

Nonparametric Tests for 2×2 Cross-Over Design

Kyuhoon Gee^a, Dongjae Kim^{1,a}

^aDepartment of Biostatistics, The Catholic University of Korea

Abstract

A 2×2 Cross-over design is widely used in clinical trials for comparison studies of two kinds of drugs or medical treatments. This design has many statistical methods such as Hills-Armitage's (1979) method or Koch's (1972) method. In this paper, we propose a nonparametric test for 2×2 Cross-over design based on a two-sample test suggested by Baumgartner *et al.* (1998). In addition, a Monte Carlo simulation study is adapted to compare the power of the proposed methods with those of previous methods.

Keywords: Cross-over design, two-sample tests, nonparametric, comparison of power.

1. 서론

의과학 분야에서 임상시험(clinical trials)을 통한 새로운 약제 또는 치료법의 개발은 매우 중요하며 일반적인 일이다. 특히 제 3상 임상시험(phase III)에서는 기존의 약제 또는 치료법과 비교연구를 하여야만 한다. 이와 같이 두 가지 처리를 비교하는 방법은 독립인 두 표본의 여러 가지 검정법들이 널리 이용된다. 그러나 이 방법은 각 개체의 치료 전 상태가 매우 다를 수 있고 또한 치료에 대한 반응 역시 매우 다를 수 있다는 어려움이 있다. 이를 보완하기 위한 한 방법으로 두 가지 처리를 같은 개체에 교차시켜 적용하는 실험계획법을 교차계획법(cross-over design)이라 한다 (Kim, 1999).

교차계획법은 같은 개체에 두 가지 처리를 모두 적용하여 측정된 반응값으로 통계적 분석을 하므로 변이가 비교적 적어 검정의 효율성이 높다는 장점이 있다. 반면에 첫 번째 적용한 처리의 효과가 두 번째 적용한 처리에까지 영향을 미칠 수도 있고, 개체의 상태가 두 번째 처리의 적용시기에서 안정되지 유지되지 않을 수 있다는 단점을 가지고 있다. 이런 이유로 특수한 의학연구에서와 같이 완치된 치료효과 보다는 짧은 기간의 치료반응에 관심을 기울이는 경우에 널리 이용한다. 예를 들어, 고혈압, 천식, 관절염과 같은 만성질환(chronic disease)의 경우처럼 완전치유라는 멀고 먼 결과보다는 비교적 단기간의 증상의 변화 또는 치료의 반응을 알고자 할 때 자주 이용한다 (Park과 Song, 1998).

가장 단순한 2×2 교차계획법은 같은 개체에 두 가지 약물 또는 치료, 즉 A와 B 처리를 모두 적용하게 된다. 처리의 적용 순서에 의해 달라질 수 있는 처리의 효과를 감안하여, 계획된 연구 대상자를 두 군으로 랜덤하게 나누어 한 군은 처리 A를 먼저 적용한 후 처리 B를 적용하고, 다른 군은 그 반대의 순서로 처리를 적용한다. 이 때 어떤 개체가 치료순서 AB군에 속하고 어떤 개체가 치료순서 BA군에 속하게 될지를 랜덤하게 결정한다 (Park과 Song, 1998).

이 2×2 교차계획 자료의 분석은 두 가지 처리의 직접 효과(treatment direct effect)의 차이 여부 검정에 주목적이 있으며, 이월 효과(carry-over effect)와 시기 효과(period effect)의 차이 유무도 검정할 수 있다.

¹ Corresponding author: Professor, Department of Biostatistics, The Catholic University of Korea, Banpo-Dong, Seocho-Gu, Seoul 137-701, Korea. E-mail: djkim@catholic.ac.kr

2×2 교차계획법을 분석하는 기존의 방법으로는 여러 가지가 있다. 대표적으로 Hills-Armitage (1979)의 방법이 있다. 이는 모수적 검정법으로 자료의 합(sum)과 차(difference)를 T-test에 적용하는 방법이며 계산 과정이 간단하고 많이 이용되고 있다. 이 외에도 비모수적 검정법으로 Koch (1972)와 Kim (1999)의 방법 등이 있으며 Hills-Armitage (1979)의 방법과 유사하게 원 자료의 합과 차를 각각의 비모수적 검정법에 적용하는 방법이다. 이와 같이 비모수적 검정법을 적용한 방법들은 모집단이 정규성 가정을 만족하지 않아도 된다는 장점이 있다.

본 논문에서는 Baumgartner 등 (1998)으로부터 제시된 자료의 순위(rank)를 이용한 이표본 검정법을 2×2 교차계획법에 적용하는 새로운 비모수적 검정 방법을 제안한다. 이 Baumgartner 등 (1998)의 이표본 검정법은 모의실험에서 대립가설의 형태나 모집단이 어떠한 분포를 따르는지에 상관없이 높은 검정력을 유지한다는 점을 보여주고 있다 (Baumgartner 등, 1998). 이러한 장점이 여러 효과의 검정을 하는 교차계획법의 특성에 유용할 것이라 사료되며 이를 소표본 모의실험을 통하여 제안한 검정법과 위에 언급된 기존의 검정법들의 검정력을 비교하며 알아보려고 한다.

