응용통계연구 제 6권 2호
1993년 9월 pp.355-370.

일원변량모형에서의 임의의 분포에 대한 MINQE 추정량의 효율성[1)]

이 장 택[2)]

요 약

일원변량모형에서 처리효과와 오차항의 분포가 임의의 분포를 따를때, MINQE 추정량의 효율성을 MSE와 편의 판정기준아래에서 ANOVA, MIVQUE, REML, ML 추정량과 비교하여 알아 본다. 그리고 역행렬의 계산이 필요하지 않는 간편한 MIVQUE 추정량, REML 추정량과 ML 추정량의 계산 방법이 주어진다. 또한 MINQE의 효율성을 높일 수 있는 사전추측값을 제공하는 방법이 제안된다. 결론적으로 MINQE 추정량은 비록 편의는 다른 추정량들에 비해서 크지만, 모든 분포에 대해서 MSE의 판정기준아래에서 가장 효율성이 높으며, 변량인자들이 따르는 분포의 첨도가 클수록 그 효율성은 증가한다.

1. 서 론

분산성분의 추정에 관한 토픽은 선형모형에 있어서 중요한 비중을 차지하고 있으며, 생물학, 산업공학, 의학등의 여러 분야에서 그 응용도를 찾아 볼 수 있다. 분산성분추정에 대한 이론및 그 응용에 관한 상세한 내용과 설명은 Searle(1971, 1987)과 Rao와 Kleffe(1988)의 책등에서 찾아 볼 수 있다.

일반적으로, 아마 가장 많이 사용되어지는 분산성분의 추정량은 분산분석표에서 얻어지는 ANOVA 추정량일 것이다. 만약 자료들이 균형적(각 셀당 관측치가 모두 같은 경우)이면, ANOVA 추정량은 최소분산 선형 불편추정량(BLUE)이 된다. 하지만 자료들이 불균형적(각 셀당 관측치가 모두 같지는 않을 경우)이면, ANOVA 추정량은 불편성을 제외한 다른 좋은 통계적 성질을 갖지 못한다. 따라서, 실질적으로 불균형 자료들이 여러 종류의 실험에서 빈번하게 발생하게 되므로, 불균형 자료들인 경우에 분산성분의 추정에 관련된 추정량에 관한 연구가 지속적으로 진행되었는데, 1960년대 말을 기점으로 크게 두 가지 분류의 추정량들에 관한 연구의 결과가 발표되기 시작되었고, 또 그 추정량들의 효율성을 인정받게 되었다. 첫째 분류가 변량인자들의 분포가 정규성을 가지고 있다는 가정이 요구되어지는 최우추정량(MLE)과 제한적 최우추정량(REMLE)이고, 다른 둘째 분류는 최소노옴 이차 불편추정량(MINQUE)이다. 이들 추정량에 관한 상세한 설명과 관계는 앞에서 언급한 바 있는 Searle(1971, 1987)과 Rao와 Kleffe(1988)의 책등에서 자세히 찾아 볼 수 있다.

이 논문에서는 불균형 일원변량모형에서의 분산성분 추정에 관한 문제를 다룰려고 한다. 불균형 일원변량모형에서의 여러 가지 분산성분 추정량의 효율성에 대한 비교에 관련된 훌륭한 논문들은 Swallow와 Monahan(1984)의 논문과 Conerly와 Webster (1987)의 논문(이 두가지 논문은 앞으로 계속 언급이 되므로 SM과 CW로 언급을 하기로 하겠다)을 들 수 있는 데, SM은 불균형 일원변량모형의 ANOVA, MIVQUE(A), MIVQUE(0), REML 그리고 ML 추정량을 평

1) 본 연구는 1992년 한국과학재단 연구지원에 의한 결과임. (KOSEF 921-0100-008-1)
2) (140-714) 서울시 용산구 한남동 산 8번지 단국대학교 전산통계학과

