

## 비열등성 시험에서 위치 방법에 기초한 표본 수 결정

김지연<sup>1</sup> · 김동재<sup>2</sup>

<sup>12</sup>가톨릭대학교 의학통계학과

접수 2013년 8월 26일, 수정 2013년 9월 23일, 게재확정 2013년 10월 25일

### 요약

임상시험에서 적당한 크기의 표본 수 결정은 통계적으로 유의한 연구결과의 도출과 연구수행의 효율적인 비용을 산출하기 위해서 중요한 사항 중의 하나이다. 기존의 비열등성 시험에서 표본 수 계산방법에는 t 검정법을 이용한 모수적 방법이 있고, Wilcoxon 순위합 검정을 이용하여 Wang 등 (2003)이 제안한 표본 수 계산방법을 Kim과 Kim (2007)이 비열등성 시험에 확장시켜 적용하여 제시한 비모수적 방법이 있다. 본 논문에서는 Orban과 Wolfe (1982)가 제안한 선형위치통계량의 검정법에 Kim (1994)이 계산한 검정력의 결과를 이용한 표본 수 계산 방법을 제안하고, 그에 따른 표본 수 계산결과를 기존에 제시된 표본 수 계산 결과와 비교하였다. 그 결과 기존의 방법들보다 본 논문에서 제안하는 방법으로 계산한 경우의 표본 수가 가장 작게 나왔다. 따라서 모집단에 대해서 구체적인 분포함수를 가정하기 힘든 경우 모수적 방법을 이용하게 되면 검정력이 떨어지거나 유의수준을 제어하지 못하는 문제점을 보완하고, 모수적 방법에 비해 표본 수가 크게 나와 시간이나 비용 면에서 효율적이지 않았던 Wilcoxon 순위합 검정을 이용한 방법의 단점을 보완할 수 있는 방법으로 본 논문에서 제시한 위치 (placement)를 이용한 표본 수 계산이 이용될 수 있다.

주요용어: 비모수적 방법, 비열등성 시험, 표본수.

### 1. 서론

신약을 개발 하였을 때 치료효과를 알아보기 위해 기존의 약과 비교하는 임상시험을 시행하여 신약의 유효성을 증명한다. 이러한 임상시험으로는 제 3상 시험 (phase III clinical trials)을 들 수 있고, 이는 제 2상 시험을 통해 효과가 있다고 판단된 신약 (new treatment)과 기존의 약 (standard treatment)의 효과 (effect)를 비교하여 유효성 (efficiency)과 안전성 (safety)을 입증하는 단계이며 제 3상에서의 임상시험을 비교임상시험이라고도 한다. 이러한 비교시험을 위한 방법으로는 우월성 시험 (superiority trials), 동등성 시험 (equivalence trials) 또는 치료적 동등성 시험 (therapeutic equivalence trials), 비열등성 시험 (non-inferiority trials) 등이 있으며 이 중에서 비열등성 시험은 새로운 치료법이 기존의 치료법에 비해 열등하지 않음을 보이는 것이 목적인 시험이다 (Lee 등, 2005).

임상시험연구에 있어서 가장 중요한 사항 중의 하나가 연구에 필요한 환자 수 (표본수)를 결정하는 문제이다. 연구대상 환자수는 연구의 설계단계에서 미리 고려되어야 하는 것임에도 불구하고 많은 연구에서 이러한 원칙이 무시되고 일정기간동안에 확보할 수 있는 정도를 대상환자수로 결정해 버리는 경우가 있는데, 이렇게 되면 임상적으로 의미있는 차이가 있는데도 검정력이 떨어져서 이를 찾아내지 못하는 경우가 많게 된다. 검정력이 작은 연구의 위험성은 유익하게 사용될 수 있는 치료법 등이 충분히 검정되지

<sup>1</sup> (137-701) 서울시 서초구 반포동 505번지, 가톨릭대학교 의학통계학과, 대학원생.

<sup>2</sup> 교신저자: (137-701) 서울시 서초구 반포동 505번지, 가톨릭대학교 의학통계학과, 교수.