2. 수리적 모형

Grizzle (1965)의 논문으로부터 2×2 교차계획법의 적절한 모형은 다음과 같다.

$$y_{ijk} = \mu + b_{ij} + p_k + \tau_l + \rho_{l'} + e_{ijk} \quad i = 1, 2; j = 1, \dots, n_i; k = 1, 2; l, l' = 1, 2; \quad (2.1)$$

μ : 전체평균

b_{ij} : i 번째 순서(sequence)의 그룹 안에서 j 번째 개체(subject)의 효과

p_k : k 번째 시기(period)의 효과

τ_l : l 번째 처리의 직접 효과(treatment effect)

$\rho_{l'}$: l' 번째 처리의 이월 효과(carry-over effect)

e_{ijk} : 측정값의 랜덤 오차

여기서, b_{ij} 는 서로 독립적이며 e_{ijk} 는 $E(e_{ijk}) = 0$, $\text{Var}(e_{ijk}) = \sigma^2$, $\text{Cov}(e_{ijk}, e_{ijk'}) = 0$ 임을 가정한다. Grizzle은 위 모형에서 두 처리의 직접 효과, 이월 효과, 시기 효과에 대한 검정법을 제시하였고 본 논문에서도 같은 질문에 대해서 논의하도록 한다. 단, 비모수적 검정법이기에 정규성 가정은 포함하지 않는다.

3. 제안된 검정법

이 절에서는 순위를 이용한 새로운 검정법에 대해 논의한다. 이 검정법은 Baumgartner 등 (1998)에 의해 제안되었으며 경험적 분포에 근거하는 비모수적 이표본 검정 방법이다.

3.1. Baumgartner 등 (1998)의 새로운 순위 검정법

X_1, X_2, \dots, X_m 와 Y_1, Y_2, \dots, Y_n 를 각각의 모집단에서 추출한 독립인 확률표본이라 하고 F 와 G 를 각 모집단을 따르는 연속적인 분포라고 하자. 또한, 각각의 분포함수는 위치모수를 제외하고 같은 분포함수를 갖는다고 고려해보자. 즉,

$$G(t) = F(t - \theta)$$

이 때 X 와 Y 의 혼합표본의 순서통계량을 Z_1, Z_2, \dots, Z_N ($N = m + n$)라 하면 경험적 분포함수(empirical distribution function) $F_N(z)$ 는 다음과 같이 정의된다.

$$F_N(z) = \begin{cases} 0, & \text{for } z < Z_1, \\ i/N, & \text{for } Z_i \leq z < Z_{i+1}, i = 1, 2, \dots, N - 1, \\ 1, & \text{for } z \geq Z_N. \end{cases}$$

Baumgartner 등 (1998)은 여기서 두 모집단의 경험적 분포함수의 차이($F_m(z) - F_n(z)$)에 분산($z(1 - z)$)을 가중한 검정통계량 B 를 제안하였고 그 식은 다음과 같다.

$$B = \frac{mn}{m+n} \int_0^1 \frac{1}{z(1-z)} (F_m(z) - F_n(z))^2 dz,$$

여기서 R_i 와 H_j 를 Z 에서 각각 표본 X 와 Y 에 해당하는 값이라 할 때 식 (3.1)은 다음과 같이 근사한다,

$$B_X = \frac{1}{m} \sum_{i=1}^m \frac{\left(R_i - \frac{m+n}{m} \times i\right)^2}{\left(\frac{i}{m+1}\right)\left(1 - \frac{i}{m+1}\right)\left(\frac{n(m+n)}{m}\right)}$$

또는

$$B_Y = \frac{1}{n} \sum_{j=1}^n \frac{\left(H_j - \frac{m+n}{n} \times j\right)^2}{\left(\frac{j}{n+1}\right)\left(1 - \frac{j}{n+1}\right)\left(\frac{m(m+n)}{n}\right)}.$$

결론적으로 귀무가설 $H_0 : \theta = 0$ 을 검정하기 위한 검정통계량 B 는 다음과 같이 나타낸다.

$$B = \frac{1}{2}(B_X + B_Y)$$

통계량 B 는 $m \rightarrow \infty, n \rightarrow \infty, m/n \rightarrow \alpha$ (임의의 상수)이면 다음과 같은 근사분포를 가지며 이를 이용하여 검정할 수 있다.

$$\psi(b) = \lim_{n,m \rightarrow \infty} \Pr(B < b) = \sqrt{\frac{\pi}{2}} \frac{1}{b} \sum_{j=0}^{\infty} \binom{-1/2}{j} (4j+1) \int_0^1 \frac{1}{\sqrt{r^3(1-r)}} \exp\left(\frac{rb}{8} - \frac{\pi^2(4j+1)^2}{8rb}\right) dr$$

여기서 $\binom{-1/2}{j} = \{(-1)^j \Gamma(j+1/2)\} / \{\Gamma(1/2)j!\}$ 이다.

3.2. 이월 효과의 동일성 검정

실제로는 이월 효과가 없음을 검정하기 원하지만 시기 효과가 있는 경우에는 이월 효과가 없음을 검정할 수 없다. 따라서 두 약물 또는 약제의 이월 효과가 같다는 사실을 검정한다. 식 (2.1)로부터 같은 환자의 두 관측값은

$$y_{ij1} + y_{ij2} = 2(\mu + b_{ij}) + (p_1 + p_2) + (\tau_1 + \tau_2) + \rho_i + (e_{ij1} + e_{ij2})$$

이고 ρ_i 는 ii' 순서에서 i 번째 처리의 이월 효과를 나타낸다. 그러므로 이월 효과의 동일성 검정

$$H_{0\rho} : \rho_1 = \rho_2$$

은 개체 안에서의 합을 3.1절에서 논의한 검정법에 적용하여 검정할 수 있다.

