The Korean Journal of Applied Statistics (2016) DOI: http://dx.doi.org/10.5351/KJAS.2016.29.4.707
29(4), 707–717

Nonparametric procedures based on aligned method and placement for ordered alternatives in randomized block design

Hyosook Kim[a] · Dongjae Kim[a,1]

[a]Department of Biomedicine · Health Science, The Catholic University of Korea

(Received March 17, 2016; Revised April 13, 2016; Accepted April 16, 2016)

Abstract

Nonparametric procedures in a randomized block design was proposed by Friedman (1937) as a general alternative as well as suggested as a test for ordered alternatives by Page (1963). These methods are used for the rank of treatments in each block. In this paper, we proposed nonparametric procedures using aligned method proposed by Hodges and Lehmann (1962) to reduce among block information and based on placement suggested by Kim (1999) in a randomized block design. We also perform a Monte Carlo study to compare the empirical powers of the proposed procedures and established method.

Keywords: aligned method, nonparametric procedures, ordered alternatives, placement, randomized block design

1. 서론

처리가 3개 이상일 때 처리 효과의 차이를 알기 위한 랜덤화 블록 계획법(randomized block design)은 연구대상을 비슷한 특성을 가진 블록(block)으로 구성한 다음, 무작위로 한 가지의 처리 수준에 대한 한 명의 연구대상을 할당하는 실험 계획법이다. 처리 간 효과 차이유무를 검정하기 위해 오차가 서로 독립이고 정규 분포를 따르는 확률변수임이 가정된다면, 모수적 방법인 분산분석법을 사용해 처리 효과들이 모두 같다는 귀무가설을 검정할 수 있다. 하지만 위 가정을 만족하지 않을 때에는 분산분석법 사용에 따른 문제점의 발생으로 인해 제 1종 오류를 제어할 수 있는 비모수적 방법을 선택해야 한다.

일반적으로 약의 복용량이 증가한다면, 그에 따라 효과도 증가하는 패턴을 보인다. 그렇기 때문에 임상시험의 약에 대한 효능 검사에서 대조(control)그룹, 소량 투여그룹과 다량 투여그룹으로 나누어 조사할 때, 그룹에 따라 약의 효과에 대한 증가 여부를 사전에 예측할 수 있다. 이런 경우, 일반대립가설을 사용하는 것보다 순서대립가설을 사용하는 것이 더 나은 결과를 낳을 수 있다. 또한 온도의 증가에 따른 처리효과의 증감이나 자극의 양에 따른 반응량의 증감 등에 대한 사전 정보가 있는 경우에도 순서대립가설을 사용하는 것이 더 효과적이다. 랜덤화 블록 모형에서 순서대립가설에 대한 비모수적 검정방법은 Page (1963)가 제안하였다. 이 방법은 블록 안에서의 순위(rank)를 이용하여 처리 별 가중치를 적용해

[1]Corresponding 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

통계량을 계산한다. 반면 같은 모형에서 각 처리의 효과가 적어도 하나가 다르다는 일반대립가설에 대한 검정법은 Friedman (1937)이 제안하였고, 이 방법 역시 블록 내 순위를 사용해 처리 간 효과의 차이를 검정한다. 그러나 Page (1963)와 Friedman (1937)이 제안한 방법은 두 가지 다 블록 내 순위를 사용해 검정하기 때문에 블록 간의 정보 손실이 있을 수 있다는 단점이 존재한다.

Orban과 Wolfe (1982)는 두 개의 처리 간 효과 차이를 검정하기 위해 위치(placement)를 이용한 비모수적 검정 방법을 제안하였다. 이는 두 처리 중 어느 한 처리에 대한 상대적 위치정보를 사용해 처리 효과 차이를 검정하는 방법으로써, 대조군의 표본 수가 처리군의 표본 수보다 클 때 더 유용하다고 알려져 있다. 또한 Kim (1999)은 이것을 확장하여 위치를 기반으로 한 대조군과 처리군의 비교 방법을 제안하였다 (Lee와 Kim, 2011).

본 논문에서는 Hodges와 Lehmann (1962)의 정렬방법을 이용한 블록 간의 정보 사용과 Kim (1999)이 일원배치 모형에 대해 제안한 대조군과 처리군 방법을 적용한 순서대립가설에 대한 비모수적 방법을 제안하였다. 그리고 제안한 검정 방법과 랜덤화 블록 계획법에서 모수적 검정법인 분산분석, 비모수적 검정법에서 Friedman (1937), Page (1963) 방법의 검정력(power) 비교를 위한 Monte Carlo 모의실험을 시행하였다.

2. 제안하는 방법

블록이 있고 처리가 k개인 모형

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

에서 Y_{ij}는 j번째 블록에서 i번째 처리의 반응 값이다. μ는 전체 평균, α_i는 i번째 처리의 효과, β_j는 j번째 블록의 효과를 나타내고, ϵ_{ij}는 오차항이며, 동일한 연속분포를 따르는 서로 독립인 확률변수로 가정한다. 블록 간 정보 손실을 줄이기 위해, 블록 간의 정보를 이용하기 위한 Hodges와 Lehmann (1962)이 제안한 정렬방법을 적용해 생성한 정렬자료는

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

이고, $\overline{Y_{\cdot j}} = \sum_{j=1}^{n} Y_{ij}/n$는 각 블록 효과의 블록 평균이다. 각 처리의 효과가 모두 동일하다는 귀무가설과 순서형의 일반적인 대립가설은 다음과 같다.

