The Korean Journal of Applied Statistics (2017) DOI: https://doi.org/10.5351/KJAS.2017.30.2.291
30(2), 291–299

Nonparametric procedures using aligned method and joint placement in randomized block design with replications

Eunjee Lee[a] · Dongjae Kim[a,1]

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

(Received February 13, 2017; Revised March 20, 2017; Accepted March 20, 2017)

Abstract

Mack and Skillings (1980) proposed nonparametric procedures in a randomized block design with replications as general alternatives. This method is used to find the difference in the treatment effect; however, it can cause a loss of inter block information using the ranking in each block. In this paper, we proposed new nonparametric procedures in a randomized block design with replications using an aligned method proposed by Hodges and Lehmann (1962) that used information of blocks and based on the joint placement suggest by Chung and Kim (2008). We also compared the power of the test of the proposed procedures and established a method through Monte Carlo simulation.

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

1. 서론

류마티스 관절염의 치료제에는 비스테로이드항염제(NSAID), 항류마티스약제(DMARDs), 스테로이드(Steroid)가 있다. 임상시험에서 치료제에 따라 환자가 느끼는 통증의 정도에 차이가 있는지 알아보기 위하여 환자의 류마티스 관절염 증세를 등급(level)에 따라 블록으로 나누고, 블록 안에서 세 치료제에 따라 무작위로 한 명의 환자를 할당하여 처방하였다. 이 때, 세 개 이상의 처리의 효과 차이를 검정하기 위한 방법의 실험 계획법에는 랜덤화 블록 계획법(randomized block design)이 있다. 여기서 처리수준의 각 블록마다 두 명 이상의 연구대상을 할당하는 경우를 반복이 있는 랜덤화 블록 계획법(randomized block design with replications)이라고 한다 (Lee와 Kim, 2012). 처리의 효과차이를 검정하기 위해서 오차가 서로 독립이고 정규분포 $N(0, \sigma^2)$를 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).

반복이 있는 랜덤화 블록 계획법에서 처리 효과의 차이를 알아보기 위한 검정의 모수적 방법으로는 분산분석법이 있고, 처리의 효과가 모두 같지는 않다는 일반대립가설에 관한 비모수적인 방법으로는

[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

Mack과 Skillings (1980)이 제안한 방법, Mack (1981)이 제안한 방법, Lee와 Kim (2012)가 제안한 방법이 있다. Mack과 Skillings (1980)의 방법은 각 블록의 반복 관측치의 평균을 이용하여 랜덤화 완전 블록 계획법에서 사용되는 Friedman (1937)의 방법을 적용한 것으로 블록 내에서 순위를 구한 후, 처리 별 블록 평균 순위 합을 이용한다. Mack (1981)의 방법은 한 블록을 일원배치모형으로 생각하여 각 블록에서 Kruskal-Wallis 검정통계량을 이용하여 검정하는 방법이다. Mack과 Skillings (1980)의 방법과 Mack (1981)의 방법은 모두 블록 내의 순위를 이용하기 때문에 블록 간 정보가 손실될 위험이 있고, 관측치의 평균을 이용하므로 모든 관측 정보를 이용하지 않아 정보 손실이 있을 수 있다 (Lee와 Kim, 2012). Lee와 Kim (2012)의 방법은 정렬방법을 이용한 후 일원배치법에서 사용되는 Kruskal-Wallis (1952)의 방법을 적용한 것인데, 정렬방법을 이용하여 생성된 일원배치모형에서 자료 전체의 순위를 구하여 각 처리에서의 평균 순위를 이용한다. Lee와 Kim (2012)의 방법 또한 관측치의 평균 순위를 이용하여 정보 손실이 있을 수 있다.

Hodges와 Lehmann (1962)은 블록 간의 정보를 이용하기 위하여 정렬방법(aligned method)를 제안하였는데, 정렬방법은 각 처리의 확률표본에서 각 블록의 효과인 블록평균을 빼어 정렬자료를 생성하는 방법이다. Oraban과 Wolfe (1982)는 두 처리 간 효과의 차이를 검정하기 위해 위치(placement)를 사용한 비모수 검정법을 제안하였다. 이 방법은 두 처리 중 어느 한 처리에 대한 상대적 위치정보를 이용하여 처리 효과의 차이를 검정하는 방법으로써 대조군의 표본 크기가 처리군의 표본 크기보다 클 때 더 유용하다고 알려져 있다 (Chung과 Kim, 2007). Chung과 Kim (2007)은 이를 확장하여 일원배치법에서 결합위치(joint placement) 방법을 제안하였다. 결합위치 방법은 세 개 이상의 처리 효과 차이를 알아보고자 할 경우, 비교하고자 하는 확률표본이 포함되어있는 처리군을 제외한 모든 처리군과 비교해 상대적인 위치를 이용하여 처리효과를 검정하는 방법이며, 각 처리의 표본크기가 동일할 때 Kruskal-Wallis 검정법과 같게 된다. Cho와 Kim (2013)은 Hodges와 Lehmann (1962)이 제안한 정렬방법을 이용하여 Chung과 Kim (2007)의 결합위치 방법을 랜덤화 블록 계획법으로 확장하였다.

