응용통계연구 (2010)
**23**(6), 1081–1091
DOI: 10.5351/KJAS.2010.23.6.1081

# Jonckheere-Terpstra 추세 검정통계량에 근거한 비모수적 층화분석법

조도연[1] · 양수[2] · 송혜향[3]

[1]가톨릭대학교 대학원 의학통계학과, [2]가톨릭대학교 간호대학
[3]가톨릭대학교 대학원 의학통계학과

(2010년 8월 접수, 2010년 10월 채택)

**요 약**

각 의료기관에서 수집될 수 있는 환자수가 한정되어 있는 질병의 경우에는 주로 다기관연구로써 임상연구가 진행된다. Jonckheere (1954)와 Terpstra (1952)의 추세 검정법으로 분석해야 하는 독립된 여러 군의 자료를 다기관에서 수집한 경우에 이질성을 고려하여 각 연구기관을 하나의 층으로 보아 층화분석법으로 분석하지 않으면 옳지 않은 결론에 도달할 수가 있다. 본 논문에서는 van Elteren (1960)이 제시한 Wilcoxon (1945) 검정통계량의 층화분석법을 이용하여 Jonckheere (1954)와 Terpstra (1952)의 추세 검정통계량에 근거한 층화분석법을 제시한다. 예제 자료에 이 층화분석법을 적용하며 효율성을 모의실험으로 알아본다.

주요용어: Jonckheere-Terpstra 추세검정, Wilcoxon 검정, 층화분석.

## 1. 서론

임상의 수많은 연구가 다기관연구(multi-center study)로 진행되고 있는데, 그 이유는 각 의료기관에서 수집 가능한 환자수가 한정되어 있기 때문이다. 비록 각 기관에서 수집되는 환자 자료는 동질적인 조건과 상태하에서 수집된다 하더라도, 여러 의료기관의 환자 자료는 서로 다른 환경과 조건으로 매우 이질적인 것이 일반적인 특징이다. 그러므로 다기관연구 자료의 분석에서 특징적인 점은 각 기관의 환자 자료는 독립된 한 세트(set)의 자료로서 분석될 만큼 충분한 수가 아니면서 또한 모든 기관의 환자 자료를 병합하여 분석하기에는 매우 이질적이라는 것이다. 따라서 이러한 다기관연구 자료의 분석에는 각 기관의 자료에서 계산된 검정통계량을 병합하여 하나의 검정통계량을 도출하는 층화분석법(stratified analysis)이 적절하다고 하겠다.

비정규분포 자료의 비모수적 검정법에 근거한 층화분석법을 Bajorski와 Petkau (1999)가 제안하였으며 저자들의 방법은 두 치료군과 세 수준 반응변수로 구성된 $2 \times 3$ 범주형 자료의 분석인 경우에 국한된다. 즉, 다발성경화증(mulitiple sclerosis) 환자를 치료군과 대조군으로 나누어 각 환자의 치료전과 비교한 치료후의 증상변화를 악화, 무변화, 호전의 세 수준으로 나눈 자료를 환자의 치료전 증상의 심각성에 따라 분석의 단계에서 층화시켜(post stratification) 치료효과를 알아보았다. 한편 연속형 자료의 경우에 두 군 비교에 대한 Wilcoxon (1945) 검정통계량을 병합하여 층화분석하는 방법을 일찍이 van Elteren (1960)이 제안하면서 근사적으로 정규분포하는 어떤 검정통계량의 선형결합도 가능하다고 설명하였고, Zhang (1996)은 이에 덧붙여 생존자료에 대한 층화분석법을 새로이 추가하였다.

[3]교신저자: (137-701) 서울시 서초구 반포동 505, 가톨릭대학교 대학원 의학통계학과, 교수. 인간유전체다형성연구소. E-mail: hhsong@catholic.ac.kr

본 논문에서는 단일 기관의 자료라면 Jonckheere (1954)와 Terpstra (1952)의 추세 검정법으로 분석해야 하는 독립된 세 군의 자료가 특히 다기관에서 수집된 임상자료의 분석방법을 다룬다(독립되게 쓰여진 두 논문을 간략히 Jonckheere-Terpstra 추세 검정법이라 부른다). 단일 기관의 자료에 적용하는 Jonckheere-Terpstra 추세 검정통계량은 독립된 세 군의 경우에 실제로 1군과 2군, 1군과 3군 그리고 2군과 3군의 Wilcoxon (1945) 검정통계량을 병합하여 검정통계량을 생성하므로 이제 Jonckheere-Terpstra 추세검정통계량에 근거한 층화분석법은 van Elteren (1960)이 제시한 Wilcoxon (1945) 검정통계량의 층화분석법과 기본적인 이론이 동일하다 하겠다. 본 논문에서 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법을 제시하며 이 층화분석법의 효율성을 모의실험으로 알아본다.

## 2. Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법

비층화 Jonckheere-Terpstra 추세 검정법을 우선 간략히 설명한 후, van Elteren (1960)이 제시한 독립된 두 군의 Wilcoxon (1945) 검정에 근거한 층화분석법을 설명한다. 앞의 두 방법에 바탕하여 Jonckheere-Terpstra 추세 검정법에 근거한 층화분석법을 제안한다.

