응용통계연구 제10권 1호
1997년 3월 pp. 177-187

이원혼합모형에서 고정효과 유의성검정에 대한 검정력 분석

이 장 택[1)]

요 약

고정효과가 하나인 이원혼합모형에서의 고정효과 유의성검정에 대한 검정력 분석을 고려한다. 고정효과 수준간의 차이를 검정하는 데 사용되는 일반화 최소제곱 F 통계량을 헨더슨의 방법 III, 사전추측값이 1인 MINQUE 추정량, 최우추정법, 제한적 최우추정법을 이용하여 구하고 이 검정통계량들의 검정력을 모의실험을 통하여 알아본다. 모의실험의 결과는 결론적으로 검정력의 측면에서 살펴본 효율성은 4가지 추정량 모두 대체적으로 비슷한 것으로 판명되었다.

1. 서 론

분산성분을 추정하는 수많은 방법들이 오랜 세월에 거처 제안되고 많은 사람들의 공감을 얻지 못하는 것들은 독자들의 관심밖으로 사라져 갔다. 그 중 불균형혼합모형에서 오늘날까지 널리 사용되는 방법은 대략 4가지 정도라고 할 수 있는 데, 이들을 살펴보면, 가장 고전적으로 사용되는 헨더슨의 방법 III과 1967년을 기점으로 새롭게 주목받기 시작한 ML, REML 그리고 MINQUE가 있다. 여기에 소개된 4가지 방법은 모두 그 나름대로의 이론적 근거가 있기 때문에 어느 것이 가장 우수한 추정량이라고는 말하지 못하며 이는 전적으로 분석자의 선택에 달려있는 문제이다. 4가지 방법에 대한 비교및 검토는 주로 모의실험을 이용하여 비교가 용이한 일원변량모형에서 분산성분의 추정에 대하여 최소자승오차(MSE)와 편의의 관점에서 비교분석 되었는데, Swallow와 Monahan(1984), Conerly와 Webster(1987), Lee(1993)와 같은 논문에서 그 결과를 찾아 볼 수 있다.

하지만, 만일 우리의 관심사가 단지 혼합모형에서 고정효과의 수준간에 유의한 차이가 있는지를 알아보는 데에 있다면, 점추정량의 관점에서 살펴본 경우와 결과가 일치할까? 하는 의문이 생긴다. 그러나 여러가지 가설에 대하여 다양한 분산성분추정량이 가설검정의 결과에 미치는 영향을 이론적으로 밝히는 일은 불가능하다. 그러므로 몇가지 형태의 가설에 대하여 여러가지 분산성분추정량을 이용한 검정통계량들의 검정력을 경험적으로 알아보는 것이 부분적으로나마 문제의 해답을 찾는 방안이 될 수 있는데, 이러한 방향으로 접근한 최초의 논문은 Rao, Sutrahar, Kim(1993)의 논문이다. 그들은 이단집락표본을 이용한 회귀분석에서 일반화 최소제곱 F 검정에 관해 논의했으며 헨더슨의 방법 III을 사용하는 경우가 분산분석추정량, Wu, Holt,

1) (140-714) 서울시 용산구 한남동 산 8번지 단국대학교 전산통계학과 부교수

Holmes(1988)의 방법을 사용하는 경우보다 검정력이 우수하다고 컴퓨터 모의실험을 통하여 밝혔다.

이 논문은 실험계획에서 가장 빈번하게 사용되는 혼합모형중의 하나인 이원혼합모형인 경우에 이단계 최소제곱분석을 고려한다. 먼저 이원혼합모형에 포함된 분산성분들을 핸더슨의 방법 III과 최우추정법, 제한적 최우추정법 그리고 사전추측값이 1인 MINQUE를 사용하여 구하고 다음으로 추정된 모수를 사용하여 고정효과에 대한 일반화 최소제곱분석을 수행한다. 본 논문의 구성은 2절에서는 혼합모형과 일반화 최소제곱 F-검정 그리고 4가지 분산성분추정량에 대하여 간략하게 서술하고, 3절에서는 모의실험의 계획및 사용된 가설의 종류, 모의실험의 과정과 4가지 추정량을 이용한 일반화 최소제곱 F 통계량의 제1종오류율과 경험적 검정력의 결과를 서술한다. 끝으로 4절에서는 본 연구의 결론이 주어진다.

