와 더불어 치료성적에 영향을 미친다. 이와 같이 치료 효과에 영향을 미칠 수 있다고 짐작되는 환자의 특성을 공변수라 부르는데, 치료 효과는 공변수의 영향을 제거한 후 비교되어야 한다. 이것이 가능한 실험방식이 랜덤화 블록 계획법(randomized block design)이다. 이는 처리가 3개 이상일 때 처리 효과의 차이 유무를 알기 위한 방법으로 연구 대상을 비슷한 특성을 가진 블록(block)으로 구분한 다음, 무작위로 한 가지의 처리 수준에 한 명의 연구대상을 할당하는 방법이다. 처리 간의 효과 차이 유무를 검정하기 위해서는 오차가 서로 독립이고 정규분포 $N(0, \sigma^2)$ 을 따르는 확률변수라는 가정이 성립된다면 모수적 방법인 분산분석법을 사용해 처리 효과들이 모두 같다는 귀무가설을 검정할 수 있다. 하지만 위의 가정이 만족하지 않을 때에는 분산분석법 사용 시 문제점이 발생하기 때문에 제 1종 오류를 제어할 수 있는 비모수적 검정법을 선택해야 한다 (Song과 Lee, 1995).

랜덤화 블록 계획법에서 처리 효과의 차이를 알아보기 위한 검정에서 모수적인 방법으로는 분산분석법을 사용하고, 비모수적 방법으로는 Friedman (1937)이 제안한 각 처리의 효과가 적어도 하나는 다르다는 일반 대립가설에 관한 검정법이 있다. Friedman (1937)이 제안한 검정법은 각 블록 내에서 k 개의 관측값들을 작은 것부터 차례로 순위를 부여하여 처리 간의 효과 차이를 검정한다. 순서 대립가설에 대한 검정법으로는 Page (1963)가 제안한 방법이 있는데, 이 검정법은 블록 안에서의 순위(rank)를 이용해 처리별 가중치를 적용하여 통계량을 계산한다. 하지만 Friedman (1937)이 제안한 방법과 Page

¹Corresponding author: Department of Biomedicine · Health Science, The Catholic University of Korea, 222, Banpo-daero, Seocho-gu, Seoul 06591, Korea. E-mail: djkim@catholic.ac.kr

(1963)가 제안한 방법은 블록 내 순위를 사용해 검정함으로써 블록 간의 정보 손실이 있다는 단점이 존재한다. Jo와 Kim (2013)은 Hodges와 Lehmann (1962)이 제안한 정렬방법을 이용하였고, Chung과 Kim (2007)의 일원배치모형 검정법을 확장하여 랜덤화 블록 계획법에서의 결합위치 방법을 제안하였다. 이 방법은 처리 효과가 크고 블록 크기가 클수록 Kruskal-Wallis (1952) 검정법보다 효율적이다.

반복이 있는 랜덤화 블록 계획법에서 일반 대립가설에 대한 검정법으로는 Mack과 Skillings (1980)이 제안한 방법과 Mack (1981)이 제안한 방법이 있고, 순서 대립가설에 대해서는 Skillings와 Wolfe (1977, 1978), Hettmansperger (1975)가 제안한 검정법이 있다. Mack과 Skillings (1980)이 제안한 검정법은 Friedman (1937)의 방법을 적용한 것인데, 이 방법은 각 블록의 처리에서 반복된 각 관측값 대신에 반복된 관측값들의 평균을 이용하여 순위를 매기기 때문에 모든 관측의 정보를 이용하지 않아 정보의 손실이 있을 수 있다. Lee와 Kim (2012)은 Kruskal-Wallis (1952) 방법과 Jonckheere (1954) 방법에 Hodges와 Lehmann (1962)이 제안한 정렬방법을 적용하여 블록 간의 정보를 이용하는 일반 대립가설에 대한 검정법을 제안하였다. Kang과 Kim (2014)은 Page (1963)가 제안한 방법을 확장하여 순서 대립가설에 대한 방법을 제안하였고, Kim과 Kim (2016)은 Hodges와 Lehmann (1962)이 제안한 정렬방법을 이용하여 블록 간의 정보를 사용하고 일원배치모형에 대해 Kim (1999)이 제안한 대조군과 처리군의 방법을 적용한 순서 대립가설에 대한 비모수적 방법을 제안하였다.