### 2.1. Jonckheere-Terpstra 추세 검정통계량

비층화 Jonckheere-Terpstra 추세 검정법은 서로 독립인 $g$개($g \geq 3$)의 군에서 수집된 연속자료의 증가(또는 감소) 추세를 알아보는 비모수적 검정법으로 순위에 기초한다. 연속자료 $X_j$는 $j$군에서의 반응변수이며, $X_j$의 누적분포함수가 $F_j(x)$일 때, 위치모수 $\tau_j(>0)$를 이용하여 여러 군의 추세를 $F_j(x) = F(x-\tau_j)$, $(j = 1,\ldots,g)$으로 표현한다. 귀무가설하에서 확률변수 $X_1,\ldots,X_g$는 동일 누적분포함수를 가진다. 대립가설하에서는 모든 $x$에 대하여 $F_1(x) \geq F_2(x) \geq \cdots \geq F_g(x)$이 성립하며, 이 관계식에서도 적어도 하나의 부등식이 존재한다. 즉 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_g \quad vs. \quad H_1 : \tau_1 \leq \tau_2 \leq \cdots \leq \tau_g \tag{2.1}$$

마찬가지로 대립가설 $H_1$에서도 적어도 하나의 부등식이 존재한다. 위의 가설을 검정하기 위한 Jonckheere-Terpstra 추세 검정통계량은 다음과 같다.

$$J = \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} U_{uv}, \tag{2.2}$$

여기서 $U_{uv}$는 두 군 $u$와 $v$의 순위자료에 근거한 Mann과 Whitney (1947) 통계량(간략히 Wilcoxon 통계량)이며, 추세 검정통계량 $J$는 결과적으로 $g(g-1)/2$개의 Wilcoxon (1945) 통계량의 병합으로 표현된 것이다. 여기서 $u$군과 $v$군에 속한 표본수가 각각 $n_u$과 $n_v$이고, $u$군에 속한 $i$번째 자료와 $v$군에 속한 $i'$번째 자료의 값을 각각 $X_u^i$와 $X_v^{i'}$이라 할 때 동순위를 고려한 $U_{uv}(1 \leq u < v \leq g)$는 다음과 같이 정의할 수 있다.

$$U_{uv} = \sum_{i=1}^{n_u} \sum_{i'=1}^{n_v} \phi(X_u^i, X_v^{i'}) \quad \text{한편,}\ \phi(X_u^i,\ X_v^{i'}) = \begin{cases} 1, & X_u^i < X_v^{i'}, \\ \dfrac{1}{2}, & X_u^i = X_v^{i'}, \\ 0, & X_u^i > X_v^{i'}. \end{cases} \tag{2.3}$$

귀무가설하에서의 $U_{uv}$의 기댓값과 Tryon과 Hettmansperger (1973)가 제시한 대표본 근사하에서의 분

산과 공분산은 다음과 같다.

$$E(U_{uv}) = \frac{n_u n_v}{2}, \qquad \mathrm{Var}(U_{uv}) = \frac{n_u n_v (n_u + n_v + 1)}{12}, \tag{2.4}$$

$$\begin{cases} \mathrm{Cov}(U_{uv}, U_{st}) = 0, \\ \mathrm{Cov}(U_{uv}, U_{us}) = \mathrm{Cov}(U_{vu}, U_{su}) = \dfrac{n_u n_v n_s}{12}, \\ \mathrm{Cov}(U_{uv}, U_{su}) = \mathrm{Cov}(U_{vu}, U_{us}) = -\dfrac{n_u n_v n_s}{12}, \quad u \neq v \neq s \neq t. \end{cases} \tag{2.5}$$

위의 기댓값과 분산, 공분산을 이용하여 구한 식 (2.2)의 Jonckheere-Terpstra 추세 검정통계량의 기댓값과 분산은 다음과 같다 (Hollander와 Wolfe, 1999). 여기서 $N$은 $\sum_{j=1}^{g} n_j$이다.

$$E(J) = \frac{N^2 - \sum\limits_{j=1}^{g} n_j^2}{4}, \tag{2.6}$$

$$\mathrm{Var}(J) = \frac{N^2(2N+3) - \sum\limits_{j=1}^{g} n_j^2(2n_j+3)}{72}. \tag{2.7}$$

증가추세 검정은 통계량 $Z = [J - E(J)]/\sqrt{\mathrm{Var}(J)}$가 대표본하에서 근사적으로 표준정규분포함을 이용하여 단측검정으로 시행하며 검정통계량 $J$의 대표본 정규분포성은 $k$-표본 $U$-통계량($k$-sample $U$-statistics) 이론에 근거한다 (Lehmann, 1998).

### 2.2. van Elteren (1960)의 Wilcoxon (1945) 검정통계량에 근거한 층화분석법

이제 van Elteren (1960)이 제시한 Wilcoxon (1945) 검정통계량에 근거한 층화분석법을 설명한다. 여기서 총 $K$개의 층(stratum)이 있고, 각 층에 정해진 두 군이 있는 경우이며, 이 두 군을 $u$군과 $v$군이라고 하자. 확률변수 $X_{uk}$ $(k = 1, \ldots, K)$는 $k$번째 층에서 $u$군의 반응변수로써 누적분포함수 $F_{uk}(x)$를 가지며, 여러 층에서 두 군의 반응변수간에 차이가 존재하는 경우에 $F_{u1}(x), \ldots, F_{uK}(x)$는 완전 일치하지 않으며 또한 $v$군의 반응변수 $X_{vk}$ $(k = 1, \ldots, K)$의 누적분포함수 $G_{v1}(x), \ldots, G_{vK}(x)$도 완전 일치하지 않으므로 모든 층에 걸친 자료는 동일 분포하는 표본자료가 아니다. 이러한 경우에 모든 층의 자료를 종합하여 분석하는 자연스러운 방법은 각 층의 검정통계량, 즉 Wilcoxon (1945) 검정통계량을 종합하는 것이다. $k$번째 층에서 두 군의 표본수는 $n_{uk}$와 $n_{vk}$이며, $k$번째 층의 총 자료수는 $N_{uvk} = n_{uk} + n_{vk}$이 된다. 두 군에서의 $i$번째 자료와 $j$번째 자료의 값을 각각 $X_{iuk}$, $X_{jvk}$라 하자.