3.3. 이월 효과가 같을 때의 처리 효과의 검정

위의 검정과정에서 이월 효과가 같다고 가정하고 처리 효과의 검정은 식 (2.1)로부터 같은 환자의 두 관측값의 차이가

$$y_{ij1} - y_{ij2} = (p_1 - p_2) + (-1)^{i+1}(\tau_1 - \tau_2) + (e_{ij1} - e_{ij2})$$

이고 순서 AB에서는 $(-1)^{i+1} = 1$, 순서 BA에서는 $(-1)^{i+1} = -1$ 이다. 그러므로 두 처리의 직접 효과의 차이가 없다는 가설

$$H_{0\tau} : \tau_1 = \tau_2$$

의 검정은 개체 안에서의 차를 3.1절에서 논의한 검정법에 적용하여 검정할 수 있다.

만약 두 치료의 이월효과가 유의하게 다르다고 밝혀진 경우에는, 두 번째 시기에 대한 관측값에는 첫 번째 치료의 이월효과가 함께 섞여 있으므로(contaminated), 두 번째 시기의 관측값은 분석에 사용할 수 없다. 결과적으로 첫 번째 시기의 관측값만으로 처리의 효과를 검정하게 된다. 그러므로 이처럼 이월효과가 있는 경우에는 두 시기의 자료를 사용하는 교차계획법의 장점이 없게 된다.

3.4. 이월 효과가 없을 때의 시기 효과의 검정

단순한 2x2 교차계획법에서는 위의 두 가지 검정으로 그치는 것이 보통이다. 그러나 연구의 특성에 따라서는 이월효과가 전혀 개입되지 않아서 $p_1 = p_2 = 0$ 임을 가정하는 것에 무리가 없는 경우가 있겠다. 이 때 치료효과에 대한 검정과 더불어 다음에 제시하는 시기효과에 대한 검정이 가능하다.

순서 AB에서는

$$y_{1j1} - y_{1j2} = (p_1 - p_2) + (\tau_1 - \tau_2) + (e_{1j1} - e_{1j2})$$

이고, 순서 BA에서는

$$y_{2j1} - y_{2j2} = (p_2 - p_1) + (\tau_1 - \tau_2) + (e_{2j1} - e_{2j2})$$

인 교차 차이를 생각해보자. 그러면 시기 효과의 차이가 없다는 가설

$$H_{0p} : p_1 = p_2$$

의 검정은 개체 안에서의 교차 차이를 3.1절에서 논의한 검정법에 적용하여 검정할 수 있다.

4. 모의실험의 계획 및 결과

본 논문에 제안한 검정법과 기존의 검정법들을 비교하기 위하여 소표본 모의실험을 실시하였다. 제안한 검정법을 비롯하여 T-test를 적용한 Hills-Armitage (1979)의 방법, Wilcoxon rank sum test를 적용한 Koch (1972)의 방법, 그리고 Placement 통계량을 적용한 Kim (1999)의 방법을 비교한다. 단, 여기서 Kim (1999)의 방법은 2가지 점수함수(정규점수함수와 지수점수함수)를 이용하여 모두 5가지의 검정 방법을 비교하였다.

세 가지 효과(이월 효과, 처리 효과, 시기 효과)에서 각각 차이가 없을 경우와 있는 경우를 고려하며 각 효과의 검정은 교차계획법의 검정 방법에 따라 시행하였다. 정확히 말하자면 먼저 이월 효과의 검정을 한 후에 처리와 시기 효과의 검정을 하는데 처리 효과 검정의 경우, 이월 효과의 검정에서 채

택할 경우에만 두 시기의 자료를 모두 이용한 처리 효과 검정(Treatment I)을 하며 이월 효과의 검정에서 기각할 경우에는 첫 번째 시기의 자료만을 이용한 처리 효과 검정(Treatment II)을 한다. 마찬가지로 시기 효과 검정의 경우, 이월 효과의 검정에서 채택할 경우에만 이월 효과가 없다고 가정하고 시기 효과의 검정을 실시한다. 위 절차에 나타나 있듯이 이월 효과의 검정 결과에 따라 처리 효과의 검정 방법이 달라지고 시기 효과의 실시 여부를 결정하기에 총 4가지의 상황, 즉 이월 효과의 차이가 없을 때 처리와 시기 효과의 차이가 유무, 이월 효과의 차이가 있을 때 처리와 시기 효과의 차이가 유무일 경우를 각각 고려하였다. 그리고 처리와 시기 효과에 대한 검정은 같은 개체에게서 측정된 수치의 차를 사용하여 검정하므로 비교적 정밀도(precision)가 높은 예민한 검정법(sensitive test)으로 효과의 차이를 $\theta_r = \theta_p = 0.5$ 로 주었지만, 이와 반대로 이월 효과에 대한 검정은 개체의 합을 이용하므로 정도가 낮아 효과의 차이를 $\theta_p = 1$ 로 하였다.

SAS 9.2에서 RANNOR, RANEXP, RANCAU를 이용하여 각각 정규분포, 지수분포, 코시분포의 난수를 생성시켰으며, RANUNI를 이용하여 역변환법을 사용해서 이중지수분포의 난수를 생성시켜 4가지의 모집단 분포를 고려하였으며 유의수준 α 는 0.05로 하였다. 마지막으로 각 검정법들의 검정력을 비교하기 위하여 10,000번 반복 실험하였다.