균자승오차(MSE)와 편의(BIAS)의 두가지 판정기준 아래에서 컴퓨터 모의실험을 이용하여, 비교 분석 하였다. CW는 불균형 일원변량모형인 경우에, 변량인자들이 정규분포를 따른다는 가정아래에서 SM이 사용하였던 불편추정량들보다 MINQE 추정량이 MSE 판정아래에서 $\sigma_a^2/\sigma_e^2 > 0.5$인 경우에 우수하다고 컴퓨터 모의실험을 이용하여 밝혔다. Westfall (1987) 은 불균형 일원 변량모형에 있어서 처리효과와 오차항의 분포가 임의의 분포를 따른다는 가정아래에서 사전추측값을 0, 1, ∞를 사용한 MINQUE, REML, ML 추정량의 대표본 분산행렬을 구하여 5가지 추정량의 효율성을 비교하였으며, 추정량들의 극한 분산은 실험계획의 모양, 분산성분들의 실제값, 그리고 모수들의 분포의 첨도에 의존한다고 밝혔다.

본 논문의 목적은 처리효과와 오차항이 임의의 분포를 따르는 경우에 MSE와 편의 판정기준 아래에서 MINQE를 포함한 6가지 추정량들의 효율성을 비교하고, Westfall (1987) 에서 고려한 추정량들보다 MINQE가 MSE의 판정 기준아래에서 효율성이 높다는 것을 밝히려고 한다. 2절에서는 일원변량모형의 일반적인 모양과 SM과 CW에서 고려되어진 여섯가지 추정량들에 대한 간략한 소개및 일원변량모형에서 역행렬의 계산이 필요하지 않는 MIVQUE, REML, 그리고 ML 추정량을 구하는 방법을 언급한다. 3절에서는 2절에서 고려된 σ_a^2와 σ_e^2의 6가지 추정량의 효율성을 모의실험을 통하여 알아보며, 4절에서는 MINQE의 효율성을 높일 수 있는 사전추측값을 간단하게 구하는 방법을 소개한다. 끝으로 5절에서는 결론과 향후과제들을 제시한다.

2. 일원변량모형과 여러가지 추정량들

반복수가 같지 않은 i번째 처리효과에 있어서 j번째 관측치에 대한 일원변량모형은 다음과 같이 나타낼 수 있다.

$$y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, \cdots, k, \quad k \ge 2, \quad j = 1, \cdots, n_i, \quad N = \sum_i^k n_i. \tag{2.1}$$

여기서 μ는 미지의 모수, a_i와 e_{ij}는 서로 독립이며, 평균이 0이고 분산이 각각 σ_a^2와 σ_e^2인 확률변수, n_i는 i번째 처리효과에 있어서의 관측치의 갯수, 그리고 N은 총 관측치의 갯수를 의미한다. 모형 (2.1)은 행렬기호로 쓰면 다음과 같이 쓸 수 있는 데,

$$y = 1_N\mu + Z_1 a + Z_2 e,$$

여기서 $a' = (a_1, \ldots, a_k)$, $e' = (e_{11}, e_{12}, \ldots, e_{kn_k})$, 1_N은 모든 원소가 1인 N-벡터, $Z_1 = \Sigma^+ 1_{n_i}$는 $N \times k$ 행렬, $Z_2 = I_N$는 $N \times N$ 항등행렬, 그리고 Σ^+는 행렬들의 직합(direct sum)을 의미한다. 따라서 y는 평균이 $1_N\mu$이고 분산행렬이

$$V = \sigma_e^2 I_N + \sigma_a^2 Z_1 Z_1'. \tag{2.2}$$

인 다변량 확률벡터이다. σ_a^2와 σ_e^2가 (2.2)식에서 각각의 추정치 $\hat{\sigma}_a^2$와 $\hat{\sigma}_e^2$로 바뀌면, 다음 식 (2.3)을 얻게 된다.

$$\tilde{V} = \hat{\sigma}_e^2 I_N + \hat{\sigma}_a^2 Z_1 Z_1'. \tag{2.3}$$

그리고, 행렬 $\tilde{P}$ S, Q,와 벡터 u를 다음과 같이 정의하자.