Wilcoxon (1945) 검정통계량은 2.1절의 $U$ 통계량으로 표현할 수도 있고, 또는 두 군을 종합한 자료에서 $v$군 자료의 순위합으로도 표현할 수가 있다. 이제 $k$번째 층의 두 군을 종합한 자료에서 $S_{1k}, S_{2k}, \ldots, S_{n_{vk}k}$가 $v$군 자료의 순위라고 할 때 Wilcoxon (1945) 검정통계량인 순위합은 다음과 같다.

$$W_{uvk} = \sum_{i=1}^{n_{vk}} S_{ik}. \tag{2.8}$$

이 Wilcoxon (1945) 검정통계량을 병합하여 층화분석 검정통계량을 생성한다. 이 $W_{uvk}$의 기댓값과 대표본하에서의 분산은 다음과 같다 (Hollander와 Wolfe, 1999).

$$E(W_{uvk}) = \frac{n_{vk}(N_{uvk}+1)}{2}, \qquad \mathrm{Var}(W_{uvk}) = \frac{n_{uk} n_{vk}(N_{uvk}+1)}{12}. \tag{2.9}$$

참고로 층화를 고려하지 않는 경우에 식 (2.8)의 $W_{uvk}$는 $W_{uv}$으로 표현되며 앞의 2.1절에서 제시한 $U_{uv}$는 $W_{uv}$와 상수관계, 즉 $W_{uv} = U_{uv} + n_v(n_v+1)/2$의 관계에 있으므로 식 (2.4)에 의해 기댓값과 분산을 유도할 수도 있다 (Hollander와 Wolfe, 1999).

이제 층화분석 검정통계량은 $W_{uvk}$에 가중상수(weight) $\lambda_k$을 곱한 후 병합하여 구한다. 즉,

$$\sum_{k=1}^{K} \lambda_k W_{uvk} > C. \tag{2.10}$$

이 만족할 때 귀무가설을 기각하게 되며, $C$는 유의수준 $\alpha$에 따라 결정되는 상수이다. 한편 가중상수는 여러 층의 자료의 위치(location)의 차이뿐만 아니라 척도(scale)의 다름을 반영하여 정하게 된다. 이제 $F_{uk}(x)$와 $G_{vk}(x)$ $(1 \le k \le K)$가 위치-척도족(location-scale family)에 속하는 분포라고 가정하면 확률분포함수는 다음과 같이 표현된다.

$$f_{uk}(x) = \frac{1}{b_k} f\left(\frac{x-a_k}{b_k}\right), \quad -\infty < a_k < +\infty,\ b_k > 0. \tag{2.11}$$

$$g_{vk}(x) = \frac{1}{b_k} g\left(\frac{x-a_k}{b_k}\right) = \frac{1}{b_k} f\left(\frac{x-a_k-\delta}{b_k}\right). \tag{2.12}$$

이러한 분포모형하에서 $\delta = 0$인 귀무가설과 $\delta > 0$인 대립가설에 대한 지역적으로 최대검정력(locally most powerful)이 보장되는 검정통계량은 다음과 같은 형태를 지니게 된다 (van Elteren, 1960, Zhang, 1996).

$$\sum_{k=1}^{K} \left(\frac{\tau_{uvk}}{N_{uvk}+1}\right) W_{uvk} > C, \tag{2.13}$$

여기서 $\tau_{uvk} = \sigma_{uv1}/\sigma_{uvk}$이며 $\sigma_{uvk}^2$는 $k$층에서 동일분산 가정하에서 $u$와 $v$ 두 군의 공통분산이다.

한편, 식 (2.10)에 의해 귀무가설 하에서 각 층의 통계량

$$\left(\frac{\tau_{uvk}}{N_{uvk}+1}\right) W_{uvk} \tag{2.14}$$

는 근사적으로 평균이 $\tau_{uvk} n_{vk}/2$, 분산이 $\tau_{uvk}^2 n_{uk} n_{vk}/[12(N_{uvk}+1)]$인 정규분포에 따름을 알 수 있다 (Hájek과 Šidák, 1967). 이제 모든 층에 대해서 가중상수를 곱하여 병합한 층화분석 검정통계량

$$W_{uv} = \sum_{k=1}^{K} \left(\frac{\tau_{uvk}}{N_{uvk}+1}\right) W_{uvk} \tag{2.15}$$

의 기댓값과 분산은 다음과 같다 (Zhang, 1996).

$$E(W_{uv}) = \sum_{k=1}^{K} \frac{\tau_{uvk} n_{vk}}{2}, \qquad \mathrm{Var}(W_{uv}) = \sum_{k=1}^{K} \frac{\tau_{uvk}^2 n_{uk} n_{vk}}{12(N_{uvk}+1)}. \tag{2.16}$$

$\tau_{uvk} = \sigma_{uv1}/\sigma_{uvk}$의 추정량으로서는 각 층의 두 군의 표본수가 작지 않다면 합병표본분산인 $\hat{\sigma}_{uvk}^2$ $(k = 1, \ldots, K)$을 분자와 분모에 대입하고 이를 식 (2.15)와 (2.16)에 대입하여 검정한다. 구체적으로 $k$번째 층에서의 합병표본분산은

$$\hat{\sigma}_{uvk}^2 = \frac{\sum_{i=1}^{n_{uk}} (x_{iuk} - \bar{x}_{uk})^2 + \sum_{j=1}^{n_{vk}} (x_{jvk} - \bar{x}_{vk})^2}{N_{uvk} - 2} \tag{2.17}$$

이고, 여기서 $\bar{x}_{uk}$와 $\bar{x}_{vk}$는 각각 $x_{1uk}, \ldots, x_{n_{uk}uk}$와 $x_{1vk}, \ldots, x_{n_{vk}vk}$의 평균이다. 이제 통계량 $Z^* = [W_{uv} - E(W_{uv})]/\sqrt{\mathrm{Var}(W_{uv})}$가 대표본하에서 근사적으로 표준 정규분포함을 이용하여 층화분석을 시행한다. 층화분석에서 근사 표준 정규분포의 대표본이란 여러 층의 표본수를 모두 합한 총 표본수가 대표본임을 의미한다 (Cochran, 1968; Palta, 1983).