2. 일반화 최소제곱 F 검정

일반적으로 혼합모형은 다음과 같이 서술할 수 있는데, 이 경우

$$\underline{y} = X\underline{\beta} + Z\underline{\gamma} + \underline{e} \tag{2.1}$$

$\underline{y}$는 알려진 $n\times 1$ 자료벡터, X는 고정효과에 관련된 $n\times p$ 계획행렬, Z는 랜덤효과에 관련된 $n\times q$ 계획행렬, $\underline{\beta}$는 고정효과로 불리는 $p\times 1$ 열벡터, $\underline{\gamma}$는 랜덤효과로 불리는 $q\times 1$ 열벡터, 그리고 $\underline{e}$는 $n\times 1$ 오차벡터이다. 또한 랜덤효과가 c개의 그룹으로 분할된다면, 행렬 Z의 분할행렬 $Z=(Z_1 | Z_2 | \cdots | Z_c)$, 벡터 $\underline{\gamma}$의 분할벡터 $\underline{\gamma}' = (\underline{\gamma_1}' | \underline{\gamma_2}' | \cdots | \underline{\gamma_c}')$를 고려하여 식(2.1)을 다음과 같은 식(2.2)로 표현할 수 있는데, 여기서 Z_i의 차수는 $n\times q_i$, $\underline{\gamma_i}$는 $q_i\times 1$ 열벡터이며 q와 q_i의 관계는 $q=\sum_{i=1}^{c} q_i$이다.

$$\underline{y} = X\underline{\beta} + Z_1\underline{\gamma_1} + Z_2\underline{\gamma_2} + \cdots + Z_c\underline{\gamma_c} + \underline{e}. \tag{2.2}$$

그리고 일반적으로 혼합모형에서 가장 많이 사용되는 가정인 $\underline{\gamma_1}, \underline{\gamma_2}, \cdots, \underline{e}$는 정규분포를 따르고 서로 독립이라는 가정을 이용하여 $\underline{\gamma_i} \sim N(\underline{0}, \sigma_i^2 I_{q_i})$, $i=1,2,\ldots c$ 와 $\underline{e} \sim N(\underline{0}, \sigma_e^2 I_n)$가 성립한다고 가정하자. 여기서 I_{q_i}은 $q_i\times q_i$ 항등행렬이며 I_n은 $n\times n$ 항등행렬이다. 이 경우 $\sigma_1^2, \sigma_2^2, \ldots \sigma_c^2, \sigma_e^2$를 우리는 분산성분이라고 부르며 따라서 $\underline{y}$는 평균이 $X\underline{\beta}$이고 분산-공분산 행렬이 $V=\sigma_e^2 I_n + \sum_{i=1}^{c} \sigma_i^2 Z_i Z_i'$인 다변량 정규벡터이다.

만일 식(2.1)에서 벡터 $\underline{y}$의 분산-공분산행렬 V가 $V=\sigma^2 I_n$일때 우리가 관심을 가지고 있는 가설의 모양이 $H_o: K'\underline{\beta} = \underline{m}$, 단 K'은 알려진 $r\times p$ 행렬로 계수는 $r(<p)$이며 $\underline{m}$은 알려진 $r\times 1$ 벡터, 이면 우리가 흔히 사용하는 표준 F-검정통계량은 다음과 같다.

$$F_s = \frac{(K'\underline{b}-\underline{m})'(K'(X'X)^{-}K)^{-1}(K'\underline{b}-\underline{m})/r}{(\underline{y}-X\underline{b})'(\underline{y}-X\underline{b})/(n-p)}. \tag{2.3}$$

식(2.3)에서 사용된 $\underline{b}=(X'X)^{-}X'\underline{y}$는 $\underline{\beta}$의 최소제곱추정량이며 $(X'X)^{-}$는 $(X'X)$의 일반화역행렬이다. 여기서 가설 H_0는 유의수준 α와 기각치 $F(r, n-p;\alpha)$를 사용하여 $F_s >$ $F(r, n-p;\alpha)$이면 기각된다. 실제로 분산-공분산행렬 V가 $V=\sigma^2 I_n$이면, 우도비 검정과 위의 가설검정법은 일치하며 또한 완전모형과 축소모형을 이용한 F-통계량과 일치한다. 하지만 분산-공분산행렬 V가 일반적인 모양이면, $\underline{\beta}$의 최소제곱추정량 $\underline{b}$는 더이상 $\underline{\beta}$의 최소선형불편추정량(BLUE)이 되지 않으며, 예를 들어서 회귀분석에서 이단집락추출을 통하여 얻어진 자료를 집락내상관계수를 무시하고 표준 F-검정통계량을 이용하면 제1종오류가 매우 크게 되는 경우가 발생한다 (Wu, Holt, Holmes; 1993).

