The Korean Journal of Applied Statistics (2017) DOI: https://doi.org/10.5351/KJAS.2017.30.2.281
30(2), 281–290

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

Soyoung Jeon[a] · Dongjae Kim[a,1]

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

(Received February 8, 2017; Revised March 9, 2017; Accepted March 20, 2017)

Abstract

Mack and Skillings (1980) proposed a nonparametric method in a randomized block design with replications. This method employs the mean of observations instead of each observation. However, it has the inherent disadvantage that there may be a loss of information. In this paper, we proposed a nonparametric method that employees an aligned method and linear placement statistics to supplement its weakness. A Monte-Carlo study is performed to compare the power of the proposed method with previous methods.

Keywords: nonparametric method, randomized block design, aligned method, linear placement statistics

1. 서론

일원배치법을 이용하여 실험할 때 처리가 다른 여러 군의 실험단위들은 처치 전에는 서로 동질적이어서 순수한 처리효과의 차이가 반응측정값의 차이로 나타나는 것이 바람직하다. 하지만 현실적으로 실험단위가 여러 요인들에 의해 균일치 않아 반응측정값의 변동이 심하게 되면 처리 효과간의 차이를 알아내기 어렵게 된다. 이러한 문제점을 해결하기 위한 한 방법이 랜덤화블록 계획법(randomized block design)이다. 랜덤화블록 계획법은 연구대상을 비슷한 특성을 가진 블록으로 구분한 뒤에 무작위로 한 가지의 처리 수준에 한 명의 연구대상을 할당하는 방법이다. 이때 각 처리 수준에서 블록마다 두명 이상의 연구대상을 할당하면 반복이 있는 랜덤화 블록 계획법이 된다 (Song과 Kim, 2015).

랜덤화 블록 계획법에서 처리들 간에 효과의 차이가 있는지 알아보기 위한 검정들 중에 모수적 방법으로는 분산분석법을 사용한다. 비모수적 방법은 Friedman (1937)이 제안한 방법과 Page (1963)가 제안한 방법이 있다. Friedman (1937)이 제안한 검정법은 각 블록 내의 관측값들을 작은 것부터 순위를 매겨서 검정하는 방법으로 처리들 중 적어도 하나는 효과가 다르다는 일반 대립가설에서 사용한다. Page (1963)가 제안한 검정법은 블록 내의 관측값들에 순위를 부여하고 처리 별 가중치를 적용해 검정하는 방법으로 순서형 대립가설에서 사용한다.

반복이 있는 랜덤화 블록 계획법에서 사용하는 모수적 검정법은 분산분석법이 있다. 비모수적 검정법으로는 일반 대립가설에서 사용하는 Mack과 Skillings (1980)가 제안한 방법과 Mack (1981)이 제안한 방

[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

법이 있고 순서형 대립가설에서 사용하는 Skillings와 Wolfe (1977)가 제안한 방법과 Hettmansperger (1975)가 제안한 방법이 있다. 여기서 Mack (1981)이 제안한 방법은 각 블록을 일원배치모형으로 생각하여 검정통계량을 계산하는 방법이다. 블록별로 Kruskal과 Wallis (1952)의 검정법을 적용한 후 각 블록의 통계량을 모두 더하여 검정한다. Mack과 Skillings (1980)가 제안한 검정법은 각 블록내의 반복 관측치들의 평균을 이용하여 Friedman (1937)이 제안한 검정통계량을 구하는 것이다. 이 검정법은 각각의 관측값을 이용하는 대신 각 블록에서의 반복된 관측값들의 평균으로 순위를 매겨 검정하는 것이기 때문에 모든 관측치의 정보를 이용하지 않게 된다. 따라서 관측치들의 정보를 손실할 수 있다는 단점이 있다. 이러한 단점을 보완하기 위해 Lee와 Kim (2011)은 Orban과 Wolfe (1982)가 제안한 위치(placement)와 Kim (1999)이 제안한 대조군과 처리군의 방법을 확장하여 새로운 검정법을 제안하였고 Lee와 Kim (2012)은 Hodges와 Lehmann (1962)의 정렬 방법을 이용한 새로운 검정법을 제안하였다. 여기서 Hodges와 Lehmann (1962)의 방법은 자료를 정렬시켜 블록효과를 없애는 방법으로 블록간의 정보를 이용하기 위해 쓰인다 (Lee와 Kim, 2012).