Orban과 Wolfe (1982)는 두 처리 간 효과 차이를 검정하기 위해 위치(placement)를 사용한 비모수적 검정법을 제안하였다. 이 방법은 두 처리 중 어느 한 처리에 대한 상대적 위치정보를 이용해 처리 효과 차이를 검정하는 방법으로써, 대조군의 블록 크기가 처리군의 블록 크기보다 클 때 더 유용하다고 알려져 있다. Chung과 Kim (2007)은 Orban과 Wolfe (1982)가 제안한 비모수적 도구인 위치를 일원배치모형으로 확장하여 결합위치(joint placement)에 대한 검정법을 제안하였다. Jeon과 Kim (2016)은 Chung과 Kim (2007)이 제안한 결합위치에 점수함수를 적용한 선형 위치 통계량(linear placement statistic)을 제안하였다.

이 논문에서는 Hodges와 Lehmann (1962)의 블록 간의 정보를 사용하는 정렬방법(aligned method)과 Jo와 Kim (2013)의 랜덤화 블록 계획법의 결합위치를 확장하여 결합위치에 점수함수(score function)를 적용한 선형 위치 통계량을 이용하여 비모수적 검정법을 제안하였다. Monte Carlo simulation을 통하여 랜덤화 블록 계획법에서 모수적 검정법인 분산분석법과 비모수적 검정법인 Friedman (1937) 방법, 본 논문에서 제안한 검정법의 검정력(power)을 비교하였다.

2. 정렬방법을 이용한 결합위치(Joint Placement)

블록이 있고 처리의 개수가 k 개인 랜덤화 블록 계획법 모형

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad (i = 1, 2, \dots, k; j = 1, 2, \dots, n)$$

에서 Y_{ij} 는 j 번째 블록에서 i 번째 처리의 반응값이고, μ 는 전체 평균, τ_i 는 처리 효과, β_j 는 블록 효과, ϵ_{ij} 는 오차항이며 동일한 연속분포를 따르는 서로 독립인 확률변수를 가정한다. 블록 간의 정보를 이용하기 위해 Hodges와 Lehmann (1962)이 제안한 정렬방법을 적용하여 생성된 정렬자료는

$$Y_{ij}^* = Y_{ij} - \bar{Y}_{\cdot j}$$

이고, 여기서

$$\bar{Y}_{\cdot j} = \sum_{i=1}^k \frac{Y_{ij}}{k}$$

는 각 블록의 효과인 블록평균이다.

모집단에 대하여 구체적인 분포함수를 가정할 수 없는 경우에 각 처리의 효과가 모두 동일하다는 귀무가설과 각 처리의 효과가 적어도 하나는 다르다는 일반 대립가설은 다음과 같다.

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

자료를 정렬시켜 블록 효과를 없애 자료 전체에 대한 결합위치(joint placemet) O_{ij} 를 다음과 같이 정의한다.

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

결합위치 O_{ij} 는 i 번째 처리의 관측값을 제외한 모든 처리에서 Y_{ij}^* 보다 작거나 같은 관측값의 개수를 이용한 $[0, 1]$ 에서 정의된 확률변수이다. 이를 이용한 선형 위치 통계량은 다음과 같다.

$$S^{\Phi_H} = \sum_{i=1}^k \sum_{j=1}^n \Phi(O_{ij})$$

여기서 $\Phi(\cdot)$ 는 $[0, 1]$ 에서 정의된 실수값인 점수함수이다.

Hong과 Lee (2014)에 의해, 귀무가설 하에서 $n \rightarrow \infty$ 이면 표준화된 통계량 \hat{S}^{Φ_H} 의 분포는 표준정규분포로 수렴하게 된다. 즉, 검정통계량 S^{Φ_H} 의 근사분포는

$$\hat{S}^{\Phi_H} = \frac{S^{\Phi_H} - E(S^{\Phi_H})}{\sqrt{\text{Var}(S^{\Phi_H})}} \xrightarrow{N \rightarrow \infty} N(0, 1)$$

이다.

선형 위치 통계량을 만드는 점수함수로 $\Phi_H(x) = |1 - 2x|$ 를 고려했고, 점수함수에 따른 선형 위치 통계량은

$$S^{\Phi_H} = \sum_{i=1}^k \sum_{j=1}^n |1 - 2O_{ij}|$$

이며, S^{Φ_H} 의 기대값은 귀무가설 하에서

$$E(S^{\Phi_H}) = \begin{cases} \frac{nk(nk - n + 2)}{2(nk - n + 1)}, & n \text{이 짝수일 때,} \\ \frac{nk(nk - n + 1)}{2n(k - 1)}, & n \text{이 홀수일 때} \end{cases}$$

이다. 한편 S^{Φ_H} 의 분산은 임의의 n 과 k 에 대하여는 구체적인 식으로 표현할 수 없고 특정한 n 과 k 에 대하여 계산할 수 있다.