본 논문에서는 Hodges와 Lehmann (1962)이 제안한 정렬 방법을 이용하여 블록 간의 정보를 이용하고, Chung과 Kim (2007)이 일원 배치모형에서 제안한 결합위치 방법을 확장하여 반복이 있는 랜덤화 블록 계획법에서 비모수적 방법을 제안하였다. 또, 제안된 검정법과 기존 비모수 검정법인 Mack과 Skillings (1980) 방법, Mack (1981) 방법, 그리고 모수적 방법인 분산분석법과 Monte Carlo simulation을 통하여 검정력을 비교하였다.

2. 방법

2.1. 모형 및 가설

반복이 있는 랜덤화 블록모형은 다음과 같다.

$$X_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ijk}, \quad (i = 1, 2, \ldots, t;\ j = 1, 2, \ldots, b;\ k = 1, 2, \ldots, r),$$

여기서 X_{ijk}는 i번째 처리에서 j번째 블록의 r번째 관측의 반응값이고, μ는 전체 평균, τ_i는 i번째 처리 효과, β_j는 j번째 블록 효과를 나타낸다. ϵ_{ijk}는 오차항이며, 동일한 연속분포를 따르는 서로 독립인 확률변수로 가정한다. 각 처리의 효과가 모두 동일하다는 귀무가설과 일반대립가설은 다음과 같다.

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_t \qquad \text{vs.} \qquad H_1 : \tau_i\text{들이 모두 같지는 않다.}$$

2.2. 새로운 제안방법

블록 간 정보를 이용하기 위해 Hodges와 Lehmann (1962)이 제안한 정렬 방법을 이용하여 생렬된 정렬

자료는

$$X^*_{ijk} = X_{ijk} - \bar{X}_{.j.}$$

이고, 여기서

$$\bar{X}_{.j.} = \sum_{i=1}^{t}\sum_{k=1}^{r}\frac{X_{ijk}}{t\cdot r}$$

는 각 블록의 효과인 블록평균이다.

자료를 정렬시켜 블록효과를 없애 자료 전체에 대한 결합위치(joint placement)를 다음과 같이 정의한다.

$$V_{ijk} = \sum_{h=1}^{t}\sum_{\substack{s=1\\ h\neq i}}^{b}\sum_{k=1}^{r}\chi\left(X^*_{hsk}, X^*_{ijk}\right), \qquad \chi(x,y) = \begin{cases} 1, & x \le y\text{일 경우},\\ 0, & \text{그 외}.\end{cases}$$

결합위치 V_{ijk}는 i번째 처리의 관측값을 제외한 혼합표본에서 X^*_{ijk}보다 작거나 같은 관측값의 개수이다. 또한 결합위치의 처리 별 평균 $\overline{V}_{i\cdot\cdot}$과 전체평균 $\overline{V}_{\cdots}$는 다음과 같다. 여기서 $N = b\cdot t\cdot r$은 전체 자료 수이다.

$$\begin{aligned}
\overline{V}_{i\cdot\cdot} &= \frac{1}{b\cdot r}\sum_{j=1}^{b}\sum_{k=1}^{r}V_{ijk}, \quad j = 1,2,\ldots,b;\ k = 1,2,\ldots,r,\\
\overline{V}_{\cdots} &= \frac{1}{N}\sum_{i=1}^{t}\sum_{j=1}^{b}\sum_{k=1}^{r}V_{ijk}\\
&= \frac{1}{N}\sum_{i=1}^{t}b\cdot r\overline{V}_{i\cdot\cdot}\\
&= \frac{1}{N}\sum_{i=1}^{t}V_{i\cdots}
\end{aligned}$$

따라서 각 군에서의 처리 별 평균 $\overline{V}_{i\cdot\cdot}$과 전체 평균 $\overline{V}_{\cdots}$의 차이를 이용한 검정통계량 H는 다음과 같이 정의된다.