본 논문에서는 관측치의 정보를 손실할 수 있다는 기존 방법의 단점을 보완하기 위하여 새로운 비모수적 검정법을 제안하고자 한다. 제안하는 검정법은 모형에 Hodges와 Lehmann (1962)의 정렬방법(aligned method)을 적용하고 Jo와 Kim (2013)의 결합위치(joint placement)를 확장하여 결합위치에 점수함수(score function)를 적용한 선형 위치 통계량(linear placement statistics)을 이용하는 방법이다. 제안하는 방법은 Hodges와 Lehmann (1962)의 정렬방법을 이용하기 때문에 블록간의 정보를 이용할 수 있고, 관측치들의 평균이 아닌 Jo와 Kim (2013)의 결합위치를 사용하기 때문에 기존 방법의 단점을 보완할 수 있다. 또한 Hong과 Lee(2014)에 의해 선형 위치 통계량의 근사분포가 표준정규분포라는 것을 알 수 있기 때문에 검정법을 이용하기에 편하다는 장점이 있다. 결합위치와 선형 위치 통계량을 이용한 다른 검정법은 Jeon과 Kim (2016)이 제시한 방법이 있다. 본문에서는 새롭게 제안한 검정법에 대해 소개하고, 제안 방법과 기존에 사용하던 비모수적 검정법인 Mack과 Skillings (1980)의 방법, Mack (1981)의 방법 그리고 모수적 검정법인 분산분석법과 모의실험을 실시하여 검정력을 비교하였다.

2. 방법

2.1. 반복이 있는 랜덤화 블록 모형 및 가설

반복이 있는 랜덤화 블록 모형은 다음과 같은 형태를 가진다.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}, \quad (i = 1, 2, \ldots, t;\ j = 1, 2, \ldots, b;\ k = 1, 2, \ldots, m),$$

여기서 Y_{ijk}는 i번째 처리에서 j번째 블록의 k번째 반응값이고, μ는 전체 평균, α_i는 i번째 처리 효과, β_j는 j번째 블록 효과, ϵ_{ijk}는 오차항을 나타낸다. 오차항은 동일한 연속분포를 따르는 서로 독립인 확률변수를 가정한다.

처리의 수가 t개이고 모집단에 대한 구체적인 분포함수를 가정할 수 없을때 각 처리마다 효과가 모두 동일하다는 귀무가설과 일반적인 대립가설은 다음과 같다.

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_t \qquad \text{vs.} \qquad H_1 : \alpha_i \text{ 들이 모두 같지는 않다.}$$

귀무가설은 처리들의 효과가 모두 같다는 것이고 대립가설은 처리들의 효과 중 적어도 하나는 다르다는 것이다.

2.2. 기존의 방법

반복이 있는 랜덤화 블록 모형에서 처리들 간에 효과의 차이가 존재하는지 검정하는 기존의 방법으로는

Mack과 Skillings (1980)가 제안한 방법이 있다. 반복이 있는 랜덤화 블록 계획법에서 각각의 블록은 t개의 처리에 m명의 연구대상이 할당되어 있고 N_j는 $t \cdot m$개의 관측값으로 되어있다. 처리의 수가 세 개 이상일 때 처리들의 효과 차이를 검정하기 위해 N_j개의 관측값으로 혼합표본을 만들고 작은 것부터 차례대로 순위를 부여한다. 다시말해,

$$R_{ijk} = N_j\text{개의 혼합표본에서 } Y_{ijk}\text{의 순위}$$

로 정의된다. 각 처리에서 블록들의 평균 순위의 합은

$$S_i = \sum_{j=1}^{b} \left[\sum_{k=1}^{m} \frac{R_{ijk}}{m} \right], \quad i = 1, 2, \ldots, t$$

이다. 일반적인 대립가설을 검정하기 위해 Mack과 Skillings (1980)가 제안한 검정통계량은

$$\mathrm{MS} = \left[\frac{12}{t(N+b)} \right] \sum_{i=1}^{t} \left[S_i - \frac{N+b}{2} \right]^2 = \left[\frac{12}{t(N+b)} \right] \sum_{i=1}^{t} S_i^2 - 3(N+b)$$

이다. 표본의 크기가 충분히 크면 Mack과 Skillings가 제안한 검정통계량은 귀무가설 하에서 근사적으로 자유도가 $t-1$인 카이제곱분포를 따르게 된다. 이 방법으로 검정을 할 때 기각역은 자유도가 $t-1$인 카이제곱분포의 상위 $100 \cdot \alpha$ 백분위수인 $\chi^2_{t-1,\alpha}$이다.

다른 기존의 방법은 Mack (1981)이 제안한 방법이 있다. Kruskal과 Wallis (1952)는 일원배치모형에서 처리들의 효과 차이에 대한 검정법을 제안했다. 반복이 있는 랜덤화 블록 모형에서 각 블록은 반복수가 m인 일원배치모형이므로 블록별로 Kruskal과 Wallis (1952)의 검정법을 적용할 수 있다. j번째 블록에서 구한 Kruskal-Wallis 검정통계량을 H_j라 하면 반복이 있는 랜덤화 블록 계획법에서 처리들 간에 효과 차이가 있는가를 검정하기 위해 Mack이 제안한 통계량은

$$H_{ma} = \sum_{j=1}^{b} H_j$$

가 된다. 표본의 크기가 충분히 크고 귀무가설을 만족하면 통계량 H_{ma}는 근사적으로 자유도가 $u \cdot (t-1)$인 카이제곱분포를 따른다. 또한 이 방법으로 검정할 때 기각역은 자유도가 $u \cdot (t-1)$인 카이제곱 분포의 상위 $100 \cdot \alpha$ 백분위수 $\chi^2_{b(t-1),\alpha}$이다.