E-mail: djkim@catholic.ac.kr

못하고 기각되어 두 번 다시 고려될 수 없다는 데 있다 (Lee 등, 2005). 또한, 임상시험에서 아무리 정확한 방법을 이용하여 연구를 수행한다 해도 표본의 크기를 너무 작게 잡으면 그 연구의 가설을 증명하기 힘들게 되며, 반면에 표본의 크기가 필요 이상으로 크면 연구의 수행이 어렵고 과다한 경비가 들게 되므로, 적당한 크기의 표본수를 산정하는 것은 통계적으로 유의한 연구결과의 도출을 위해서, 그리고 동시에 연구수행의 효율적인 비용을 산출하기 위해서 중요하다 (Shin 등, 1997).

기존에는 비열등성 시험에서의 표본 수를 계산하기 위해 이전의 임상시험 자료들의 값을 이용하여 정규분포 한다는 가정 하에 모수적 방법으로 구하였다. 하지만 실제로는 모집단의 분포가 정규분포에 가까운 경우가 많은 것이지 데이터가 정규모집단에서 얻어지는 경우는 없다. 또한 표본 통계량의 분포가 정규분포에 수렴하는 속도가 아주 느리거나, 데이터를 가운데 이상점 (outlier)이 있는 경우에는 신뢰수준이나 유의수준을 제어하기 힘들고 검정력도 급격히 떨어지게 된다. 이와 같이 모집단에 대하여 구체적인 분포함수를 가정하는 것이 힘들 경우에는 모집단 분포에 대한 가정을 약화시킬 수 있는 비모수적 방법을 이용하는 것이 대안이 된다.

Kim과 Kim (2007)은 Wang 등 (2003)이 제안한 비모수적 방법인 Wilcoxon 순위합 검정을 이용한 표본 수 공식을 비열등성 시험에 확장시켜서 고정된 유의수준과 검정력 하에서 표본 수를 구하고, 모수적 방법의  $t$ 검정을 이용하여 계산한 표본 수와 비교하였다. 그 결과  $t$ 검정을 이용하였을 때보다 Wilcoxon 순위합 검정을 이용하였을 때가 표본수가 더 커지므로, 이는 연구 수행의 시간이나 비용이 많이 든다는 단점이 된다.

본 논문에서는 모수적 방법에서의 정규 분포 가정을 약화시켜 검정력이 떨어지거나 유의수준을 제어하지 못하는 문제점을 보완하고, 모수적 방법에 비해 표본 수가 더 크게 나와 시간이나 비용을 많이 들게 하는 Wilcoxon 순위합 검정을 이용한 방법을 보완하기 위하여 Orban과 Wolfe (1982)가 제안한 비모수적 방법인 선형위치통계량을 이용한 검정법과 Kim (1994)이 계산한 선형위치통계량을 이용한 검정법의 검정력을 이용하여 비열등성 시험을 위한 표본 수 계산방법을 제시한다. 또한 유의수준과 검정력을 고정하여 본 논문에서 제안하는 방법과 기존의 방법들을 이용해 구한 표본 수 계산 결과를 비교하여 결론을 제시한다.

## 2. 방법

확률표본  $X_i, i = 1, \dots, m$ 과  $Y_j, j = 1, \dots, n$ 을 각각 대조군의 누적분포함수  $F(x)$ 와 치료군의 누적분포함수  $G(y) = F(y - \theta)$ 에서 추출된 확률변수라고 하자. 이 때  $\theta$ 는 대조군과 치료군의 치료효과 차이를 나타내는 위치모수이며, 비열등성 시험에서의 가설은 다음과 같다.

$$H_0 : \mu_S - \mu_T \geq \delta \text{ vs } H_1 : \mu_S - \mu_T < \delta$$

여기서  $\delta$ 는 실험 전에 임상적인 정보를 이용하여 결정되어지는 양의 실수로 비열등성의 한계라고 하고,  $\mu_S$ 와  $\mu_T$ 는 각각 대조군과 치료군의 평균 반응 값이다.