$$\tilde{P} = \tilde{V}^{-1}(I - 1_N (1'_N \tilde{V}^{-1} 1_N)^{-1} 1'_N \tilde{V}^{-1}), \tag{2.4}$$

$$S = \{s_{ij}\} = \mathrm{tr}(\tilde{P} Z_i Z_i' \tilde{P} Z_j Z_j'), \quad i, j = 1, 2, \tag{2.5}$$

$$Q = \{q_{ij}\} = \mathrm{tr}(\tilde{V}^{-1} Z_i Z_i' \tilde{V}^{-1} Z_j Z_j'), \quad i, j = 1, 2, \tag{2.6}$$

$$u = \{u_i\} = y \tilde{P} Z_i Z_i' \tilde{P} y, \quad i = 1, 2. \tag{2.7}$$

아울러 계산상의 편리함을 위해 우리는 다음과 같은 기호들을 앞으로 사용하기로 한다. 식 (2.8)에서 사용된 r은 σ_a^2/σ_e^2의 사전추측값이다.

$$m_i = n_i/(1+n_i r), \quad m = 1/\sum_{i=1}^{k} m_i. \tag{2.8}$$

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i, \quad \bar{y}_{..} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}/N, \quad \tilde{y}_{..} = (\sum_{i=1}^{k} m_i \bar{y}_{i.})/(\sum_{i=1}^{k} m_i). \tag{2.9}$$

이제 널리 사용되는 여섯가지 추정량들을 살펴 보기로 하자.

2.1 분산분석 추정량 (ANOVA Estimator)

분산분석 추정량들은 분산분석표에서 제곱평균의 기대값을 이용하여 구한 추정량들이다. 분산분석 추정량들은 다음과 같이 주어지는 데,

$$\hat{\sigma}_e^2 = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2}{(N-k)}, \quad \hat{\sigma}_a^2 = \frac{[\sum_i n_i (\bar{y}_{i.} - \bar{y}_{..})^2 - (k-1)\hat{\sigma}_e^2]}{(N - \sum_i n_i^2/N)}, \tag{2.10}$$

여기서 $\bar{y}_{i.}$와 $\bar{y}_{..}$는 각각 i번째 처리효과의 평균과 전체 평균을 의미한다.

2.2 제한적 최우추정량 (REMLE)

REML 추정량 $\hat{\sigma}^2(r) = (\hat{\sigma}_a^2, \hat{\sigma}_e^2)'$은 모형 (2.1)에서 처리효과와 오차항이 정규분포를 따른다는 가정아래에서 (2.11)식을 이용하여 구한 반복수렴해이다.

$$\hat{\sigma}^2(r) = S^{-1} u. \tag{2.11}$$

(2.11)식에 나타나는 행렬과 벡터표현의 원소들을 간단히 계산하기 위해서 다음과 같은 표기들을 사용하기로 한다. 이 표기들은 Swallow과 Searle(1978)의 논문에 있는 표기들을 약간 수정하여 사용한 것들이다.

$$\begin{aligned}
s_{11}\sigma_e^4 &= \sum m_i^2 - 2m\sum m_i^3 + m^2(\sum m_i^2)^2, \\
s_{12}\sigma_e^4 &= \sum(m_i^2/n_i) - 2m\sum(m_i^3/n_i) + m^2\sum m_i^2 \sum(m_i^2/n_i), \\
s_{22}\sigma_e^4 &= N - k + \sum(m_i^2/n_i^2) - 2m\sum(m_i^3/n_i^2) + m^2(\sum m_i^2/n_i)^2 \\
u_1\sigma_e^4 &= \sum m_i^2(\bar{y}_{i.} - m\sum m_i \bar{y}_{i.})^2, \\
u_2\sigma_e^4 &= (\sum\sum y_{ij}^2 - \sum n_i \bar{y}_{i.}^2) + \sum(m_i^2/n_i)(\bar{y}_{i.} - m\sum m_i \bar{y}_{i.})^2.
\end{aligned}$$

계산상의 편리를 위해서 $i,j = 1, 2$에 대하여 $s_{ij}^* = s_{ij}\sigma_e^4$와 $u_i^* = u_i\sigma_e^4$를 정의하자. 그러면, σ_a^2와 σ_e^2의 REML 추정량의 첫번째 단계의 추정량은 식 (2.11)으로부터 각각 다음 식

(2.12)과 (2.13)와 같이 구하여 진다.

$$\hat{\sigma}_a^2 = (s_{22}^* u_1^* - s_{12}^* u_2^*)/(s_{11}^* s_{22}^* - s_{12}^{*2}). \tag{2.12}$$

$$\hat{\sigma}_e^2 = (s_{11}^* u_2^* - s_{12}^* u_1^*)/(s_{11}^* s_{22}^* - s_{12}^{*2}). \tag{2.13}$$

그러므로, 각 σ_a^2/σ_e^2의 사전추측값 r은 그 전단계의 $\hat{\sigma}_a^2/\hat{\sigma}_e^2$를 사용하며, 우리는 적당한 수렴판정이 만족될 때까지 반복과정을 계속한다.

