The Korean Communications
in Statistics Vol. 6, No. 1, 1999
pp. 151-157

Distribution-Free Tests for Cross-Over Design Data

Dongjae Kim[1)]

Abstract

Distribution-free tests are proposed for AB/BA 2*2 cross-over design data based on placements introduced by Orban and Wolfe(1982). In this paper, we suggest the homogeneity test for carry-over effects and also suggest the test for direct effects and the test for period effects under the same carry-over effects. The properties such as iterative asymptotic distribution for the proposed tests are discussed.

1. 서 론

의과학 분야에서 임상시험(clinical trials)을 통한 새로운 약제 또는 치료법의 개발은 매우 중요하며 일반적인 일이다. 특히 제 3상 임상시험(phase III clinical trials)에서는 기존의 약제 또는 치료법과 비교연구를 하여야만 한다. 이와 같이 두 가지 처리를 비교하는 방법은 독립인 두 표본의 여러 가지 검정법들(two-sample tests)이 널리 이용된다. 그러나 이 방법은 각 개체의 치료전 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.

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 2*2 교차계획 자료의 분석은 두 가지 처리의 직접 효과(direct effect)의 차이 여부 검정에 주 목적이 있으며, 이월 효과(carry-over effect)와 시기 효과(period effect)의 차이 유무도 검정할 수 있다. 분석방법으로는 Hills-Armitage(1979)방법이 가장 많이 이용되고 있다. 이는 모수적 방법으로 원 자료의 합과 차를 이용하여 독립된 두 표본의 t-검정(unpaired t-test)을 적용하는 방법이다. 또한 비모수적 방법으로는 Wilcoxon rank-sum 검정을 적용한 Koch(1972)의 방법이 흔히 쓰인다.

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 안정되게 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 만성질환의 경우처럼 완전치유라는 멀고 먼 결과보다는

---

1) Assistant Professor, Department of Biostatistics, The Catholic University of Korea, Seoul, 137-701, Korea

비교적 단기간의 증상의 변화 또는 치료의 반응을 알고자 할 때 자주 이용한다.

본 연구에서는 Grizzle(1965)의 2*2 교차계획의 모형을 2절에서 제시하고, Orban과 Wolfe(1982)가 제안한 선형 placement 통계량을 적용하여 2*2 교차계획 자료의 이월 효과의 동일성, 두 처리의 직접 효과와 시기 효과의 차이 유무에 대한 검정법을 3절에서 제안하였다. 4절에서는 제안된 검정통계량의 반복점근분포(iterative asymptotic distribution)등의 성질과 활용 방안에 대하여 논의하였다.

2. 수리적 모형

Grizzle(1965)의 논문으로부터 2*2 교차계획법의 적절한 모형은

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

이고, 여기서 μ는 총 평균, b_{ij}는 i번째 순서(AB 또는 BA)에서의 j번째 환자의 랜덤 효과, p_k는 k번째 시기 효과, τ_l은 l번째 처리의 직접 효과, $\rho_{l'}$은 l'번째 처리의 이월 효과(carry-over effect) 그리고 e_{ijk}는 측정값의 랜덤 오차이다. 이 경우에서 b_{ij}와 e_{ijk}는 서로 독립이고 각각 평균이 0인 정규분포를 따른다고 가정한다. Grizzle은 이 모형에서 두 처리의 직접 효과, 이월 효과 그리고 시기 효과에 대한 검정법을 제안하였다. 이 논문에서는 같은 질문에 대한 검정법에 대하여 논의하며, 단지 정규성 가정만을 포함하지 않는다.

3. 제안된 검정법

이 절에서는 선형 placement 통계량을 이용한 비모수적 검정법에 대하여 논의한다. 이는 Orban과 Wolfe(1982)가 처음 제안한 통계량으로 서로 다른 요인 조합에서 다른 양상의 변이(variability)를 보이거나 이상점이 있는 경우에 적합한 비모수적 검정법이다.