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_k \quad \text{vs.} \quad H_1 : \alpha_1 \le \alpha_2 \le \alpha_3 \le \cdots \le \alpha_k.$$

2.1. Updated Control Group 검정법

각 처리 군에서 i번째 처리의 j번째 블록 표본 Y_{ij}의 Updated Placement(UP) U_{ij}는

$$\begin{aligned} N_i U_{ij} &= \sum_{h=1}^{i-1} \sum_{s=1}^{n} \chi(Y_{hs}, Y_{ij}) \\ &= [(i-1)\text{번째까지의 처리군에서 } Y_{ij}\text{보다 작거나 같은 표본의 개수}] \end{aligned}$$

라고 정의한다. 여기서 $i = 2, \ldots, k;\ j = 1, \ldots, n;\ \chi(x, y)$는 $x \le y$일 때 1, 그 외의 경우는 0이고, $N_i = n_1 + \cdots + n_{i-1} = n \times (i-1)$이다.

가설 검정을 위한 Kim (1999)이 제안한 방법에서 UP를 이용하는 방법을 적용한 검정통계량은

$$S_i = \sum_{j=1}^{n} \phi_{N_i}(U_{ij}),$$
$$S_{up} = \sum_{i=2}^{k} S_i = \sum_{i=2}^{k} \sum_{j=1}^{n} \phi_{N_i}(U_{ij})$$

로 정의된다. 이 때 ϕ_{N_i}는 $[0,1]$의 범위에서 정의된 실수 값인 점수함수(score function)이고, 선형 위치 통계량을 만드는 특별한 점수함수로는 정규점수함수(normal score function) $\phi_{N_i}(x) = \Phi^{-1}(x)$($\Phi^{-1}$는 표준정규분포 누적분포함수의 역함수)와 지수점수함수(exponential score function) $\phi_{N_i}(x) = -\ln(1-x)$가 있다.

본 논문에서 제안하는 각각의 점수함수를 이용한 통계량은

$$S_{up}^{NS} = \sum_{i=2}^{k} \sum_{j=1}^{n} \Phi^{-1}\left[\frac{N_i U_{ij} + 1}{N_i + 2}\right],$$
$$S_{up}^{E} = -\sum_{i=2}^{k} \sum_{j=1}^{n} \ln\left[1 - \frac{N_i U_{ij}}{N_i + 1}\right]$$

이며, 여기서 Kim 등 (2011)이 제안한 표준화된 S_{up} 통계량 S_{up}^*의 근사분포는

$$S_{up}^* = \frac{S_{up} - E(S_{up})}{\sqrt{n}} \sim N\left(0, \sigma_{up}^2\right)$$

이다. 이 때, $n \to \infty,\ \rho_n = 1/i \to \rho_i,\ 0 < \rho_i < \infty,\ \tilde{\rho}_i = 1/(1 + \rho_1 + \cdots + \rho_{i-1})$이고, 분산 σ_{up}^2는

$$\sigma_{up}^2 = \sum_{i=2}^{k} \tilde{\rho}_i(\tilde{\rho}_i + 1)\sigma_\phi^2, \quad \sigma_\phi^2 = \int_0^1 \phi_{N_i}^2(x)dx - \left(\int_0^1 \phi_{N_i}(x)dx\right)^2$$

이다. 또한 서로 독립인 각 블록의 S_{up}^* 통계량을 이용하는 각각의 점수함수를 적용한 검정통계량은

$$S_{up}^{NS} = \sum_{i=2}^{k} \sum_{j=1}^{n} \frac{\Phi^{-1}\left[\frac{N_i U_{ij}+1}{N_i+2}\right]}{\sqrt{n}} \sim N\left(0, \sigma_{up,NS}^2\right),$$
$$S_{up}^{E} = -\sum_{i=2}^{k} \sum_{j=1}^{n} \frac{\ln\left[1 - \frac{N_i U_{ij}}{N_i+1}\right] - E\left(S_{up}^E\right)}{\sqrt{n}} \sim N\left(0, \sigma_{up,E}^2\right)$$

이며, 이를 이용하여 검정할 수 있다.

2.2. Fixed Control Group 검정법

각 처리 군에서 i번째 처리의 j번째 블록 표본 Y_{ij}의 Fixed Placement(FP) U_{ij}^*는

$$NU_{ij}^* = [Y_{1j}\text{가 } Y_{ij}\text{보다 작거나 같은 표본의 개수}]$$

로 정의한다. 여기서 $i = 2, \ldots, k;\ j = 1, \ldots, n;\ N = n$이다.

가설 검정을 위해 Kim (1999)이 제안한 방법 중 FP를 이용하는 방법을 적용한 검정통계량은

$$S_f = \sum_{i=2}^{k} \sum_{j=1}^{n} \phi_N(U_{ij}^*)$$

로 정의된다. 이 때, ϕ_N는 $[0,1]$의 범위에서 정의된 실수 값인 점수함수(score function)이고, 선형 위치 통계량을 만드는 특별한 점수함수로는 Updated Control Group 검정법에서 사용했던 것과 동일하게 정규 점수함수(normal score function) $\phi_N(x) = \Phi^{-1}(x)$(Φ^{-1}는 표준정규분포 누적분포함수의 역함수)와 지수점수함수(exponential score function) $\phi_N(x) = -\ln(1-x)$가 있다.

