The Korean Journal of Applied Statistics (2015) DOI: http://dx.doi.org/10.5351/KJAS.2015.28.6.1181
28(6), 1181–1189

Nonparametric Method in One-way Layout for Umbrella Alternatives based on Placement

Hyejung Lee[a] · Dongjae Kim[a,1]

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

(Received October 12, 2015; Revised November 5, 2015; Accepted November 11, 2015)

Abstract

The treatment effect in clinical tests depending on dose of the drug; however, it can show a decreasing trend in fixed dose level due to side effects. The trend is known as an umbrella pattern; in addition, the method for the umbrella alternative is quite useful when the tendency is predicted in advance. In this paper, we propose a nonparametric method of umbrella alternatives for a one-way layout by using linear placement described in Orban and Wolfe (1982). The Monte Carlo simulation is adapted to compare the power of proposed procedure with previous methods.

Keywords: one-way layout, placement, umbrella alternatives, nonparametric

1. 서론

일원배치모형(one-way layout)은 셋 이상의 모집단 간 평균을 비교하기 위한 실험계획법이다. 처리 효과의 차이 유무를 검정하기 위해서는 오차가 서로 독립이고 정규분포를 따르는 확률변수라는 가정이 성립한다면, 모수적 방법인 분산분석법을 사용하여 처리들이 모두 같다는 귀무가설을 검정할 수 있다. 그러나 이 가정을 만족하지 않는 경우에는 제 1종 오류를 제어할 수 있는 비모수적 검정법을 선택해야 한다.

일원배치모형의 대립가설에는 여러 가지 형태가 있는데 우선 각 처리의 효과가 적어도 하나는 다르다는 일반대립가설이 있다. 또한 압력의 세기에 따른 처리 효과의 증감이나 약의 복용량에 따른 수치의 증감 등에 대한 사전 정보가 있는 경우 순서대립가설(ordered alternative; $H_1 : \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$)을 사용한다. 그러나 약의 복용량을 증가시키면 효과가 증가하다가 약의 독성에 따른 부작용으로 인해 약의 효과가 감소하는 경향을 살펴볼 수 있다. 이러한 경향을 우산형 패턴(umbrella pattern)이라고 하며, 처리 효과가 증가의 추세를 보이다가 감소의 추세로 바뀌는 처리 용량 수준을 우산형 패턴의 정점이라 한다 (Mack과 Wolfe, 1981).

일원배치모형에서 일반대립가설에 관한 대표적인 비모수 검정법으로 관측값 대신에 혼합표본에서의 순위(rank)를 이용하는 Kruskal과 Wallis (1952)가 제안한 방법과 Chung과 Kim (2007)이 제안한 결합위치를 이용한 비모수 검정법이 있다. 순서대립가설에 관해서는 Mann과 Whitney가 제안한 U통계량

[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

을 확장하여 순서대립가설에 적용한 Jonckeere (1954)와 Terpstra (1952)가 제안한 검정법이 있다. 또한 일원배치모형에서 updated 위치와 fixed 위치를 이용하여 순서대립가설에 적용한 Kim (1999)이 제안한 검정법이 있다. 우산형 대립가설을 위한 비모수 검정법은 많은 방법이 제안되었다. Mack과 Wolfe (1981)는 k개의 표본 문제에 대해 Mann과 Whitney (1947)의 통계량을 이용하여 검정하였다. Hettmansperger와 Norton (1987)은 k개의 표본 문제에 대해 순위 통계량을 사용한 비모수적인 방법을 제시하였고, Chen과 Wolfe (1990)는 Mack과 Wolfe (1981)가 제안한 검정법을 확장하였다. Chen과 Wolfe (1993)는 우산형 패턴하에서 0용량 대조군보다 처리 효과가 좋은 적어도 하나의 처리 용량이 있는지를 검정하기 위한 방법을 제시하였다. Alvo (2008)는 Spearman distance, Kendall distance를 이용하여 검정하는 방법을 제시하였고, Bhat (2009)는 우산형 패턴의 정점을 아는 경우에 U통계량을 이용한 검정법을 제시하였다.