만일 식(2.1)에서 벡터 $\underline{y}$의 분산-공분산행렬 V의 구조가 알려져 있으면, $\underline{y}^* = V^{-1/2}\underline{y}$, $X^* = V^{-1/2}X$, $Z^* = V^{-1/2}Z$, $\underline{e}^* = V^{-1/2}\underline{e}$ 와 같은 변환을 통하여 가우스-마르코프 모형인 $\underline{y}^* = X^*\underline{\beta} + Z^*\underline{\gamma} + \underline{e}^*$를 얻게 된다. 그리고 표준 F-검정은 변환된 자료에 적용할 수 있게 되고 가설 $H_o: K'\underline{\beta} = \underline{m}$에 대하여 다음과 같은 일반화 F-검정통계량을 얻게 된다.

$$F_g = \frac{(K'\underline{b}_g - \underline{m})'(K'(X'V^{-1}X)^{-}K)^{-1}(K'\underline{b}_g - \underline{m})/r}{(\underline{y} - X\underline{b}_g)'V^{-1}(\underline{y} - X\underline{b}_g)/(n-p)}. \tag{2.4}$$

단, 식(2.4)에서 $\underline{b}_g = (X'V^{-1}X)^{-}X'V^{-1}\underline{y}$는 $\underline{\beta}$의 일반화 최소제곱추정량이다. 하지만 대부분의 실험자는 V의 구조를 알고 있는 경우는 거의 없고, 따라서 V는 주어진 데이타로 부터 추정되어야 한다. 그러므로 V는 분산성분의 함수이므로 분산성분이 추정되어지면, 보편화된 방법은 추정치의 값을 모수의 값으로 이용하여 고정인자에 관한 일반화 최소제곱분석을 시행하게 되며, 이것을 경험적일반화 최소제곱분석(EGLS)이라고 한다. 여기서 우리는 EGLS를 수행할 때 과연 분산성분추정량을 어떤 것을 이용하는 것이 검정력의 측면에서 가장 우수할까?, 또 그 결과는 실험계획의 모양에 종속될까? 등의 질문을 생각하게 된다. 이 질문에 대한 답을 일반적으로 하는 것은 매우 어려운 일로 생각되며 따라서 부분적인 답을 알아보기 위하여 먼저 분산성분추정량의 종류를 알아보고, 혼합모형에서 가장 간단한 예인 고정인자가 한개인 이원변량모형인 경우에 모의실험을 통하여 알아보기로 하겠다. 이제 널리 사용되는 네가지 추정량들을 간략하게 살펴 보기로 하자. 네가지 추정량에 관한 자세한 설명과 상호관계는 Searle(1979, 1987)과 Rao와 Kleffe(1988)의 책등에서 찾아볼 수 있다.

2.1 핸더슨의 방법 III 추정량

핸더슨의 방법 III은 모든 혼합모형에 적용될 수 있으며 이 방법을 이용하여 구해진 추정량은 모두 불편추정량이다. 이 방법은 비록 고정효과모형의 분석에 널리 사용되어지는 제곱합의 원리를 이용하지만 수학적으로 통계적 최적성질이 밝혀진 것은 아직 없다. 또한 이 방법에 의하여 구하여지는 추정량은 유일하지 않으며 여러가지 추정량중에서 어느 것을 사용하는 것이 가장 좋다는 것도 알려져 있지 않다. 이 방법에 의한 분산성분의 추정값은 음수가 될 수도 있으나 보통 음수가 되면 0으로 간주한다.

2.2 최우추정량 (MLE)

혼합모형의 랜덤효과와 오차항이 정규분포를 따른다는 가정아래에서 분산성분을 최우추정법을 이용하여 구한다. 최우추정 방정식은 각 분산성분의 값에 초기값을 주어서 적당한 수렴판정이 만족될 때까지 반복과정을 계속하여 구할 수 있다. 분산성분의 추정값은 항상 0이상으로 나온다.

2.3 제한적 최우추정량 (REMLE)

제한적 최우추정법 역시 혼합모형의 랜덤효과와 오차항이 정규분포를 따른다는 가정을 이용하기 때문에 최우추정법과 비슷하나, 고정효과 부분과 랜덤효과 부분을 따로 분리하여 추정하는 것이 다르다. 제한적 최우추정량은 최우추정량과 마찬가지로 반복수렴해이며 분산성분이 0 이상이란 조건을 이용하여 구하기 때문에 추정값은 항상 0이상으로 나온다.

2.4 최소노옴 불편추정량 (MINQUE)

최소노옴 불편추정량(MINQUE)은 Rao(1971)에 의해서 처음으로 사용되어졌는데, 그후 여러 학자들에 의해 그 성질들이 자세히 연구되었다. MINQUE는 랜덤효과와 오차항의 분포와 무관하게 사용될 수 있다는 장점이 있으며, MINQUE를 이용하는 경우에는 분산성분비율의 초기값이 필요한데 초기값을 1로 사용하는 경우가 가장 흔한 경우이다. MINQUE와 REMLE는 서로 밀접한 관련성이 있는 데, REMLE의 첫번째 단계의 값이 하나의 MINQUE 추정치가 될 수 있다.