또한, Fixed Control Group 검정법에서 제안하는 각각의 점수함수를 이용한 통계량은

$$S_f^{NS} = \sum_{i=2}^{k} \sum_{j=1}^{n} \Phi^{-1} \left[\frac{NU_{ij}^* + 1}{N+2} \right],$$

$$S_f^{E} = -\sum_{i=2}^{k} \sum_{j=1}^{n} \ln \left[1 - \frac{NU_{ij}^*}{N+1} \right]$$

이고, 여기서 Kim 등 (2011)의 논문에서 제안된 표준화된 S_f 통계량 S_f^*의 근사분포는

$$S_f^* = \frac{S_f - E(S_f)}{\sqrt{n}} \sim N\left(0, \sigma_f^2\right)$$

이다. 이 때, $n \to \infty,\ \rho_i = 1,\ 0 < \rho_i < \infty$이고, 분산 σ_f^2는

$$\sigma_f^2 = \sum_{i=2}^{k} \rho_i \left(\sum_{i=2}^{k} \rho_i + 1 \right) \sigma_\phi^2, \quad \sigma_\phi^2 = \int_0^1 \phi_N^2(x)dx - \left(\int_0^1 \phi_N(x)dx \right)^2$$

이다. Fixed Control Group 검정법에서 제안하는 각각의 점수함수에 따른 S_f^*검정통계량은 다음과 같으며, 이를 이용하여 검정할 수 있다.

$$S_f^{NS} = \sum_{i=2}^{k} \sum_{j=1}^{n} \frac{\Phi^{-1} \left[\frac{NU_{ij}^*+1}{N+2} \right]}{\sqrt{n}} \sim N\left(0, \sigma_{f,NS}^2\right),$$

$$S_f^{E} = -\sum_{i=2}^{k} \sum_{j=1}^{n} \frac{\ln \left[1 - \frac{NU_{ij}^*}{N+1} \right] - E\left(S_f^E\right)}{\sqrt{n}} \sim N\left(0, \sigma_{f,E}^2\right).$$

3. 모의실험 및 결과

본 논문에서는 정렬방법과 위치에 따른 검정통계량에 근거해 새로이 제시한 제안방법 4가지와 기존의 모수적 방법인 ANOVA, 일반대립가설의 비모수적 방법인 Friedman (1937)과 순서형 대립가설의 비모수적 방법인 Page (1963)을 비교하였다. 모집단의 분포로는 정규분포, 비대칭적인 지수분포, 정규분포보다 꼬리가 두꺼운 코시분포, 정규분포보다 꼬리가 약간 두꺼운 이중지수분포를 채택하였다. SAS 9.3을 이용하여 정규분포는 RANNOR함수, 지수분포는 RANEXP함수, 코시분포는 RANCAU함수로 난수를 추출, 데이터를 생성하였다. 이중지수분포는 RANNUI함수를 이용해 추출한 난수를 역변환 하는 방법으로 데이터를 생성하였다. 또한 유의수준 α는 0.05로 고려하였다.

처리 개수는 각각 4개와 6개, 각 처리마다의 블록 수는 5개, 10개일 경우를 선택하였으며, α를 0.05로 보정하기 위해 확률화 검정을 사용하였다. 위 조건을 토대로 각각의 통계량에 대한 검정력 비교를 위해

Table 3.1. Monte Carlo power estimates($\alpha = 0.05$, treatment = 4, block = 5)

DIST	α_1	α_2	α_3	α_4	ANO	PAGE	FRI	FNS	FE	UNS	UE
	0	0	0	0	0.0503	0.0526	0.0352	0.0497	0.0512	0.0499	0.0499
	0	0	0	0.5	0.0987	0.1656	0.0618	0.0415	0.0673	0.0793	0.1114
	0	0.5	0.5	0.5	0.0903	0.1569	0.0587	0.1343	0.1909	0.1117	0.1393
Normal	0	0.5	0.5	1	0.1814	0.3533	0.1126	0.1826	0.2595	0.2484	0.3003
	0	1	1	0.5	0.2284	0.1326	0.1509	0.2785	0.3667	0.1301	0.1501
	0	0.5	1	1.5	0.4040	0.6800	0.2667	0.3363	0.4463	0.5690	0.6298
	0.5	0	1	1.5	0.3886	0.0040	0.2554	0.0375	0.0703	0.0394	0.0135
	0	0	0	0	0.0407	0.0525	0.0328	0.0366	0.0410	0.0414	0.0449
	0	0	0	0.5	0.1002	0.2234	0.0862	0.0479	0.0799	0.0855	0.1159
	0	0.5	0.5	0.5	0.1015	0.2497	0.1130	0.2048	0.2548	0.1694	0.1819
Exponential	0	0.5	0.5	1	0.2138	0.5374	0.2120	0.2798	0.3423	0.3560	0.3637
	0	1	1	0.5	0.2896	0.1807	0.2878	0.3835	0.4441	0.1870	0.2096
	0	0.5	1	1.5	0.4792	0.8330	0.4311	0.4490	0.5196	0.6760	0.6641
	0.5	0	1	1.5	0.4802	0.6581	0.4388	0.0431	0.1108	0.2275	0.3645
	0	0	0	0	0.0157	0.0552	0.0331	0.0501	0.0507	0.0508	0.0505
	0	0	0	0.5	0.0214	0.1000	0.0430	0.0534	0.0587	0.0557	0.0614
	0	0.5	0.5	0.5	0.0181	0.0976	0.0423	0.0629	0.0819	0.0562	0.0665
Cauchy	0	0.5	0.5	1	0.0264	0.1769	0.0552	0.0755	0.1061	0.0814	0.0987
	0	1	1	0.5	0.0281	0.0962	0.0626	0.0886	0.1235	0.0661	0.0783
	0	0.5	1	1.5	0.0406	0.3120	0.0894	0.1106	0.1500	0.1397	0.1636
	0.5	0	1	1.5	0.0403	0.2277	0.0870	0.0509	0.0658	0.0820	0.1036
	0	0	0	0	0.0426	0.0516	0.0328	0.0479	0.0481	0.0489	0.0468
	0	0	0	0.5	0.0729	0.1400	0.0594	0.0524	0.0743	0.0700	0.0910
	0	0.5	0.5	0.5	0.0726	0.1453	0.0573	0.1157	0.1677	0.0954	0.1159
Double exponential	0	0.5	0.5	1	0.1198	0.2958	0.0937	0.1523	0.2147	0.1827	0.2197
	0	1	1	0.5	0.1486	0.1231	0.1194	0.2164	0.2925	0.1137	0.1361
	0	0.5	1	1.5	0.2436	0.5494	0.1896	0.2671	0.3643	0.3916	0.4474
	0.5	0	1	1.5	0.2433	0.3992	0.1872	0.0518	0.0936	0.1620	0.2432