2.3. 정렬방법을 이용한 결합위치(joint placement)

반복이 있는 랜덤화 블록 모형에서 블록 간의 정보를 이용하기 위해 Hodges와 Lehmann (1962)의 정렬방법을 적용하여 정렬자료를 만든다. 정렬자료를 구하는 방법은 아래와 같다.

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

여기서

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

은 각 블록의 평균이다.

자료를 정렬시킨 후에 결합위치 V_{ijk}를 구한다. 결합위치를 구하는 방법은 다음과 같다.

$$V_{ijk} = \frac{1}{N - b \cdot m} \sum_{h=1}^{t} \sum_{\substack{s=1 \\ h \neq i}}^{b} \sum_{l=1}^{m} \chi\left(Y^*_{hsl}, Y^*_{ijk}\right), \qquad \chi(x,y) = \begin{cases} 1, & x \leq y \text{일 경우}, \\ 0, & \text{그 외}, \end{cases}$$

여기서 N은 전체 표본의 수이다. 결합위치 V_{ijk}는 i번째 처리의 관측값을 제외한 혼합표본에서 Y^*_{ijk}보다 작거나 같은 관측값의 개수를 이용한 확률변수이다. 결합위치를 이용한 선형 위치 통계량은 다음과 같이 구할 수 있다.

$$S_c = \sum_{i=1}^{t} \sum_{j=1}^{b} \sum_{k=1}^{m} \Phi(V_{ijk}).$$

위 식에서 $\Phi(\cdot)$는 $[0,1]$에서 정의된 실변수인 점수함수(score function)이다.

Hong과 Lee (2014)에 의해, 귀무가설 하에서 $N \to \infty$을 만족하면 표준화된 통계량 $\hat{S}_c$의 분포는 표준정규분포로 수렴하게 된다.

즉, 선형위치 통계량의 근사분포는

$$\hat{S}_c = \frac{S_c - E(S_c)}{\sqrt{\mathrm{Var}(S_c)}} \xrightarrow{N \to \infty} N(0,1)$$

가 된다. 본 논문에서는 선형 위치 통계량을 구하기 위한 점수함수로 $\Phi(x) = |1 - 2x|$을 사용한다. 따라서 점수함수를 사용한 선형 위치 통계량은 다음과 같이 구한다.

$$S_c = \sum_{i=1}^{t} \sum_{j=1}^{b} \sum_{k=1}^{m} \left|1 - 2V_{ijk}\right|.$$

이때 귀무가설 하에서 S_c의 기대값은 N이 전체 표본의 수일 때

$$E(S_c) = \begin{cases} \dfrac{N(N - b \cdot m + 2)}{2(N - b \cdot m + 1)}, & N - b \cdot m \text{이 짝수일 때}, \\[2ex] \dfrac{N(N - b \cdot m + 1)}{2(N - b \cdot m)}, & N - b \cdot m \text{이 홀수일 때} \end{cases}$$

이다. S_c의 분산은 임의의 t와 b와 m에 대해서는 구체적인 식으로 나타낼 수 없고 특정한 t와 b와 m에 대해서는 계산할 수 있다.

3. 모의실험 계획 및 결과

본 논문에서는 반복이 있는 랜덤화 블록 모형에서 정렬자료를 만든 후 결합위치를 이용해 점수함수를 적용시킨 검정통계량에 근거한 새로운 검정법과 기존의 검정법들과의 검정력을 비교하였다. 비교한 기존의 방법들은 모수적 방법인 분산분석법(ANOVA)과 비모수적 방법인 Mack과 Skillings (1980)의 방법, Mack (1981)의 방법이다.

기존의 3가지 방법과 본 논문에서 제시한 새로운 방법을 비교하기 위해 SAS를 사용해 모의실험을 시행하였다. 모집단의 분포로는 정규분포, 지수분포, Cauchy분포, 이중지수분포를 사용하였다. 각각의 난수 생성은 RANNOR, RANEXP, RANCAU함수를 사용하였고 이중지수분포는 RANUNI함수와 역변환기법을 사용하여 생성하였다. 생성된 난수를 이용하여 계산된 검정통계량이 기각역에 속하는지를 판단하는 과정을 10,000번 반복하는 Monte Carlo Study를 실시하였다.

처리의 수는 4개와 6개일 경우를 비교하였고 블록의 수는 5개와 7개일 경우를 고려하였다. 또한 각 처리의 블록 내에서의 반복수는 2번으로 설정하였고 유의수준은 0.05로 해서 각 통계량들의 검정력을 비교하였다. 각 분포에서 검정력들을 비교한 결과는 처리의 수가 4개이고 블록의 수가 5개일 때는 Table 3.1에, 처리의 수가 4개이고 블록의 수가 7개일 때는 Table 3.2에 정리하였다. 또한 처리의 수가 6개이고 블록의 수가 5개일 때는 Table 3.3에, 처리의 수가 6개이고 블록의 수가 7개일 때는 Table 3.4에 정리하였다.