이 논문에서는 일원배치모형에서 대조군의 크기가 점점 커질수록 더 유용한 Orban과 Wolfe (1982)가 제안한 선형 위치 통계량(linear placement statistic)을 이용하여 우산형 대립가설에 대한 검정을 제안하였다. 선형 위치 통계량의 점수함수(score function)는 정규 점수함수(normal score function)와 지수 점수함수(exponential score function)를 사용했다. 모의실험을 통하여 모수적 방법인 분산분석과 우산형 대립가설에 대해 Mack과 Wolfe (1981)가 제안한 방법, 이 논문에서 제안한 두 가지 검정법들의 검정력(power)을 비교하였다.

2. 제안하는 방법

처리의 수가 k개인 일원배치모형

$$Y_{ij} = \mu + \theta_i + \epsilon_{ij}, \quad i = 0, 1, \ldots, k;\ j = 1, 2, \ldots, n_i$$

에서 μ는 전체 평균을 나타내고 θ_i는 i번째 처리의 효과, ϵ_{ij}는 오차항을 나타내며 동일한 연속분포를 따르는 서로 독립인 확률변수로 가정한다. $Y_{i1}, Y_{i2}, \ldots, Y_{in_i};\ i = 0, 1, \ldots, k$를 i번째 처리에서 추출한 확률표본이라 하고, 전체표본의 크기는 $N = \sum_{i=0}^{k} n_i$이라 하자.

2.1. 순서대립가설에서의 Kim (1999) 검정법

Kim (1999)은 순서대립가설에서 위치를 이용하여 비모수적 검정법을 검정하였다. 순서형 대립가설은 다음과 같다.

$$H_1 : \theta_0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k.$$

0용량 대조군을 포함한 i번째 용량 수준보다 작은 모든 용량 수준들의 조합군과 비교하는 i번째 용량 수준의 j번째 관찰값, Y_{ij}의 updated 위치를 아래와 같이 정의한다.

$$m_i U_{ij} = \left[\text{number of } \ Y'_{sj}s \leq Y_{ij}\right], \quad i = 1, \ldots, k;\ j = 1, \ldots, n_i;\ s = 0, \ldots, i-1, \tag{2.1}$$

여기서 m_i는 i번째 용량 수준보다 작은 모든 용량 수준의 총 표본수를 나타낸다. 즉 $m_i = \sum_{s=0}^{i-1} n_s$이다. 이를 이용한 선형 위치 통계량은 다음과 같다.

$$S_{up} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \phi_{m_i}(U_{ij}), \tag{2.2}$$

여기서 $\phi_{m_i}(\bullet)$는 $[0,1]$의 범위에서 정의된 실수값인 점수함수(score function)이다. 선형 위치 통계량을 만드는 특수한 점수함수로는 정규점수함수(normal score function) $\phi_{m_i}(\bullet) = \Phi^{-1}(x)$($\Phi^{-1}$는 표준정규분포 누적분포함수의 역함수)와 지수점수함수(exponential score function) $\phi_{m_i}(\bullet) = -\ln(1-x)$가 있다. 또한 정규점수함수를 이용한 선형 위치 통계량 S_{up}^{NS}와 지수점수함수를 이용한 선형 위치 통계량 S_{up}^{E}은 다음과 같다.

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

귀무가설 하에서 S_{up}^{NS}와 S_{up}^{E}의 평균과 분산은 아래와 같다.

$$E\left(S_{up}^{NS}\right) = 0,$$
$$\mathrm{Var}\left(S_{up}^{NS}\right) = \sum_{i=1}^{k}\frac{n_i(m_i+n_i+1)v_{m_i+1}}{(m_i+1)(m_i+2)}, \quad v_{m_i+1} = \sum_{l=1}^{m_i+1}\left[\Phi^{-1}\left(\frac{l}{m_i+2}\right)\right]^2,$$
$$E\left(S_{up}^{E}\right) = \sum_{i=1}^{k}\frac{n_i[(m_i+1)\ln(m_i+1)-\ln[(m_i+1)!]]}{m_i+1},$$
$$\mathrm{Var}\left(S_{up}^{E}\right) = \sum_{i=1}^{k}\frac{n_i(m_i+n_i+1)\left[q_{m_i+1}-(m_i+1)\bar{\phi}_{m_i}^2\right]}{(m_i+1)(m_i+2)}, \quad q_{m_i+1} = \sum_{l=1}^{m_i+1}\left[\ln\left(\frac{l}{m_i+1}\right)\right]^2.$$