위와 같은 방식에 따라 도출된 결과를 표로 정리하였으며 특히 처리와 시기 효과의 검정의 경우 이월 효과의 검정에 따라 실험을 시행한 결과가 10,000번 반복을 만족하지 않으므로 각 검정력에 반복수를 같이 표시하였다.

Table 1에서는 대표본으로 $m = 30, n = 30$ 일 때의 결과를, Table 2에서는 소표본으로 $m = 10, n = 10$ 일 때의 결과를 나타내었다.

1) 대표본의 경우($m = n = 30$)

먼저 각 효과의 차이가 없을 때($\theta_p = \theta_r = \theta_p = 0$)유의수준을 잘 제어하고 있는지 확인해 보면 대체적으로 모든 검정법들이 어느 분포에서나 $\alpha = 0.05$ 를 만족하고 있지만 첫 번째 시기의 자료만을 이용한 처리 효과의 검정(Treatment II)에서는 어느 검정법이나 유의수준이 전혀 제어되지 못하는 결과를 보이고 있다. 또한, 모집단의 분포가 코시분포인 경우엔 모수적 검정법인 Hills-Armitage (1979)의 방법이 유의수준을 잘 제어하지 못하는 것을 보이고 있으며 본 논문에서 제안된 검정법의 경우 대부분의 상황에서 유의수준 $\alpha = 0.05$ 보다는 약간씩 높은 경향을 보이고 있다.

다음으로 이월 효과의 차이가 있는 경우($\theta_p = 1, \theta_r = \theta_p = 0$)를 확인해 보자. 두 시기의 자료를 이용한 처리 효과의 검정(Treatment I)이나 시기 효과의 검정에서는 이월 효과로 인하여 유의수준을 제어하지 못하고 있다는 것을 볼 수 있다. 첫 번째 시기의 자료만을 이용한 처리 효과의 검정(Treatment II)에서는 유의수준을 어느 정도 잘 제어하고 있지만 모집단의 분포와 검정법에 따라서 다른 결과를 보이고 있다. 특히, 코시분포의 경우에는 0.1에 가까운 값들을 나타내며 유의수준이 제어되지 못하고 있다.

이제 각 상황에서의 검정력을 살펴보도록 하면 이월 효과의 차이가 없을 때 처리(Treatment I)와 시기 효과의 검정력 그리고 이월 효과의 검정력은 정규분포에서는 모수적 방법인 Hills-Armitage (1979)의 방법이 어느 효과의 검정에서나 가장 높은 검정력을 보이지만 그 외 이중지수분포, 코시분포, 지수분포에서는 제안된 검정법이 가장 높은 검정력을 보이고 있다. 그리고 이월 효과의 차이가 있을 때 처리 효과(Treatment II)의 검정력은 정규분포에서는 지수점수함수를 이용한 Kim (1999)의 방법, 이중지수분포에서는 제안된 검정법, 코시분포에서는 Koch (1972)의 방법, 지수분포에서는 다시 제안된 검정법이 각각 가장 높은 검정력을 보이고 있다.

2) 소표본의 경우($m = n = 10$)

소표본의 경우에도 제 1종 오류를 제어하는 측면은 대표본의 경우와 비슷한 양상을 보인다. 단, 제

Table 1: The simulation result with large sample size $m = n = 30$ ($\alpha = 0.05$)