### 2.1. Wilcoxon 순위합 검정에서의 표본수

Wilcoxon 순위합 검정의 검정통계량  $W$ 는 대조군과 치료군의 통합순위에서 치료군의 순위합으로 다음과 같이 정의된다.

$$\begin{aligned}
 W &= \sum_{j=1}^n R_j \\
 &= \sum_{j=1}^n \left( \sum_{i=1}^n I\{Y_j \geq Y_i\} + \sum_{i=1}^m I\{Y_j \geq X_i\} \right) \\
 &= \frac{n(n+1)}{2} + \sum_{i=1}^m \sum_{j=1}^n I\{Y_j \geq X_i\}
 \end{aligned}$$

여기서  $R_j$ 는 치료군의 통합순위로  $X_1, \dots, X_m$ 과  $Y_1, \dots, Y_n$  중  $Y_j$ 의 순위를 의미한다. 그리고 귀무가설 하에서 Wilcoxon 순위합 검정의 검정통계량

$$W^* = \frac{W - E(W)}{\sqrt{\text{var}(W)}} = \frac{W - \frac{1}{2}n(n+m+1)}{\frac{1}{12}mn(m+n+1)}$$

은 근사적으로 표준 정규분포하게 되고 만약  $|W^*| \geq z_\alpha$ 를 만족하면 유의수준  $\alpha$ 하에서 귀무가설을 기각한다. 여기서  $z_\alpha$ 는  $W^*$ 의 귀무가설 하에서 분포에 대한  $(100 \times \alpha)\%$  백분위수이다.

일반적으로  $W$ 의 평균과 분산은

$$\mu_W = \frac{n(n+1)}{2} + mnp_1$$

$$\sigma_W^2 = mnp_1(1-p_1) + mn(m-1)(p_2-p_1^2) + mn(n-1)(p_3-p_1^2)$$

이고, 여기서  $p_1, p_2, p_3$ 는 다음과 같이 정의된다.

$$\begin{aligned}
 p_1 &= P(Y_1 \geq X_1) \\
 p_2 &= P(Y_1 \geq X_1 \text{ and } Y_1 \geq X_2) \\
 p_3 &= P(Y_1 \geq X_1 \text{ and } Y_2 \geq X_1)
 \end{aligned}$$

따라서 Kim과 Kim (2007)에 의해 계산된 비열등성 시험에서 Wilcoxon 순위합검정을 이용한 표본 수 계산은 아래와 같다.

대립가설이 참일 경우의 Wilcoxon 순위합검정의 검정력 (*power*)은

$$\begin{aligned}
 \text{power} &= P(|W^*| > z_\alpha) \\
 &\approx P(W^* > z_\alpha) \\
 &= P\left( \frac{W - n(n+1)/2 - mnp_1}{\sigma_W} > \frac{z_\alpha \sqrt{mn(m+n+1)/12} + mn(1/2 - p_1 - \delta)}{\sigma_W} \right)
 \end{aligned}$$

이고, 두 군의 표본 수가 동일할 때 검정력이 가장 크므로  $n = m$ 으로 놓으면 검정력은 다음과 같다.

$$\text{power} = 1 - \Phi\left( \frac{z_\alpha \sqrt{1/6} - \sqrt{n}(1/2 - p_1 - \delta)}{\sqrt{p_2 + p_3 - 2p_1^2}} \right)$$

여기서  $\Phi(\cdot)$ 는 표준정규분포의 분포함수이다.

한 그룹 당 표본 수는

$$n = \left( \frac{z_\alpha \sqrt{1/6} + z_\beta \sqrt{p_2 + p_3 - 2p_1^2}}{1/2 - p_1 - \delta} \right)^2 \quad (2.1)$$

이다.

Wilcoxon 순위합검정의 검정력에서  $p_1, p_2, p_3$ 는 모집단의 분포에 따라 다음과 같이 계산된다.