3. 모의실험의 계획 및 결과

본 논문에서는 결합위치를 이용한 선형 위치 통계량에 근거하여 새롭게 제시한 검정법과 기존의 검정법들을 비교하기 위해 모의실험을 시행하였다. 모수적인 방법으로는 F -통계량을 이용한 분산분석

법(ANOVA)을 사용하고, 비모수적인 방법으로는 Friedman (1937)이 제안한 방법을 사용하였다. 모집단의 분포는 대칭분포인 정규분포, 이중지수분포, Cauchy 분포와 비대칭분포인 지수분포를 채택하였다. 정규분포, Cauchy 분포, 지수분포의 난수생성은 SAS의 RANNOR 함수, RANCAU 함수, RANEXP 함수를 이용하였고, 이중지수분포는 RANUNI 함수를 이용하여 역변환 방법으로 난수를 생성하였다.

처리의 수는 4개와 6개일 경우를 선택하였고, 블록의 수는 5개와 10개일 경우를 고려하였다. 유의수준 α 는 0.05로 하고 여러 가지 처리 조합에 대하여 검정력을 비교하였다. 각 검정통계량의 검정력을 비교하기 위해 10,000번 반복하는 Monte Carlo simulation을 시행하였다. 모의실험 결과는 처리의 수준 개수에 따라 나누어 처리가 4개일 때의 결과는 Table 3.1, 처리가 6개일 때의 결과는 Table 3.2에 정리하였다.

처리의 수가 4개이고 각각의 처리 효과가 모두 같을 때 유의수준이 0.05를 잘 만족하는지 살펴보면, 모집단의 분포가 정규분포인 경우 블록이 5개, 10개일 때, ANOVA는 0.048, 0.047로 0.05에 근사한 값을 얻었고 제안한 방법은 0.05, 0.049의 값을 얻었다. 이를 통해 ANOVA와 제안한 방법이 제 1종 오류를 잘 제어하고 있음을 알 수 있다. 지수분포와 이중지수분포에서는 제안한 방법이 0.05에 근사한 값을 얻었고, ANOVA는 0.042, 0.045로 0.04에 근사한 값을 얻었다. 이는 제안한 방법은 정규분포일 때와 마찬가지로 제 1종 오류를 잘 제어하고 있으나, ANOVA는 정규분포의 결과에 비해 제 1종 오류를 제어함에 있어 어려움이 있다는 사실을 알 수 있다. Cauchy 분포의 경우 ANOVA는 블록에 따라 0.018, 0.016의 값을 얻었는데, 정규성 가정을 만족하지 못할 때 제 1종 오류를 제어하기가 어렵다는 이전의 내용과 같은 결과이다. 또한 제안한 방법은 블록의 수가 5개일 때 0.0394, 블록의 수가 10개일 때 0.0452로 0.04에 근사한 값을 보였는데 다른 분포에 비해서 조금 보수적이라는 것을 알 수 있다. Friedman (1937)이 제안한 방법은 모든 분포에서 블록의 수가 5개일 때는 0.064, 0.07, 0.064, 0.069의 값을, 블록의 수가 10개일 때는 0.051, 0.051, 0.053, 0.057의 값을 얻었는데, Friedman (1937)의 검정법은 블록 크기가 작을 때 블록 크기가 큰 경우보다 제 1종 오류를 제어함에 있어 어려움이 있다는 점을 알 수 있다.

처리의 수가 6개일 때 정규분포를 보면 ANOVA는 블록의 수에 따라 0.047, 0.053으로 0.05에 근접한 값을 얻었고, 제안한 방법 또한 블록의 수에 따라 0.05, 0.048으로 0.05에 근사한 값을 얻은 것으로 확인할 수 있다. 지수분포와 이중지수분포의 경우 제안한 방법은 0.05에 근사한 값을 얻었고, ANOVA는 0.04에 근사한 값을 얻었는데, 이는 처리의 수가 4개일 때와 비슷한 결과로 다른 방법에 비해 조금 보수적이라는 것을 알 수 있다. Cauchy 분포에서 ANOVA는 0.017, 0.015의 값으로 나타난 것으로 보아 제 1종 오류를 제어함에 있어 어려움이 있다는 결과를 보인다. Friedman (1937)이 제안한 방법은 모든 분포에서 블록의 수가 5개일 때는 0.0558, 0.0609, 0.0611, 0.0548, 블록의 수가 10개일 때는 0.0541, 0.0508, 0.0532, 0.0531의 값을 얻었다. 블록의 수가 10개인 경우에 처리의 수가 4개이고 블록의 수가 5개일 때보다 작은 값을 보이지만, 이와 비슷하게 블록의 크기가 작을 때 블록 크기가 큰 경우보다 다소 높은 값이 나타난 것을 알 수 있다.