### 2.3. Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법

이제 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법을 설명한다. 식 (2.2)에서 제시된 $k$번째 층에서의 $J_k$통계량은 $U$통계량 또는 순위합을 이용하여 표현될 수 있다.

$$J_k = \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} U_{uvk} = \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( W_{uvk} - \frac{n_{vk}(n_{vk}+1)}{2} \right). \tag{2.18}$$

이제 모든 층에 대해 병합한 $J^*$통계량은 다음과 같다.

$$\begin{aligned} J^* &= \sum_{k=1}^{K} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) J_k \\ &= \sum_{k=1}^{K} \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) U_{uvk} \\ &= \sum_{k=1}^{K} \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) \left( W_{uvk} - \frac{n_{vk}(n_{vk}+1)}{2} \right). \end{aligned} \tag{2.19}$$

$J^*$의 기댓값과 분산은 식 (2.4), (2.5) 또는 (2.9)를 이용하면 다음과 같다.

$$\begin{aligned} E(J^*) &= \sum_{k=1}^{K} \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) \left( E(W_{uvk}) - \frac{n_{vk}(n_{vk}+1)}{2} \right) \\ &= \sum_{k=1}^{K} \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) \left( \frac{n_{uk}n_{vk}}{2} \right). \end{aligned} \tag{2.20}$$

$$\begin{aligned} \mathrm{Var}(J^*) &= \mathrm{Var} \left[ \sum_{k=1}^{K} \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) U_{uvk} \right] \\ &= \sum_{k=1}^{K} \left[ \sum_{u=1}^{g-1} \sum_{v=u+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right)^2 \mathrm{Var}(U_{uvk}) \right. \\ &\quad \left. + \sum_{u=1}^{g-1} \sum_{\substack{v=u+1 \\ (u,v)\neq}}^{g} \sum_{\substack{s=1 \\ (s,t)}}^{g-1} \sum_{t=s+1}^{g} \left( \frac{\tau_{uvk}}{N_{uvk}+1} \right) \left( \frac{\tau_{stk}}{N_{stk}+1} \right) \mathrm{Cov}(U_{uvk}, U_{stk}) \right], \end{aligned} \tag{2.21}$$

여기서의 $\mathrm{Var}(U_{uvk})$와 $\mathrm{Cov}(U_{uvk}, U_{stk})$는 다음과 같다.

$$\mathrm{Var}(U_{uvk}) = \frac{n_{uk}n_{vk}(N_{uvk}+1)}{12}, \quad 1 \leq u < v \leq g, \tag{2.22}$$

$$\begin{cases} \mathrm{Cov}(U_{uvk}, U_{stk}) = 0, \\ \mathrm{Cov}(U_{uvk}, U_{usk}) = \mathrm{Cov}(U_{vuk}, U_{suk}) = \dfrac{n_{uk}n_{vk}n_{sk}}{12}, \\ \mathrm{Cov}(U_{uvk}, U_{vsk}) = \mathrm{Cov}(U_{vsk}, U_{uvk}) = -\dfrac{n_{uk}n_{vk}n_{sk}}{12}, \quad u \neq v \neq s \neq t,\ k = 1, \ldots, K. \end{cases} \tag{2.23}$$

표 3.1. Yang (1992)의 자료

| 센터A | | | 센터B | | | 센터C | | |
|---|---|---|---|---|---|---|---|---|
| 대조군 | 치료 I군 | 치료 II군 | 대조군 | 치료 I군 | 치료 II군 | 대조군 | 치료 I군 | 치료 II군 |
| 9 | 19 | 22 | 13 | 10 | 24 | 8 | 24 | 25 |
| 5 | 16 | 27 | 10 | 16 | 19 | 19 | 19 | 26 |
| 15 | 28 | 24 | 19 | 15 | 26 | 18 | 26 | 26 |
| 13 | 21 | 24 | 10 | 18 | 20 | 20 | 26 | 27 |
| 11 | 25 | 19 | 13 | 22 | - | 10 | 25 | 27 |
| 14 | 13 | 20 | 8 | - | - | - | 19 | 28 |
| 23 | - | 24 | - | - | - | - | 25 | 28 |
| 8 | - | - | - | - | - | - | - | - |
| 10 | - | - | - | - | - | - | - | - |

이제 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법 검정통계량은 $Z^* = [J^* - E(J^*)]/\sqrt{\mathrm{Var}(J^*)}$가 대표본하에서 근사적으로 표준정규분포함을 이용하여 증가추세를 단측검정으로 시행한다. 이 때 가중상수 $\tau_{uvk} = \sigma_{uv1}/\sigma_{uvk}$의 분자와 분모의 추정량으로서 각 층의 표본수가 작지 않다면 2.2절에서와 같이 각 층에서의 합동표본분산으로 대치하면 된다.