이때 updated 위치에 정규점수함수와 지수점수함수를 이용한 표준화한 선형 위치 통계량은 각각

$$S_{up}^{NS^*} = \frac{S_{up}^{NS}-E\left(S_{up}^{NS}\right)}{\sqrt{\mathrm{Var}\left(S_{up}^{NS}\right)}}, \tag{2.3}$$
$$S_{up}^{E^*} = \frac{S_{up}^{E}-E\left(S_{up}^{E}\right)}{\sqrt{\mathrm{Var}\left(S_{up}^{E}\right)}} \tag{2.4}$$

이고, 귀무가설 하에서 식 (2.3)과 식 (2.4)는 근사적으로 표준정규분포한다.

2.2. 우산형 대립가설에서 제안한 검정법

우산형 대립가설의 형태는 다음과 같다.

$$H_1 : \theta_0 \le \theta_1 \le \cdots \le \theta_{p-1} \le \theta_p \ge \theta_{p+1} \ge \cdots \ge \theta_{k-1} \ge \theta_k. \tag{2.5}$$

updated 위치를 이용한 검정 방법의 순서형 대립가설에서의 통계량을 우산형 대립가설로 확장하기 위해서 p(peak)를 중심으로 왼쪽 순서형 대립가설($H_L : \theta_0 \le \theta_1 \le \cdots \le \theta_{p-1} \le \theta_p$)과 오른쪽 순서형 대립가설($H_R : \theta_p \ge \theta_{p+1} \ge \cdots \ge \theta_{k-1} \ge \theta_k$)을 나눌 수 있다. 또한 위의 순서형 대립가설의 통계량 (2.2)를 두 개로 나뉜 왼쪽과 오른쪽 대립가설에 각각 적용하였다.

왼쪽 순서형 대립가설에 대한 위치(placement) U_{ij}^L는

$$m_i^L U_{ij}^L = \left[\text{number of } Y_{sj}'s \le Y_{ij}\right], \quad i=1,\ldots,p;\ j=1,\ldots,n_i;\ s=0,\ldots,i-1$$
$$= [0\text{번째부터 } (i-1)\text{번째 처리군까지 } Y_{ij}\text{보다 작거나 같은 표본의 개수}]$$

라고 정의한다. 여기서 m_i^L은 i번째 용량 수준보다 작은 모든 총 표본수 $m_i^L = \sum_{s=0}^{i-1} n_s$이다.
또한 오른쪽 순서형 대립가설에 대한 위치(placement) $U_{k-i,j}^R$는

$$\begin{aligned} m_{k-i}^R U_{k-i,j}^R &= \left[\text{number of } Y'_{k-s,j} s \le Y_{k-i,j}\right], \quad i=1,\ldots,k-p;\ j=1,\ldots,n_i;\ s=0,\ldots,i-1 \\ &= [(k-i+1)\text{번째부터 } k\text{번째 처리군까지 } Y_{k-i,j}\text{보다 작거나 같은 표본의 개수}] \end{aligned}$$

라고 정의한다. 여기서 $m_{k-i}^R = n_k + n_{k-1} + \cdots + n_{k-i+1} = \sum_{s=0}^{i-1} n_{k-s}$이다.
대조군과 처리군의 위치를 이용하는 왼쪽과 오른쪽 순서형 대립가설에 대한 통계량은 각각

$$S_{up}^L = \sum_{i=1}^{p}\sum_{j=1}^{n_i} \phi_{m_i}\left(U_{ij}^L\right),$$

$$S_{up}^R = \sum_{i=1}^{k-p}\sum_{j=1}^{n_{k-i}} \phi_{m_{k-i}}\left(U_{k-i,j}^R\right)$$

와 같이 정의된다. 서로 독립인 U_{ij}^L와 $U_{k-i,j}^R$의 통계량을 이용하는 우산형 대립가설의 검정통계량은

$$S_{up} = S_{up}^L + S_{up}^R = \sum_{i=1}^{p}\sum_{j=1}^{n_i} \phi_{m_i}\left(U_{ij}^L\right) + \sum_{i=1}^{k-p}\sum_{j=1}^{n_{k-i}} \phi_{m_{k-i}}\left(U_{k-i,j}^R\right)$$