대조군과 치료군의 분포가 각각  $N(\mu_X, \sigma_X^2), N(\mu_Y, \sigma_Y^2)$ 을 따르는 경우  $p_1, p_2, p_3$ 는 다음과 같이 계산된다 (simonoff 등, 1986).

$$p_1 = \Phi \left( \frac{\mu_Y - \mu_X}{\sqrt{(\sigma_X^2 + \sigma_Y^2)}} \right)$$

$$p_2 = \frac{p_1^2}{p_1^2 - p_1 + 1}$$

$$p_3 = \frac{p_1^2}{p_1^2 - p_1 + 1}$$

또한 대조군과 치료군의 분포가 평균  $\mu_X, \mu_Y$ 이고 분산  $\sigma_X^2, \sigma_Y^2$ 인 이종지수 분포를 따르는 경우  $p_1, p_2, p_3$ 는

$$p_1 = 1 + \left( \frac{\sigma_X}{4(\sigma_X - \sigma_Y)} - \frac{1}{2} \right) e^{\frac{\mu_X - \mu_Y}{\sigma_Y}} + \left( \frac{\mu_X - \mu_Y}{4\sigma_Y} \right) e^{\frac{\mu_X \sigma_Y - \mu_Y \sigma_X}{\sigma_X \sigma_Y}} - \frac{\sigma_X}{4(\sigma_X + \sigma_Y)} e^{\frac{\mu_X - \mu_Y}{\sigma_X}}$$

$$p_2 = \frac{p_1^2}{p_1^2 - p_1 + 1}$$

$$p_3 = \frac{p_1^2}{p_1^2 - p_1 + 1}$$

이다.

## 2.2. 선형위치통계량을 이용한 검정법에서의 표본수

선형위치통계량을 이용한 검정법에서 위치 (placement) 확률변수  $U_1, U_2, \dots, U_n$ 은 다음과 같이 정의된다 (Orban과 Wolfe, 1982).

$$mU_i = [\text{number of } X' \text{'s} \leq Y_i], \quad i = 1, 2, \dots, n$$

여기서  $m$ 은 대조군의 표본수이다. 그리고 선형위치통계량은

$$S_{n,m}^P = \sum_{i=1}^n \varphi_m(U_i)$$

이고, 여기서  $\varphi_m(x)$ 는  $[0,1]$ 에서 정의된 실수 값을 가지는 점수함수이다. 다음을 만족할 때 유의수준  $\alpha$  하에서 귀무가설은 기각된다.

$$P(S_{n,m}^P \geq s_\alpha) = \alpha$$

여기서  $s_\alpha$ 는  $S_{n,m}^P$ 의 귀무가설하에서의 분포에 대한  $(100 \times \alpha)\%$  백분위수이다. 본 논문에서는 균일 점수함수와 지수 점수함수를 이용한 선형위치통계량을 채택하여 계산하고, 이 때 점수함수는 다음과 같다.

$$\begin{aligned} \varphi(x) &= x && \text{균일 점수함수} \\ \varphi(x) &= -\ln(1-x) && \text{지수 점수함수} \end{aligned}$$

그리고 선형위치통계량의 평균과 분산은

$$\begin{aligned} \bar{\varphi} &= E(\varphi[F(Y_i)]) = \int \varphi[F(y)]dG(y) \\ V_\varphi &= Var(\varphi[F(Y_i)]) = \int [\varphi[F(y)] - \bar{\varphi}]^2 G(y) \end{aligned}$$

이고, Kim (1994)이 계산한 선형위치통계량을 이용한 검정의 검정력은

$$\begin{aligned} \beta(\Theta) &= P(S_{n,m}^P > s_\alpha : H_1 \text{ is true}) \\ &\approx P\left(Z \geq \frac{n\bar{\varphi}_0 + z_\alpha \sqrt{nV_{\varphi_0}} - n\bar{\varphi}_\theta}{\sqrt{nV_{\varphi_0}}}\right) \\ &= 1 - \Phi\left(z_\alpha + \frac{\sqrt{n}(\bar{\varphi}_0 - \bar{\varphi}_\theta)}{\sqrt{V_{\varphi_0}}}\right) \end{aligned}$$

이다. 여기서  $\bar{\varphi}_0$ 와  $V_{\varphi_0}$ 는 각각 귀무가설하에서의 평균과 분산이고  $\bar{\varphi}_\theta$ 는 대립가설하에서의 평균이며,  $\Phi(\cdot)$ 는 표준정규분포의 분포함수이다.