## 3. 예제자료의 분석 및 모의실험

### 3.1. 예제자료의 분석

본 논문에서 분석한 예제 자료는 정신분열증(schizophrenia) 환자를 세 다른 센터(A, B, C)에서 모집하여 항정신분열증 약제(antipsychotic drugs)로 인해 배변이 수월하지 않은 총 56명을 2주간의 사전시행(pre-trial)후에 각 센터별로 세 군, 즉 대조군, 치료 I군, 치료 II군에 할당한 자료이다 (Yang, 1992). 대조군은 수분섭취만을 처방받는다. 치료 I군은 매일 일정량의 수분섭취와 15분간의 운동을 처방받으며, 치료 II군은 치료 I군에 덧붙여 식이섬유질을 처방받는다. 정상적인 배변은 하루에 적어도 한 번 이상이며 이를 성공으로 간주한다.

센터 A에서는 총 22명의 환자가 모집되었으며 대조군에 9명, 치료 I군에 6명, 치료 II군에 7명이 랜덤할당되었다. 센터 B에서는 총 15명의 환자가 대조군 6명, 치료 I군 5명, 치료 II군 4명이 랜덤할당되고 센터 C에서는 대조군 5명, 치료 I군 7명, 치료 II군 7명이 랜덤할당되었다.

반응 변수는 총 4주간의 치료기간 중 하루에 적어도 한 번 이상 배변에 성공한 날짜수이며 표 3.1에 각 센터별 환자들의 4주간 배변에 성공한 날짜수가 제시되었다. 분석의 목적은 대조군과 비교하여 치료 I군, 치료 II군의 순서로 배변이 점진적으로 더욱 수월한가를 검정하는것이며 따라서 배변에 성공한 날짜수가 증가하는가에 대한 검정이 된다. 표 3.1의 자료를 살펴보면 모든 센터에서 대조군, 치료 I군, 치료 II군의 순서로 배변일수가 증가추세에 있음을 볼 수 있다. 센터별로 각 군마다 Shapiro-Wilk 검정통계량을 이용한 정규성 검정을 한 결과 센터 C의 치료 I군에서의 자료가 정규분포를 따르지 않는다 ($p$-value $= 0.016$).

표 3.1의 자료를 이용하여 모수적 방법과 비모수적 방법으로 추세 검정을 하였다.

모수적 방법으로는 회귀직선의 기울기로 추세를 알아보는 방법이며 세 군을 나타내는 순위변수에는 각각 0, 1, 2의 값을 부여하였다. 층화분석으로는 세 센터를 두 지시변수로서 회귀직선에 추가하여 층에 따른 영향을 공변량으로 조정한 후 기울기에 대해 유의성을 알아보는 방법이다.

비모수적 방법은 2장에서 설명한 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법을 말한

표 **3.2.** Yang (1992) 자료의 분석결과

| | | 모수적 방법 | | | | 비모수적 방법 | | |
|---|---|---|---|---|---|---|---|---|
| | | 추정량 | 표준오차 | $t$ (df) | $p$-value | $J$ 통계량 | $Z$ | $p$-value |
| 비층화분석 | 기울기 | 5.742 | 0.728 | 7.88 (54) | < .0001 | 894.500 | 5.619 | < .0001 |
| 층화분석 | 기울기 | 5.436 | 0.665 | 8.17 (52) | < .0001 | 26.968 | 6.152 | < .0001 |
| | 센터B | −1.296 | 1.361 | −0.95 (52) | 0.345 | | | |
| | 센터C | 3.627 | 1.280 | 2.83 (52) | 0.006 | | | |

표 **3.3.** 정규분포 자료에 근거한 모수적, 비모수적 방법의 제 1종 오류와 검정력

| 총 표본수($N$) | 54명 | | 90명 | | 180명 | |
|---|---|---|---|---|---|---|
| $(\tau_1, \tau_2, \tau_3)$ | 모수적 방법 | 비모수적 방법 | 모수적 방법 | 비모수적 방법 | 모수적 방법 | 비모수적 방법 |
| $K$=1 | $n_{jk} = 18$ | | $n_{jk} = 30$ | | $n_{jk} = 60$ | |
| (0, 0, 0) | 0.054 | 0.052 | 0.056 | 0.053 | 0.047 | 0.051 |
| (0, 0, 0.5) | 0.325 | 0.426 | 0.461 | 0.564 | 0.751 | 0.801 |
| (0, 0.5, 1) | 0.827 | 0.871 | 0.962 | 0.977 | 1.000 | 1.000 |
| (0, 1, 1) | 0.837 | 0.879 | 0.974 | 0.987 | 1.000 | 1.000 |
| $K$=2 | $n_{jk} = 9$ | | $n_{jk} = 15$ | | $n_{jk} = 30$ | |
| (0, 0, 0) | 0.054 (0.031)+ | 0.046 (0.028) | 0.058 (0.036) | 0.049 (0.028) | 0.056 (0.031) | 0.048 (0.032) |
| (0, 0, 0.5) | 0.324 (0.230) | 0.424 (0.341) | 0.496 (0.405) | 0.587 (0.520) | 0.786 (0.693) | 0.848 (0.795) |
| (0, 0.5, 1) | 0.838 (0.773) | 0.879 (0.848) | 0.972 (0.949) | 0.979 (0.967) | 0.999 (0.999) | 1.000 (1.000) |
| (0, 1, 1) | 0.815 (0.743) | 0.870 (0.836) | 0.968 (0.947) | 0.977 (0.967) | 1.000 (1.000) | 1.000 (1.000) |
| $K$=3 | $n_{jk} = 6$ | | $n_{jk} = 10$ | | $n_{jk} = 20$ | |
| (0, 0, 0) | 0.050 (0.013) | 0.054 (0.017) | 0.045 (0.007) | 0.049 (0.011) | 0.045 (0.013) | 0.055 (0.015) |
| (0, 0, 0.5) | 0.311 (0.112) | 0.410 (0.238) | 0.467 (0.273) | 0.576 (0.377) | 0.772 (0.563) | 0.830 (0.689) |
| (0, 0.5, 1) | 0.837 (0.654) | 0.871 (0.766) | 0.962 (0.898) | 0.977 (0.932) | 1.000 (0.998) | 1.000 (1.000) |
| (0, 1, 1) | 0.822 (0.646) | 0.870 (0.748) | 0.963 (0.890) | 0.973 (0.928) | 0.999 (1.000) | 1.000 (1.000) |
| $K$=6 | $n_{jk} = 3$ | | $n_{jk} = 5$ | | $n_{jk} = 10$ | |
| (0, 0, 0) | 0.059 (0.000) | 0.051 (0.001) | 0.053 (0.000) | 0.056 (0.000) | 0.056 (0.000) | 0.053 (0.001) |
| (0, 0, 0.5) | 0.307 (0.009) | 0.403 (0.031) | 0.468 (0.019) | 0.601 (0.057) | 0.772 (0.019) | 0.821 (0.239) |
| (0, 0.5, 1) | 0.833 (0.163) | 0.824 (0.296) | 0.973 (0.478) | 0.973 (0.644) | 1.000 (0.932) | 1.000 (0.978) |
| (0, 1, 1) | 0.806 (0.138) | 0.806 (0.284) | 0.963 (0.455) | 0.960 (0.599) | 0.999 (0.926) | 0.999 (0.968) |