2.3 최우추정량 (MLE)

ML 추정량 $\hat{\sigma}^2(m) = (\hat{\sigma}_a^2, \hat{\sigma}_e^2)'$은 REML 추정량과 마찬가지로, 모형 (2.1)에서 처리효과와 오차항이 정규분포를 따른다는 가정아래에서 (2.14)식을 이용하여 구한 반복수렴해이다.

$$\hat{\sigma^2}(m) = Q^{-1} u. \tag{2.14}$$

각각의 i, j = 1, 2에 대하여 q_{ij}^*, q_{ij}, u_i^*, 그리고 u_i사이의 관계를 다음과 같이 두면,

$$q_{ij}^* = q_{ij}\sigma_e^4, \quad u_i^* = u_i \sigma_e^4.$$

ML 추정량을 구하기 위해 필요한 식(2.14)의 행렬과 백터에 포함된 원소들은 다음과 같은 식으로 표시된다.

$$q_{11}^* = \sum m_i^2, \quad q_{12}^* = \sum(m_i^2/n_i), \quad q_{22}^* = N - 2r\sum m_i + r^2 \sum m_i^2,$$

$$u_1^* = \sum m_i^2(\bar{y}_{i.} - m\sum m_i \bar{y}_{i.})^2,$$

$$u_2^* = (\sum\sum y_{ij}^2 - \sum n_i \bar{y}_{i.}^2) + \sum(m_i^2/n_i)(\bar{y}_{i.} - m\sum m_i \bar{y}_{i.})^2.$$

식 (2.14)으로 부터 σ_a^2의 ML 추정량의 첫번째 단계의 추정량은

$$\hat{\sigma}_a^2 = (q_{22}^* u_1^* - q_{12}^* u_2^*)/(q_{11}^* q_{22}^* - q_{12}^{*2}) \tag{2.15}$$

이며, σ_e^2의 ML 추정량의 첫번째 단계의 추정량은

$$\hat{\sigma}_e^2 = (q_{11}^* u_2^* - q_{12}^* u_1^*)/(q_{11}^* q_{22}^* - q_{12}^{*2}) \tag{2.16}$$

이다. 윗 식들을 이용하여 적당한 수렴판정이 만족될 때까지 반복과정을 계속하면 우리는 각 분산성분들의 최우추정량들을 구할 수 있다.

2.4 최소분산 이차 불편추정량 (MIVQUE)

최소분산 이차 불편추정량(MIVQUE)은 Rao(1971)에 의해서 처음으로 사용되어졌는데, 그후 Swallow와 Searle(1978)에 의해 그 성질들이 자세히 연구되었다. MIVQUE는 그 응용면에서 두가지가 일반적으로 고려되어진다. 그 한가지가 식 (2.12)과 식 (2.13)을 반복없이 사용하고, ANOVA 분산분석 추정치 $\hat{\sigma}_a^2$와 $\hat{\sigma}_e^2$를 이용하여 σ_a^2/σ_e^2의 사전추측값으로 $r = \hat{\sigma}_a^2/\hat{\sigma}_e^2$을 사용하는 MIVQUE(A)인데, MIVQUE(A)는 REML 추정량들을 구하는 첫번째 단계의 값이다.

다른 한가지는 SAS의 프로시듀어 VARCOMP의 디폴트 추정량으로서 유명한 MIVQUE(0)인데, σ_a^2와 σ_e^2의 사전추측값으로써, $\sigma_a^2=0$, $\sigma_e^2=1$을 사용하고, 식 (2.12)와 (2.13)을 반복없이 이용하여 구할 수 있다. 이 추정량들을 구할 때, 계산속도가 빠르고 아주 큰 데이타를 다룰 수

있다는 장점은 있지만, MIVQUE(0)는 변량인자들이 정규분포를 따른다는 가정아래에서 σ_a^2 를 추정하는데 그 정밀도가 떨어지며, σ_e^2 를 추정하는 데는 비록 자료의 불균형도가 심하지 않더라도 그 정밀도가 매우 나빠진다(Swallow and Monahan, 1984).