3.1. 선형 placement 통계량

$X_{11}, X_{12}, \cdots, X_{1m}$ 과 $X_{21}, X_{22}, \cdots, X_{2n}$ 을 각각 대조군과 실험군에서 추출한 독립인 확률표본이라 하고 각각의 분포함수는 위치모수에 의하여 특징지어지는 같은 형태의 알려지지 않는 연속함수라고 가정한다. 즉

$$\text{대조군 : } X_{11},\ X_{12},\ \cdots,\ X_{1m} \sim F(x - \eta_1)$$
$$\text{실험군 : } X_{21},\ X_{22},\ \cdots,\ X_{2n} \sim F(x - \eta_2)$$

이 때 확률변수 $U_1,\ U_2,\ \cdots,\ U_n$을 다음과 같이 정의한다.

$$mU_j = [X_{2j} \text{ 보다 적거나 같은 } X_{1i} \text{ 들의 개수}] \tag{2}$$

이 때 U_j 를 X_{2j} 의 X_1들 중의 placement라 부른다. 이 placement를 이용한 두 모집단에서 위치모수 검정을 위한 분포무관 선형 placement 통계량은 다음과 같다.

$$S_{n,m} = \sum_{j=1}^{n} \varphi_m(U_j). \tag{3}$$

여기서 점수함수(score function) $\varphi_m(x)$ 는 [0, 1]에서 정의된 실변수 함수이다.

선형 placement 통계량의 몇 가지 예제를 고려해 보자. 첫째로 두 모집단에서 위치모수 검정을 위한 Mann-Whitney-Wilcoxon 통계량과 항등함수 $\varphi_m(x) = x$ 를 이용한 선형 placement 통계량이 유일하게 같다(Orban 과 Wolfe(1982)). 둘째로 정규점수함수(Normal score function)를 이용하면 새로운 통계량

$$S^{NS}_{n,m} = \sum_{j=1}^{n} \Phi^{-1}[(mU_j + 1)/(m+2)]$$

이 만들어진다. 여기서 $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp(-u^2/2)\, du$ 이다.

선형 placement 통계량은 두 모집단에서 실험군과 대조군의 할당에 대하여 symmetric 하지 않다. 그러나 한 집단의 표본크기가 다른 집단의 표본크기에 비해 월등히 클 경우에도 우리가 동의할 수 있는 할당방법이 있다. placement의 정의 (2)의 본질은 대조군의 경험분포(empirical distribution)에 대한 실험군의 자료에 대한 위치를 나타내므로 표본크기가 큰 모집단을 대조군으로 할당하는 것이 자연스러운 것이다.

선형 placement 통계량을 이용한 가설 $H_0 : \eta_1 = \eta_2$ 를 검정하기 위하여 적절한 기각역을 정해주어야 한다. 물론 대립가설 H_1 에 따라 기각역의 형태는 달라져야 하겠지만 점수함수 φ_m 의 선택의 자유로움으로 인하여 일반성을 잃지 않고 통계량 $S_{n,m}$ 에 근거한 유의수준 α인 기각역은 다음과 같은 형태라고 가정할 수 있다.

$$\text{reject } H_0 : \eta_1 = \eta_2 \quad iff \quad S_{n,m} \geq S_\alpha$$

여기서 S_α 는 귀무가설 하에서의 $S_{n,m}$ 의 분포의 제 100(1 - α) 백분위수이다.

3.2. 이월 효과의 동일성 검정

실제로는 이월 효과가 없음을 검정하기 원하지만 시기 효과가 있는 경우에는 이월 효과가 없음을 검정할 수 없다. 따라서 두 약물의 이월 효과가 같다는 사실을 검정한다. 식 (1)으로부터 같은 환자의 두 관측값은

$$y_{ij1} + y_{ij2} = 2(\mu + b_{ij}) + (p_1 + p_2) + (\tau_1 + \tau_2) + 2\rho_i + (e_{ij1} + e_{ij2}) \tag{4}$$

이고 ρ_i는 ii' 순서에서 i번째 처리의 이월 효과를 나타낸다. 그러므로 이월 효과의 동일성 검정

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

은 환자 안에서의 합을 3.1절에서 논의한 두 모집단의 선형 placement 통계량을 이용한 검정법에 적용하여 검정할 수 있다.