이며 귀무가설 $H_0 : \theta_0 = \theta_1 = \cdots = \theta_k$를 검정하기 위한 기각역은 $S_{up} \ge S_\alpha$이다. 여기서 S_α는 귀무가설 하에 $P_0\left[S_{up} \ge S_\alpha\right] = \alpha$를 만족하는 상수이다.
각각의 점수함수로 정규점수함수(normal score function)와 지수점수함수(exponential score function)에 따른 통계량은 다음과 같다.

$$\begin{aligned} S_{up}^{NS} &= \sum_{i=1}^{p}\sum_{j=1}^{n_i} \Phi^{-1}\left(\frac{m_i^L U_{ij}^L + 1}{m_i^L + 2}\right) + \sum_{i=1}^{k-p}\sum_{j=1}^{n_{k-i}} \Phi^{-1}\left(\frac{m_{k-i}^R U_{k-i,j}^R + 1}{m_{k-i}^R + 2}\right), \\ S_{up}^{E} &= -\sum_{i=1}^{p}\sum_{j=1}^{n_i} \ln\left(1 - \frac{m_i^L U_{ij}^L}{m_i^L + 1}\right) - \sum_{i=1}^{k-p}\sum_{j=1}^{n_{k-i}} \ln\left(1 - \frac{m_{k-i}^R U_{k-i,j}^R}{m_{k-i}^R + 1}\right). \end{aligned}$$

귀무가설 하에서 S_{up}^{NS}와 S_{up}^E의 평균과 분산은 아래와 같다.

$$\begin{aligned} E\left(S_{up}^{NS}\right) &= 0, \\ \mathrm{Var}\left(S_{up}^{NS}\right) &= \sum_{i=1}^{p} \frac{n_i\left(m_i^L + n_i + 1\right) v_{m_i^L+1}}{\left(m_i^L+1\right)\left(m_i^L+2\right)} + \sum_{i=1}^{k-p} \frac{n_{k-i}\left(m_{k-i}^R + n_{k-i} + 1\right) v_{m_{k-i}^R+1}}{\left(m_{k-i}^R+1\right)\left(m_{k-i}^R+2\right)}, \\ E\left(S_{up}^{E}\right) &= \sum_{i=1}^{p} \frac{n_i\left[\left(m_i^L+1\right)\ln\left(m_i^L+1\right) - \ln\left[\left(m_i^L+1\right)!\right]\right]}{m_i^L+1} \\ &\quad + \sum_{i=1}^{k-p} \frac{n_{k-i}\left[\left(m_{k-i}^R+1\right)\ln\left(m_{k-i}^R+1\right) - \ln\left[\left(m_{k-i}^R+1\right)!\right]\right]}{m_{k-i}^R+1}, \\ \mathrm{Var}\left(S_{up}^{E}\right) &= \sum_{i=1}^{p} \frac{n_i(m_i^L + n_i + 1)\left[q_{m_i^L+1} - \left(m_i^L+1\right)\bar{\phi}^2_{m_i^L}\right]}{\left(m_i^L+1\right)\left(m_i^L+2\right)} \\ &\quad + \sum_{i=1}^{k-p} \frac{n_{k-i}\left(m_{k-i}^R + n_{k-i} + 1\right)\left[q_{m_{k-i}^R+1} - \left(m_{k-i}^R+1\right)\bar{\phi}^2_{m_{k-i}^R}\right]}{\left(m_{k-i}^R+1\right)\left(m_{k-i}^R+2\right)}. \end{aligned}$$

이때 updated 위치에 정규점수함수와 지수점수함수를 이용한 표준화한 선형 위치 통계량은 각각

$$S_{up}^{NS^*} = \frac{S_{up}^{NS} - E\left(S_{up}^{NS}\right)}{\sqrt{\mathrm{Var}\left(S_{up}^{NS}\right)}},$$
$$S_{up}^{E^*} = \frac{S_{up}^{E} - E\left(S_{up}^{E}\right)}{\sqrt{\mathrm{Var}\left(S_{up}^{E}\right)}}$$

이고, 귀무가설 하에서 위의 두 식은 근사적으로 표준정규분포한다.