2.5 최소노움 이차 추정량 (MINQE)

MINQE는 최소노움 이차 추정량(Minimum Norm Quadratic Estimator)을 의미하는데, MINQUE 이론에서 불편성을 제거한 추정량이다. σ_a^2와 σ_e^2의 MINQE의 일반적인 모양은 Ahrens et al.(1981)에 의해서 다음과 같이 주어지며, r은 σ_a^2/σ_e^2의 사전추측값을 의미한다.

$$\hat{\sigma}_a^2 = r^2 \sum_{i=1}^{k} m_i^2 \, (\bar{y}_{i.} - \tilde{y}_{..})^2 \,/\, k. \tag{2.17}$$

$$\hat{\sigma}_e^2 = [\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^{k} (m_i^2/n_i)(\bar{y}_{i.} - \tilde{y}_{..})^2] \,/\, N. \tag{2.18}$$

모수의 추정에서 MSE를 줄이는 것이 가장 중요한 관점이라면, MINQE의 효율성은 정규성의 가정아래에서 다른 추정량들에 비하여 월등히 뛰어 나다. CW에서도 알 수 있듯이, $\sigma_a^2/\sigma_e^2 > 1$이면 MINQE의 MSE는 MLE의 MSE보다 작으며, 또한 $\sigma_a^2/\sigma_e^2 > 0.5$일때 MINQE의 MSE는 다른 어떤 불편추정량보다 뛰어나다.

3. 모의실험의 결과분석

불균형 일원변량모형에 대한 분산성분 추정의 결과분석은 다양한 불균형 자료의 형태를 고려해야 함으로 굉장히 복잡한 양상을 가지게 되나, 다행스럽게도 N-형식만을 고려하면 충분한 것으로 알려져 있다. N-형식은 Swallow와 Searle(1978)의 논문에서 처음으로 사용되었으며, 이후 대부분의 일원변량 불균형 실험계획의 분산성분추정에 관한 연구논문은 N-형식을 사용하였다. 이 연구에서도 N-형식을 사용할 것이며, N-형식에 관한 일반적인 모양은 표 2에 제시되어 있다.

여러 가지 추정량들을 계산할 때, REML 추정량과 ML 추정량을 구할 때는, 식 (2.3) 에 있어서 $\hat{\sigma}_a^2$와 $\hat{\sigma}_e^2$의 시초값은 모두 분산분석 추정량을 사용하였으며, 두 추정량 모두 수렴조건을 만족하기 위한 최대 반복횟수는 50번으로 계산하였다. 또한 수렴판정의 기준은 일반적으로 주관적인 문제이나, SAS에서 기본적으로 사용하는 k번째와 $k+1$번째 반복이 다음 조건을 만족할 때, k번째에서 추정량이 수렴되었다고 정의하였다.

$$| \hat{\sigma}^2_{k+1} - \hat{\sigma}^2_{k} | < 10^{-8}.$$

그리고, MINQE의 사전추측값은 가장 많이 사용되는 값 1을 사용하였으며, 4절의 결과를 이용하여 사전추측값을 설정하면 효율성은 더욱 증가한다.

각각의 N-형식과 급내상관계수 $\rho = \sigma_a^2/(\sigma_a^2+\sigma_e^2)$들의 값에 대하여, 1000개에서 3000개의 자료가 만들어졌다. Monte Carlo 연구에는 균등, 정규, 이중지수분포 그리고 오염정규분포가 사용되었는데, 이때 오염정규분포(첨도 9)의 확률밀도함수는 다음과 같이 주어 진다.

$$f_{X}(x) = \frac{w_1}{\sqrt{2\pi}\sigma_1}\exp(-\frac{x^2}{2\sigma_1^2}) + \frac{w_2}{\sqrt{2\pi}\sigma_2}\exp(-\frac{x^2}{2\sigma_2^2}),$$

여기서 $w_1=0.1$, $w_2=0.9$, $\sigma_1^2=(1+3\sqrt{3})\sigma^2$, $\sigma_2^2=(1-\sqrt{3}/3)\sigma^2$, 그리고 σ^2는 오염정규확률변수의 분산이다. 따라서 일원변량모형에는 분산성분이 2가지가 있으므로, 모두 16가지 분포의 조합이 고려되었다.