Distribution	effect				Test				
	θ_ρ	θ_τ	θ_p		T	W	$P_{(m)}$	$P_{(e)}$	B
Normal	0	0	0	Carry-over	0.0497	0.0482	0.0474	0.0468	0.0556
				Treatment I (replications)	0.0500 (9503)	0.0485 (9518)	0.0499 (9526)	0.0515 (9532)	0.0540 (9444)
				Treatment II (replications)	0.3099 (497)	0.2801 (482)	0.3038 (474)	0.3162 (468)	0.2752 (556)
				Period (replications)	0.0484 (9503)	0.0457 (9518)	0.0459 (9526)	0.0471 (9532)	0.0507 (9444)
	0	0.5	0.5	Carry-over	0.0497	0.0482	0.0474	0.0468	0.0556
				Treatment I (replications)	0.7659 (9503)	0.7425 (9518)	0.7509 (9526)	0.7527 (9532)	0.7495 (9444)
				Treatment II (replications)	0.5292 (497)	0.5124 (482)	0.5253 (474)	0.7308 (468)	0.5234 (556)
				Period (replications)	0.7688 (9503)	0.7445 (9518)	0.7523 (9526)	0.7536 (9532)	0.7549 (9444)
	1	0	0	Carry-over	0.7672	0.7376	0.7446	0.7485	0.7461
				Treatment I (replications)	0.7599 (2328)	0.7256 (2624)	0.7420 (2554)	0.5992 (2515)	0.7341 (2539)
				Treatment II (replications)	0.0343 (7672)	0.0342 (7376)	0.0344 (7446)	0.0493 (7485)	0.0415 (7461)
				Period (replications)	0.7732 (2328)	0.7530 (2624)	0.7608 (2554)	0.7483 (2515)	0.7641 (2539)
1	0.5	0.5	Carry-over	0.7672	0.7376	0.7446	0.7485	0.7461	
			Treatment I (replications)	0.0533 (2328)	0.0488 (2624)	0.0489 (2554)	0.0541 (2515)	0.0536 (2539)	
			Treatment II (replications)	0.5858 (7672)	0.5639 (7376)	0.5673 (7446)	0.5874 (7485)	0.5722 (7461)	
			Period (replications)	1.0000 (2328)	0.9996 (2624)	1.0000 (2554)	1.0000 (2515)	1.0000 (2539)	
Double Exp	0	0	0	Carry-over	0.0508	0.0520	0.0498	0.0511	0.0559
				Treatment I (replications)	0.0499 (9492)	0.0466 (9480)	0.0483 (9502)	0.0494 (9489)	0.0559 (9441)
				Treatment II (replications)	0.2913 (508)	0.2692 (520)	0.2610 (498)	0.2681 (511)	0.2934 (559)
				Period (replications)	0.0461 (9492)	0.0444 (9480)	0.0465 (9502)	0.0486 (9489)	0.0523 (9441)
	0	0.5	0.5	Carry-over	0.0508	0.0520	0.0498	0.0511	0.0559
				Treatment I (replications)	0.4864 (9492)	0.5186 (9480)	0.5009 (9502)	0.5104 (9489)	0.5343 (9441)
				Treatment II (replications)	0.4705 (508)	0.4750 (520)	0.4739 (498)	0.6771 (511)	0.4902 (559)
				Period (replications)	0.4857 (9492)	0.5224 (9480)	0.5000 (9502)	0.5113 (9489)	0.5329 (9441)
	1	0	0	Carry-over	0.4850	0.5244	0.5029	0.5123	0.5345
				Treatment I (replications)	0.4858 (5150)	0.5301 (4756)	0.5083 (4971)	0.3834 (4877)	0.5386 (4655)
				Treatment II (replications)	0.0524 (4850)	0.0465 (5244)	0.0471 (5029)	0.0718 (5123)	0.0533 (5345)
				Period (replications)	0.4759 (5150)	0.5160 (4756)	0.4900 (4971)	0.4931 (4877)	0.5278 (4655)
1	0.5	0.5	Carry-over	0.4850	0.5244	0.5029	0.5123	0.5345	
			Treatment I (replications)	0.0507 (5150)	0.0477 (4756)	0.0491 (4971)	0.0459 (4877)	0.0606 (4655)	
			Treatment II (replications)	0.5012 (4850)	0.5620 (5244)	0.5377 (5029)	0.5528 (5123)	0.5678 (5345)	
			Period (replications)	0.9604 (5150)	0.9750 (4756)	0.9674 (4971)	0.9592 (4877)	0.9744 (4655)	

(Continued)

Cauchy	0	0	0	Carry-over	0.0196	0.0497	0.0474	0.0527	0.0567
				Treatment I (replications)	0.0207 (9804)	0.0498 (9503)	0.0508 (9526)	0.0481 (9473)	0.0557 (9433)
				Treatment II (replications)	0.2041 (196)	0.2877 (497)	0.2869 (474)	0.2865 (527)	0.2892 (567)
				Period (replications)	0.0224 (9804)	0.0530 (9503)	0.0507 (9526)	0.0490 (9473)	0.0592 (9433)
	0	0.5	0.5	Carry-over	0.0196	0.0497	0.0474	0.0527	0.0567
				Treatment I (replications)	0.0333 (9804)	0.1751 (9503)	0.1479 (9526)	0.1694 (9473)	0.1913 (9433)
				Treatment II (replications)	0.1888 (196)	0.3722 (497)	0.3924 (474)	0.5123 (527)	0.3862 (567)
				Period (replications)	0.0340 (9804)	0.1739 (9503)	0.1498 (9526)	0.1674 (9473)	0.1935 (9433)
	1	0	0	Carry-over	0.0319	0.1694	0.1437	0.1651	0.1898
				Treatment I (replications)	0.0365 (9681)	0.1760 (8306)	0.1567 (8563)	0.0962 (8349)	0.1956 (8102)
				Treatment II (replications)	0.1285 (319)	0.0986 (1694)	0.1162 (1437)	0.1454 (1651)	0.1064 (1898)
				Period (replications)	0.0336 (9681)	0.1723 (8306)	0.1476 (8563)	0.1583 (8349)	0.1917 (8102)
1	0.5	0.5	Carry-over	0.0319	0.1694	0.1437	0.1651	0.1898	
			Treatment I (replications)	0.0201 (9681)	0.0484 (8306)	0.0493 (8563)	0.0460 (8349)	0.0539 (8102)	
			Treatment II (replications)	0.2759 (319)	0.4906 (1694)	0.4760 (1437)	0.4815 (1651)	0.4863 (1898)	
			Period (replications)	0.0729 (9681)	0.4952 (8306)	0.4141 (8563)	0.4051 (8349)	0.5490 (8102)	
Exponential	0	0	0	Carry-over	0.0510	0.0474	0.0496	0.0504	0.0561
				Treatment I (replications)	0.0532 (9490)	0.0493 (9526)	0.0507 (9504)	0.0504 (9496)	0.0538 (9439)
				Treatment II (replications)	0.3098 (510)	0.2764 (474)	0.2601 (496)	0.3095 (504)	0.2870 (561)
				Period (replications)	0.0465 (9490)	0.0465 (9526)	0.0444 (9504)	0.0414 (9496)	0.0528 (9439)
	0	0.5	0.5	Carry-over	0.0510	0.0474	0.0496	0.0504	0.0561
				Treatment I (replications)	0.7678 (9490)	0.8664 (9526)	0.8299 (9504)	0.8139 (9496)	0.8773 (9439)
				Treatment II (replications)	0.5059 (510)	0.6076 (474)	0.5806 (496)	0.7500 (504)	0.8556 (561)
				Period (replications)	0.7654 (9490)	0.8707 (9526)	0.8360 (9504)	0.8158 (9496)	0.8825 (9439)
	1	0	0	Carry-over	0.7759	0.8701	0.8464	0.7164	0.9414
				Treatment I (replications)	0.7398 (2241)	0.8345 (1299)	0.7773 (1536)	0.7521 (2836)	0.8635 (586)
				Treatment II (replications)	0.0358 (7759)	0.0380 (8701)	0.0366 (8464)	0.0543 (7164)	0.0500 (9414)
				Period (replications)	0.7488 (2241)	0.8630 (1299)	0.7982 (1536)	0.7236 (2836)	0.8891 (586)
1	0.5	0.5	Carry-over	0.7759	0.8701	0.8464	0.7164	0.9414	
			Treatment I (replications)	0.0585 (2241)	0.0508 (1299)	0.0449 (1536)	0.0384 (2836)	0.0700 (586)	
			Treatment II (replications)	0.6097 (7759)	0.8197 (8701)	0.7873 (8464)	0.6540 (7164)	0.9475 (9414)	
			Period (replications)	0.9996 (2241)	1.0000 (1299)	0.9993 (1536)	0.9972 (2836)	0.9983 (586)	