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{t}b\cdot r\left(\overline{V}_{i\cdot\cdot} - \overline{V}_{\cdots}\right)^2.$$

검정통계량 H는 귀무가설 하에서 자유도가 $t-1$인 카이제곱분포를 근사적으로 따르게 되고, 기각역은 자유도가 $t-1$인 카이제곱분포의 상위 100α 백분위 수인 $\chi^2(t-1,\alpha)$이 된다. 귀무가설 하에서는 처리 별 평균 $\overline{V}_{i\cdot\cdot}$과 전체 평균 $\overline{V}_{\cdots}$의 값이 차이가 적을 것으로 기대되어 검정통계량 H의 값이 작아지고 이는 귀무가설을 지지하게 된다. 반대로 대립가설 하에서는 처리 별 평균 $\overline{V}_{i\cdot\cdot}$과 전체 평균 $\overline{V}_{\cdots}$의 값의 차이가 클 것으로 기대 되므로 귀무가설을 기각하게 된다.

Chung과 Kim (2007)에 의해, 각 처리별 표본의 크기가 동일하게 n인 경우 Kruskal-Wallis 통계량과 H 통계량은 상수 배의 관계에 놓여있기 때문에 검정통계량은 같다고 할 수 있다. 그리고 각 처리 별 표본크기가 동일할 때 결합위치를 이용한 검정방법과 Kruskal-Wallis가 제안한 검정법은 같은 검정방법이 된다.

Table 3.1. Monte Carlo power estimates in randomized block design with replications ($\alpha = 0.05$, treatment = 3, block = 5, replications = 2)

Dist	τ_1	τ_2	τ_3	H	M-S	M	A
Normal	0	0	0	0.0485	0.0466	0.0468	0.0495
	0	0.3	0.8	0.2942	0.2751	0.6498	0.3053
	0	0.5	1.0	0.4266	0.3918	0.6951	0.4478
	0	1.0	1.0	0.5471	0.5051	0.7726	0.5710
	0	1.0	0	0.5494	0.5104	0.7662	0.5699
	0	1.2	1.0	0.6478	0.6011	0.8150	0.6709
	0	1.5	0.8	0.7868	0.7353	0.8778	0.8034
	0	1.5	1.5	0.8985	0.8572	0.9300	0.9111
Exponential	0	0	0	0.0471	0.0505	0.0470	0.0449
	0	0.3	0.8	0.4426	0.4852	0.7595	0.3518
	0	0.5	1.0	0.5903	0.6262	0.8302	0.5043
	0	0	1.0	0.6716	0.6815	0.8590	0.6101
	0	1.0	0	0.7202	0.7410	0.8797	0.6095
	0	1.2	1.0	0.7564	0.7601	0.8950	0.7005
	0	1.5	0.8	0.8784	0.8823	0.9508	0.8155
	0	1.5	1.5	0.9052	0.8956	0.9548	0.8845
Double-Exponential	0	0	0	0.0521	0.0505	0.0470	0.0492
	0	0.3	0.8	0.2061	0.2140	0.5966	0.1888
	0	0.5	1.0	0.2933	0.2967	0.6642	0.2642
	0	0	1.0	0.3669	0.3682	0.7069	0.3297
	0	1.0	0	0.3742	0.9772	0.6984	0.3390
	0	1.2	1.0	0.4411	0.4429	0.7450	0.4106
	0	1.5	0.8	0.5683	0.5585	0.8063	0.5260
	0	1.5	1.5	0.6763	0.6631	0.8471	0.6378
Cauchy	0	0	0	0.0236	0.0468	0.0359	0.0179
	0	0.3	0.8	0.0536	0.1185	0.5291	0.0326
	0	0.5	1.0	0.0677	0.1450	0.5648	0.0378
	0	0	1.0	0.0835	0.1773	0.5889	0.0442
	0	1.0	0	0.0823	0.1776	0.5810	0.0470
	0	1.2	1.0	0.0931	0.1982	0.5968	0.0497
	0	1.5	0.8	0.1213	0.2620	0.6335	0.0645
	0	1.5	1.5	0.1615	0.3265	0.6816	0.0833

H = method using aligned method and joint placement; M-S = Mack-Skillings' method; M = Mack's method; A = ANOVA.