난수생성은 통계팩키지 SAS의 난수들을 이용하였는데, 균등분포 난수는 RANUNI를, 정규분포와 오염정규분포 난수는 RANNOR를, 이중지수분포 난수는 역분포함수와 RANUNI를 사용하여 생성되었고, 기술통계량도 SAS를 이용하여 구하였다. 종합적으로, 16가지 분포의 조합에 대해 N-형식이 모두 10가지가 사용되고, 또 사용되어진 급내상관계수 ρ는 0.1부터 0.9까지 0.1의 간격으로 계산되었다. 또한, 일반성을 잃지 않고 일원변량모형에서 두가지 가정 $\sigma_a^2+\sigma_e^2=1$과 $\mu=0$를 사용하였다.

다음 표 1은 16가지 분포의 조합에 있어서 σ_a^2에 대한 각 추정량들의 추정된 MSE의 평균및 표준편차이다. 표 1에서 K_a는 처리효과의 첨도, K_e는 오차항의 첨도를 의미하며, 첨도가 -1.2는 균등분포, 0는 정규분포, 3은 이중지수분포, 그리고 9는 오염정규분포를 의미한다. 또한 효율성은 각 추정량의 추정된 MSE를 추정된 MIVQUE(A)의 MSE로 나눈 것을 의미하며, 표 1, 표 2, 표 4, 그리고 표 5에서 언급되는 효율성의 우수성은 평균적으로 우수하다는 뜻으로 모든 가능한 조합의 경우에 항상 우수하다는 말은 아니다.

3.1 $\hat{\sigma}_a^2$의 MSE와 편의들

표 1로 부터 이끌어 낼 수 있는 중요한 결론들은 다음과 같다.

1. 16가지 분포의 조합에 대해서 ANOVA, REML 추정량, 그리고 MIVQUE(A)는 MSE 판정기준 아래에서 서로 비슷한 효율성을 가지고 있다. ANOVA 추정량은 K_a가 감소하고 K_e가 증가할수록 효율성은 증가하며, 고정된 K_e의 값에 대해서는 K_a가 작아질수록 효율성은 커진다. REML 추정량은 K_a와 K_e가 감소할수록 효율성은 높아지나 그 정도는 심하지 않다. 반면 MIVQUE(A)는 REML 추정량보다 대체로 약간 우수하다. MIVQUE(0)는 SM의 논문에서 알려진 바 대로 효율성은 많이 떨어진다.
2. ML은 위에서 언급된 4가지 추정량보다는 우수하나, MINQE는 모든 경우에 있어서 단연 효율성이 뛰어나다. MINQE의 효율성은 K_a가 커질수록 증가한다. 그리고 고정된 K_a에 대해서 K_e가 커질수록 MINQE 추정량의 효율성은 증가한다.
3. REML과 ML 추정량은 비록 정규성의 가정하에 구해진 추정량이지만, 효율성은 변량인자들의 분포의 종류에 큰 영향을 받지 않는다.

표 2에서는 σ_a^2에 대한 각 추정량들의 효율성과 N-형식과의 관련성을 나타낸다. 여기서 나오는 평균과 표준편차는 데이타 144개를 이용하여 구하였다. 다음은 표 2로부터 이끌어 낼 수 있는 중요한 결론들이다.

1. ML 추정량과 MINQE 추정량의 효율성은 처리효과의 수준의 수가 증가할수록 효율성은 떨어진다. 또한 ML 추정량은 전반적으로 실험계획의 불균형성이 심할수록 효율성은 떨어지나, MINQE 추정량은 그 반대이다.

2. ANOVA와 REML 추정량의 효율성은 처리효과의 수준의 수가 증가할수록 그리고 실험계획의 불균형성이 심할수록 효율성은 떨어지나, 그 차이는 심하지 않다. MIVQUE(0)는 모든 디자인에 대해서 효율성이 떨어지며, 처리효과의 수준의 수와 디자인의 불균형성이 커질수록 효율성이 떨어지나, 불균형성에 더욱 민감하다.