+ θ_p : Carry-over effect, θ_T : Treatment effect, θ_p : Period effect
 + Treatment I : Test of treatment effect that use the data of two periods when the test of carry-over effect is accepted
 + Treatment II : Test of treatment effect that use the data of first period when the test of carry-over effect is rejected
 + Period: Test of period effect when the test of carry-over effect is accepted
 + T: Hills-Armitage (1979)'s test, W: Koch (1972)'s test, $P_{(n)}$: Kim (1999)'s test that use normal score function,
 $P_{(e)}$: Kim (1999)'s test that use exponential score function, B: Proposed test

Table 2: The simulation result with small sample size $m = n = 10$ ($\alpha = 0.05$)

Distribution	effect				T	W	Test		
	θ_ρ	θ_τ	θ_ρ				$P_{(m)}$	$P_{(c)}$	B
Normal	0	0	0	Carry-over	0.0546	0.0470	0.0488	0.0538	0.0596
				Treatment I (replications)	0.0542 (9454)	0.0460 (9530)	0.0463 (9512)	0.0460 (9462)	0.0589 (9404)
				Treatment II (replications)	0.2729 (546)	0.2489 (470)	0.2295 (488)	0.2454 (538)	0.2450 (596)
				Period (replications)	0.0546 (9454)	0.0484 (9530)	0.0459 (9512)	0.0500 (9462)	0.0606 (9404)
				Carry-over	0.0546	0.0470	0.0488	0.0538	0.0596
				Treatment I (replications)	0.3184 (9454)	0.2804 (9530)	0.2788 (9512)	0.3498 (9462)	0.3017 (9404)
	0	0.5	0.5	Treatment II (replications)	0.4158 (546)	0.3979 (470)	0.4057 (488)	0.6190 (538)	0.3943 (596)
				Period (replications)	0.3195 (9454)	0.2826 (9530)	0.2811 (9512)	0.3494 (9462)	0.3114 (9404)
				Carry-over	0.3296	0.2953	0.2977	0.3786	0.3249
				Treatment I (replications)	0.3121 (6704)	0.2854 (7047)	0.2788 (7023)	0.1600 (6214)	0.3118 (6751)
				Treatment II (replications)	0.0689 (3296)	0.0660 (2953)	0.0648 (2977)	0.0893 (3786)	0.0736 (3249)
				Period (replications)	0.3110 (6704)	0.2790 (7047)	0.2772 (7023)	0.3482 (6214)	0.3035 (6751)
	1	0	0	Carry-over	0.3296	0.2953	0.2977	0.3786	0.3249
				Treatment I (replications)	0.0537 (6704)	0.0447 (7047)	0.0444 (7023)	0.0470 (6214)	0.0579 (6751)
				Treatment II (replications)	0.4075 (3296)	0.3749 (2953)	0.3729 (2977)	0.4247 (3786)	0.3857 (3249)
				Period (replications)	0.8361 (6704)	0.7949 (7047)	0.7937 (7023)	0.8296 (6214)	0.8095 (6751)
				Carry-over	0.0530	0.0457	0.0477	0.0497	0.0558
				Treatment I (replications)	0.0488 (9470)	0.0418 (9543)	0.0423 (9523)	0.0457 (9503)	0.0507 (9442)
	0	0	0	Treatment II (replications)	0.3019 (530)	0.2495 (457)	0.2537 (477)	0.2998 (497)	0.2688 (558)
				Period (replications)	0.0490 (9470)	0.0455 (9543)	0.0466 (9523)	0.0471 (9503)	0.0584 (9442)
				Carry-over	0.0530	0.0457	0.0477	0.0497	0.0558
				Treatment I (replications)	0.1963 (9470)	0.1900 (9543)	0.1844 (9523)	0.2446 (9503)	0.2151 (9442)
				Treatment II (replications)	0.3887 (530)	0.3457 (457)	0.3438 (477)	0.5936 (497)	0.3799 (558)
				Period (replications)	0.1969 (9470)	0.1903 (9543)	0.1866 (9523)	0.2481 (9503)	0.2188 (9442)
Double Exp	1	0	0	Carry-over	0.1968	0.1924	0.1862	0.2487	0.2175
				Treatment I (replications)	0.1905 (8032)	0.1805 (8076)	0.1797 (8138)	0.0941 (7513)	0.2107 (7825)
				Treatment II (replications)	0.1032 (1968)	0.0936 (1924)	0.0929 (1862)	0.1242 (2487)	0.1048 (2175)
				Period (replications)	0.1947 (8032)	0.1881 (8076)	0.1842 (8138)	0.2442 (7513)	0.2175 (7825)
				Carry-over	0.1968	0.1924	0.1862	0.2487	0.2175
				Treatment I (replications)	0.0494 (8032)	0.0420 (8076)	0.0425 (8138)	0.0453 (7513)	0.0515 (7825)
	1	0.5	0.5	Treatment II (replications)	0.4121 (1968)	0.4002 (1924)	0.4023 (1862)	0.4371 (2487)	0.4161 (2175)
				Period (replications)	0.5700 (8032)	0.5661 (8076)	0.5586 (8138)	0.6215 (7513)	0.6070 (7825)