각 처리의 효과가 모두 동일할 때 유의수준을 보면 대부분 0.05에서 크게 벗어나지 않았지만 다른 방법들에 비해 유의수준이 작게 나타나는 것들이 있었다. 정규분포에서는 처리가 4개이고 블록이 7개일 때 Mack과 Skillings의 방법의 유의수준이 0.04로 다른 방법들에 비해 작게 나타났다. 지수분포에서는 처리가 4개이고 블록이 7개일 때 분산분석법의 유의수준이 0.04로 작게 나타났다. 또한 처리가 6개이고 블록이 5개일 때 Mack의 방법의 유의수준이 0.04로 비교적 작았다. 이중지수분포에서는 처리가 4개이고 블록이 7개일 때 분산분석법의 유의수준이 0.04로 작게 나타났다. 또한 처리가 6개이고 블록이 5개일 때 Mack의 방법의 유의수준이 0.04로 작았다. Cauchy분포에서는 모든 모형에서 새롭게 제안한 방법과 분산분석법의 유의수준이 0.01과 0.03 사이의 값으로 작게 나타났다. 또한 처리가 6개이고 블록이 5개일 때 Mack의 방법의 유의수준이 0.03으로 비교적 작았다. 이를 제외한 나머지 다른 방법들은 모두 0.05에 가까운 값들을 얻은 것으로 보아 1종 오류를 제어하는 데 문제가 없음을 알 수 있다.

검정력을 살펴보면 정규분포에서는 모형에 상관없이 분산분석법의 검정력이 가장 높았다. 새롭게 제안한 방법의 검정력은 Mack과 Skillings의 방법보다는 낮았지만 Mack의 방법보다 높았다. 처리의 수가 4개 또는 6개이고 블록의 수가 5개일 때 대립가설이 증가했다가 완만해지는 형태에서 새롭게 제안한 방법과 Mack과 Skillings의 방법 그리고 분산분석법의 검정력이 각각 0.14, 0.16, 0.17로 비교적 차이가 크지 않았다. 처리의 수가 4개이고 블록의 수가 7개일 때에도 대립가설이 증가했다가 완만해지는 형태에서 세 분석법의 검정력이 각각 0.18, 0.19, 0.22로 다른 대립가설 형태에 비해 차이가 적었다. 처리의 수가 6개이고 블록의 수가 5개일 때는 대립가설이 큰 폭으로 감소하는 형태에서 세 방법의 검정력이 0.89, 0.96, 0.98로 비교적 차이가 적었다. 처리의 수가 6개이고 블록의 수가 7개일 때는 작은 폭으로 증가하는 형태의 대립가설에서 각각의 검정력이 0.14, 0.15, 0.17로 비슷하게 나타났다.

지수분포에서는 전체적으로 Mack과 Skillings의 방법의 검정력이 가장 높았고 새롭게 제안한 검정법이 두 번째로 높았다. 또한 모형에 상관없이 대립가설이 증가했다가 완만해지는 형태에서 두 검정법의 검정력이 큰 차이를 보였다. 모형별로 검정력을 살펴보면, 처리의 수가 4개이고 블록의 수가 5개일 때는 대립가설이 증가했다가 감소하고 다시 증가하는 형태에서 새롭게 제안한 방법과 Mack과 Skillings의 방법의 검정력이 각각 0.84와 0.88로 비교적 차이가 적게 나타났다. 또한 대립가설이 감소했다가 증가하는 형태에서는 검정력이 각각 0.86과 0.91로 차이가 크지 않았다. 처리의 수가 4개이고 블록의 수가 7개일 때는 대립가설이 완만했다가 증가하는 형태에서 각각의 검정력이 0.37과 0.39로 차이가 적었고 대립가설이 감소했다가 증가하는 형태에서는 검정력이 0.96과 0.97로 매우 비슷했다. 처리의 수가 6개이고 블록의 수가 5개일 때는 대립가설이 큰 폭으로 감소하는 형태일 때 검정력이 각각 0.98과 0.99로 매우 비슷했다. 처리의 수가 6개이고 블록의 수가 7개일 경우에는 대립가설이 완만했다가 증가하고 다시 완만해지는 형태에서 검정력이 두 방법 모두 0.99로 거의 같았다.