검정력을 비교해보면 처리의 수와 관계없이 정규분포에서 블록 크기가 클수록 모수적 방법인 ANOVA의 검정력이 높다는 것을 알 수 있다. 하지만 정규분포를 제외한 지수분포, 이중지수분포, Cauchy 분포에서는 ANOVA가 제안된 다른 비모수 방법들에 비해 대체로 낮은 검정력을 보인다.

처리의 수가 4개일 때, 모집단의 분포가 지수분포인 경우 블록의 수가 5개일 때 대부분의 패턴에서 제안한 방법이 ANOVA보다 높거나 비슷한 결과를 보였고, 블록의 수가 10개일 때는 8개 패턴의 모든 대립 가설에서 제안한 방법의 검정력이 ANOVA보다 더 높게 나타났다. 블록 크기에 관계없이 급격하게 증가했다가 급격하게 낮아지는 패턴에서는 Friedman (1937)의 검정법보다 제안한 방법의 검정력이 더 높

Table 3.1. Monte Carlo power estimates in randomized block design ($\alpha = 0.05$, treatment = 4)

Dist.	block	τ_1	τ_2	τ_3	τ_4	A	F	O	
Normal	5	0	0.0	0.0	0.0	0.0484	0.0649	0.0502	
		0	0.3	0.5	1.0	0.1813	0.1927	0.1499	
		0	0.5	1.5	1.5	0.5430	0.5083	0.4718	
		0	1.5	0.3	2.0	0.7586	0.7102	0.6793	
		0	2.0	1.5	0.0	0.8266	0.7749	0.7695	
		0	0.8	0.8	0.3	0.1587	0.1699	0.1373	
	10	0	1.0	1.0	0.0	0.3511	0.3470	0.3214	
		0	0.3	1.5	2.0	0.4816	0.4559	0.3934	
		0	0.0	0.0	0.0	0.0475	0.0511	0.0490	
		0	0.3	0.5	1.0	0.4025	0.3418	0.3257	
		0	0.5	1.5	1.5	0.9152	0.8396	0.8304	
		0	1.5	0.3	2.0	0.9903	0.9658	0.9626	
	Exponential	5	0	2.0	1.5	0.0	0.9969	0.9813	0.9814
			0	0.8	0.8	0.3	0.3600	0.3093	0.3078
			0	1.0	1.0	0.0	0.6952	0.5898	0.5889
			0	0.3	0.8	0.8	0.3678	0.3089	0.2954
			0	0.0	0.0	0.0	0.0425	0.0700	0.0492
			0	0.3	0.5	1.0	0.2231	0.3140	0.2228
10		0	0.5	1.5	1.5	0.5936	0.6589	0.6140	
		0	1.5	0.3	2.0	0.7726	0.8053	0.7986	
		0	2.0	1.5	0.0	0.8237	0.8377	0.8511	
		0	0.8	0.8	0.3	0.2016	0.3042	0.2231	
		0	1.0	1.0	0.0	0.3930	0.4978	0.4641	
		0	0.3	1.5	2.0	0.5671	0.6255	0.5626	
Double exponential		5	0	0.0	0.0	0.0	0.0450	0.0516	0.0502
			0	0.3	0.5	1.0	0.4512	0.5781	0.5113
			0	0.5	1.5	1.5	0.8946	0.9307	0.9205
			0	1.5	0.3	2.0	0.9755	0.9888	0.9873
			0	2.0	1.5	0.0	0.9858	0.9935	0.9946
			0	0.8	0.8	0.3	0.4064	0.5336	0.4460
	10	0	1.0	1.0	0.0	0.7113	0.8040	0.7875	
		0	0.3	0.8	0.8	0.4026	0.5331	0.4493	
		0	0.0	0.0	0.0	0.0417	0.0642	0.0506	
		0	0.3	0.5	1.0	0.1194	0.1594	0.1237	
		0	0.5	1.5	1.5	0.3159	0.3631	0.3301	
		0	1.5	0.3	2.0	0.4832	0.5144	0.4889	
	Cauchy	5	0	2.0	1.5	0.0	0.5402	0.5692	0.5669
			0	0.8	0.8	0.3	0.1122	0.1501	0.1207
			0	1.0	1.0	0.0	0.1998	0.2471	0.2222
			0	0.3	1.5	2.0	0.4752	0.5148	0.4913
			0	0.0	0.0	0.0	0.0451	0.0537	0.0509
			0	0.3	0.5	1.0	0.2383	0.2641	0.2396
10		0	0.5	1.5	1.5	0.6278	0.6463	0.6401	
		0	1.5	0.3	2.0	0.8331	0.8363	0.8390	
		0	2.0	1.5	0.0	0.8833	0.8894	0.8969	
		0	0.8	0.8	0.3	0.2092	0.2296	0.2182	
		0	1.0	1.0	0.0	0.4095	0.4464	0.4487	
		0	0.3	0.8	0.8	0.1529	0.1751	0.1612	
ANOVA		5	0	0.0	0.0	0.0	0.0185	0.0692	0.0394
			0	0.3	0.5	1.0	0.0263	0.1031	0.0545
			0	0.5	1.5	1.5	0.0367	0.1412	0.0788
			0	1.5	0.3	2.0	0.0751	0.2401	0.1411
			0	2.0	1.5	0.0	0.0810	0.2603	0.1630
			0	0.8	0.8	0.3	0.0284	0.0972	0.0495
	10	0	1.0	1.0	0.0	0.0367	0.1412	0.0788	
		0	0.3	1.5	2.0	0.0705	0.2386	0.1390	
		0	0.0	0.0	0.0	0.0168	0.0571	0.0452	
		0	0.3	0.5	1.0	0.0292	0.1349	0.0877	
		0	0.5	1.5	1.5	0.0577	0.3007	0.1931	
		0	1.5	0.3	2.0	0.0855	0.4300	0.2878	
	Friedman	10	0	2.0	1.5	0.0	0.0967	0.4682	0.3333
			0	0.8	0.8	0.3	0.0276	0.1178	0.0712
			0	1.0	1.0	0.0	0.0406	0.2076	0.1342
			0	0.3	0.8	0.8	0.0281	0.1242	0.0807