3. 모의실험계획

3.1 모의실험계획

모의실험에서 우리는 고정인자가 한개인 이원변량모형인 경우에 다음과 같은 2가지 종류의 혼합모형을 고려하였다. 2가지 모형은 교호작용이 없는 이원혼합모형, 이원지분혼합모형이며 모두 고정인자가 한개만 포함되어 있다. 또한 2가지 모형은 각 효과들의 수준의 값, 오차항의 분산성분에 대한 랜덤효과에 관련된 분산성분의 비율, 그리고 실험계획의 불균형정도가 고려되어져야 하는 데 이상을 고려하여 다음과 같은 실험계획을 선택하였다.

<u>모형 1: 교호작용이 없는 이원혼합모형</u>

$$y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}.$$

$$i = 1, 2, \cdots, a; \quad j = 1, 2, \cdots, b; \quad k = 1, 2, \cdots, n_{ij};$$

$$\beta_j \sim N(0, \sigma_\beta^2); \quad \varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2);$$

$$\beta_j, \varepsilon_{ijk}: \text{ 서로 독립}$$

$$\sigma_\beta^2 / \sigma_\varepsilon^2 = 0.5, 1, 4; \quad \sigma_\varepsilon^2 = 1.$$

실험계획: $(a, b) = (3, 5)$.

$$A_{35}[\mathrm{I}] = (2,\ 2,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ 4,\ 4,\ 4),$$
$$A_{35}[\mathrm{II}] = (2,\ 2,\ 3,\ 3,\ 4,\ 2,\ 3,\ 3,\ 4,\ 4,\ 3,\ 3,\ 4,\ 3,\ 3),$$
$$A_{35}[\mathrm{III}] = (1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ 15).$$

모형1에서 제안된 실험계획은 수준조합마다의 관측치의 수의 차이를 고려한 불균형정도를 고려하여 $A_{35} = (n_{11}, n_{12}, n_{13}, n_{14}, n_{15}, n_{21}, n_{22}, \cdots, n_{35})$와 같은 표기를 사용하였으며, 따라서 [I]이라고 표시된 계획은 불균형정도가 약하고 [II]로 표시된 계획은 적당한 정도의 불균형정도를 지니고 있으며 [III]으로 표시된 계획은 불균형정도가 심한 실험계획이다. 모형2는 이원지분 혼합모형으로 불균형정도를 나타내기 위하여 b_i와 n_{ij}를 동시에 고려하여야 하는 데 사용된 실험계획은 두 인자를 동시에 고려한 $(b_1, b_2, \cdots, b_i:\ n_{11}, n_{12}, \cdots, n_{ab_a})$로 표기하였다. 따라서 $N_3[\mathrm{I}]$와 $N_3[\mathrm{II}]$는 $N_3[\mathrm{III}]$과 $N_3[\mathrm{IV}]$와 비교하여 b_i에 대하여 불균형정도가 약하며, $N_3[\mathrm{II}]$와 $N_3[\mathrm{IV}]$는 $N_3[\mathrm{I}]$와 $N_3[\mathrm{III}]$에 비하여 n_{ij}에 대하여 불균형정도가 약한 실험계획이다.

<u>모형 2: 이원지분 혼합모형</u>

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{k(ij)}.$$
$$i = 1,\ 2,\ \cdots,\ a\ ;\quad j = 1,\ 2,\ \cdots,\ b_i\ ;\quad k = 1,\ 2,\ \cdots,\ n_{ij}\ ;$$
$$\beta_{j(i)} \sim N(0, \sigma^2_\beta)\ ;\quad \epsilon_{k(ij)} \sim N(0, \sigma^2_\epsilon)\ ;$$
$$\beta_{j(i)}\ ,\ \epsilon_{k(ij)}\ :\ \text{서로 독립}$$
$$\sigma^2_\beta / \sigma^2_\epsilon = 0.5,\ 1,\ 4\ ;\quad \sigma^2_\epsilon = 1.$$

실험계획: $a = 3$.

$$N_3[\mathrm{I}] = (2,\ 3,\ 4 : 1,\ 4,\ 1,\ 3,\ 4,\ 1,\ 3,\ 4,\ 6),$$
$$N_3[\mathrm{II}] = (2,\ 3,\ 4 : 2,\ 3,\ 2,\ 3,\ 3,\ 2,\ 2,\ 3,\ 3),$$
$$N_3[\mathrm{III}] = (1,\ 5,\ 8 : 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8),$$
$$N_3[\mathrm{IV}] = (1,\ 5,\ 8 : 2,\ 3,\ 2,\ 2,\ 3,\ 3,\ 3,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 4).$$

3.2 사용된 가설의 종류

우리는 일반성을 잃지 않고 다음과 같은 스칼라가설과 벡터가설 두가지 가설에 기본을 두고 시뮬레이션을 수행하였는데 고려되어진 실험계획은 모두 고정효과의 수준이 3이므로 검정가능한 (testable) 다음과 같은 가설을 고려하였다.