3.3. 이월 효과가 같을 때의 직접 효과의 비교 검정

위의 검정과정에서 이월 효과가 같다고 가정하고 직접 효과의 비교 검정은 식 (1)으로부터 같은 환자의 두 관측값의 차이가

$$y_{ij1} - y_{ij2} = (p_1 - p_2) + (-1)^{i+1}(\tau_1 - \tau_2) + (e_{ij1} - e_{ij2}) \tag{5}$$

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 환자 안에서의 차를 3.1절에서 논의한 두 모집단의 선형 placement 통계량을 이용한 검정법에 적용하여 검정할 수 있다.

3.4. 이월 효과가 같을 때의 시기 효과의 비교 검정

위의 검정과정에서 이월 효과가 같다고 가정하고 시기 효과의 비교 검정을 생각하여 보자. 만일 $\rho_1 = \rho_2$가 성립한다면 같은 환자의 두 관측값의 차이가 식 (5)와 같이 성립하게 된다. 이 때 순서 AB에서는 $y_{ij1} - y_{ij2}$, 순서 BA에서는 $y_{ij2} - y_{ij1}$ 의 교차 차이(cross-over differences)를 고려하자. 그러면 시기 효과의 차이가 없다는 가설

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

의 검정은 환자 안에서의 교차 차이를 3.1절에서 논의한 두 모집단의 선형 placement 통계량을 이용한 검정법에 적용하여 검정할 수 있다.

4. 제안된 검정법의 성질과 결론

이 절에서는 이월 효과, 처리의 직접 효과 및 시기 효과를 검정하기 위하여 제안된 통계량 $S_{n,m}$ 의 반복점근분포(iterative asymptotic distribution)를 Orban과 Wolfe(1982)와 Kim(1994)의 결과를 이용하여 알아보았다.

먼저 식 (2)에서 정의된 placement U_j들의 벡터를 $m\,\mathrm{U} = (mU_1, \cdots, mU_n)$ 이라 하고 $\mathrm{r} = (r_1, r_2, \cdots, r_n)$을 j 값을 갖는 r_i 들이 n_j 개, $0 \le n_j \le n$, $n_0 + \cdots + n_m = n$, $j = 0, 1, \cdots, m$ 을 만족하는 정수의 벡터라 하면 단순한 조합연산을 이용하여 다음과 같은 결과를 얻을 수 있다.

정리 1. 귀무가설 하에서

$$P_0[m\,\mathrm{U} = \mathrm{r}] = \frac{m!\cdot\prod_{j=0}^{m}(n_j!)}{(m+n)!}$$

여기서 벡터 r 은 j 값을 갖는 r_i 들이 n_j 개, $0 \le n_j \le n$, $n_0 + \cdots + n_m = n$, $j = 0, 1, \cdots, m$ 을 만족하는 정수 벡터이고 이를 만족시키지 못하는 벡터 r 에 대해서는 확률이 0이다.

이를 이용하여 귀무가설하에서의 선형 placement 통계량 $S_{n,m}$ 의 평균과 분산을 구하면

$$E_0(S_{n,m}) = n\,\overline{\varphi}_m$$

$$Var_0(S_{n,m}) = \frac{n(m+n+1)}{(m+1)(m+2)}\left[\sum_{i=0}^{m}\varphi_m^2\left(\frac{i}{m}\right) - (m+1)\,\overline{\varphi}_m^{\,2}\right]$$

이 되며, 여기서 $\overline{\varphi}_m = \sum_{i=0}^{m}\varphi_m\left(\frac{i}{m}\right)\Big/(m+1)$ 이다.

대조군의 표본크기 m이 무한히 커지고 그 후 실험군의 표본크기 n이 무한히 커질 경우, 즉 반복적인 개념 ($m\to\infty$ 이고 그 후 $n\to\infty$)에서의 선형 placement 통계량 $S_{n,m}$ 의 점근분포에 대하여 알아보자. 이는 Orban과 Wolfe(1982)과 Kim(1994)의 결과가 그대로 적용된다. 즉 점수 함수 φ_m 에 대한 두 가지 가정과 분포함수에 대한 가정 하에 다음과 같은 결과를 얻을 수 있다.