이중지수분포에서는 새롭게 제안한 방법과 Mack과 Skiilings의 방법의 검정력이 다른 검정법들에 비해 높았고 두 검정력이 대부분 비슷하게 나타났다. 처리의 수가 4개 또는 6개이고 블록의 수가 5개일 때 모든 대립가설 형태에서 Mack과 Skillings의 방법의 검정력이 가장 컸지만 새롭게 제안한 방법과 큰 차이를 보이지는 않았다. 처리의 수가 4개이고 블록의 수가 7개일 때는 대립가설이 증가했다가 완만해지는 형태와 증가했다가 감소하고 다시 증가하는 형태에서 새롭게 제안한 방법의 검정력이 각각 0.15와

Table 3.1. Monte Carlo power estimates: $\alpha = 0.05$, treatment = 4, block = 5, replication = 2

Dist	α_1	α_2	α_3	α_4	S	MS	M	Anova
Normal	0	0	0	0	0.0501	0.0466	0.0508	0.0522
	0	0.5	0.5	0.5	0.1432	0.1680	0.0766	0.1755
	0	0.3	0.5	0.8	0.2073	0.2404	0.0974	0.2666
	0	0.5	0.5	1.0	0.2974	0.3589	0.1243	0.3925
	1.2	0.8	0.4	0	0.4634	0.5365	0.1927	0.5922
	0	0.8	0	1.2	0.6005	0.6852	0.2628	0.7385
	0.5	0	1.0	1.5	0.6551	0.7425	0.2918	0.7981
Exponential	0	0	0	0	0.0504	0.0523	0.0461	0.0484
	0	0.5	0.5	0.5	0.1596	0.2909	0.1200	0.1854
	0	0.3	0.5	0.8	0.3642	0.4726	0.1753	0.3054
	0	0.5	0.5	1.0	0.5051	0.6121	0.2347	0.4332
	1.2	0.8	0.4	0	0.6947	0.7898	0.3645	0.6310
	0	0.8	0	1.2	0.8404	0.8815	0.4616	0.7518
	0.5	0	1.0	1.5	0.8647	0.9104	0.5204	0.8094
Double exponential	0	0	0	0	0.0516	0.0523	0.0461	0.0498
	0	0.5	0.5	0.5	0.1141	0.1288	0.0655	0.1066
	0	0.3	0.5	0.8	0.1806	0.1978	0.0853	0.1587
	0	0.5	0.5	1.0	0.2390	0.2629	0.1028	0.2128
	1.2	0.8	0.4	0	0.3857	0.4127	0.1492	0.3421
	0	0.8	0	1.2	0.4997	0.5118	0.1869	0.4434
	0.5	0	1.0	1.5	0.5413	0.5781	0.2164	0.5128
Cauchy	0	0	0	0	0.0212	0.0477	0.0501	0.0157
	0	0.5	0.5	0.5	0.0322	0.0844	0.0567	0.0190
	0	0.3	0.5	0.8	0.0427	0.1072	0.0645	0.0245
	0	0.5	0.5	1.0	0.0555	0.1379	0.0680	0.0279
	1.2	0.8	0.4	0	0.1053	0.2375	0.0946	0.0382
	0	0.8	0	1.2	0.0998	0.2384	0.0960	0.0428
	0.5	0	1.0	1.5	0.1135	0.2579	0.1027	0.0450

S = linear placement statistics using aligned method and joint placement; MS = Mack and Skillings' method; M = Mack's method.

Table 3.2. Monte Carlo power estimates: $\alpha = 0.05$, treatment = 4, block = 7, replication = 2

Dist	α_1	α_2	α_3	α_4	S	MS	M	Anova
Normal	0	0	0	0	0.0528	0.0449	0.0497	0.0466
	0	0	0	0.5	0.1805	0.2015	0.0731	0.2282
	0	0.5	0.5	0.5	0.1891	0.1992	0.0818	0.2217
	0	0.3	0.5	0.8	0.3098	0.3519	0.1092	0.3974
	1.2	0.8	0.4	0	0.6301	0.6967	0.2284	0.7738
	0	0.8	0	1.2	0.6584	0.7248	0.2540	0.7946
	0.5	0	1.0	1.5	0.8262	0.8843	0.3801	0.9303
Exponential	0	0	0	0	0.0503	0.0480	0.0485	0.0439
	0	0	0	0.5	0.3744	0.3991	0.1248	0.2306
	0	0.5	0.5	0.5	0.2233	0.3878	0.1339	0.2530
	0	0.3	0.5	0.8	0.5138	0.6273	0.2116	0.4214
	1.2	0.8	0.4	0	0.8609	0.9182	0.4534	0.7814
	0	0.8	0	1.2	0.8869	0.9307	0.4780	0.8001
	0.5	0	1.0	1.5	0.9624	0.9792	0.6433	0.9189
Double exponential	0	0	0	0	0.0506	0.0480	0.0485	0.0433
	0	0	0	0.5	0.1594	0.1606	0.0728	0.1335
	0	0.5	0.5	0.5	0.1537	0.1529	0.0818	0.1335
	0	0.3	0.5	0.8	0.2504	0.2543	0.0934	0.2160
	1.2	0.8	0.4	0	0.5294	0.5475	0.1674	0.4771
	0	0.8	0	1.2	0.6770	0.6741	0.2256	0.5970
	0.5	0	1.0	1.5	0.7096	0.7294	0.2652	0.6630
Cauchy	0	0	0	0	0.0257	0.0516	0.0516	0.0177
	0	0	0	0.5	0.0406	0.0919	0.0546	0.0192
	0	0.5	0.5	0.5	0.0450	0.0937	0.0579	0.0207
	0	0.3	0.5	0.8	0.0577	0.1268	0.0666	0.0241
	1.2	0.8	0.4	0	0.1053	0.2375	0.0946	0.0382
	0	0.8	0	1.2	0.1498	0.3173	0.1103	0.0433
	0.5	0	1.0	1.5	0.1658	0.3519	0.1170	0.0488

S = linear placement statistics using aligned method and joint placement; MS = Mack and Skillings' method; M = Mack's method.