(가) 스칼라가설: $H_0: 2\alpha_1 - \alpha_2 - \alpha_3 = 0$을 이용하여 $H_1: 2\alpha_1 - \alpha_2 - \alpha_3 = 1$ (검정력1)과 $H_1: 2\alpha_1 - \alpha_2 - \alpha_3 = 2$ (검정력2)인 두가지 경우에 검정력을 구한다.

(나) 벡터가설: $H_0:\begin{cases}\alpha_1-\alpha_3=0\\ \alpha_2-\alpha_3=0\end{cases}$ 을 이용하여 $H_1:\begin{cases}\alpha_1-\alpha_3=1\\ \alpha_2-\alpha_3=1\end{cases}$(검정력1)과

$H_1:\begin{cases}\alpha_1-\alpha_3=2\\ \alpha_2-\alpha_3=0\end{cases}$ (검정력2)인 두가지 경우에 검정력을 구한다.

그리고 가설검정의 경험적 검정력은 다음과 같이 구할 수 있는데,

[단계1] 각 모형에 대하여 두가지 상이한 대립가설을 이용하여 정규난수를 각각 만든다.

[단계2] 분산성분을 4가지 방법을 이용하여 계산한뒤 EGLS 검정통계량을 계산한다.

[단계3] 각각의 검정통계량값을 이용하여 유의수준 $\alpha=0.05$와 $\alpha=0.1$ 아래에서 채택여부를 판가름한다.

[단계4] 모든 시행이 끝난 뒤 경험적 검정력을 총기각수/총반복수로 구한다.

또한 귀무가설이 참인 경우에 귀무가설을 기각하게 되는 제1종오류률도 위에서 고려된 단계와 비슷하게 각 모형에 대하여 구할 수 있다.

3.3 모의실험의 과정

난수생성은 통계팩키지 SAS/IML을 이용하였으며, REML 추정량과 ML 추정량을 구할 경우에 두 추정량 모두 수렴조건을 만족하기 위한 최대반복횟수는 30번으로 제한하였다. 또한 수렴판정의 기준은 일반적으로 주관적인 문제이나, SAS의 VARCOMP 프로시듀어에서 사용하는 k번째와 $k+1$번째 반복이 $|\hat{\theta}_{k+1}-\hat{\theta}_k|<10^{-8}$을 만족할 때, k번째에서 추정량이 수렴하였다고 정의하였다. 또한 MINQUE추정량을 구할 경우에는 사전추측값이 필요한데 가장 많이 사용하는 값 1을 사용하였다. 두가지 모형에 있어서 고려된 여러가지 경우에 대하여 각각 3000개의 자료가 생성되었다. 그리고 정규분포의 난수는 통계팩키지 SAS의 난수생성함수 RANNOR를 이용하여 만들어졌다. 또한 분산성분을 추정하는 데 사용되는 네가지 방법 모두 $n\times n$ 역행렬의 계산을 포함한 많은 계산량을 필요로 하는 데, $n\times n$ 역행렬의 계산을 피하기 위하여 Lee와 Kim(1988), Kim과 Lee(1988)의 W-변환의 재구성 기법을 이용하였다. W-변환의 재구성 기법은 네가지 방법을 이용하는 경우에 요구되어지는 계산량을 훨씬 줄여 줌으로써 컴퓨터 메모리 부족현상을 막는 데 상당한 도움을 준다. 또한 모형1과 모형2에서 제1종오류율을 구할때는 고정효과 α_i의 값은 모두 0을 이용하였고, 검정력1을 구하는 경우에 모형1과 모형2에서 사용된 α_i의 값은 일반성을 잃지 않고 $\alpha_1=0.5, \alpha_2=0.5, \alpha_3=-0.5$를 사용하였다. 그리고 검정력2를 구하는 경우에는 스칼라가설인 경우에는 $\alpha_1=1,\ \alpha_2=1,\ \alpha_3=-1$, 벡터가설에는 $\alpha_1=1, \alpha_2=-1, \alpha_3=-1$을 이용하였다.