이 식을 정리하면

$$\begin{aligned} \frac{\sqrt{n}(\bar{\varphi}_0 - \bar{\varphi}_\theta)}{\sqrt{V_{\varphi_0}}} &= \sqrt{n} \left( \frac{\int \varphi(F(y))dF(y - \delta) - \int \varphi(F(y))dF(y - \theta)}{\sqrt{\int (\varphi(F(y)) - \bar{\varphi}_0)^2 dF(y)}} \right) \\ &= \sqrt{n} \left( \frac{\int \varphi(F(y + \delta))dF(y) - \int \varphi(F(y + \theta))dF(y)}{\sqrt{\int (\varphi(F(y)) - \bar{\varphi}_0)^2 dF(y)}} \right) \\ &\approx \frac{(\delta - \theta)\sqrt{n} \int (\varphi F)'(y)dF(y)}{\sqrt{\int (\varphi(F(y)) - \bar{\varphi}_0)^2 dF(y)}} \end{aligned}$$

이고, 이를 이용하여 검정력을 구하면

$$\beta(\Theta) \approx 1 - \Phi\left(z_\alpha + \frac{(\delta - \theta)\sqrt{n} \int (\varphi F)'(y)dF(y)}{\sqrt{\int (\varphi(F(y)) - \bar{\varphi}_0)^2 dF(y)}}\right)$$

이고, 각 군의 표본수는

$$n = \left[ \frac{(z_\alpha + z_\beta) \sqrt{\int (\varphi(F(y)) - \bar{\varphi}_0)^2 dF(y)}}{(\theta - \delta) \int (\varphi F)'(y) dF(y)} \right]^2 \quad (2.2)$$

이다. 여기서  $(\varphi F)'(t) = d\varphi[F(y)]/dt$ 이다.

### 2.3. 모수적 방법의 표본수

대조군과 치료군의 표본 수가 동일할 때 검정력이 가장 크므로  $n = m$ 으로 놓으면 검정력은

$$power = \Phi \left( \frac{\delta - \theta}{\frac{\theta_S^2 + \theta_T^2}{n}} - z_\alpha \right)$$

이고, 각 군의 표본 수는

$$n = \frac{2\sigma^2(z_\alpha + z_\beta)^2}{(\theta - \delta)^2} \quad (2.3)$$

이다.

## 3. 표본 수 비교 및 결과

본 논문에서는 비열등성 시험을 위한 기존의 방법인  $t$  검정법을 이용한 표본 수 계산방법과 Wilcoxon 순위합검정을 이용한 표본수 계산방법을 새로 제안한 선형위치통계량을 이용한 검정법에서의 표본 수 계산방법과 비교하였다.

먼저, 유의수준과 검정력은  $\alpha = 0.05$ ,  $1 - \beta = 0.8$ 인 경우로 고정하고, 대조군과 치료군의 치료효과 차이  $\mu_S - \mu_T$ 에서 비열등성의 한계  $\delta$ 와의 차이인  $\Delta$ 는 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.8, 1로 변화시켜 계산하였다. Wilcoxon 순위합검정에서의 표본 수는 식 (2.1), 제안한 선형위치통계량을 이용한 검정법에서의 표본 수는 식 (2.2), 모수적 방법을 이용한 표본 수는 식 (2.3)을 이용하여 계산하였고, 모집단의 분포로는 이용가능성이 많은 정규분포와 이중지수분포를 채택하였다. 검정력 0.9를 가정한 위치방법에 따른 표본수의 결과는 검정력 0.8인 본 논문의 경우와 유사하였으며, 또한 지수분포 등 기타분포 경우들의 검정력 계산 결과는 좋은 장점을 발견할 수가 없어서 추가비교를 제시하지 않았다. 이에 따른 표본 수 계산 결과를 Table 3.1로 정리하였다.

각 표에서  $\Delta$ 가 작을수록, 또한  $\sigma^2$ 이 커질수록 표본 수가 커지는 경향을 보인다. 또한 세가지 방법을 비교하면 Wilcoxon 순위합 검정을 이용한 경우 표본수가 가장 크고 그 다음으로 모수적 방법을 이용한 경우이고 위치 (placement) 방법의 경우가 가장 작게 나왔다.