3. 모의실험의 계획 및 결과

반복이 있는 랜덤화 블록 계획법에서 모수적인 검정 방법으로는 분산분석법이 있고, 비모수적 검정 방법은 Mack과 Skilling (1980)의 방법과 Mack (1891)의 방법이 있다. 따라서 본 논문에서는 새로 제안한 방법과 위의 3가지 방법의 검정력을 비교하기 위해 SAS를 이용하여 모의실험을 시행하였다. 모집단의 분포는 정규분포, 지수분포, 이중지수분포, Cauchy분포를 선택하였다. SAS에서 난수생성하기 위해 정규분포는 RANNOR함수, 지수분포는 RANEXP함수, Cauchy분포는 RANCAU함수, 이중지수분포는 RANUNI함수와 역변환기법을 이용하였다. 생성된 난수를 실제 표본으로 간주하여 계산된 검정통계량이 기각역에 속하는지를 판단하는 과정을 10,000번 반복하는 Monte Carlo Study를 하였다. 처리의 수는 3개와 5개이고, 블록의 수는 5와 7인 경우를 비교하며 각 처리의 블록 내 반복수는 2번, 유의수준은 0.05로 여러 조합에 대해 검정력을 비교하였다. 각 분포에서 방법별 검정력의 비교 결과를 처리의 수가 3인 경우, 블록의 수가 5개일 때의 결과는 Table 3.1, 블록의 수가 7일 때의 결과는 Table 3.2, 처리의 수가 5인 경우, 블록의 수가 5일 때의 결과는 Table 3.3에, 블록의 수가 7일 때의 결과를 Table 3.4에 정리하였다.

처리의 효과가 모두 같은 경우 제 1종 오류가 잘 제어 되는지 살펴보면, 모집단의 분포가 정규분포인 경우 모든 처리와 블록에서 Mack과 Skillings의 방법, Mack의 방법, 제안방법, ANOVA 모두 유의수준이 0.0464에서 0.0512사이의 값을 가지고, 지수분포인 경우 처리의 수가 5이고, 블록의 수가 7일 때 제안방법의 유의수준이 0.0416인 경우를 제외하고 모든 방법의 유의수준이 0.0440에서 0.0515사이의 값을 갖는다. 또한 이중지수분포인 경우에도 모든 처리와 블록에서 모든 방법들의 유의수준이 0.0464에

Table 3.2. Monte Carlo power estimates in randomized block design with replications ($\alpha = 0.05$, treatment = 3, block = 7, replications = 2)

Dist	τ_1	τ_2	τ_3	H	M-S	M	A
Normal	0	0	0	0.0509	0.0472	0.0464	0.0500
	0	0	0.5	0.2329	0.1962	0.0976	0.2409
	0	0.3	0.8	0.4170	0.3616	0.1334	0.4334
	0	0	0.8	0.5305	0.4612	0.1600	0.5425
	0	0.5	1.0	0.5982	0.5150	0.1918	0.6119
	0	1.0	0	0.7201	0.6370	0.2429	0.7392
	0	1.2	0.8	0.7774	0.7037	0.2607	0.7919
	0	1.5	0	0.9735	0.9464	0.5602	0.9793
Exponential	0	0	0	0.0456	0.0488	0.0515	0.0465
	0	0	0.5	0.3495	0.3869	0.1491	0.2644
	0	0.3	0.8	0.5972	0.6268	0.2563	0.4742
	0	0	0.8	0.7203	0.7376	0.2974	0.5727
	0	0.5	1.0	0.7668	0.7794	0.3600	0.6437
	0	1.0	0	0.8854	0.8843	0.4200	0.7665
	0	1.2	0.8	0.8754	0.8742	0.4917	0.8015
	0	1.5	0	0.9938	0.9912	0.7380	0.9711
Double-Exponential	0	0	0	0.0490	0.0488	0.0515	0.0486
	0	0	0.5	0.1620	0.1536	0.0857	0.1433
	0	0.3	0.8	0.2837	0.2719	0.1166	0.2478
	0	0	0.8	0.3578	0.3419	0.1329	0.3137
	0	0.5	1.0	0.4143	0.3957	0.1504	0.3635
	0	1.0	0	0.5168	0.4971	0.1815	0.4607
	0	1.2	0.8	0.5642	0.5428	0.2121	0.5056
	0	1.5	0	0.8493	0.8177	0.3804	0.8020
Cauchy	0	0	0	0.0243	0.0472	0.0488	0.0158
	0	0	0.5	0.0446	0.0947	0.0717	0.0241
	0	0.3	0.8	0.0602	0.1324	0.0788	0.0314
	0	0	0.8	0.0710	0.1561	0.0846	0.0343
	0	0.5	1.0	0.0760	0.1727	0.0888	0.0330
	0	1.0	0	0.1019	0.2209	0.1056	0.0446
	0	1.2	0.8	0.1167	0.2508	0.1133	0.0504
	0	1.5	0	0.2082	0.4211	0.1641	0.0861

H = method using aligned method and joint placement; M-S = Mack-Skillings' method; M = Mack's method; A = ANOVA.