ANO = ANOVA for randomized block design, PAGE = Page 제안방법, FRI = Freidman 제안방법,
FNS = Fixed Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
FE = Fixed Control Group 검정법에서 exponential score function를 이용한 선형위치통계량,
UNS = Updated Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
UE = Updated Control Group 검정법에서 exponential score function를 이용한 선형위치통계량.

10,000번 반복하는 Monte Carlo 모의실험을 실시하였다. 또한 모의실험 결과를 처리 수준 개수와 블록 개수에 따라 나누어 처리가 4개이고, 블록이 5개일 때의 결과는 Table 3.1, 처리가 4개, 블록이 10개일 때의 결과는 Table 3.2, 처리가 6개이고 블록 개수가 5개인 경우의 결과는 Table 3.3, 처리 개수 6개, 블록 개수 10개인 경우의 결과를 Table 3.4에 나타내었다.

각각의 처리효과가 모두 같은 경우에서 제 1종 오류를 제어할 수 있는가를 살펴보면, 처리 개수가 4개, 6개이고 각 조건에 따른 블록의 개수가 5개, 10개일 때, 분석분석법 ANOVA의 유의수준이 정규분포에서 0.05 정도의 결과를 보인 것은 자명하다. 또한 모수적 방법은 Table 3.1부터 Table 3.4의 모든 결과표에서 지수분포와 이중지수분포에 관한 결과를 살펴보았을 때도 대체적으로 0.043과 0.045에 근접한 값을 얻은 것으로 보아 0.05에 근사한 값을 갖는다고 볼 수 있었다. 하지만 정규분포의 결과에 비해 제 1종 오류를 제어하는 데는 어려움이 있다는 사실을 알 수 있다. 특히, 각 표에 나타난 코시분포의 결

Table 3.2. Monte Carlo power estimates($\alpha = 0.05$, treatment = 4, block = 10)

DIST	α_1	α_2	α_3	α_4	ANO	PAGE	FRI	FNS	FE	UNS	UE
Normal	0	0	0	0	0.0525	0.0445	0.0466	0.0491	0.0493	0.0498	0.0496
	0	0	0	0.5	0.1728	0.2126	0.1301	0.0696	0.0955	0.1153	0.1529
	0	0.5	0.5	0.5	0.1705	0.2179	0.1323	0.2489	0.3094	0.2199	0.2378
	0	0.5	0.5	1	0.3900	0.5419	0.2990	0.3948	0.4658	0.4969	0.5286
	0	1	1	0.5	0.5122	0.1836	0.3911	0.5548	0.6211	0.3264	0.3522
	0	0.5	1	1.5	0.7938	0.9101	0.6530	0.6992	0.7628	0.8792	0.8992
	0.5	0	1	1.5	0.7944	0.7401	0.6578	0.0967	0.1777	0.3991	0.5654
Exponential	0	0	0	0	0.0433	0.0439	0.0453	0.0476	0.0542	0.0459	0.0482
	0	0	0	0.5	0.1769	0.3367	0.2297	0.0751	0.1059	0.1310	0.1552
	0	0.5	0.5	0.5	0.1787	0.3470	0.2393	0.3378	0.3559	0.3050	0.2736
	0	0.5	0.5	1	0.4282	0.7777	0.5213	0.4881	0.5037	0.6157	0.5642
	0	1	1	0.5	0.5635	0.2515	0.6289	0.6449	0.6404	0.4185	0.4102
	0	0.5	1	1.5	0.8018	0.9772	0.8529	0.7457	0.7433	0.9153	0.8721
	0.5	0	1	1.5	0.8039	0.9000	0.8558	0.1151	0.2204	0.4310	0.5618
Cauchy	0	0	0	0	0.0140	0.0450	0.0485	0.0467	0.0463	0.0471	0.0445
	0	0	0	0.5	0.0187	0.1134	0.0677	0.0441	0.0517	0.0479	0.0528
	0	0.5	0.5	0.5	0.0216	0.1183	0.0699	0.0703	0.0916	0.0675	0.0752
	0	0.5	0.5	1	0.0280	0.2410	0.1061	0.0873	0.1175	0.1044	0.1148
	0	1	1	0.5	0.0292	0.1027	0.1328	0.1072	0.1420	0.0804	0.0992
	0	0.5	1	1.5	0.0475	0.4602	0.2101	0.1397	0.1798	0.1882	0.2026
	0.5	0	1	1.5	0.0468	0.3347	0.2124	0.0518	0.0720	0.0889	0.1074
Double exponential	0	0	0	0	0.0477	0.0444	0.0453	0.0512	0.0498	0.0492	0.0462
	0	0	0	0.5	0.1053	0.1856	0.1103	0.0576	0.0810	0.0920	0.1141
	0	0.5	0.5	0.5	0.1010	0.1762	0.1019	0.1654	0.2158	0.1510	0.1648
	0	0.5	0.5	1	0.2184	0.4443	0.2189	0.2583	0.3251	0.3255	0.3546
	0	1	1	0.5	0.2960	0.1601	0.2956	0.3757	0.4473	0.2301	0.2552
	0	0.5	1	1.5	0.4993	0.7884	0.4925	0.4860	0.5590	0.6712	0.6854
	0.5	0	1	1.5	0.4924	0.6034	0.4845	0.0827	0.1416	0.2731	0.3676