표 1. σ_a^2의 추정량들의 MSE에 대한 효율성

Ka	*Ke*	추정량	평균	표준편차	*Ka*	*Ke*	추정량	평균	표준편차
-1.2	-1.2	ANOVA	1.254	0.390	3	-1.2	ANOVA	1.326	0.436
		REML	1.070	0.087			REML	1.140	0.380
		MIVQUE(0)	1.827	1.059			MIVQUE(0)	1.823	0.949
		ML	0.724	0.154			ML	0.686	0.195
		MINQE	0.651	0.424			MINQE	0.399	0.146
-1.2	0	ANOVA	1.197	0.425	3	0	ANOVA	1.313	0.412
		REML	1.088	0.123			REML	1.152	0.325
		MIVQUE(0)	1.727	1.095			MIVQUE(0)	1.804	0.893
		ML	0.728	0.154			ML	0.693	0.186
		MINQE	0.645	0.440			MINQE	0.407	0.142
-1.2	3	ANOVA	1.078	0.451	3	3	ANOVA	1.180	0.437
		REML	1.121	0.161			REML	1.208	0.389
		MIVQUE(0)	1.518	1.046			MIVQUE(0)	1.575	0.882
		ML	0.731	0.160			ML	0.702	0.200
		MINQE	0.604	0.426			MINQE	0.376	0.142
-1.2	9	ANOVA	0.936	0.469	3	9	ANOVA	1.084	0.478
		REML	1.152	0.183			REML	1.222	0.397
		MIVQUE(0)	1.285	0.997			MIVQUE(0)	1.427	0.888
		ML	0.730	0.179			ML	0.701	0.206
		MINQE	0.554	0.420			MINQE	0.363	0.159
0	-1.2	ANOVA	1.280	0.401	9	-1.2	ANOVA	1.380	0.512
		REML	1.091	0.152			REML	1.212	0.510
		MIVQUE(0)	1.808	0.978			MIVQUE(0)	1.876	1.041
		ML	0.700	0.163			ML	0.689	0.213
		MINQE	0.516	0.249			MINQE	0.335	0.100
0	0	ANOVA	1.223	0.424	9	0	ANOVA	1.361	0.521
		REML	1.110	0.178			REML	1.224	0.529
		MIVQUE(0)	1.710	1.000			MIVQUE(0)	1.848	1.041
		ML	0.701	0.162			ML	0.690	0.214
		MINQE	0.504	0.252			MINQE	0.330	0.096
0	3	ANOVA	1.121	0.443	9	3	ANOVA	1.253	0.495
		REML	1.136	0.200			REML	1.266	0.514
		MIVQUE(0)	1.535	0.961			MIVQUE(0)	1.668	0.968
		ML	0.707	0.168			ML	0.697	0.216
		MINQE	0.483	0.256			MINQE	0.324	0.107
0	9	ANOVA	0.987	0.445	9	9	ANOVA	1.130	0.523
		REML	1.158	0.206			REML	1.227	0.498
		MIVQUE(0)	1.315	0.886			MIVQUE(0)	1.459	0.959
		ML	0.713	0.184			ML	0.699	0.228
		MINQE	0.452	0.266			MINQE	0.302	0.122

표 3은 (K_a, K_e)=(3,3)인 경우에 각 추정량들의 효율성을 나타내고 있다. 분산성분 σ_a^2에 대한 MINQE의 효율성은 놀랍게도 실험계획 P8에서 ρ=0.1인 경우를 제외하고, 모든 경우에서 가장 우수하다. 따라서, K_a의 값이 증가하면 증가할수록 MINQE는 다른 추정량들보다 MSE 관점에서 효율성이 뛰어나다. 그러나, 다른 추정량들의 효율성은 SM과 CW의 결과들과 비슷하다. MINQE는 ML 추정량과 함께 편의가 큰 추정량이다. 이러한 경향은 K_a의 값이 증가하더라도 뚜렷하게 개선되지 않는다. 편의에 관한 결과들은 SM과 CW의 결과에 유사하므로 생략하였다.