Table 3.3. Monte Carlo power estimates in randomized block design with replications ($\alpha = 0.05$, treatment = 5, block = 5, replications = 2)

Dist	τ_1	τ_2	τ_3	τ_4	τ_5	H	M-S	M	A
Normal	0	0	0	0	0	0.0512	0.0470	0.0496	0.0494
	0	0	0	0	0.5	0.1517	0.1329	0.0744	0.1534
	0	0	0.5	0.5	0.5	0.2726	0.1897	0.0863	0.2185
	0.5	0	0	0	0.5	0.2659	0.1886	0.0809	0.2143
	0	0.5	0.5	0.5	0	0.2737	0.1872	0.0857	0.2196
	0	0.5	1.0	0.5	0	0.5458	0.4199	0.1561	0.4850
	0	1.5	1.0	1.5	1.0	0.8246	0.7741	0.3114	0.8456
	0	1.0	0.3	1.5	1.5	0.9201	0.8812	0.4334	0.9257
Exponential	0	0	0	0	0	0.0466	0.0476	0.0488	0.0507
	0	0	0	0	0.5	0.2292	0.2416	0.1083	0.1629
	0	0	0.5	0.5	0.5	0.4265	0.3818	0.1602	0.2434
	0.5	0	0	0	0.5	0.4297	0.3873	0.1548	0.2404
	0	0.5	0.5	0.5	0	0.4244	0.3389	0.1583	0.2413
	0	0.5	1.0	0.5	0	0.7524	0.7065	0.3052	0.5172
	0	1.5	1.0	1.5	1.0	0.9028	0.8927	0.5243	0.8387
	0	1.0	0.3	1.5	1.5	0.9677	0.9630	0.6761	0.9140
Double-Exponential	0	0	0	0	0	0.0514	0.0476	0.0488	0.0490
	0	0	0	0	0.5	0.1213	0.1144	0.0721	0.1041
	0	0	0.5	0.5	0.5	0.2037	0.1492	0.0818	0.1269
	0.5	0	0	0	0.5	0.2009	0.1507	0.0803	0.1308
	0	0.5	0.5	0.5	0	0.2015	0.1506	0.0791	0.1319
	0	0.5	1.0	0.5	0	0.3949	0.3170	0.1304	0.2708
	0	1.5	1.0	1.5	1.0	0.6072	0.5879	0.2272	0.5366
	0	1.0	0.3	1.5	1.5	0.7341	0.7065	0.2980	0.6433
Cauchy	0	0	0	0	0	0.0179	0.0462	0.0514	0.0182
	0	0	0	0	0.5	0.0255	0.0709	0.0578	0.0190
	0	0	0.5	0.5	0.5	0.0482	0.0820	0.0622	0.0230
	0.5	0	0	0	0.5	0.0446	0.0857	0.0642	0.0197
	0	0.5	0.5	0.5	0	0.0451	0.0875	0.0648	0.0205
	0	0.5	1.0	0.5	0	0.0729	0.1423	0.0816	0.0262
	0	1.5	1.0	1.5	1.0	0.0926	0.2605	0.1118	0.0398
	0	1.0	0.3	1.5	1.5	0.1200	0.3185	0.1321	0.0478

H = method using aligned method and joint placement; M-S = Mack-Skillings' method; M = Mack's method; A = ANOVA.