ANO = ANOVA for randomized block design, PAGE = Page 제안방법, FRI = Freidman 제안방법,
FNS = Fixed Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
FE = Fixed Control Group 검정법에서 exponential score function를 이용한 선형위치통계량,
UNS = Updated Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
UE = Updated Control Group 검정법에서 exponential score function를 이용한 선형위치통계량.

과는 모의실험한 모든 값들이 각 표의 순서대로 0.0157, 0.0140, 0.0162, 0.0153의 값으로 나타나는 것을 알 수 있는데, 이것은 정규성가정을 만족하지 못할 경우, 제 1종 오류를 제어하기가 힘들다는 이전 내용과 일치하는 결과이다. Friedman (1937)이 제안한 방법은 표본의 크기가 작은 경우에서 0.0352와 0.0325의 값들을 얻었고, 표본크기가 큰 경우 0.0466과 0.0432의 값을 얻은 것을 확인할 수 있다. 이는 이 방법이 표본크기가 작은 경우에 다른 방법에 비해 조금 더 보수적이라는 것을 알 수 있는 결과이다. 반면, Page (1963)가 제안한 방법은 정규분포, 지수분포와 이중지수분포에서 대체로 제 1종 오류를 제어하는데 있어 큰 문제를 보이지 않고 있지만, 코시분포에서의 값들이 0.0552, 0.0450, 0.0547, 0.0521로 다소 높게 나타난 것을 알 수 있다. 그러나 본 논문에서 제안된 방법들은 표본크기가 가장 작은 경우에서의 지수분포가 제 1종 오류를 제어함에 있어 약간 보수적인 결과를 보이는 것을 제외하고, 각각의 분포에서 실험유의수준으로 제 1종 오류 제어에 큰 문제가 없음을 알 수 있다.

Table 3.3. Monte Carlo power estimates($\alpha = 0.05$, treatment = 6, block = 5)