3.4 EGLS의 경험적 검정력

제안된 실험계획을 바탕으로 모의실험을 통하여 구한 경험적 검정력을 관측치로 하여 검정력의 변화에 작용하는 인자들을 알아볼 수 있는데, 이 사실을 알아보기 위하여 모형1과 모형2에

대하여 각각 다원배치모형을 고려하여 유의성검정을 하여 보았다. 먼저 모형1에 관한 고려된 인자는 4가지로써 A:가설의 종류, B:자료의 불균형도, C:분산성분의 비율값, D:추정량의 종류를 이용하였는데, 여기서 가설의 종류란 스칼라가설인가 아니면 벡터가설인가를 의미하며 종속변수는 경험적 검정력이 된다. 모형1을 이용하여 모의실험의 결과로 부터 나온 자료를 이용하여 SAS를 활용하여 분산분석을 한 결과 유의수준 5%에서 유의한 인자는 다중비교를 실시하였는데, 집단간의 평균값의 차이를 보기 위하여 매우 보수적인 Scheffe와 정반대인 Duncan방법을 사용한 결과, 분류결과는 모두 같았다.

다음 <표1>은 모형1에 대한 가설의 종류, 분산성분의 비율값, 자료의 불균형도, 추정량의 종류에 대한 다중비교및 경험적 검정력의 평균값이다. <표1>에서 사용된 기호 **와 *, -는 같은 문자로 표기된 집단간에는 차이가 없다는 것을 의미하며 다른 문자로 표기된 집단간에는 차이가 있다는 것을 의미하는 데, **표기는 *표기보다 또 *표기는 -표기보다 숫자값이 유의수준 5%에서 큰 경우를 의미한다.

<표 1> 모형1을 이용한 여러 가지 인자에 대한 경험적 검정력

인자	수준	유의수준 (5%)			유의수준 (10%)		
		1종오류율	검정력1	검정력2	1종오류율	검정력1	검정력2
가설의종류	스칼라	0.0529*	0.1871-	0.5404-	0.1027-	0.2849-	0.6474-
	벡터	0.0533*	0.5935*	0.8857*	0.1059*	0.6895*	0.9262*
불균형도	상	0.0531*	0.5373*	0.7894**	0.1028-	0.6133*	0.8484**
	중	0.0502-	0.3158-	0.7074*	0.1027-	0.4248-	0.7821*
	하	0.0560**	0.3180-	0.6422-	0.1074*	0.4235-	0.7299-
분산성분 비율	0.5	0.0523*	0.4114*	0.8556**	0.1060*	0.5020*	0.9058**
	1	0.0531*	0.7213**	0.7894*	0.1037*	0.7625**	0.8539*
	4	0.0539*	0.2320-	0.4941-	0.1032*	0.3272-	0.6007-
추정량의 종류	헨더슨III	0.0533*	0.3907*	0.7134*	0.1046*	0.4876*	0.7870*
	MINQUE1	0.0533*	0.3908*	0.7133*	0.1046*	0.4875*	0.7870*
	MLE	0.0527*	0.3891*	0.7120*	0.1034*	0.4860*	0.7861*
	REMLE	0.0533*	0.3908*	0.7134*	0.1045*	0.4876*	0.7871*

<표1>을 이용하면 모형1에 대하여 모의실험결과로부터 우리는 다음과 같은 중요한 결론을 얻을 수 있다.

1. 인자가 가설의 종류인 경우에는 두가설에 있어서 제1종오류율인 경우는 비슷하나 나머지 검정력1과 검정력2인 경우에는 벡터가설의 검정력이 훨씬 커짐을 알 수 있는 데, 벡터가설이 스칼라가설보다 α_i들에 대한 조건이 강하기 때문에 검정력이 클 것이라는 직관과 일치한다.
2. 불균형정도는 제1종오류율과 검정력1,2 모두 불균형정도가 가장 약한 경우가 가장 나쁘게 나타났다. 또한 두 검정력의 경우에는 불균형도가 심할 때가 더욱 검정력이 크게 나타났

다.

3. 분산성분의 비율인 경우에는 제1종오류율인 경우에는 유의하지 않은 것으로 나타났으나 두가지 검정력인 경우에는 분산성분의 비가 가장 큰 경우가 가장 검정력이 나쁜 것으로 나타났다.
4. 제1종오류율과 모든 검정력은 4가지 분산성분추정량의 종류에는 영향을 받지 않는 것으로 판명되었다. 그리고 사용된 4가지 방법을 이용한 일반화 최소제곱검정법은 모두 유의수준 α검정법에 가까움을 알수 있다.
5. 모형1에서 고려된 각 경우에 약간의 차이는 있지만 가설의 모양, 분산성분의 비율값, 자료의 불균형정도는 서로 교호작용이 존재하는 경우가 많기 때문에 검정력을 높이기 위한 최적실험계획을 제시하는 일은 매우 어려운 일이라고 할 수 있다.