Table 3.4. Monte Carlo power estimates in randomized block design with replications ($\alpha = 0.05$, treatment = 5, block = 7, replications = 2)

Dist	τ_1	τ_2	τ_3	τ_4	τ_5	H	M-S	M	A
Normal	0	0	0	0	0	0.0512	0.0507	0.0474	0.0511
	0	0	0.3	0	0.3	0.1253	0.1162	0.0630	0.1287
	0	0	0	0	0.5	0.2012	0.1874	0.0756	0.2110
	0	0.5	0.5	0.5	0	0.2976	0.2724	0.0964	0.3142
	0.5	0	0	0	0.5	0.3014	0.2716	0.0988	0.3140
	0	0.3	0.5	0.8	1.0	0.5895	0.5401	0.1703	0.6117
	0	0.5	1.0	0.5	0	0.6416	0.5927	0.1853	0.6671
	0	0	1.0	0.8	1.2	0.9106	0.8752	0.3580	0.9224
Exponential	0	0	0	0	0	0.0416	0.0469	0.0464	0.0457
	0	0	0.3	0	0.3	0.1938	0.2348	0.0973	0.1324
	0	0	0	0	0.5	0.3288	0.3704	0.1233	0.2255
	0	0.5	0.5	0.5	0	0.4991	0.5492	0.1859	0.3352
	0.5	0	0	0	0.5	0.5068	0.5579	0.1739	0.3339
	0	0.3	0.5	0.8	1.0	0.8105	0.8400	0.3418	0.6320
	0	0.5	1.0	0.5	0	0.8653	0.8810	0.3702	0.6869
	0	0	1.0	0.8	1.2	0.9762	0.9776	0.6203	0.9130
Double-Exponential	0	0	0	0	0	0.0472	0.0469	0.0464	0.0474
	0	0	0.3	0	0.3	0.0935	0.0957	0.0640	0.0865
	0	0	0	0	0.5	0.1478	0.1499	0.0719	0.1283
	0	0.5	0.5	0.5	0	0.2101	0.2091	0.0890	0.1779
	0.5	0	0	0	0.5	0.2121	0.2169	0.0824	0.1726
	0	0.3	0.5	0.8	1.0	0.4065	0.4087	0.1289	0.3376
	0	0.5	1.0	0.5	0	0.4543	0.4506	0.1465	0.3807
	0	0	1.0	0.8	1.2	0.7304	0.7232	0.2476	0.6335
Cauchy	0	0	0	0	0	0.0140	0.0473	0.0515	0.0148
	0	0	0.3	0	0.3	0.0226	0.0693	0.0577	0.0182
	0	0	0	0	0.5	0.0254	0.0836	0.0602	0.0201
	0	0.5	0.5	0.5	0	0.0289	0.1066	0.0637	0.0239
	0.5	0	0	0	0.5	0.0295	0.1049	0.0641	0.0194
	0	0.3	0.5	0.8	1.0	0.0506	0.1810	0.0795	0.0269
	0	0.5	1.0	0.5	0	0.0571	0.1982	0.0839	0.0267
	0	0	1.0	0.8	1.2	0.0989	0.3346	0.1157	0.0355

H = method using aligned method and joint placement; M-S = Mack-Skillings' method; M = Mack's method; A = ANOVA.

서 0.0521사이의 값을 가지므로 정규분포, 지수분포, 이중지수분포의 유의수준은 모두 0.05에 근사한 값을 갖는다. Cauchy분포인 경우 모의실험 결과 값들이 다른 분포들에 비해 유의수준 0.05에서 벗어났다. Cauchy분포의 표 순서대로 처리의 수가 3인 경우, 블록의 수가 5일 때의 유의수준의 값은 0.0236, 0.0468, 0.0359, 0.0179로 Mack과 Skillings의 방법을 제외하고 Mack의 방법, 제안방법, ANOVA의 제 1종 오류를 잘 제어하지 못하는 것이 나타나며, 블록의 수가 7일 때의 유의수준의 값은 0.0243, 0.0472, 0.0488, 0.0158로 제안방법과 ANOVA의 1종오류를 잘 제어하지 못한 것이 나타났다. 처리의 수가 5인 경우, 블록의 수가 5일 때 제안방법과 ANOVA의 유의수준이 0.0179, 0.0182이고, 블록이 7일 때 0.0140, 0.0148인 것으로 보아 제안방법과 ANOVA는 Cauchy분포에서 다른 검정법들에 비해 보수적인 검정법임을 알 수 있다. Cauchy분포에서는 다른 분포의 결과와 비교하였을 경우 1종 오류를 제어하는데 있어 어려움이 있었다.