DIST	α_1	α_2	α_3	α_4	α_5	α_6	ANO	PAGE	FRI	FNS	FE	UNS	UE
Normal	0	0	0	0	0	0	0.0477	0.0475	0.0325	0.0495	0.0500	0.0500	0.0500
	0	0.5	1.5	0.8	0.5	0.5	0.3491	0.0909	0.2260	0.2731	0.3431	0.1298	0.1347
	0	1	1.5	1.5	2	2	0.7127	0.9120	0.5213	0.8383	0.8713	0.9345	0.9120
	0	0.5	1.5	0.8	1	1.5	0.4726	0.5885	0.3409	0.5003	0.5716	0.5891	0.5899
	0	0.5	1	1	1	1	0.2565	0.4392	0.1632	0.4004	0.4646	0.4497	0.4406
	0	0	0	0	0.5	1	0.2380	0.4333	0.1550	0.0736	0.1064	0.2227	0.3016
	0	0.5	1	1	1	0.5	0.2427	0.2233	0.1590	0.3268	0.3923	0.2792	0.2870
	0.8	0.5	0	0.5	1	1.5	0.3633	0.3264	0.2480	0.0439	0.0349	0.0886	0.1608
Exponential	0	0	0	0	0	0	0.0409	0.0493	0.0347	0.0486	0.0497	0.0491	0.0564
	0	0.5	1.5	0.8	0.5	0.5	0.3946	0.1179	0.3803	0.4042	0.4289	0.2138	0.2061
	0	1	1.5	1.5	2	2	0.7464	0.9606	0.7044	0.8141	0.8018	0.9408	0.8774
	0	0.5	1.5	0.8	1	1.5	0.5280	0.7501	0.5428	0.5892	0.5993	0.6981	0.6088
	0	0.5	1	1	1	1	0.2823	0.6102	0.3084	0.5156	0.5193	0.5877	0.4934
	0	0	0	0	0.5	1	0.2813	0.6247	0.2737	0.0968	0.1404	0.2755	0.3062
	0	0.5	1	1	1	0.5	0.2740	0.3088	0.3136	0.4461	0.4627	0.3985	0.3503
	0.8	0.5	0	0.5	1	1.5	0.4171	0.4324	0.4395	0.0332	0.0239	0.0486	0.1248
Cauchy	0	0	0	0	0	0	0.0162	0.0547	0.0354	0.0486	0.0499	0.0495	0.0498
	0	0.5	1.5	0.8	0.5	0.5	0.0303	0.0740	0.0826	0.0863	0.1159	0.0633	0.0729
	0	1	1.5	1.5	2	2	0.0423	0.4518	0.1368	0.1968	0.2497	0.2423	0.2397
	0	0.5	1.5	0.8	1	1.5	0.0336	0.2667	0.0943	0.1150	0.1568	0.1344	0.1426
	0	0.5	1	1	1	1	0.0235	0.2023	0.0619	0.0966	0.1334	0.1044	0.1125
	0	0	0	0	0.5	1	0.0243	0.2078	0.0643	0.0503	0.0630	0.0718	0.0826
	0	0.5	1	1	1	0.5	0.0249	0.1197	0.0652	0.0865	0.1180	0.0795	0.0867
	0.8	0.5	0	0.5	1	1.5	0.0280	0.1548	0.0812	0.0462	0.0414	0.0499	0.0532
Double exponential	0	0	0	0	0	0	0.0429	0.0465	0.0347	0.0534	0.0491	0.0526	0.0537
	0	0.5	1.5	0.8	0.5	0.5	0.1959	0.0779	0.1613	0.2031	0.2722	0.1045	0.1232
	0	1	1.5	1.5	2	2	0.4246	0.7811	0.3537	0.6292	0.6784	0.7507	0.7221
	0	0.5	1.5	0.8	1	1.5	0.2628	0.4699	0.2250	0.3482	0.4259	0.4066	0.4167
	0	0.5	1	1	1	1	0.1418	0.3563	0.1206	0.2838	0.3521	0.3168	0.3207
	0	0	0	0	0.5	1	0.1414	0.3537	0.1162	0.0695	0.0985	0.1610	0.2083
	0	0.5	1	1	1	0.5	0.1380	0.1846	0.1208	0.2341	0.3025	0.1974	0.2183
	0.8	0.5	0	0.5	1	1.5	0.2075	0.2620	0.1767	0.0447	0.0384	0.0749	0.1153

ANO = ANOVA for randomized block design, PAGE = Page 제안방법, FRI = Freidman 제안방법,
FNS = Fixed Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
FE = Fixed Control Group 검정법에서 exponential score function를 이용한 선형위치통계량,
UNS = Updated Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
UE = Updated Control Group 검정법에서 exponential score function를 이용한 선형위치통계량.

표본에 크기에 상관없이 각 방법들의 검정력이 대체로 비슷한 결과를 보이고 있지만, 표본의 크기가 클수록 정규분포에서는 모수적 방법인 분산분석의 검정력이 대체로 높다는 것을 알 수 있었다. 하지만 정규분포하지 않은 모수적 방법인 분산분석은 제안된 다른 비모수 방법들에 비해 대체로 검정력이 낮게 나타났다. 대립가설의 형태가 같은 크기로 증가하는 순서형, 완만한 형태를 보이다 증가하는 패턴의 순서형, 계단 형태의 패턴을 보이며 증가하는 순서형인 경우, 분포에 상관없이 본 논문에서 제안한 방법들이 Page (1963)가 제안한 방법보다는 전반적으로 낮았다. 하지만 다른 비모수 방법인 Friedman (1937)이 제안한 방법이나 모수적 방법 ANOVA보다는 검정력이 높게 나타났다. 주목할 점은 본 논문

Table 3.4. Monte Carlo power estimates($\alpha = 0.05$, treatment = 6, block = 10)