모집단의 분포를 고려하여 표본수를 비교해보면 Wilcoxon 순위합검정 방법에서 이중지수분포를 가정한 경우가 표본수가 가장 크고, 그 다음으로는 Wilcoxon 순위합검정 방법에서 정규분포를 가정한 경우이고, 그 다음은 모수적 방법으로 계산하였을 경우이고, 네번째로는 위치 방법에서 정규분포를 가정하고 지수점수함수로 계산하였을 경우이다. 그 다음으로는 위치 방법에서 균일 점수함수로 계산한 경우이고, 위치 방법에서 이중지수분포를 가정하고 지수점수함수를 적용하였을 경우는 이와 근소한 차이를 보이며 작게 나왔다. 마지막으로 위치 방법에서 이중지수분포를 가정하고 균일점수함수를 적용하였을 경우의 표본수가 가장 작게 나왔다.

**Table 3.1** Sample size calculation:  $\alpha = 0.05$ , power  $(1 - \beta) = 0.8$

$\Delta$	$\sigma_X^2 = \sigma_Y^2$	placement				Wilcoxon		t test
		normal		double exp.		normal	double exp.	
		uniform	exp.	uniform	exp.			
0.1	1	648	758	413	644	1345	1709	1237
	2	1295	1516	825	1287	2661	3380	2474
	3	1943	2274	1237	1931	3971	5047	3710
0.15	1	288	337	184	286	610	774	550
	2	576	674	367	572	1199	1522	1100
	3	864	1011	550	858	1785	2267	1649
0.2	1	162	190	104	161	350	444	310
	2	324	379	207	322	684	868	619
	3	486	569	310	483	1015	1289	928
0.25	1	104	122	66	103	228	290	198
	2	208	243	132	206	443	563	396
	3	311	364	198	309	657	834	594
0.3	1	72	85	46	72	161	206	138
	2	144	169	92	143	312	397	275
	3	216	253	138	215	461	586	413
0.35	1	53	62	34	53	121	154	101
	2	106	124	68	106	232	296	202
	3	159	186	101	158	343	435	303
0.4	1	41	48	26	41	94	121	78
	2	81	95	52	81	180	230	155
	3	122	143	78	121	265	337	232
0.5	1	26	31	17	26	62	81	50
	2	52	61	33	52	119	151	99
	3	78	91	50	78	174	221	149
0.6	1	18	22	12	18	45	58	35
	2	36	43	23	36	84	108	69
	3	54	64	35	54	123	157	104
0.8	1	11	12	7	11	27	36	20
	2	21	24	13	21	50	65	39
	3	31	36	20	31	72	93	58
1	1	7	8	5	7	18	25	13
	2	13	16	9	13	34	44	25
	3	20	23	13	20	48	62	38

\*\*exp.=exponential, double exp.=double exponential

#### 4. 결론 및 토의

본 논문에서는 비열등성 시험에서 표본 수를 계산하는 방법으로 모수적 방법과 비모수적 방법을 사용하여 각각의 결과들을 비교하였다. 유의수준과 검정력을 고정하고, 기존의 방법들과 본 논문에서 제안하는 방법을 비교한 결과 모수적 방법과 Wilcoxon 순위합 검정 방법보다도 본 논문에서 제안한 위치 (placement) 방법에서 표본 수가 가장 작게 나왔고, 그 중에서도 이중지수분포를 가정하고 균일점수함수를 적용해 계산한 경우가 표본 수가 가장 작음을 보았다. 본 논문에서 제안한 방법에 의한 표본크기가 모수적 방법에 비해 거의 절반의 크기로 나왔는데 이는 제시된 표본 수 계산 공식에서의 분산추정량이 모수적 방법에서의 계산에 이용된 분산에 비해 상당히 작게 나오기 때문에 그 결과 표본 수가 절반 가까이 작아지게 된 것이라고 할 수 있다. 그리고 본 논문에서 제안하는 방법에 사용된 균일점수함수는 일정한 함수로 가중치가 필요 없을 경우에 쓰며, 지수점수함수는  $x$ 가 점점 커짐에 따라 함수값이 작아지는 형태로 가중치를 적게 주고 싶은 경우에 사용하는 것이 좋다. 일반적으로는 균일점수함수를 주로 사용

한다. 또한 본 논문에서는 정규분포와 이중지수분포를 다루었는데, 이중지수분포가 정규분포일 때 보다 본 논문에서 제안한 방법이 기존의 방법보다 더 좋은 효과를 보였다. 따라서 모집단의 분포가 꼬리가 두텁고 넓게 퍼져있는 특성을 보이는 경우에는 본 논문에서 제안한 방법을 쓰는 게 더 좋은 결과를 보인다.