A: ANOVA for Randomized Block Design; F: Friedman 제안 방법; O: 결합위치 방법에 검수함수를 이용한 선형 위치 통계량.

다는 것을 알 수 있다. 이종지수분포의 경우, 블록의 수가 5개일 때의 검정력은 모든 패턴의 대립가설에서 Friedman (1937) 검정법이 가장 높고, 제안한 방법, ANOVA 순으로 낮다. 블록의 수가 10개일 때는 큰 폭으로 증가했다가 큰 폭으로 감소하고 다시 큰 폭으로 증가하여 정점을 찍는 패턴의 대립가설과 증가하다가 정점을 찍고 감소하는 패턴에서 Friedman (1937)의 검정법보다 제안한 방법의 검정력이 더 높게 나타났고, 나머지 5개의 패턴에서는 Friedman (1937) 검정법, 제안한 방법, ANOVA 순으로 검정

Table 3.2. Monte Carlo power estimates in randomized block design ($\alpha = 0.05$, treatment = 6)

Dist.	block	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	A	F	O	
Normal	5	0	0.0	0.0	0.0	0.0	0.0	0.0479	0.0558	0.0504	
		0	0.0	1.5	1.5	0.0	0.0	0.7388	0.6779	0.5796	
		0	1.0	0.0	0.0	1.0	0.0	0.3736	0.3439	0.2770	
		0	0.3	0.5	0.8	1.0	0.3	0.1981	0.1894	0.1591	
		0	0.5	0.3	0.5	0.3	0.8	0.1207	0.1256	0.0986	
		0	1.0	0.8	0.8	0.3	0.0	0.2760	0.2626	0.2135	
	10	0	1.0	0.5	0.5	1.0	0.5	0.2106	0.2094	0.1553	
		0	1.0	1.0	0.3	0.3	1.5	0.4376	0.4133	0.3318	
		0	0	0.0	0.0	0.0	0.0	0.0532	0.0541	0.0484	
		0	0.0	1.5	1.5	0.0	0.0	0.9911	0.9699	0.9231	
		0	1.0	0.0	0.0	1.0	0.0	0.7622	0.6771	0.5752	
		0	0.3	0.5	0.8	1.0	0.3	0.4250	0.3697	0.2845	
	Exponential	5	0	0.5	0.3	0.5	0.3	0.8	0.2372	0.2113	0.1622
			0	1.0	0.8	0.8	0.3	0.0	0.5942	0.5191	0.4265
			0	1.5	1.0	1.0	0.5	0.5	0.7707	0.8049	0.5702
			0	0.8	0.8	0.3	0.3	0.0	0.4210	0.3595	0.2850
			0	0	0.0	0.0	0.0	0.0	0.0441	0.0609	0.0482
			0	0.0	1.5	1.5	0.0	0.0	0.7674	0.8261	0.8802
10		0	1.0	0.0	0.0	1.0	0.0	0.4214	0.5348	0.5548	
		0	0.3	0.5	0.8	1.0	0.3	0.2222	0.3395	0.2474	
		0	0.5	0.3	0.5	0.3	0.8	0.1296	0.2135	0.1364	
		0	1.0	0.8	0.8	0.3	0.0	0.3133	0.4506	0.3356	
		0	1.0	0.5	0.5	1.0	0.5	0.2344	0.3527	0.2426	
		0	1.0	1.0	0.3	0.3	1.5	0.5042	0.6304	0.5208	
Double exponential		5	0	0	0.0	0.0	0.0	0.0	0.0424	0.0508	0.0500
			0	0.0	1.5	1.5	0.0	0.0	0.9818	0.9962	0.9986
			0	1.0	0.0	0.0	1.0	0.0	0.7785	0.9126	0.9164
			0	0.3	0.5	0.8	1.0	0.3	0.4615	0.6549	0.5592