Table 3.3. Monte Carlo power estimates: $\alpha = 0.05$, treatment = 6, block = 5, replication = 2

Dist	α_1	α_2	α_3	α_4	α_5	α_6	S	MS	M	Anova
Normal	0	0	0	0	0	0	0.0483	0.0487	0.0480	0.0522
	0.3	0.5	1.0	1.0	1.0	1.0	0.2117	0.2900	0.0980	0.3250
	0	0.2	0.4	0.6	0.8	1.0	0.2958	0.3973	0.1233	0.4484
	1.2	1.0	0.8	0.6	0.4	0.2	0.3017	0.4020	0.1216	0.4533
	0	0.5	1.0	1.0	0.5	0	0.4298	0.5556	0.1831	0.6167
	0	0	0	1.2	1.2	1.2	0.8192	0.9141	0.4409	0.9426
	2.0	1.6	1.2	0.8	0.4	0	0.8946	0.9698	0.5608	0.9849
Exponential	0	0	0	0	0	0	0.0455	0.0506	0.0429	0.0484
	0.3	0.5	1.0	1.0	1.0	1.0	0.2783	0.5414	0.1905	0.3579
	0	0.2	0.4	0.6	0.8	1.0	0.5255	0.7083	0.2611	0.4866
	1.2	1.0	0.8	0.6	0.4	0.2	0.5159	0.7017	0.2659	0.4875
	0	0.5	1.0	1.0	0.5	0	0.7042	0.8385	0.3699	0.6490
	0	0	0	1.2	1.2	1.2	0.9395	0.9800	0.6780	0.9283
	2.0	1.6	1.2	0.8	0.4	0	0.9857	0.9953	0.8003	0.9707
Double exponential	0	0	0	0	0	0	0.0443	0.0506	0.0429	0.0498
	0.3	0.5	1.0	1.0	1.0	1.0	0.1900	0.2190	0.0801	0.1811
	0	0.2	0.4	0.6	0.8	1.0	0.2621	0.2955	0.1094	0.2490
	1.2	1.0	0.8	0.6	0.4	0.2	0.2613	0.3044	0.1052	0.2507
	0	0.5	1.0	1.0	0.5	0	0.3698	0.4146	0.1335	0.3381
	0	0	0	1.2	1.2	1.2	0.7385	0.7555	0.2955	0.6631
	2.0	1.6	1.2	0.8	0.4	0	0.7953	0.8591	0.3790	0.7904
Cauchy	0	0	0	0	0	0	0.0139	0.0466	0.0398	0.0169
	0.3	0.5	1.0	1.0	1.0	1.0	0.0308	0.1103	0.0550	0.0247
	0	0.2	0.4	0.6	0.8	1.0	0.0388	0.1368	0.0669	0.0245
	1.2	1.0	0.8	0.6	0.4	0.2	0.0337	0.1357	0.0691	0.0241
	0	0.5	1.0	1.0	0.5	0	0.0503	0.1769	0.0754	0.0255
	0	0	0	1.2	1.2	1.2	0.1203	0.3493	0.1243	0.0442
	2.0	1.6	1.2	0.8	0.4	0	0.1424	0.4288	0.1371	0.0498

S = linear placement statistics using aligned method and joint placement; MS = Mack and Skillings' method; M = Mack's method.

Table 3.4. Monte Carlo power estimates: $\alpha = 0.05$, treatment = 6, block = 7, replication = 2