3. 모의실험 및 결과

본 논문에서는 updated 위치를 이용한 정규점수함수와 지수점수함수 검정통계량에 근거하여 새롭게 제시한 방법과 기존의 검정법들의 비교를 위해 모수적 방법과 비모수적 방법과의 차이를 비교해보았다. 모수적인 방법으로는 F-통계량을 이용한 분산분석법(ANOVA)을 사용하고, 비모수적인 방법으로는 Mack과 Wolfe (1981)가 제안한 검정법을 사용하였다. 이 방법은 우산형 대립가설에 대한 기존방법들 중 일반적인 우산형 패턴에서 가장 우수한 검정력을 보였기 때문이다. 모집단의 분포로는 정규분포, 비대칭적인 지수분포, 정규분포보다 꼬리가 두꺼운 Cauchy 분포, 정규분포보다 꼬리가 약간 두꺼운 이중지수분포를 고려하였다. SAS를 이용하여 정규분포의 난수생성은 RANNOR 함수를 이용하였다. 지수분포의 난수생성은 RANEXP 함수, Cauchy 분포의 난수생성은 RANCAU 함수를 이용하였다. 이중지수분포는 RANUNI 함수를 이용하여 역변환 방법으로 난수를 생성하였다. 또한 유의수준 α는 0.05로 하였다.

처리의 수는 5개인 경우를 선택하였고 각 표본의 크기는 10으로 같은 경우를 고려하였다. 그리고 유의수준 α를 0.05로 보정하기 위해 확률화 검정을 사용하였다. 이러한 조건에서 각 방법들의 검정력을 비교하기 위해 10,000번 반복하는 Monte Carlo 모의실험을 실시하였다. 이 결과를 Table 3.1과 Table 3.2로 정리하였다.

각 처리의 효과가 모두 동일할 경우에 유의수준이 0.05를 만족하는지 살펴보면, 분산분석법의 실험유의수준은 정규분포, 지수분포인 Table 3.1에서 0.0546, 0.0485, 그리고 이중지수분포와 Cauchy 분포인 Table 3.2에서 0.0480, 0.0165의 값들을 얻었다. Cauchy 분포일 때, 분산분석법은 1종 오류를 제어하지 못함을 의미한다. 또한 Mack과 Wolfe (1981)가 제안한 분석법의 실험유의수준은 정규분포, 지수분포인 Table 3.1에서 0.0541, 0.0521 이중지수분포와 Cauchy 분포인 Table 3.2에서 0.0490, 0.0521의 값들을 얻었다. 각 분포에서 실험유의수준으로 1종 오류를 제어함을 볼 수 있다. 본 논문에서 제안한 방법들 또한 각 분포에서의 실험유의수준으로 1종 오류를 제어하는데 0.05를 약간 상회하지만 큰 문제가 없음을 보여주고 있다.

검정력을 살펴보면 대립가설의 형태에 따라 다름을 알 수 있다. 우선 정규분포에서는 순서형 대립가설에서 평균의 차이가 클수록 분산분석의 효율이 좋음을 알 수 있다. 순서형 대립가설에서 평균 차이가 큰 경우, Cauchy 분포를 제외한 모든 분포에서 F-통계량의 검정력만 월등하게 높게 나왔다. 이는 평균 차이가 크고 우산형 대립가설과는 형태가 많이 달라 이런 결과를 보였다. 초반에 커지다가 중간부터 작아지는 우산형 대립가설에서는 평균의 차이가 클수록 Mack과 Wolfe (1981)가 제안한 방법의 효율이 좋음을 알 수 있다. 처음에 작았다가 중간부분에서 커져 평행을 이루다가 마지막에 작아지는 우산형 대립가설의 형태에서 updated 위치 검정법의 효율이 가장 좋다. 계속 평행하다가 마지막에 줄어드는 대립가설에서도 updated 위치 검정법의 효율이 좋았다. 또한 중간 정도까지 평행하다가 작아져서 또다시 평행하는 대립가설에서도 평균의 차이가 클수록 분산분석의 효율이 좋음을 알 수 있다.