			0	0.5	0.3	0.5	0.3	0.8	0.2568	0.4141	0.3085
			0	1.0	0.8	0.8	0.3	0.0	0.6264	0.7992	0.6993
	10	0	1.5	1.0	1.0	0.5	0.5	0.7859	0.9064	0.8315	
		0	0.8	0.8	0.3	0.3	0.0	0.4472	0.6562	0.5798	
		0	0	0.0	0.0	0.0	0.0	0.0448	0.0611	0.0495	
		0	0.0	1.5	1.5	0.0	0.0	0.4456	0.4921	0.4334	
		0	1.0	0.0	0.0	1.0	0.0	0.2080	0.2533	0.2120	
		0	0.3	0.5	0.8	1.0	0.3	0.1200	0.1512	0.1266	
	Cauchy	5	0	0.5	0.3	0.5	0.3	0.8	0.0805	0.1070	0.0814
			0	1.0	0.8	0.8	0.3	0.0	0.1626	0.2050	0.1756
			0	1.0	0.5	0.5	1.0	0.5	0.1207	0.1256	0.0986
			0	1.0	1.0	0.3	0.3	1.5	0.2529	0.3061	0.2552
			0	0	0.0	0.0	0.0	0.0	0.0471	0.0532	0.0514
			0	0.0	1.5	1.5	0.0	0.0	0.8198	0.8594	0.8376
10		0	1.0	0.0	0.0	1.0	0.0	0.4498	0.5114	0.4904	
		0	0.3	0.5	0.8	1.0	0.3	0.2269	0.2798	0.2574	
		0	0.5	0.3	0.5	0.3	0.8	0.1349	0.1639	0.1550	
		0	1.0	0.8	0.8	0.3	0.0	0.3266	0.3931	0.3841	
		0	1.5	1.0	1.0	0.5	0.5	0.4596	0.5228	0.4835	
		0	0.8	0.8	0.3	0.3	0.0	0.2225	0.2780	0.2598	
ANOVA		5	0	0	0.0	0.0	0.0	0.0	0.0178	0.0548	0.0508
			0	0.0	1.5	1.5	0.0	0.0	0.0492	0.2154	0.1602
			0	1.0	0.0	0.0	1.0	0.0	0.0263	0.1341	0.0946
			0	0.3	0.5	0.8	1.0	0.3	0.0212	0.0917	0.0673
			0	0.5	0.3	0.5	0.3	0.8	0.0209	0.0749	0.0612
			0	1.0	0.8	0.8	0.3	0.0	0.0237	0.1087	0.0855
	10	0	1.0	0.5	0.5	1.0	0.5	0.0222	0.0982	0.0736	
		0	1.0	1.0	0.3	0.3	1.5	0.0322	0.1450	0.1116	
		0	0	0.0	0.0	0.0	0.0	0.0156	0.0531	0.0527	
		0	0.0	1.5	1.5	0.0	0.0	0.0489	0.4189	0.3020	
		0	1.0	0.0	0.0	1.0	0.0	0.0302	0.2249	0.1598	
		0	0.3	0.5	0.8	1.0	0.3	0.0215	0.1358	0.1064	
	Friedman	10	0	0.5	0.3	0.5	0.3	0.8	0.0209	0.0931	0.0814
			0	1.0	0.8	0.8	0.3	0.0	0.0281	0.1888	0.1403
			0	1.5	1.0	1.0	0.5	0.5	0.0322	0.2278	0.1626
			0	0.8	0.8	0.3	0.3	0.0	0.0221	0.1307	0.1061

A: ANOVA for Randomized Block Design; F: Friedman 제안 방법; O: 결합위치 방법에 검수함수를 이용한 선정 위치 통계량.

력이 낮았다. Cauchy 분포에서는 블록의 크기에 관계없이 모든 패턴의 대립가설에서 Friedman (1937) 검정법의 검정력이 가장 높게 나타났고, 제안한 방법, ANOVA 순으로 낮은 검정력을 보였다.