Dist	α_1	α_2	α_3	α_4	α_5	α_6	S	MS	M	Anova
Normal	0	0	0	0	0	0	0.0540	0.0461	0.0511	0.0500
	0.7	0.8	0.9	1.0	1.1	1.2	0.1434	0.1540	0.0755	0.1762
	0.3	0.5	1.0	1.0	1.0	1.0	0.3394	0.4039	0.1337	0.4615
	0	0.2	0.4	0.6	0.8	1.0	0.4700	0.5594	0.1672	0.6291
	1.2	1.0	0.8	0.6	0.4	0.2	0.4685	0.5553	0.1767	0.6217
	0	0.5	1.0	1.0	0.5	0	0.6354	0.7383	0.2491	0.7964
	0	0	0	1.2	1.2	1.2	0.9536	0.9837	0.5988	0.9926
Exponential	0	0	0	0	0	0	0.0497	0.0484	0.0539	0.0504
	0.7	0.8	0.9	1.0	1.1	1.2	0.2145	0.3393	0.1147	0.1805
	0.3	0.5	1.0	1.0	1.0	1.0	0.4344	0.7011	0.2614	0.4904
	0	0.2	0.4	0.6	0.8	1.0	0.7385	0.8712	0.3633	0.6487
	1.2	1.0	0.8	0.6	0.4	0.2	0.7318	0.8660	0.3602	0.6465
	0	0.5	1.0	1.0	0.5	0	0.8758	0.9489	0.4994	0.8053
	0	0	0	1.2	1.2	1.2	0.9929	0.9990	0.8272	0.9860
Double exponential	0	0	0	0	0	0	0.0512	0.0484	0.0539	0.0497
	0.7	0.8	0.9	1.0	1.1	1.2	0.1280	0.1243	0.0651	0.1085
	0.3	0.5	1.0	1.0	1.0	1.0	0.3039	0.3033	0.1091	0.2437
	0	0.2	0.4	0.6	0.8	1.0	0.4165	0.4248	0.1421	0.3467
	1.2	1.0	0.8	0.6	0.4	0.2	0.4124	0.4197	0.1344	0.3450
	0	0.5	1.0	1.0	0.5	0	0.5603	0.5692	0.1822	0.4774
	0	0	0	1.2	1.2	1.2	0.9073	0.9041	0.4100	0.8347
Cauchy	0	0	0	0	0	0	0.0171	0.0477	0.0517	0.0182
	0.3	0.5	1.0	1.0	1.0	1.0	0.0487	0.1385	0.0706	0.0222
	0	0.2	0.4	0.6	0.8	1.0	0.0589	0.1760	0.0821	0.0249
	1.2	1.0	0.8	0.6	0.4	0.2	0.0619	0.1762	0.0816	0.0258
	0	0.5	1.0	1.0	0.5	0	0.0851	0.2410	0.0910	0.0279
	0	0	0	1.2	1.2	1.2	0.1892	0.4846	0.1637	0.0404
	0.7	0.8	0.9	1.0	1.1	1.2	0.0242	0.0707	0.0569	0.0175

S = linear placement statistics using aligned method and joint placement; MS = Mack and Skillings' method; M = Mack's method.

0.67로 다른 검정법들에 비해 크게 나타났다. 처리의 수가 6개이고 블록의 수가 7개일 때는 대립가설이 증가했다가 완만해지는 형태와 완만했다가 증가하고 다시 완만해지는 형태에서 제안방법의 검정력이 각각 0.30과 0.90으로 다른 검정법들보다 크게 나타났다.

Cauchy분포에서는 다른 분포에 비해 검정력이 전체적으로 작게 나타났다. 또한 전체적으로 Mack과 Skillings의 방법의 검정력이 새롭게 제안한 방법과 Mack의 방법보다 컸다. 처리의 수가 4개이고 블록의 수가 5개일 때 대립가설이 증가했다가 완만해지는 형태에서 새롭게 제안한 방법과 Mack과 Skillings의 방법의 검정력이 0.03과 0.08로 다른 대립가설 형태에 비해 차이가 적었다. 처리의 수가 4개이고 블록의 수가 7개일 때도 증가했다가 완만해지는 대립가설 형태에서 검정력이 각각 0.04와 0.09로 비교적 차이가 적었다. 처리의 수가 6개이고 블록의 수가 5개일 때는 증가했다가 완만해지는 형태에서 두 검정력이 0.03과 0.11로 다른 대립가설 형태에서보다 비슷하게 나타났다. 처리의 수가 6개이고 블록의 수가 7개일 때는 점점 증가하는 대립가설 형태에서 검정력이 각각 0.02와 0.07로 비교적 비슷했다.

모의실험의 결과를 보면, 정규분포일 때를 제외하고 제안하는 방법이 분산분석법보다 검정력이 높았다. 따라서 정규분포일 때를 제외하면 제안하는 방법이 분산분석법보다 효율적이다. 또한 제안하는 방법이 Cauchy분포일 때를 제외하면 Mack의 방법보다 검정력이 높았다. Cauchy분포에서도 두 검정법의 검정력의 차이가 크지 않았다. 제안하는 방법의 검정통계량의 근사분포가 표준정규분포라는 장점을 고려하면 검정력의 차이가 크지 않을 때도 제안하는 방법을 사용하는 것이 더 효율적이다. 따라서 모든 분포에서 Mack의 방법보다 제안하는 방법을 사용하는 것이 더 바람직하다. Mack과 Skillings의 방법은 제안하는 방법보다 검정력이 높은 경우가 많았다. 하지만 Mack과 Skillings의 방법은 관측치의 정보를 손실할 수 있다는 단점이 있기 때문에 기존의 방법보다 검정력이 크게 높지 않을 경우에는 기존의 방법을 사용하는 것이 더 효율적이다.