Table 3.1. Monte Carlo power estimates ($\alpha = 0.05$)

분포	θ_0	θ_1	θ_2	θ_3	θ_4	F	M&W	UNS	UES
	0	0	0	0	0	0.0546	0.0541	0.0497	0.0506
	0	0.1	0.2	0.3	0.3	0.0866	0.0689	0.0654	0.0674
	0	0.5	1	1.5	1.5	0.8994	0.1238	0.1641	0.1973
	0	0.2	0.4	0.2	0	0.1127	0.2525	0.2448	0.2281
	0	0.5	1	0.5	0	0.4980	0.7961	0.7895	0.7572
Normal	0.1	0.5	0.9	0.1	0	0.4106	0.6697	0.6253	0.5863
	0	0.5	1	0	0	0.5563	0.7898	0.7302	0.6963
	0	1	1	1	0	0.7617	0.7745	0.8465	0.8668
	0.5	0.5	0.5	0	0	0.2309	0.1467	0.1377	0.1197
	0.5	0.5	0.5	0.5	0	0.0640	0.1480	0.1564	0.1618
	0	0.3	0	0.3	0	0.1093	0.0520	0.0684	0.0868
	0	0	0	0	0	0.0485	0.0521	0.0505	0.0517
	0	0.1	0.2	0.3	0.3	0.0653	0.0618	0.0625	0.0631
	0	0.5	1	1.5	1.5	0.6967	0.1139	0.1402	0.1509
	0	0.2	0.4	0.2	0	0.0844	0.2166	0.1979	0.1675
	0	0.5	1	0.5	0	0.3197	0.7062	0.6460	0.5367
Exponential	0.1	0.5	0.9	0.1	0	0.2599	0.5755	0.4819	0.3849
	0	0.5	1	0	0	0.3611	0.6998	0.5764	0.4710
	0	1	1	1	0	0.5284	0.6756	0.7030	0.6613
	0.5	0.5	0.5	0	0	0.1548	0.1276	0.1110	0.0908
	0.5	0.5	0.5	0.5	0	0.1156	0.1324	0.1330	0.1283
	0	0.3	0	0.3	0	0.0821	0.0560	0.0746	0.0866

F: ANOVA for one-way layout; M&W: Mack과 Wolfe가 제안한 방법;
UNS: 정규점수함수를 이용한 선형 위치 통계량; UES: 지수점수함수를 이용한 선형 위치 통계량.

지수분포에서도 대립가설의 형태에 따라 효율성이 많이 달라지는 것을 알 수 있다. 정규분포와 같이 순서형 대립가설의 형태에서 평균의 차이가 클수록 분산분석의 효율이 가장 좋았고, 초반에 커지다가 중간부터 작아지는 우산형 대립가설에서 Mack과 Wolfe (1981) 검정법의 효율이 좋다. 처음에 작았다가 중간부분에서 커져 평행을 이루다가 마지막에 작아지는 대립가설의 형태에서 정규점수함수를 이용한 updated 위치 검정법의 효율이 좋음을 알 수 있다. 계속 평행하다가 마지막에 줄어드는 대립가설에서도 정규점수를 이용한 updated 위치 검정법의 효율이 좋았다. 중간까지 평행하다가 작아져서 또다시 평행하는 대립가설에서 분산분석의 효율이 좋음을 알 수 있다.

이중지수분포에서는 순서형 대립가설에서 평균의 차이가 작을 때에는 지수점수함수를 이용한 updated 위치 검정법의 효율이 좋았고, 평균의 차이가 클 때에는 분산분석의 효율이 가장 좋음을 알 수 있고, 초반에 커지다가 중간부터 작아지는 우산형 대립가설에서 updated 위치 검정법의 효율이 가장 좋았다. 처음에 작았다가 중간부분에서 커져 평행을 이루다가 마지막에 작아지는 대립가설과 초반에 커지다가 중간부터는 작아지는 대립가설에서도 updated 위치 검정법의 효율성이 좋음을 알 수 있다. 중간까지 평행하다가 작아져서 또 다시 평행하는 대립가설에서 분산분석의 효율이 좋음을 알 수 있다.