(Continued)

Cauchy	0	0	0	Carry-over	0.0232	0.0472	0.0481	0.0515	0.0585
				Treatment I (replications)	0.0205 (9768)	0.0437 (9528)	0.0458 (9519)	0.0471 (9485)	0.0557 (9415)
				Treatment II (replications)	0.1552 (232)	0.2394 (472)	0.2370 (481)	0.2680 (515)	0.2684 (585)
				Period (replications)	0.0166 (9768)	0.0396 (9528)	0.0404 (9519)	0.0463 (9485)	0.0503 (9415)
	0	0.5	0.5	Carry-over	0.0232	0.0472	0.0481	0.0515	0.0585
				Treatment I (replications)	0.0326 (9768)	0.0803 (9528)	0.0769 (9519)	0.1149 (9485)	0.0988 (9415)
				Treatment II (replications)	0.2069 (232)	0.2987 (472)	0.2973 (481)	0.4699 (515)	0.3094 (585)
				Period (replications)	0.0291 (9768)	0.0785 (9528)	0.0758 (9519)	0.1133 (9485)	0.0988 (9415)
	1	0	0	Carry-over	0.0348	0.0870	0.0822	0.1168	0.1073
				Treatment I (replications)	0.0332 (9652)	0.0830 (9130)	0.0811 (9178)	0.0439 (8832)	0.1035 (8927)
				Treatment II (replications)	0.1121 (348)	0.1540 (870)	0.1521 (822)	0.1772 (1168)	0.1696 (1073)
				Period (replications)	0.0295 (9652)	0.0785 (9130)	0.0764 (9178)	0.1113 (8832)	0.0991 (8927)
	1	0.5	0.5	Carry-over	0.0348	0.0870	0.0822	0.1168	0.1073
				Treatment I (replications)	0.0208 (9652)	0.0448 (9130)	0.0467 (9178)	0.0471 (8832)	0.0570 (8927)
				Treatment II (replications)	0.2443 (348)	0.3563 (870)	0.3516 (822)	0.4255 (1168)	0.3653 (1073)
				Period (replications)	0.0646 (9652)	0.1828 (9130)	0.1701 (9178)	0.2267 (8832)	0.2229 (8927)
Exponential	0	0	0	Carry-over	0.0506	0.0455	0.0443	0.0515	0.0564
				Treatment I (replications)	0.0449 (9494)	0.0405 (9545)	0.0408 (9557)	0.0430 (9485)	0.0510 (9436)
				Treatment II (replications)	0.2925 (506)	0.2725 (455)	0.2731 (443)	0.2990 (515)	0.2801 (564)
				Period (replications)	0.0460 (9494)	0.0445 (9545)	0.0445 (9557)	0.0445 (9485)	0.0547 (9436)
	0	0.5	0.5	Carry-over	0.0506	0.0455	0.0443	0.0515	0.0564
				Treatment I (replications)	0.3557 (9494)	0.3913 (9545)	0.3767 (9557)	0.4425 (9485)	0.4364 (9436)
				Treatment II (replications)	0.4387 (506)	0.4659 (455)	0.4515 (443)	0.7049 (515)	0.5018 (564)
				Period (replications)	0.3594 (9494)	0.3845 (9545)	0.3720 (9557)	0.4374 (9485)	0.4267 (9436)
	1	0	0	Carry-over	0.3577	0.3864	0.3577	0.3669	0.4348
				Treatment I (replications)	0.3361 (6423)	0.3538 (6136)	0.3408 (6423)	0.2262 (6331)	0.4064 (5652)
				Treatment II (replications)	0.0626 (3577)	0.0523 (3864)	0.0576 (3577)	0.0970 (3669)	0.0609 (4348)
				Period (replications)	0.3484 (6423)	0.3760 (6136)	0.3545 (6423)	0.3900 (6331)	0.4091 (5652)
	1	0.5	0.5	Carry-over	0.3577	0.3864	0.3577	0.3669	0.4348
				Treatment I (replications)	0.0425 (6423)	0.0388 (6136)	0.0364 (6423)	0.0360 (6331)	0.0495 (5652)
				Treatment II (replications)	0.4677 (3577)	0.5520 (3864)	0.5295 (3577)	0.5233 (3669)	0.6012 (4348)
				Period (replications)	0.8314 (6423)	0.8504 (6136)	0.8284 (6423)	0.8394 (6331)	0.8756 (5652)

+ θ_p : Carry-over effect, θ_T : Treatment effect, θ_p : Period effect
 + Treatment I : Test of treatment effect that use the data of two periods when the test of carry-over effect is accepted
 + Treatment II : Test of treatment effect that use the data of first period when the test of carry-over effect is rejected
 + Period: Test of period effect when the test of carry-over effect is accepted
 + T : Hills-Armitage (1979)'s test, W : Koch (1972)'s test, $P_{(n)}$: Kim (1999)'s test that use normal score function,
 $P_{(e)}$: Kim (1999)'s test that use exponential score function, B : Proposed test

안된 검정법을 제외한 비모수적 검정법들은 유의수준이 전체적으로 0.05보다 조금씩 낮아 보수적인 성향을 나타내고 있다.