DIST	α_1	α_2	α_3	α_4	α_5	α_6	ANO	PAGE	FRI	FNS	FE	UNS	UE
Normal	0	0	0	0	0	0	0.0504	0.0487	0.0432	0.0496	0.0501	0.0499	0.0499
	0	0.5	1.5	0.8	0.5	0.5	0.7154	0.1207	0.5867	0.5086	0.5857	0.2990	0.3250
	0	1	1.5	1.5	2	2	0.9846	0.9966	0.9420	0.9899	0.9912	0.9990	0.9977
	0	0.5	1.5	0.8	1	1.5	0.8713	0.8572	0.7660	0.8093	0.8465	0.8967	0.8930
	0	0.5	1	1	1	1	0.5532	0.7058	0.4400	0.6888	0.7273	0.7762	0.7505
	0	0	0	0	0.5	1	0.5484	0.6935	0.4293	0.1119	0.1650	0.3796	0.4889
	0	0.5	1	1	1	0.5	0.5250	0.3611	0.4185	0.5751	0.6337	0.5338	0.5379
	0.8	0.5	0	0.5	1	1.5	0.7545	0.5442	0.6283	0.0423	0.0320	0.1162	0.2332
Exponential	0	0	0	0	0	0	0.0468	0.0526	0.0450	0.0476	0.0509	0.0416	0.0500
	0	0.5	1.5	0.8	0.5	0.5	0.7334	0.1868	0.8469	0.6219	0.6048	0.4090	0.3751
	0	1	1.5	1.5	2	2	0.9689	0.9994	0.9774	0.9781	0.9609	0.9989	0.9874
	0	0.5	1.5	0.8	1	1.5	0.8648	0.9552	0.9319	0.8353	0.8039	0.9330	0.8496
	0	0.5	1	1	1	1	0.5799	0.8711	0.6700	0.7595	0.7185	0.8534	0.7276
	0	0	0	0	0.5	1	0.5789	0.8925	0.7284	0.1491	0.2032	0.4561	0.4694
	0	0.5	1	1	1	0.5	0.5538	0.5017	0.6933	0.6755	0.6474	0.6510	0.5479
	0.8	0.5	0	0.5	1	1.5	0.7573	0.7354	0.8747	0.0381	0.0325	0.0589	0.1799
Cauchy	0	0	0	0	0	0	0.0153	0.0521	0.0416	0.0518	0.0490	0.0519	0.0500
	0	0.5	1.5	0.8	0.5	0.5	0.0275	0.0899	0.1801	0.1122	0.1386	0.0864	0.0949
	0	1	1.5	1.5	2	2	0.0572	0.7158	0.3606	0.3131	0.3434	0.3919	0.3630
	0	0.5	1.5	0.8	1	1.5	0.0352	0.4250	0.2406	0.1767	0.2074	0.2082	0.1982
	0	0.5	1	1	1	1	0.0238	0.3169	0.1367	0.1414	0.1667	0.1583	0.1569
	0	0	0	0	0.5	1	0.0238	0.3292	0.1378	0.0610	0.0701	0.0968	0.0944
	0	0.5	1	1	1	0.5	0.0246	0.1791	0.1323	0.1231	0.1505	0.1231	0.1236
	0.8	0.5	0	0.5	1	1.5	0.0310	0.2499	0.1932	0.0522	0.0456	0.0642	0.0653
Double exponential	0	0	0	0	0	0	0.0483	0.0482	0.0450	0.0497	0.0506	0.0439	0.0480
	0	0.5	1.5	0.8	0.5	0.5	0.4154	0.1167	0.4300	0.3555	0.4228	0.2073	0.2406
	0	1	1.5	1.5	2	2	0.7948	0.9626	0.7967	0.8921	0.8952	0.9600	0.9356
	0	0.5	1.5	0.8	1	1.5	0.5539	0.7356	0.5813	0.5978	0.6456	0.6863	0.6735
	0	0.5	1	1	1	1	0.2960	0.5720	0.3095	0.4837	0.5348	0.5451	0.5247
	0	0	0	0	0.5	1	0.3007	0.5712	0.3099	0.0922	0.1399	0.2502	0.3052
	0	0.5	1	1	1	0.5	0.2854	0.2928	0.3070	0.4008	0.4557	0.3595	0.3666
	0.8	0.5	0	0.5	1	1.5	0.4327	0.4447	0.4587	0.0461	0.0342	0.0868	0.1466

ANO = ANOVA for randomized block design, PAGE = Page 제안방법, FRI = Freidman 제안방법,
FNS = Fixed Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
FE = Fixed Control Group 검정법에서 exponential score function를 이용한 선형위치통계량,
UNS = Updated Control Group 검정법에서 normal score function를 이용한 선형위치통계량,
UE = Updated Control Group 검정법에서 exponential score function를 이용한 선형위치통계량.

에서 제안한 방법 중 Updated Control Procedures(UCP)방법이 같은 패턴에서 Fixed Control Procedures(FCP)방법에 비해 더 좋은 검정력을 보였다는 것이다. 그러나 대립가설의 형태가 순서형이지만 처리평균이 완만한 형상을 보이는 경우에서는, 본 논문에서 제시한 방법이 정규분포에서 검정력이 제일 좋다는 것을 알 수 있었다. 또한 같은 패턴에서 정규분포 하지 않은 경우를 살펴보면, 처리 개수가 작고 표본의 크기가 작을 때, 본 논문에서 제안한 방법의 검정력이 높게 나타나지만, 그렇지 않은 경우는 여전히 Page (1963)가 제안한 방법의 검정력이 제일 좋게 나타난다는 사실을 알 수 있다. 대립가설의 형태가 감소하다 증가하는 순서형일 경우에서는 본 논문에서 제안한 모든 방법들의 검정력이 대체로 비교

한 기존의 모든 방법들보다 낮게 나타남으로써, 이러한 형태의 순서형 대립가설에서는 제안방법들의 효율이 낮다는 사실을 알 수 있다. 하지만 처리개수가 6개인 경우의 패턴이었던 대립가설의 형태가 증가하다 감소한 후, 다시 증가하는 모양의 경우, 본 논문에서 제안한 UCP방법의 검정력이 정규분포일 때 모두 좋게 나타났다. 그 외의 분포에서는 Page (1963)가 제안한 방법보다는 검정력이 낮게 나타났지만, FCP, UCP방법 모두 모수적 방법이나 Friedman (1937)이 제안한 방법에 비해서는 일반적으로 높게 나타났다는 것을 알 수 있다. 대립가설이 우산형 패턴의 형태를 보이는 경우, 비교 방법 중 본 논문에서 제안한 FCP방법이 분포에 상관없이 전반적으로 검정력이 제일 높게 나왔음을 알 수 있다. 특히, 정점이 평행한 우산형 패턴에서는 FCP방법 중 지수점수함수를 적용한 방법이 분포와 표본크기에 상관없이 제일 좋게 나타나, 이런 대립가설의 형태를 보이는 경우에서는 FE방법이 더 효율적이라고 볼 수 있다. 그러나 우산형 패턴 형태의 반대의 경우인 U자형의 패턴에서는 본 논문에서 제안한 방법의 검정력이 기존의 방법들에 비해 전반적으로 낮게 나타난 것을 알 수 있다. 즉, U자 패턴 형태에서는 본 논문에서 제시한 방법들에 비해 기존의 방법을 사용하는 것이 더 효율적일 것이다.