결과적으로 본 논문에서 제시된 선형위치통계량을 이용한 방법은 정규분포 뿐 아니라 이중지수분포까지 사용하여 정규분포를 벗어나는 범위까지 다룰 수 있으므로 모수적 방법에 비해 분포의 제약에서 좀더 자유로울 수 있게 되는 장점이 있고, 기존의 비모수적 방법인 Wilcoxon 순위합 검정을 이용한 방법이 시간과 비용 면에서 모수적 방법에 비해 효율적이지 않았던 단점을 해결할 수 있었다. 그러나 본 논문에서 제시한 표본 수 계산에서의 분산추정량이 기존의 방법에 비해 작게 추정된 것은 선형위치통계량의 근사분포에 영향을 받았다고 볼 수 있다. 이에 보다 정밀하게 추정할 수 있는 다른 선형위치통계량의 점근분포를 추후 연구해 볼 필요성이 있다고 생각한다. 이를 바탕으로 모집단에 대하여 구체적인 분포 함수를 가정하기 힘든 경우에 적절한 크기의 표본 수를 구하기 위하여 본 논문에서 제시한 방법을 이용하는 것은 통계적으로 유의한 연구결과의 도출을 위해서, 그리고 연구수행의 효율적인 비용을 산출하기 위해서 좋은 방법이 될 수 있다.

## References

- Kim, D. (1994). A comparison of distribution-free two-sample procedure based on placements or ranks. *Journal of the Korean Statistical Society*, **23**, 135-149.
- Kim, D. W. and Kim, D. (2007). Sample size comparison for non-inferiority trials. *Journal of Korean Data & Information Science Society*, **18**, 411-418.
- Lee, J. W., Park, M. R. and Yu, H. N. (2005). *Statistical method for life science research*, Free academy, Seoul.
- Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placements. *Journals of the American Statistical Association*, **77**, 666-671.
- Shin, Y. S. and Ahn, Y. O. (1997). *Medical research methodology*, Seoul National University Press, Seoul.
- Simonoff, J. S., Hochberg, Y. and Reiser, B. (1986). Alternative estimation procedures for  $\Pr(X < Y)$  in categorized data. *Biometrics*, **42**, 895-907.
- Wang, H., Chen, B. and Chew, S. C. (2003). Sample size determination based on rank tests in clinical trials. *Journal of Biopharmaceutical Statistics*, **13**, 735-751.



## Sample size determination based on placements for non-inferiority trials

Jiyeon Kim<sup>1</sup> · Dongjae Kim<sup>2</sup>

<sup>12</sup>Department of Statistics, The Catholic University of Korea

Received 26 August 2013, revised 23 September 2013, accepted 25 October 2013

### Abstract

In clinical research, sample size determination is one of the most important things. There are parametric method using t-test and non-parametric method suggested by Kim and Kim (2007) based on Wilcoxon's rank sum test for determining sample size in non-inferiority trials. In this paper, we propose sample size calculation method based on placements method suggested by Orban and Wolfe (1982) and using the power calculated by Kim (1994) in non-inferiority trials. We also compare proposed sample size with that using Kim and Kim (2007)'s formula and that of t-test for parametric methods. As the result, sample size calculated by proposed method based on placements is the smallest. Therefore, proposed method based on placements is better than parametric methods in case that it's hard to assume specific distribution function for population and also more efficient in terms of time and cost than method based on Wilcoxon's rank sum test.

*Keywords:* Non-inferiority trial, placement, sample size.

---

<sup>1</sup> Researcher, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea.

<sup>2</sup> Corresponding author: Professor, Department of Biostatistics, The Catholic University of Korea, Seoul 137-701, Korea. E-mail: djkim@catholic.ac.kr