4. 결론 및 고찰

본 논문에서는 반복이 있는 랜덤화블록 모형에 대한 검정을 위하여 비모수적 방법으로 정렬방법과 결합위치를 이용한 선형위치통계량을 제안하였다. 이 통계량은 Hodges와 Lehmann (1962)의 정렬방법을 적용하여 정렬자료를 만든 후에 Chung과 Kim (2007)이 제안한 결합위치(joint Placement)를 $[0,1]$의 범위에서 정의된 점수함수에 적용하여 구한다. 모의실험을 실시하여 새롭게 제안한 검정법의 검정력을 정규분포, 지수분포, 이중지수분포, Cauchy분포에서 비모수적 방법인 Mack의 방법, Mack과 Skillings의 방법, 그리고 모수적 방법인 분산분석법의 검정력과 비교하였다.

모의실험의 결과에서 유의수준을 보면, Cauchy분포를 제외한 나머지 분포에서는 거의 모든 방법들이 1종오류를 잘 제어했다. 검정력을 보면, 처리의 수가 4개 또는 6개이고 블록의 수가 5개일 때 정규분포에서는 분산분석법의 검정력이 가장 컸고 나머지 분포에서는 Mack과 Skillings의 방법의 검정력이 가장 컸다. 다만 이중지수분포일 때 나머지 분포에 비해 새롭게 제안한 방법과 Mack과 Skillings의 방법의 검정력의 차이가 줄어든다는 것을 알 수 있었다. 처리의 수가 4개 또는 6개이고 블록의 수가 7개일 때 정규분포에서는 분산분석법의 검정력이 가장 컸고 Cauchy분포에서는 Mack과 Skillings의 방법의 검정력이 가장 컸다. 지수분포에서는 대부분의 대립가설 형태에서 Mack과 Skillings의 방법의 검정력이 가장 컸지만 어떠한 형태에서는 새롭게 제안한 방법과 Mack과 Skillings의 방법의 검정력이 비슷하게 나타났다. 이중지수분포에서는 새롭게 제안한 방법과 Mack과 Skillings의 방법의 검정력이 대부분의 대립가설 형태에서 비슷했고 몇몇 형태에서는 새롭게 제안한 방법이 더 크게 나타났다.

본 논문에서 제안한 검정법은 모든 모형에서 Mack의 검정법보다 더 효율적이다. 블록의 수가 5개일 때는 Mack과 Skilling의 방법보다 효율적이라고 할 수 없지만 블록의 수가 7개일 때는 Mack과 Skilling의

방법과 비슷하거나 더 효율적이라고 할 수 있다. 따라서 새롭게 제안한 방법은 블록의 수가 적을 때보다 많은 경우에서 효율이 좋다는 것을 알 수 있었다.

추후에 정렬방법과 선형위치통계량을 이용한 검정방법에 대해 연구가 진행된다면 본 논문에서 제안한 점수함수의 형태 이외에 다른 함수형태를 적용할 수 있는지 더 연구해 볼 수 있을 것이다. 본 논문에서 제안한 검정법은 검정통계량이 표준정규분포로 수렴한다는 장점이 있기 때문에 또 다른 실험설계를 위한 유용한 비모수 검정법이 될 것이라고 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, I. 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, *Journal of the Korean Statistical Society*, **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 & Information Science Society*, **24**, 95–103.

Kim, D. (1999). A class of distribution-free treatments versus control tests based on placements, *Far East Journal of Theoretical Statistics*, **3**, 19–33.

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.

Lee, M. and Kim, D. (2012). Nonparametric method using an alignment method in a randomized block design with replications, *Communications for Statistical Applications and Methods*, **19**, 77–84.

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

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). Ordered hypotheses for multiple treatments: a significance test for linear ranks, *Journal of the American Statistical Association*, **58**, 216–230.

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.

Song, H. and Kim, D. (2015). *Understanding Statistics*, Cheong moon gak, Gyeonggi-do.

반복이 있는 랜덤화블록 모형에서 정렬방법과 선형위치통계량을 이용한 비모수 검정법

전소영[a] · 김동재[a,1]

[a]가톨릭대학교 의생명 · 건강과학과

(2017년 2월 8일 접수, 2017년 3월 9일 수정, 2017년 3월 20일 채택)

요 약

반복이 있는 랜덤화블록 모형에서의 비모수적 검정 방법에는 Mack과 Skillings (1980)가 제안한 방법이 있다. 이 방법은 각각의 관측값을 사용하는 대신 각 블록에서의 반복된 관측값들의 평균을 사용하여 검정하는 방법이다. 따라서 관측치들의 정보를 손실할 수 있다는 단점이 있다. 본 논문에서는 이러한 단점을 보완하기 위해 정렬방법과 선형위치통계량을 이용한 비모수 검정법을 제안하였다. 또한 몬테카를로 모의실험(Monte-Carlo Study)을 통하여 기존의 방법과 제안한 방법의 검정력을 비교하였다.

주요용어: 비모수 검정법, 랜덤화블록 모형, 정렬방법, 선형위치통계량

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