정리 2. (Kim(1994)) 몇 가지 가정 하에서 선형 placement 통계량 $S_{n,m}$ 의 반복점근분포

($m \to \infty$ 이고 그 후 $n \to \infty$)는 정규분포를 따른다. 즉

$$\frac{S_{n,m} - n\overline{\varphi}}{\sqrt{nV_\varphi}} \xrightarrow{d} N(0,1)$$

여기서 $\overline{\varphi} = E(\varphi[F(X_2)]) = \int \varphi[F(y)]\,dF(y)$

$$V_\varphi = Var(\varphi[F(X_2)]) = \int (\varphi[F(y)] - \overline{\varphi})^2\, dF(y)$$

이다.

정리 2를 이용하면 주어진 선형 placement 통계량 $S_{n,m}$ 의 근사적인 기각역과 검정력 함수를 구할 수 있다.

이제 본 논문에서 제안된 검정법을 기존의 모수적 방법인 Hills-Armitage 검정법과 비모수적 방법인 Koch의 검정법과 비교하여 보자. 검정법의 비교를 제 1종 오류 제어와 검정력의 측면에서 보면 Hills-Armitage 검정법은 실제로 원 자료의 합이나 차를 이용하여 두 모집단의 t-검정을 실시하는 것으로서 모집단의 가정에 민감하게 된다. 즉 자료가 모집단의 정규성 가정을 만족시키지 못할 경우에는 제 1종 오류를 제대로 제어하지 못하고 검정력 또한 상당히 떨어진다. 반면에 Koch의 검정법과 제안된 검정법은 비모수적 검정법이므로 모집단의 분포와 관계없이 제 1종 오류를 제어하며, 자료가 정규모집단으로부터 추출되었다 하더라도 Hills-Armitage 검정법에 비하여 검정력이 많이 떨어지지는 않는다. Koch의 검정법은 원 자료의 합이나 차를 이용한 Wilcoxon rank-sum 검정이고 제안된 검정법은 같은 상황에서의 placement를 이용한 검정법이므로 이들의 비교는 Kim(1994)의 결과에서 보듯이 반복적인 개념 ($m \to \infty$ 이고 그 후 $n \to \infty$)에서의 극한 분포가 동일하다. 또한 소 표본 모의실험의 결과로서 두 검정법의 검정력은 일반적으로 같다고 볼 수 있다. 그러나 위치 모수의 차이가 클 경우와 대조군의 표본 크기가 실험군의 표본크기에 비하여 상대적으로 클 때, 그리고 모집단의 분포가 Cauchy 또는 double exponential 분포를 따를 때 placement를 이용한 검정법의 검정력이 좋다. 그러므로 제안된 검정법은 두꺼운 꼬리를 갖는 대칭 모집단을 따를 때, 대조군의 표본크기가 실험군에 비하여 상대적으로 클 때 적용하는 것이 바람직할 것이다. 마지막으로 만일 이월 효과의 동일성 검정에서 귀무가설이 기각될 경우에는 3.3과 3.4에서 제안된 검정법은 적용할 수 없다. 이런 경우에는 Hills-Armitage 검정법과 마찬가지로 두 번째 시기에서 측정한 자료를 모두 버리고, 첫 번째 시기에서 측정된 자료만을 이용하여 직접 효과의 차이 유무를 Orban-Wolfe의 검정법을 적용한다.

참 고 문 헌

[1] Grizzle, J. E. (1965). The two-period change over design and its use in clinical trials.

Biometrics, Vol. 21, 467-480.

[2] Hills, A. V. and Armitage, P. (1979). The two-period cross-over clinical trial. *British Journal of Clinical Pharmacology*, Vol. 8, 7-20.

[3] Kim, D. (1994). A comparison of distribution-free two-sample procedures based on placements or ranks. *Journal of the Korean Statistical Society*, Vol. 23, 135-150.

[4] Koch, G. G. (1972). The use of non-parametric methods in the statistical analysis of the two-period change over design. *Biometrics*, Vol. 28, 577-584.

[5] Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placements. *Journal of the American Statistical Association*, Vol. 77, 666-671.