Cauchy 분포에서도 대립가설의 형태에 따라 효율성이 많이 달라지는 것을 알 수 있는데, 순서형 대립가설에서 평균의 차이가 작을 때에는 Mack과 Wolfe (1981)의 효율이 좋고, 평균의 차이가 클 때에는 지수점수함수를 이용한 updated 위치 검정법의 효율이 좋았다. 초반에 커지다가 중간부터 작아지는 우산형 패턴에서 Mack과 Wolfe (1981)의 효율이 가장 좋음을 알 수 있다. 처음에 작았다가 중간부분에

Table 3.2. Monte Carlo power estimates ($\alpha = 0.05$)

분포	θ_0	θ_1	θ_2	θ_3	θ_4	F	M&W	UNS	UES
Double exponential	0	0	0	0	0	0.0480	0.0521	0.0505	0.0517
	0	0.1	0.2	0.3	0.3	0.0633	0.0617	0.0623	0.0642
	0	0.5	1	1.5	1.5	0.6102	0.1099	0.1383	0.1678
	0	0.2	0.4	0.2	0	0.1092	0.4603	0.3865	0.2559
	0	0.5	1	0.5	0	0.6024	0.6301	0.6696	0.6967
	0.1	0.5	0.9	0.1	0	0.2235	0.5599	0.4974	0.4400
	0	0.5	1	0	0	0.3135	0.6656	0.5909	0.5256
	0	1	1	1	0	0.4481	0.6496	0.7022	0.7015
	0.5	0.5	0.5	0	0	0.1308	0.1306	0.1082	0.1003
	0.5	0.5	0.5	0.5	0	0.0995	0.1313	0.1313	0.1374
	0	0.3	0	0.3	0	0.1071	0.0618	0.1077	0.1219
Cauchy	0	0	0	0	0	0.0165	0.0490	0.0475	0.0506
	0	0.1	0.2	0.3	0.3	0.0168	0.0574	0.0549	0.0571
	0	0.5	1	1.5	1.5	0.0447	0.0865	0.0923	0.1045
	0	0.2	0.4	0.2	0	0.0171	0.1399	0.1264	0.1170
	0	0.5	1	0.5	0	0.0266	0.3872	0.3482	0.3027
	0.1	0.5	0.9	0.1	0	0.0245	0.3141	0.2640	0.2229
	0	0.5	1	0	0	0.0280	0.3798	0.3125	0.2622
	0	1	1	1	0	0.0350	0.3704	0.3872	0.3772
	0.5	0.5	0.5	0	0	0.0204	0.0957	0.0827	0.0745
	0.5	0.5	0.5	0.5	0	0.0190	0.0987	0.0963	0.0976
	0	0.3	0	0.3	0	0.0179	0.0498	0.0593	0.0742

F: ANOVA for one-way layout; M&W: Mack과 Wolfe가 제안한 방법;
UNS: 정규점수함수를 이용한 선형 위치 통계량; UES: 지수점수함수를 이용한 선형 위치 통계량.

서 커져 평행을 이루다가 마지막에 작아지는 대립가설과 초반에 평행하다가 중간부터 점점 줄어드는 대립가설의 형태에서 updated 위치 검정법의 효율성이 좋음을 알 수 있다. 또한 중간 정도까지 평행하다가 작아져서 또다시 평행하는 대립가설과 계속 평행하다가 마지막에 줄어드는 대립가설에서는 Mack과 Wolfe (1981)의 효율이 좋았다.

4. 결론 및 고찰

본 논문에서는 일원배치모형에서 우산형 대립가설을 검정하기 위한 비모수적 방법으로 위치를 이용한 통계량을 제안하였다. 이 통계량은 Orban과 Wolfe (1982)의 논문에서 사용된 선형 위치(linear placement)에 정규점수함수(normal score function)와 지수점수함수(exponential score function)를 이용하여 만들어졌다.

모의실험을 통하여 이 검정방법을 정규분포, 지수분포, Cauchy 분포 그리고 이중지수분포에서 모수적 검정방법과 여러 비모수적 방법을 사용한 경우를 비교한 결과 분포에 따라 효율성이 다름을 알 수 있었다. 모의실험의 전체적인 결과를 살펴보면 순서형 대립가설에서는 분산분석법의 검정력이 가장 높았다. 이는 순서형 대립가설에서 평균 차이가 큰 경우, Cauchy 분포를 제외한 모든 분포에서 F-통계량의 검정력만 월등하게 높게 나왔는데, 평균 차이가 크고 우산형 대립가설과는 형태가 많이 달라 이런 결과를 보였다. 우산형 패턴에서 정점이 명확한 우산형 대립가설에서는 Mack과 Wolfe (1981)의 검정력이 가장 높았으며, 우산형 패턴에서 정점이 평행하는 우산형 대립가설에서는 정규점수함수를 이용했을 때의