처리의 수가 6개일 때의 지수분포에서는 블록의 크기에 관계없이 제안한 방법이 ANOVA보다 높은 검

정력을 보였고, 크게 증가하여 완만한 형태를 보이다가 다시 감소하는 패턴과 증가와 감소를 반복하는 패턴의 대립가설에서는 제안한 방법의 검정력이 Friedman (1937)의 검정법보다 높게 나타났다. 이 이중지수분포인 경우에는 블록의 수가 5개일 때 증가했다가 완만한 형태를 유지한 후 다시 감소하는 패턴과 증가했다가 일정한 크기로 감소와 증가를 반복하는 패턴의 대립가설에서 ANOVA와 Friedman (1937)의 검정법보다 제안한 방법이 낮은 검정력을 보였고, 이를 제외한 나머지 5개의 패턴의 대립가설에서는 Friedman (1937) 검정법, 제안한 방법, ANOVA 순으로 검정력이 낮았다. 블록의 수가 10개일 때의 검정력은 모든 패턴의 대립가설에서 Friedman (1937) 검정법, 제안한 방법, ANOVA 순으로 낮게 나타났다. Cauchy 분포의 경우, 처리의 수가 4개일 때와 같이 모든 패턴의 대립가설에서 검정력이 Friedman (1937) 검정법, 제안한 방법, ANOVA 순으로 낮았다.

4. 결론 및 고찰

본 논문에서는 랜덤화 블록 계획법을 검정하기 위한 비모수적 방법으로 결합위치에 점수함수를 적용한 검정통계량을 제안하였다. 이 통계량은 Hodges와 Lehmann (1962)가 제안한 정렬방법을 사용하여 블록 간의 정보를 이용하고, Jo와 Kim (2013)의 논문에서 제안된 결합위치를 $[0,1]$ 의 범위에서 정의된 점수함수를 이용하여 만들었다. 결합위치를 이용한 검정법은 정규분포, 이중지수분포, Cauchy 분포, 지수분포에서 모수적 검정법인 ANOVA와 비모수적 검정법인 Friedman (1937)의 검정법을 사용하여 얻은 검정력 비교를 Monte Carlo simulation을 통하여 시행했다.

모의실험의 전체적인 결과를 보면, ANOVA는 정규분포에서 제 1종 오류를 잘 제어했지만, 나머지 3개의 모집단 분포에서는 잘 제어하지 못했음을 알 수 있다. Friedman (1937)의 검정법은 블록의 크기가 작을 때보다 블록의 크기가 클 때 제 1종 오류를 잘 제어하고 있다. 본 논문에서 제안한 방법은 제 1종 오류를 대체로 잘 제어하고 있지만, Cauchy 분포에서 처리의 수가 작을 때 제 1종 오류를 제어함에 있어 어려움이 있다는 것을 알 수 있다. 검정력 결과를 보면, 정규분포에서는 처리의 수와 관계없이 모수적 방법이 제안한 방법보다 높게 나타났지만, 나머지 3개의 분포에서는 대부분의 대립가설의 검정력이 모수적 방법보다 제안한 방법에서 높게 나타났다. 처리의 수가 작고 블록의 수가 큰 경우 지수분포와 이중지수분포의 일부 대립가설 형태에서 Friedman (1937)의 검정법보다 제안한 검정법의 검정력이 비슷하거나 더 높게 나왔음을 알 수 있다. 처리의 개수가 클 때는 블록의 수와 관계없이 지수분포의 일부 대립가설의 형태에서 제안한 방법이 Friedman (1937)의 검정법보다 높은 효율을 보였다. 이중지수분포에서는 일부 대립가설의 형태에서 모수적 방법이 제안한 방법보다 높은 검정력을 보였고 Friedman (1937)의 검정력이 가장 높게 나왔음을 알 수 있다. Cauchy 분포에서는 처리의 개수와 블록의 수에 관계없이 Friedman (1937)의 검정력이 제안한 검정법보다 높았다.

따라서 본 논문에서 제안한 방법은 처리의 수가 적을 때는 블록의 크기가 큰 경우, 처리의 수가 클 때는 블록의 크기에 관계없이 효율이 좋다는 것을 알 수 있다. 분포와 효율적인 형태의 대립가설에 따라서 본 논문에서 제안한 검정방법을 사용한다면 효율적인 분석이 될 수 있다. 그러나 랜덤화 블록 모형에는 미지의 블록 효과가 존재하기 때문에 비모수적 방법의 장점인 분포무관의 성질은 유지하면서 블록 간의 정보 손실을 가져온다는 문제점이 있다.