"+" 괄호 안에는 비층화분석법에 의한 결과가 제시되었다.

다. 검정결과를 표 3.2에 제시하였다.

표 3.2에 제시된 모수적 방법의 결과를 살펴보면 비층화분석에서 회귀식의 기울기가 양수로써 대조군에서 치료 I군, 치료 II군의 방향으로 배변일수의 증가추세가 있으며 유의하다. 모수적 층화분석의 결과에서도 모든 센터에서 증가추세가 있으며 그러나 기울기가 층화분석의 경우 약간 감소하였다. 세 센터의 지시변수에 의한 검정결과에서 센터 C의 환자는 센터 A의 환자보다도 배변일수가 유의하게 크게 나타남을 알 수 있다.

비모수적 방법의 결과를 살펴보면 비층화분석과 층화분석의 경우 모두 증가추세가 있는 것으로 나타나며 층화분석의 경우가 더욱 유의하다.

### 3.2. 모의실험 계획 및 결과

이제 모의실험으로 층화분석의 효율성을 알아본다.

정규분포와 비정규분포 자료의 생성과정은 다음과 같다.

표 **3.4.** 비정규분포 자료에 근거한 모수적, 비모수적 방법의 제 1종 오류와 검정력

| 총 표본수($N$) | 54명 | | 90명 | | 180명 | |
|---|---|---|---|---|---|---|
| $(\tau_1, \tau_2, \tau_3)$ | 모수적 방법 | 비모수적 방법 | 모수적 방법 | 비모수적 방법 | 모수적 방법 | 비모수적 방법 |
| $K$=1 | $n_{jk} = 18$ | | $n_{jk} = 30$ | | $n_{jk} = 60$ | |
| (0, 0, 0) | 0.028 | 0.032 | 0.041 | 0.050 | 0.045 | 0.044 |
| (0, 0, 0.5) | 0.051 | 0.142 | 0.069 | 0.182 | 0.090 | 0.270 |
| (0, 0.5, 1) | 0.104 | 0.302 | 0.137 | 0.433 | 0.241 | 0.688 |
| (0, 1, 1) | 0.114 | 0.309 | 0.150 | 0.436 | 0.264 | 0.710 |
| $K$=2 | $n_{jk} = 9$ | | $n_{jk} = 15$ | | $n_{jk} = 30$ | |
| (0, 0, 0) | 0.024 (0.022)+ | 0.034 (0.034) | 0.049 (0.036) | 0.056 (0.000) | 0.041 (0.036) | 0.038 (0.029) |
| (0, 0, 0.5) | 0.043 (0.041) | 0.115 (0.103) | 0.051 (0.050) | 0.163 (0.140) | 0.088 (0.085) | 0.268 (0.250) |
| (0, 0.5, 1) | 0.099 (0.094) | 0.271 (0.271) | 0.142 (0.135) | 0.421 (0.385) | 0.247(0.243) | 0.716 (0.683) |
| (0, 1, 1) | 0.093 (0.091) | 0.263 (0.246) | 0.131 (0.128) | 0.410 (0.378) | 0.267 (0.261) | 0.690 (0.654) |
| $K$=3 | $n_{jk} = 6$ | | $n_{jk} = 10$ | | $n_{jk} = 20$ | |
| (0, 0, 0) | 0.030 (0.030) | 0.037 (0.025) | 0.042 (0.032) | 0.047 (0.039) | 0.031 (0.028) | 0.039 (0.027) |
| (0, 0, 0.5) | 0.040 (0.037) | 0.135 (0.091) | 0.068 (0.064) | 0.203 (0.157) | 0.101 (0.084) | 0.280 (0.219) |
| (0, 0.5, 1) | 0.122 (0.118) | 0.303 (0.252) | 0.162 (0.151) | 0.434 (0.359) | 0.248 (0.238) | 0.714 (0.601) |
| (0, 1, 1) | 0.128 (0.112) | 0.282 (0.219) | 0.147 (0.142) | 0.410 (0.320) | 0.258 (0.247) | 0.704 (0.585) |
| $K$=6 | $n_{jk} = 3$ | | $n_{jk} = 5$ | | $n_{jk} = 10$ | |
| (0, 0, 0) | 0.023 (0.018) | 0.031 (0.016) | 0.049 (0.016) | 0.048 (0.016) | 0.043 (0.023) | 0.047 (0.016) |
| (0, 0, 0.5) | 0.036 (0.028) | 0.135 (0.086) | 0.060 (0.068) | 0.162 (0.065) | 0.082 (0.065) | 0.294 (0.135) |
| (0, 0.5, 1) | 0.162 (0.086) | 0.329 (0.152) | 0.184 (0.097) | 0.458 (0.182) | 0.268 (0.162) | 0.772 (0.328) |
| (0, 1, 1) | 0.160 (0.082) | 0.307 (0.144) | 0.184 (0.099) | 0.434 (0.188) | 0.263 (0.167) | 0.760 (0.329) |

"+" 괄호 안에는 비층화분석법에 의한 결과가 제시되었다.