검정력이 높게 나타났다. 이는 즉 우산형 패턴의 정점부분이 평행할 때 제안한 방법이 가장 좋은 방법이라 할 수 있다. 순서형 대립가설이나 중간정도에서 평행하다가 줄어드는 대립가설, 커졌다 작아졌다를 반복하는 대립가설에서는 정규점수함수를 이용했을 때보다 지수점수함수를 이용했을 때 더 높은 검정력을 보여주었다. 반면에 우산형 대립가설에서는 지수점수함수를 이용했을 때보다 정규점수함수를 이용했을 때 더 높은 검정력을 보여주었다.

본 논문에서 제안된 검정법은 모든 분포에서 Mack과 Wolfe (1981)의 방법의 검정력과 유사한 결과를 보이고 있다. 또한 우산형 패턴의 정점 부분이 평행할 때에는 Mack과 Wolfe (1981)가 제안한 검정법보다 더 효율적이라 말할 수 있다. 따라서 이런 경우에는 본 논문에서 제시한 위치를 이용한 검정방법을 사용하는 것이 더 좋은 방법이 될 것이다. 본 논문에서 제안하는 방법으로 우산형 대립가설을 검정하고자 할 때에는 용량별 표본 수가 10개 이상으로 충분히 큰지 확인이 필요할 것이다. 또한 처리군의 용량수준이 대조군과의 효과 차이가 작은 그룹이 존재한다면 매우 낮은 검정력을 갖는 문제점이 보인다. 이는 추후에 더 보완해야 할 것으로 보인다.

References

Alvo, M. (2008). Nonparametric tests of hypotheses for umbrella alternatives, *Canadian Journal of Statistics*, **36**, 143–156.

Bhat, S. V. (2009). Simple K-sample rank tests for umbrella alternatives, *Research Journal of Mathematics and Statistics*, **1**, 27–29.

Chen, Y. I. and Wolfe, D. A. (1990). A study of distribution-free tests for umbrella alternatives, *Biometrical Journal*, **32**, 48–57.

Chen, Y. I. and Wolfe, D. A. (1993). Nonparametric procedures for comparing umbrella pattern treatment effects with a control in a one-way layout, *Biometrics*, 455–465.

Chung, T. S. and Kim, D. (2007). Nonparametric method using placement in one-way layout, *The Korean Communications in Statistics*, **14**, 551–560.

Hettmansperger, T. P. and Norton, R. M. (1987). Tests for patterned alternatives in k-sample problems, *Journal of the American Statistical Association*, **82**, 292–299.

Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives, *Biometrika*, **41**, 133–145.

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.

Mack, G. A. and Wolfe, D. A. (1981). K-sample rank tests for umbrella alternatives, *Journal of the American Statistical Association*, **76**, 175–181.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other, *The Annals of Mathematical Statistics*, **18**, 50–60.

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.

Terpstra, T. J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking, *Indagationes Mathematicae*, **14**, 327–333.

일원배치법에서 Umbrella Alternatives에 대한 위치를 이용한 비모수 검정법

이혜정[a] · 김동재[a,1]

[a]가톨릭대학교 의생명 · 건강과학과

(2015년 10월 12일 접수, 2015년 11월 5일 수정, 2015년 11월 11일 채택)

요 약

임상시험에서 약의 복용량에 따라 처리 효과가 증가하다가, 부작용으로 인해 일정 용량 수준에서 감소하는 추세를 보일 수 있다. 이러한 경향을 우산형 패턴이라 하며, 우산형 패턴의 대립가설에 대한 검정은 이러한 경향이 사전에 예측 가능할 때 유용하다. 이 논문에서는 Orban과 Wolfe (1982)가 제안한 선형 위치(linear placement)를 이용하여 일원배치법에서 우산형 대립가설의 비모수적 검정법을 제안하였다. 또한 Monte Carlo 모의실험을 통하여 기존의 방법들과 검정력(power)을 비교하였다.

주요용어: 일원배치법, 위치, 우산형 대립가설, 비모수 방법

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