모형2를 분석하기 위하여 5가지 인자 A:가설의 종류, B: b_i에 대한 불균형정도, C:분산성분의 비율값, D: n_{ij}에 대한 불균형정도, E:추정량의 종류, 종속변수는 각 인자에 대응되는 경험적 검정력을 이용하여 모형1과 비슷하게 다원배치모형을 고려하여 유의성검정을 하여 보았다. <표2>는 모형2에 대한 가설의 종류, 분산성분의 비율값, 자료의 두가지 형태의 불균형정도, 추정량의 종류에 대한 다중비교및 경험적 검정력의 평균값을 보여준다. 이경우 사용된 다중비교법은 모형1과 동일하게 사용하였는데, <표2>에서 유의수준5%에서 추정량의 종류만 다르게 제시될 뿐 모두 같은 결과를 제공하였다. <표2>에서 유의수준5%에서 추정량의 종류인 경우의 다중비교는 Scheffe방법을 이용하였다.

<표 2> 모형2를 이용한 여러 가지 인자에 대한 경험적 검정력

인자	수준	유의수준 (5%)			유의수준 (10%)		
		1종오류율	검정력1	검정력2	1종오류율	검정력 1	검정력2
가설의종류	스칼라	0.0437*	0.0756-	0.1693-	0.0882*	0.1374-	0.2633-
	벡터	0.0430*	0.3137*	0.3946*	0.0881*	0.4252*	0.5060*
불균형도 (b_i)	상	0.0458*	0.2276*	0.1837-	0.0924*	0.3145*	0.2766-
	하	0.0409-	0.1617-	0.3801*	0.0839-	0.2482-	0.4927*
불균형도 (n_{ij})	상	0.0419-	0.2067*	0.2621-	0.0860*	0.2917*	0.3603-
	하	0.0447*	0.1826*	0.3017*	0.0902*	0.2710*	0.4090*
분산성분 비율	0.5	0.0451*	0.2703**	0.3838**	0.0905*	0.3657**	0.4947**
	1	0.0430*	0.2113*	0.3136*	0.0886*	0.3033*	0.4256*
	4	0.0419*	0.1024-	0.1483-	0.0853*	0.1750-	0.2336-
추정량의 종류	헨더슨III	0.0435*	0.1744*	0.2942*	0.0852*	0.2501*	0.3894*
	MINQUE1	0.0406*	0.1933*	0.2676*	0.0851*	0.2820*	0.3725*
	MLE	0.0477*	0.2150*	0.2949*	0.0963**	0.3076*	0.4002*
	REMLE	0.0415*	0.1960*	0.2710*	0.0860*	0.2855*	0.3765*

다음은 <표2>를 이용한 모형2에 대한 모의실험에 근거한 중요한 결론들이다.

1. 가설의 종류에 따른 검정력의 비교는 두가설에 있어서 제1종오류율인 경우는 비슷하나 나머지 검정력1과 검정력2인 경우에는 모형1과 마찬가지로 벡터가설의 검정력이 훨씬 커짐을 알 수 있다.
2. 불균형정도에서는 b_i에 대한 불균형정도는 각 경우에 따라 효율성의 차이가 있다고 할 수 있으나 n_{ij}에 대한 불균형정도는 대체적으로 불균형정도가 낮을수록 검정력의 측면에서 효율성이 높아진다는 사실을 알 수 있다.
3. 분산성분의 비율인 경우에는 모형1과 마찬가지로 제1종오류율인 경우에는 유의하지 않고 두가지 검정력인 경우에는 분산성분의 비가 가장 큰 경우가 가장 검정력이 나쁜 것으로 나타났다.
4. 분산성분추정량의 선택에 대하여 제1종오류율과 모든 검정력은 검정력2인 경우에 제1종오류율만 MLE를 사용하는 것이 유의하게 효율성이 높은 것으로 판명되었으나 다른 경우에는 4가지 분산성분추정량의 종류에 영향을 받지 않는 것으로 판명되었다. 하지만 4가지 분산성분추정량을 이용한 모형1의 제1종오류율보다 대체적으로 나쁘게 나타났으며 모형1은 4가지 추정량 모두 제1종오류율을 유의수준보다 크게 추정하였으나 모형2인 지분실험계획법에서는 제1종오류율을 유의수준보다 작게 추정한다. 한편 모형2에서는 통계적으로 유의한 정도는 아니지만 다른 추정량에 비하여 MLE를 사용한 경우의 효율성이 높다는 사실을 알 수 있다.
5. 모형1과 마찬가지로 경우에 따라서 약간의 차이는 있지만 가설의 모양, 분산성분의 비율값, 자료의 불균형정도는 서로 교호작용이 존재하는 경우가 많기 때문에 검정력을 높이기 위한 최적실험계획을 계획하는 일은 역시 매우 힘들것으로 간주된다.