추후에 결합위치를 이용한 검정방법에 대한 연구가 진행된다면 본 논문에서 제안한 점수함수를 이용한 선형 위치 통계량 이외에도 이차 함수 형태를 띠는 점수함수를 적용한 선형 위치 통계량이 있을 것으로 보인다. 또한 이차 함수 형태 이외에도 적용할 수 있는 함수 형태가 더 존재하는지는 연구를 통해 살펴볼 필요가 있다. 제안한 방법의 성질을 유도하는 방법 또한 추후에 연구 과제를 통해 얻을 것으로 기대된다.

References

- Chung, T. and Kim, D. (2007). Nonparametric method using placement in one-way layout, *The Korean Communications in Statistics*, **14**, 551–560.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journals of the American Statistical Association*, **32**, 675–701.
- Hettmansperger, T. P. (1975). Nonparametric inference for ordered alternatives in a randomized block design, *Psychometrika*, **40**, 53–62.
- Hodges, J. L. and Lehmann, E. L. (1962). Rank methods for combination of independent experiments in analysis of variance, *The Annals of Mathematical Statistics*, **33**, 482–497.
- Hong, Y. and Lee, S. (2014). Kruskal-Wallis one-way analysis of variance based on linear placements, *Korean Mathematical Society*, **51**, 701–716.
- Jeon, K. and Kim, D. (2016). Nonparametric method in one-way layout based on joint placement, *The Korean Journal of Applied Statistics*, **29**, 729–739.
- Jo, S. and Kim, D. (2013). Nonparametric procedures using aligned method and joint placement in randomized block design, *Journal of the Korean Data and Information Science Society*, **24**, 95–103.
- Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives, *Biometrika*, **41**, 133–145.
- Kang, Y. and Kim, D. (2014). Nonparametric method for ordered alternative in randomized block design, *The Korean Journal of Applied Statistics*, **27**, 61–70.
- Kim, D. (1999). A class of distribution-free treatments versus control tests based on placements, *Far East Journal of Theoretical Statistics*, **3**, 19–33.
- Kim, H. and Kim, D. (2016). Nonparametric procedures based on aligned method and placement for ordered alternatives in randomized block design, *The Korean Journal of Applied Statistics*, **29**, 707–717.
- Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis, *Journals of the American Statistical Association*, **47**, 583–621.
- Lee, M. and Kim, D. (2012). Nonparametric method using an alignment method in a randomized block design with replications, *The Korean Journal of Applied Statistics*, **19**, 77–84.
- Mack, G. A. (1981). A quick and easy distribution-free test for main effects in a two-factor ANOVA, *Communications in Statistics - Simulation and Computation*, **10**, 571–591.
- Mack, G. A. and Skillings, J. H. (1980). A Friedman-type rank test for main effects in a two-factor ANOVA, *Journal of the American Statistical Association*, **75**, 947–951.
- Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placement, *Journals of the American Statistical Association*, **77**, 666–671.
- Page, E. B. (1963). 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.
- Skillings, J. H. and Wolfe, D. A. (1977). Testing for ordered alternatives by combining independent distribution-free block statistics, *Communications in Statistics - Theory and Methods*, **6**, 1453–1463.
- Skillings, J. H. and Wolfe, D. A. (1978). Distribution-free tests for ordered alternatives in a randomized block design, *Journal of the American Statistical Association*, **73**, 427–431.
- Song, H. and Lee, H. (1995). *Design of Clinical Experiments using SAS*, Freeac, Seoul.

랜덤화 블록 계획법에서 정렬방법과 선형위치통계량을 이용한 비모수 검정법

한진주^a · 김동재^{a,1}

^a가톨릭대학교 의생명 · 건강과학과

(2016년 9월 30일 접수, 2016년 11월 14일 수정, 2016년 11월 23일 채택)

요약

랜덤화 블록 계획법을 검정하는 비모수적 방법에는 일반적인 대립가설에서 Friedman (1937)이 제안한 검정법이 있다. 이 방법은 처리 효과의 차이를 알아보기 위한 검정법으로 블록 내 순위를 사용해 검정함으로써 블록 간 정보의 손실이 있을 수 있다. 본 논문에서는 Hodges와 Lehmann (1962)이 제안한 정렬방법을 이용하여 블록 간 정보 손실을 줄이고, Jo와 Kim (2013)이 제안한 랜덤화 블록 계획법의 결합위치 방법을 확장하여 결합위치에 점수함수를 적용한 새로운 비모수적 방법을 제시하였다. 또한 Monte carlo simulation을 통하여 기존의 검정 방법과 제안한 검정법의 검정력을 비교하였다.

주요용어: 비모수, 랜덤화 블록 계획법, 정렬방법, 결합위치

¹교신저자: (06591) 서울특별시 서초구 반포대로 222, 가톨릭대학교 의생명 · 건강과학과.
E-mail: djkim@catholic.ac.kr