처리의 수가 3개인 경우, 블록의 수가 5개일 때의 결과에서 모든 분포, 7개의 대립가설 형태에서 모두 Mack의 방법의 검정력이 다른 방법들에 비해 높았다. 즉, Mack의 방법은 분포와 대립가설의 형태에 상관 없이 표본의 크기가 작은 경우 검정력이 높게 나타남을 알 수 있다. 블록의 수가 7개일 때의 결과에서는 모집단의 분포가 정규분포인 경우 모든 대립가설 형태에서 ANOVA의 검정력이 가장 높았고, 제안방법, Mack과 Skillings의 방법, Mack의 방법의 순으로 검정력이 높았다. 지수분포에서는 기울기가 낮은 순서형 대립가설 형태와 완만하다 증가하는 대립가설 형태에서 Mack과 Skillings의 방법이 제안방법보다 근소하게 검정력이 높았다. 나머지 대립가설 형태에서는 제안방법과 Mack과 Skillings의 검정력이 가장 높았고, ANOVA, Mack의 방법 순으로 검정력의 차이가 있었다. 이중지수분포에서는 모든

대립가설 형태에서 제안방법의 검정력이 가장 높았고, Mack과 Skillings의 방법, ANOVA, Mack의 방법 순으로 검정력의 차이가 있었다. 특히 정점이 1.5인 우산형 대립가설 형태에서는 제안방법의 검정력이 0.8493으로 Mack과 Skillings의 검정력인 0.8177에 비해 약 0.032의 차이가 났다. Cauchy분포에서는 모든 대립가설 형태에서 Mack과 Skillings의 방법의 검정력이 가장 높았다.

처리의 수가 5개인 경우, 블록의 수가 5개일 때의 결과에서 모집단의 분포가 정규분포일 때, 증가와 감소를 반복하는 대립가설 형태에서는 ANOVA의 검정력이 가장 높았고, 완만하다 증가하는 대립가설의 형태와 크게 증가하다 감소한 후 다시 증가하여 일정한 대립가설 형태에서는 제안방법과 ANOVA의 검정력이 가장 높았다. 나머지 대립가설 형태에서는 제안방법의 검정력이 가장 높았고, ANOVA, Mack과 Skillings의 방법, Mack의 방법 순으로 검정력이 차이가 났다. 우산형 대립가설 형태에서 제안방법의 검정력이 0.5458로 ANOVA의 검정력 0.4850에 비해 약 0.06 차이가 났고, 증가한 후 일정하다 다시 감소하는 대립가설 형태에서 제안방법의 검정력이 0.2737로 ANOVA의 검정력 0.2196과 약 0.054 차이가 있었다. 지수분포에서는 완만하다 증가하는 대립가설 형태에서 Mack과 Skillings의 방법의 검정력이 가장 높았다. 나머지 6개의 대립가설 형태에서는 제안방법이 모두 검정력이 높았고, Mack과 Skilligs의 방법, ANOVA, Mack의 방법 순으로 검정력이 높았다. 특히 증가한 후 일정하다 다시 감소하는 대립가설 형태에서 제안방법의 검정력이 0.4244로 Mack과 Skillings의 방법의 검정력에 비해 약 0.086 높았다. 또, 우산형 대립가설 형태에서도 제안방법의 검정력이 0.7524로 Mack과 Skillings의 방법의 검정력에 비해 약 0.046 차이가 났다. 이중지수분포에서는 모든 대립가설 형태에서 제안방법의 검정력이 가장 높았고, Mack과 Skillings의 방법, ANOVA, Mack의 방법 순으로 검정력의 차이가 있었다. 특히 제안방법이 Mack과 Skillings의 방법에 비해 감소한 후 완만하다 다시 증가하는 형태의 대립가설과 증가한 후 일정하다 다시 감소하는 대립가설 형태에서는 검정력이 약 0.05 높았고, 우산형 대립가설 형태에서는 검정력이 약 0.078 높았다. Cauchy분포에서는 Mack과 Skillings의 방법의 검정력이 가장 높았다. 블록의 수가 7개일 때의 결과에서는 모집단의 분포가 정규분포인 경우, 모든 대립가설 형태에서 검정력은 ANOVA, 제안방법, Mack과 Skillings의 방법, Mack의 방법의 순서로 높았다. 지수분포인 경우, 완만하다 증가와 감소를 반복하는 대립가설의 형태에서 제안방법과 Mack과 Skillings의 방법의 검정력이 비슷하게 높았지만 나머지 대립가설의 형태에서는 Mack과 Skillings의 방법, 제안방법, ANOVA, Mack의 방법 순으로 검정력의 차이가 있었다. 이중지수분포에서는 모든 대립가설의 형태에서 제안방법과 Mack과 Skillings의 방법의 검정력이 모두 비슷하였고 ANOVA, Mack의 방법 순으로 검정력의 차이가 있었다. Cauchy분포에서는 Mack과 Skillings의 방법의 검정력이 가장 높았다.