4. 결론 및 고찰

본 논문에서는 랜덤화 블록 계획법에서 순서형 대립가설에 대해 분포무관 검정법을 제안하였다. 이 통계량은 Hodges와 Lehmann (1962)가 제안한 정렬방법을 사용해 블록 간의 정보를 이용하도록 한 후, Orban과 Wolfe (1982)의 논문에서 사용된 위치(placement)를 사용해 만들어졌으며, 모수적 방법과 비모수적 방법의 대표적 검정법과의 검정력 비교를 몬테카를로 모의실험(Monte-Carlo simulation)을 통해 수행했다.

제안한 검정 방법을 정규분포, 지수분포, 코시분포 그리고 이중지수분포에서 모수적 검정법과 기존의 비모수적 Friedman (1937)과 Page (1963) 검정법, 정규점수함수와 지수점수함수를 각 방법에 적용한 경우와 비교하였다. 그 결과, 분포와 처리의 패턴 형태에 따른 효율성이 다르게 나타남을 알 수 있었다. 본 논문에서 제안한 방법들은 모의실험의 전체적인 결과를 살펴보았을 때, 점수함수로 지수점수함수를 사용한 경우의 검정력이 정규점수함수를 적용한 경우보다 전반적으로 높은 검정력을 보여주었다. 그러나 표본의 크기가 클수록 UCP방법, FCP방법과 점수함수의 선택에 따른 검정력의 차이가 크지 않음을 알 수 있었다. 또한 처리 개수가 클 때는 블록 수에 상관없이 대체로 정규분포에서의 다양한 패턴에서, 처리 개수와 블록의 수가 작을 때는 우산형 패턴과 정점이 완만한 순서형 대립가설의 형태에서 제안한 방법의 효율이 좋게 나타났다. 비록 일반적인 형태의 순서형 대립가설에서의 효율성은 Page (1963)가 제안한 방법의 효율성보다는 크게 높지 못하지만, 모수적 방법과 Friedman (1937) 검정법의 통계량보다는 더 효율적인 것이라고 말할 수 있다.

따라서 본 논문에서 제안한 방법은 대표본인 경우보다는 소표본일 때, 처리의 개수가 많을 때보다 적은 경우에서 효율이 좋다는 것을 알 수 있었다. 그에 따라 랜덤화 블록 모형에서 보다 효과적인 형태의 대립가설을 선택한다면, 분포에 관계없이 본 논문에서 제시한 정렬방법과 위치를 이용한 검정방법을 사용하는 것이 효율적인 분석이 될 수 있다. 그러나 랜덤화 블록 모형에서 블록 간의 정보를 이용한다는 점에서는 검정력의 효율이 좋게 나타나지만, 미지의 블록 효과가 존재할 수 있기 때문에 비모수검정법의 장점인 분포무관의 성질을 유지하며 블록 간 정보를 추출하는 것에 어려움이 따를 수 있는 문제가 있다.

References

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.

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.

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, D., Lee, S., and Wang, W. (2011). The asymptotic behavior of linear placement statistics, *Statistics and Probability Letters*, **81**, 326–336.

Lee, S. and Kim, D. (2011). Nonparametric procedures using placement in randomized block design with replications, *Journal of the Korean Data and Information Science Society*, **22**, 1105–1112.

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**, 666–671.

Page, E. B. (1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks, *Journals of the American Statistical Association*, **58**, 216–230.

랜덤화 블록 모형에서 정렬방법과 위치를 이용한 순서형 대립가설에 대한 비모수 검정법

김효숙[a] · 김동재[a,1]

[a]가톨릭대학교 의생명 · 건강과학과

(2016년 3월 17일 접수, 2016년 4월 13일 수정, 2016년 4월 16일 채택)

요 약

랜덤화 블록 계획법을 검정하는 비모수 방법은 일반 대립가설에서 Friedman (1937), 순서형 대립가설에서 Page (1963)가 제안한 방법이 있다. 이 방법은 각 블록 내 처리 간 순위를 이용해 처리 간의 차이를 검정하는 방법이다. 본 논문에서는 Hodges와 Lehmann (1962)이 제안한 정렬방법을 이용하여 블록 간 정보의 손실을 줄이고, Orban과 Wolfe (1982)가 제안한 위치를 확장하여, Kim (1999)이 제안한 대조군과 처리군의 방법을 이용하여 랜덤화 블록 모형에서 새로운 비모수 검정 방법을 제안하였다. 또한 Monte Carlo 모의실험을 통해 제안방법과 기존의 검정 방법을 비교하였다.

주요용어: 랜덤화 블록 계획법, 순서형 대립가설, 정렬방법, 위치, 비모수 방법

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