앞의 2.2절에서 설명하였듯이 확률변수 $X_{jk}$는 $k$번째 층에서 $j$군 $(j = 1, 2, 3)$의 반응변수이며 자료수는 $n_{jk}$라 하자. 총 자료수는 $N = \sum\sum_{jk} n_{jk}$이며 $N = 54$, 90, 180으로 정한다. $K$는 1, 2, 3, 6을 선택하여 $K = 1$인 경우가 층이 없는 경우이며, $K = 2$, $K = 3$과 $K = 6$은 각각 두 센터, 세 센터와 여섯 센터가 있는 경우이다. $k$번째 층에서의 $j$군의 표본수 $n_{jk}$는 모두 같도록 생성한다.

정규분포 자료는 $X_{jk} = \tau_j + \epsilon_{jk}$로서 오차항인 $\epsilon_{jk}$는 분포가 $N(v_k, \sigma^2)$을 따르도록 생성하고 $v_k = k - 1$로써 여러 층의 반응 변수의 분포에 차이가 있게 생성하였다. 귀무가설하에서는 각 층에서 동일 누적분포함수를 가지도록 정하여 $\tau_j = 0$ ($j$=1, 2, 3)이 되며, 대립가설하에서는 여러가지 형태의 추세를 검정하기 위하여 $(\tau_1, \tau_2, \tau_3) =$ (0, 0, 0.5), (0, 0.5, 1) 또는 (0, 1, 1)로 두었다.

비정규분포 자료는 위의 $\epsilon_{jk}$의 분포가 $(1 - c)N(v_k, \sigma^2) + cN(3 + v_k, (f\sigma)^2)$가 되도록 형성하였으며 $c = 0.5$, $\sigma^2 = 1$, $f = 5$ 로 가정하였다. 즉, $n_{jk} = 30$일 때 15개의 자료는 $N(v_k, \sigma^2)$에서 생성되며 나머지 15개의 자료는 $N(3 + v_k, (f\sigma)^2)$에서 생성된다.

모의실험의 각 경우에 대해서 1,000번 반복하여 제 1종 오류와 검정력을 비교하며 그 결과를 표 3.3과 3.4에 제시한다. 모의 실험의 프로그램은 SAS를 이용하였다.

표 3.3의 정규분포 자료의 모수적인 방법과 비모수적인 방법의 제 1종 오류를 살펴보면 비층화분석의 경우에는($K = 1$) 비모수적방법의 제 1종 오류가 정해진 유의수준 0.05에 더욱 근접하며 $K = 2$, 3, 6인 경우에는 비층화분석의 제 1종 오류는 층화의 수가 더 많아질수록 유의수준 0.05에 도달하지 못하는 경향을 보인다. 검정력 결과를 살펴보면 (0, 0, 0.5)나 (0, 1, 1)과 같이 어느 두 군 사이에서 추세가 급격히 생기는 경우에는 비모수적 방법인 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법이 높은 것을 알 수 있으며 선형적인 추세인 (0, 0.5, 1)에서는 모수적 방법이 더 높게 나타날 것이라는 예상과

는 달리 다른 결과를 보였다. 전체적으로 표 3.3을 살펴보면 같은 갯수의 층화($K$)에서 총 표본수($N$)가 많아질수록 검정력이 높아지는 경향을 보였다. 마찬가지로 같은 총 표본수에서 층화의 수가 많아질수록 조금 더 높아지는 경향을 보였지만 층화의 수가 3개 이상부터는 검정력의 변화가 거의 없는 것으로 나타났다. $K = 6$, $N = 54$인 경우에는 오히려 검정력이 더 떨어지는 경향을 나타났는데 이러한 결과는 $n_{jk}$의 수가 너무 작은 극소수 표본이므로 정규분포를 잘 형성하지 못하는 문제 때문이다. 그리고 층화분석에 비해 비층화분석의 검정력이 현저히 낮다는 것을 알 수 있다.

표 3.4에 제시된 비정규분포 자료의 층화분석 결과를 살펴보면 제 1종 오류수준이 $K = 2, 3$의 54명, $K = 6$의 180명에서 0.05에 미치지 못함을 알 수 있다. 그러나 표 3.3과 마찬가지로 이러한 몇 가지 경우를 제외하고는 제 1종 오류수준이 비슷함을 알 수 있다. 비층화분석의 제 1종 오류는 유의수준 0.05에 도달하지 못한다. 검정력결과는 전체적으로 비정규분포에서의 검정력이 정규분포의 경우보다 낮은 것을 알 수 있으며 모수적 방법과 비모수적 방법의 검정력 비교에서는 정규분포에서와 마찬가지로 비모수적 방법이 더 높은 경향을 보인다. 총 표본수와 층화가 많을 수록 검정력이 높아지는 결과를 보이며 정규분포에서와는 달리 극소수 표본의 문제가 발생하지 않는다. 두 방법 모두에서 군 간의 선형적인 추세의 검정력이 증가 추세가 급격한 경우의 검정력보다 두 배 이상 높아지는 것을 알 수 있다. 마찬가지로 비층화분석의 검정력은 매우 낮다.