이월 효과의 차이가 있는 경우($\theta_p = 1, \theta_r = \theta_p = 0$) 첫 번째 시기의 자료만을 이용한 처리 효과의 검정(Treatment II)에서도 대표본과 마찬가지로 모집단의 분포와 검정법에 따라 제 1종 오류를 범할 확률이 다르게 나타나고 있지만 대표본의 경우보다 좀 더 높은 경향을 보이고 있다.

소표본에서 각 상황에서의 검정력은 대표본과는 다른 결과를 보인다. 대부분의 상황에서 어느 효과의 검정에서나 지수점수함수를 이용한 Kim (1999)의 방법이 가장 높은 검정력을 보이고 있다. 다만 모집단의 분포가 지수분포인 경우에 이월 효과와 처리 효과의 검정력은 제안된 검정법이 가장 높은 검정력을 가지고 있는 것을 볼 수 있다.

5. 결론 및 고찰

첫 번째 시기의 자료만을 이용한 처리 효과의 검정(Treatment II)의 경우 이월 효과의 검정에서 각각한 자료들을 이용한 검정이기에 유의수준이 제어되지 못한다는 것을 알 수 있다. 즉, 이월 효과의 검정은 개체의 합을 이용한 검정이므로 그 유의한 차이가 'Treatment II'의 검정에서도 나타나는 것으로 볼 수 있다.

모의실험을 통한 비교 결과, 제 1종 오류를 제어하는 것과 검정력의 측면에서 보면 Hills-Armitage (1979)의 방법은 모집단의 가정에 민감하게 된다. 즉 꼬리가 두터운 분포일수록 제 1종 오류를 제어하지 못하고 있으며 검정력도 떨어지는 사실을 확인할 수 있다. 반면에 비모수적 검정법들의 경우 모집단의 분포에 상관없이 유의수준을 잘 제어하고 있으며 모집단의 분포가 정규분포인 경우에도 검정력이 Hills-Armitage (1979)의 방법에 비해 뒤떨어지지 않는다.

지수점수함수를 이용한 Kim (1999)의 방법은 표본수가 작을 경우 다른 검정법들에 비해 높은 검정력을 가지고 있다. 또한 첫 번째 시기의 자료만을 이용한 처리 효과의 검정(Treatment II)에서도 표본수에 관계없이 다른 검정들에 비해 높게 나타나고 있다. 그러므로 지수점수함수를 이용한 Kim (1999)의 방법은 매우 예민한 검정법일 것이라 사료되며 표본이 적거나 기대 효과가 미약할 경우에 적용하면 좋을 것이다. 한 가지 유의할 점으로는 소표본의 경우 모든 검정법의 검정력이 0.3~0.4정도에 불과하므로 본 논문에서 제안한 검정법이나 지수점수함수를 이용한 Kim (1999)의 방법을 이용하더라도 검정에 서 원하는 결과를 얻기 힘든 면이 있다.

본 논문에서 제안된 검정법은 표본이 충분하다면 꼬리가 두꺼운 분포에서 검정력이 좋을 뿐만 아니라 정규분포에서도 모수적 검정법인 Hills-Armitage (1979)의 방법의 검정력과 유사한 결과를 보이고 있다. 또한 모집단의 분포가 지수분포를 따를 때 전반적으로 제안된 검정법의 검정력이 좋다. 따라서 제안된 검정법은 모집단에 대한 정보가 없다면 효율적인 검정법이 될 것이며 특히, 모집단의 분포가 대칭이 아닌 분포를 따를 때 적용하면 바람직할 것이다. 그러나 제안된 검정법은 유의수준이 0.05보다 약간씩 높은 결과를 보이므로 검정통계량의 근사과정에 보완이 필요할 것으로 보인다.

References

- Baumgartner, W., Weiß, P. and Schindler, H. (1998). A nonparametric test for the general two-sample problem, *Biometrics*, **54**, 1129–1135.
- Grizzle, J. E. (1965). The two-period change over design and its use in clinical trials, *Biometrics*, **21**, 467–480.
- Hills, A. V. and Armitage, P. (1979). The two-period cross-over clinical trial, *British Journal of Clinical Pharmacology*, **8**, 7–20.

- Kim, D. (1999). Distribution-free tests for cross-over design data, *The Korean Communications in Statistics*, **6**, 151–157.
- Kim, D., Lee, S. C. and Wang, W. (2011). The asymptotic behavior of linear placements statistics, *Statistics and Probability Letters*, **81**, 326–336.
- Koch, G. G. (1972). The use of none-parametric methods in the statistical analysis of the two-period change over design, *Biometrics*, **28**, 577–584.
- Park, Y. K. and Song, H. H. (1998). *Analysis of Repeated Measures and Cross-Over Design*, Freecademy, Korea.

2012년 9월 19일 접수; 2012년 10월 24일 수정; 2012년 11월 7일 채택