한편 모형1에 교호작용을 첨가한 모형도 고려하여 보았으나 결과는 모형1에 대한 결과와 거의 유사한 형태로 나타났다. 또한 고려된 모형의 모수를 바꾸어서 모의실험을 하여 보았으나 주관심사인 분산성분추정량의 검정력에 미치는 영향은 거의 비슷하다는 결론을 내릴 수 있었다.

4. 결론

다음은 3절에서 언급된 본 연구의 결론이다. 하지만 이 결론은 제한된 두가지 이원혼합모형에 국한된 모의실험의 결과에서 성립하는 사항이지 수학적으로 모든 혼합모형에 대하여 항상 성립한다는 것은 아니며, 여기에 대한 보다 정확한 답을 하기 위하여 일반적인 혼합모형에 대한 검정력분석을 위한 보다 폭넓은 연구가 진행되어져야 하겠다.

1. 검정력에 영향을 미치는 인자는 가설의 종류, 자료의 불균형도, 분산성분의 비율이 있으며, 벡터가설을 활용하는 것이 두가지 모형 모두 검정력이 커진다. 하지만 자료의 불균형정도와 분산성분의 비율값이 검정력에 미치는 영향은 일반적으로 다른 인자들과의 교호작용이 존재하기 때문에 매우 복잡한 형태로 나타나는 것으로 판명되었다.

2. 가장 큰 관심사였던 어떠한 분산성분추정량을 사용하는 것이 검정력의 측면에서 효율적인가 하는 질문은 두가지 모형 모두 통계적으로 유의할 만큼 분산성분추정량의 영향을 받지 않는 것으로 판명되었으나, 지분실험계획법인 경우에는 최우추정량을 사용하는 것이 가장 효율적인 것으로 나타났다. 그리고 사용된 4가지 추정량을 이용한 검정법은 이원요인실험계획법에서는 대체적으로 유의수준 α검정법에 가까운 것으로 판명되었으나 이원지분실험계획법에서는 유의수준보다 다소 검정력이 낮게 나타났다.

참 고 문 헌

[1] Lee, J. T. (1993). 일원변량모형에서의 임의의 분포에 대한 MINQE 추정량의 효율성, 「응용통계연구」, 6권 2호, 355-370.

[2] Conerly, M. D. and Webster, J. T. (1987). MINQE for the One-Way Classification, *Technometrics,* Vol.29(2), 229-236.

[3] Kim, B. C. and Lee, J. T. (1988). An Algorithm for Computing MINQUE in the Mixed Model, *American Statistical Association 1989 Proceeding of the Statistical Computing Section,* 165-169.

[4] Lee, J. T. and Kim, B. C. (1988). A New Approach for the W-matrix, *Journal of Statistical Computation and Simulation,* Vol. 29, 241-254.

[5] Rao, C. R. (1971). Estimation of Variance Components - MINQUE Theory, *Journal of Multivariate Analysis,* Vol. 1, 257-275.

[6] Rao, C. R. and Kleffe, J. (1988). *Estimation of Variance Components and Applications,* Amsterdam: North-Holland.

[7] Rao, J. N. K., Sutrahar, B. C., and Kim, Y. (1993). Generalized Least Squares F test in Regression Analysis With Two-Stage Cluster Samples, *Journal of the American Statistical Association,* Vol. 88, 1388-1391.

[8] Searle, S. R. (1987). *Linear Models for Unbalanced Data,* John Wiley & Sons, New York.

[9] Searle, S. R. (1979). Notes on Variance Component Estimation: A Detailed Account of Maximum Likelihood and Kindred Methodology, *Report Number BU-673M, Biometric Unit.* Ithaca, New York: Cornell University Press.

[10] Swallow, W. H., and Monahan, J. F. (1984). Monte Carlo Comparison of ANOVA, MINQUE, REML, and ML Estimators of Variance Components, *Technometrics,* Vol. 26, 47 - 57.

[11] Wu, C. F. J., Holt, D., and Holmes, D. J. (1988). The Effect of Two-Stage Sampling on the F Statistics, *Journal of the American Statistical Association,* Vol. 83, 150-159.

Power Analysis of Testing Fixed Effects with Two Way Classification

Jang Taek Lee[1)]

Abstract

This article considers the power performance of the tests in unbalanced two way mixed linear models with one fixed factor. The generalized least squares (GLS) F statistic testing no differences among the effects of the levels of the fixed factor is estimated using Henderson's method III, minimum norm quadratic unbiased estimator (MINQUE) with prior guess 1, maximum likelihood (ML) and resticted maximum likelihood (REML). We investigate the power performance of these test statistics. It can be shown, through simulation, that the GLS F statistics using four estimators produce similar type I error rates and power performance.

1) (140-714) Associate Professor, Department of Computer Science and Statistics, Dankook University, Seoul, Korea.