## 4. 결론 및 고찰

본 논문에서는 Jonckheere-Terpstra 추세 검정통계량에 근거한 층화분석법을 제시하면서 모수적 방법과의 차이를 비교해 보았다. 앞에서 제안한 층화분석법은 다기관연구자료에서 서로 다른 층에서의 분포의 차이로 인해 초래될 수 있는 추세 검정통계량의 편의(bias)를 줄일 수 있다는 장점이 있을뿐만 아니라 현실적으로는 각 연구기관에서 수집될 수 있는 자료의 수가 너무 작은 경우나 또는 비정규분포하는 경우에 층화분석법을 적용하면 더욱 검정력이 높아짐을 모의실험을 통해 알아보았다. 이질성이 존재하여 층화분석법으로 분석되어야 하는 자료를 비층화분석법으로 분석하게 되면 제 1종 오류와 검정력이 매우 낮아 잘못된 결론으로 이끌어 질 수 있다.

Cochran (1968)과 Palta (1983)는 연속변수의 그룹화로 층화시키는 경우에 여러 군의 평균 비교에서 층화분석은 공분산분석으로 검정하는 것과 동일하여 일반적으로 적용하는 분산분석보다도 검정력이 높다고 밝혔다. 이 저자들은 또한 2개 또는 3개의 층으로써 효율성을 높이는 목적에 도달할 수 있음을 모의실험으로 제시해 보이면서 대표본하에서 근사 정규분포를 이용하는 조건은 여러 층을 합한 총 표본수임을 명시하였고 특히 Palta (1983)는 대표본의 예로써 총 표본수가 64개인 경우만을 모의실험에서 살펴보았다. 본 논문에서는 총 표본수가 최소 54명부터 180명까지를 검토하였으며, 특히 정규분포의 경우 각 층에서 3명인 극소수 표본수 $n_{jk} = 3$의 모의실험에서는 제 1종 오류가 조절되지 못하는 결과를 볼 수 있었다. 이러한 극소수 표본수의 문제로 인해 층화수가 증가함에도 불구하고 검정력의 증가를 볼 수 없었다.

Zhang (1996)은 생존자료의 층화분석법을 제시하였는데 이러한 생존율의 비교에서는 다수 공변수가 감안되어야 하므로 여러 공변수들의 상관성을 고려한 층화분석법이 적절하다 하겠다. 따라서 본 논문에서 다룬 비모수적 다변량 추세검정법에 근거한 층화분석법에서도 상관성이 있는 여러 공변수의 상관성을 감안한 층화분석법을 제시하는 것이 장차 필요하다 하겠다.

## 참고문헌

Bajorski, P. and Petkau, J. (1999). Nonparametric two-sample comparisons of changes on ordinal responses, *Journal of the American Statistical Association*, **94**, 970–978.

Cochran, W. G. (1968). The effectivness of adjustment by sub-classification in removing bias in observational studies, *Biometerics*, **24**, 295–313.

van Elteren, P. H. (1960). On the combination of independent two sample tests of wilcoxon, *Bulletin of International Statistical Institute*, **37**, 351–361.

Hájek, J. and S̃idák, Z. (1967). *Theory of Rank Test*, Academic Press, New York.

Hollander, M. and Wolfe, D. A. (1999). *Nonparametric Statistical Methods*, John Wiley & Sons,New York.

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

Lehmann, E. L. (1998). *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, San Francisco.

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

Palta, M. (1983). The effects of a stratified analysis of censored data, *Communication in Statistics - Simulation and Computation*, **12**, 273–290.

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.

Tryon, P. V. and Hettmansperger T. P. (1973). A class of non-parametric tests for homogeneity aganist ordered alternatives, *The Annals of Statistics*, **1**, 1061–1070.

Wilcoxon, F. (1945). Individual comparisons by ranking methods, *Biometrics*, **1**, 80–83.

Yang, S. (1992). Effects of fluid intake, dietary fiber supplement and abdominal muscle exercises on antipsychotic drug-induced constipation in schizophrenics, *Journal of Catholic Medical College*, **45**, 1501–1514.

Zhang, Z. (1996). Weighted combination of wilcoxon tests with interlaboratory lifetime data, *Sankhyāa Indian Journal of Statistics Series A*, **58**, 311–327.

# A Nonparametric Stratified Test Based on the Jonckheere-Terpstra Trend Statistic

Do-Yeon Cho[1] · Soo Yang[2] · Hae-Hiang Song[3]

[1]Department of Biostatistics, Graduate School, The Catholic University of Korea
[2]College of Nursing, The Catholic University of Korea
[3]Department of Biostatistics, Graduate School, The Catholic University of Korea

(Received August 2010; accepted October 2010)

**Abstract**

Clinical trials are often carried out as multi-center studies because the patients enrolled for a trial study are very limited in one particular hospital. In these circumstances, the use of an ordinary Jonckheere (1954) and Terpstra (1952) test for testing trend among several independent treatment groups is invalid. We propose a the stratified Jonckheere-Terpstra test based on van Elteren (1960)'s stratified test of Wilcoxon (1945) statistics and an application of our method is demonstrated through example data. A simulation study compares the efficiency of stratified and unstratified Jonckheere-Terpstra trend tests.

Keywords: Jonckheere-Terpstra trend test, wilcoxon test, stratified analysis.

[3]Corresponding author: Professor, Department of Biostatistics, Graduate School, Integrated Research Center for Genome Polymorphism, The Catholic University of Korea, 505 Banpo-Dong, Seocho-Gu, Seoul 137-701, Korea. E-mail: hhsong@catholic.ac.kr