4. 결론 및 고찰

본 논문에서는 반복이 있는 랜덤화 블록 계획법을 검정하기 위한 비모수적 방법으로 정렬방법과 결합위치를 이용한 검정통계량을 제안하였다. 이 통계량은 Hodges와 Lehmann (1962)가 제안한 정렬방법을 사용하여 블록 간의 정보를 이용하고, Chung과 Kim (2007)이 제안한 결합위치를 확장하였다. 정렬방법과 결합위치를 이용한 검정법을 정규분포, 지수분포, 이중지수분포, Cauchy분포에서 모수적 검정법인 ANOVA와 비모수적 검정법인 Mack과 Skillings (1980)과 Mack (1981)의 검정력을 Monte Carlo simulation을 통하여 비교하였다.

본 논문에서 제안한 방법은 정규분포, 지수분포, 이중지수분포에서 제 1종 오류를 비교적 잘 제어할 수 있었지만, Cauchy분포에서는 다른 분포의 결과와 비교하였을 경우 1종 오류를 제어하는데 있어 어려움이 있었다. Cauchy분포에서 제안방법은 다른 비모수 검정법들에 비해 상당히 보수적인 검정법임을 알 수 있다. 모의실험 결과 표본의 수가 작은 경우 모든 분포에서 Mack의 방법이 효율이 좋다고 기대되고, 표본의 수가 큰 경우 모집단의 분포가 이중지수분포일 때 제안방법이 기존의 방법들보다 우수하

다고 기대된다. 또 정규분포에서도 제안방법의 검정력이 ANOVA와 비슷하거나 결과가 나왔으므로 제안방법은 모집단의 분포가 대칭이고 꼬리가 두껍지 않은 분포에서 기존의 방법들에 비해 좋은 효율성을 가진다고 할 수 있다. 하지만 Cauchy분포에서는 Mack과 Skillings의 방법의 검정력이 가장 높았다. Cauchy분포에서 제안방법의 검정력이 낮은 이유는 유의수준이 Cauchy분포에서 작아 매우 보수적인 검정법임에서 기인한 것으로 생각된다. 지수분포에서는 처리의 수가 5이고 블록의 수가 5인 경우 제안방법을 이용하는 것이 효율이 높을 것으로 예상된다. 또한 대립가설의 형태에서는 우산형 대립가설인 경우 제안방법을 이용하는 것이 효율이 높을 것으로 기대된다.

반복이 있는 랜덤화 블록 모형에서는 미지의 블록 효과가 존재하기 때문에 비모수적 방법의 장점인 분포무관 성질은 유지하면서 블록 간의 정보 손실을 가져온다는 문제점이 있다. 추후 반복이 있는 랜덤화 블록 계획법에서 비모수 검정법에 대한 연구가 진행된다면 본 논문에서 Cauchy분포에서 제 1종 오류를 제어함에 있어 어려움이 있던 부분과 블록 간의 정보손실을 보완할 수 있는 검정법이 있는지 연구를 통해 살펴볼 필요가 있다.

References

Cho, 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.

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.

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.

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, *Journal of the American Statistical Association*, **77**, 666–671.

Song, H. and Lee, H. (1995). *Design of Clinical Experiments Using SAS*, Free Academy, Seoul.

반복이 있는 랜덤화 블록 계획법에서 정렬방법과 결합위치를 이용한 비모수 검정법

이은지[a] · 김동재[a,1]

[a]가톨릭대학교 의생명 · 건강과학과

(2017년 2월 13일 접수, 2017년 3월 20일 수정, 2017년 3월 20일 채택)

요 약

반복이 있는 랜덤화 블록 계획법을 검정하는 비모수 검정방법에는 Mack과 Skillings (1980), Mack (1981)가 제안한 방법이 있다. 본 논문에서는 Hodges와 Lehmann (1962)의 정렬 방법과 Chung과 Kim (2007)이 제안한 결합위치 검정법을 확장하여 반복이 있는 랜덤화 블록 모형에서 새로운 비모수적 방법을 제시하였다. 또한 모의실험을 통해 모수적 방법과 기존의 비모수적 방법과의 검정력을 비교하였다.

주요용어: 비모수, 랜덤화 블록 계획법, 정렬방법, 결합위치

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