표 2. N-형식에 따른 σ_a^2의 추정량들의 MSE에 대한 효율성

N-형식	추정량	평 균	표준편차
P1=(3,5,7)	ANOVA	1.026	0.059
	REML	1.011	0.016
	MIVQUE(0)	1.163	0.131
	ML	0.490	0.076
	MINQE	0.333	0.111
P2=(1,5,9)	ANOVA	1.027	0.349
	REML	1.149	0.107
	MIVQUE(0)	1.271	0.522
	ML	0.498	0.076
	MINQE	0.270	0.148
P4=(3,3,5,5,7,7)	ANOVA	1.033	0.069
	REML	1.014	0.018
	MIVQUE(0)	1.221	0.166
	ML	0.731	0.058
	MINQE	0.511	0.189
P5=(1,1,5,5,9,9)	ANOVA	1.074	0.312
	REML	1.169	0.158
	MIVQUE(0)	1.411	0.501
	ML	0.791	0.073
	MINQE	0.461	0.245
P7=(1,1,1,1,13,13)	ANOVA	1.381	0.660
	REML	1.415	0.522
	MIVQUE(0)	2.243	1.237
	ML	0.819	0.104
	MINQE	0.418	0.329
P8=(3,3,3,5,5,5,7,7,7)	ANOVA	1.047	0.058
	REML	1.014	0.025
	MIVQUE(0)	1.261	0.148
	ML	0.822	0.046
	MINQE	0.621	0.265
P9=(1,1,1,5,5,5,9,9,9)	ANOVA	1.111	0.279
	REML	1.143	0.177
	MIVQUE(0)	1.452	0.465
	ML	0.888	0.091
	MINQE	0.590	0.337
P11=(1,1,1,1,1,1,1,19,19)	ANOVA	1.207	0.916
	REML	1.086	0.724
	MIVQUE(0)	1.516	1.846
	ML	0.524	0.126
	MINQE	0.359	0.515
P12=(2,10,18)	ANOVA	1.802	0.246
	REML	1.543	0.076
	MIVQUE(0)	3.295	0.370
	ML	0.963	0.068
	MINQE	0.564	0.097
P13=(3,15,27)	ANOVA	1.230	0.222
	REML	1.067	0.085
	MIVQUE(0)	1.546	0.341
	ML	0.533	0.063
	MINQE	0.401	0.089

표 4는 급내상관계수 ρ의 여러가지 값에 대하여, σ_a^2의 각 추정량들의 MSE에 대한 효율성을 표시하고 있으며, 다음은 표 4로부터 얻을 수 있는 중요한 결론들이다.

표 3. σ_a^2의 추정량들의 효율성: $(K_a, K_e)=(3,3)$인 경우

N-형식	추정량	ρ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
P1	ANOVA	0.966	0.977	1.090	1.069	1.075	1.086	1.090	1.044	1.055
	MIVQUE(0)	1.042	1.044	1.297	1.249	1.262	1.268	1.279	1.194	1.199
	REML	1.019	1.015	1.011	1.009	1.006	1.003	1.002	1.001	1.000
	ML	0.347	0.419	0.455	0.484	0.488	0.499	0.507	0.501	0.512
	MINQE	0.269	0.234	0.248	0.277	0.287	0.307	0.335	0.329	0.365
P2	ANOVA	0.482	0.461	0.591	1.181	1.221	1.230	1.393	1.218	1.395
	MIVQUE(0)	0.428	0.448	0.613	1.501	1.515	1.549	1.784	1.527	1.763
	REML	1.230	1.195	1.133	1.119	1.074	1.064	1.062	1.075	1.534
	ML	0.376	0.438	0.456	0.477	0.496	0.528	0.518	0.520	0.487
	MINQE	0.134	0.110	0.129	0.194	0.235	0.296	0.314	0.333	0.301
P4	ANOVA	0.964	0.996	1.057	1.074	1.050	1.152	0.923	0.993	1.132
	MIVQUE(0)	1.012	1.130	1.267	1.303	1.281	1.478	1.046	1.142	1.432
	REML	1.022	1.027	1.019	1.013	1.010	1.006	1.005	1.002	1.066
	ML	0.608	0.692	0.722	0.718	0.738	0.739	0.719	0.731	0.746
	MINQE	0.559	0.365	0.360	0.368	0.413	0.464	0.426	0.488	0.489
P5	ANOVA	0.592	0.712	0.844	1.063	1.212	1.194	1.162	1.301	1.317
	MIVQUE(0)	0.567	0.822	0.971	1.355	1.661	1.551	1.356	1.709	1.748
	REML	1.462	1.349	1.219	1.147	1.117	1.069	1.059	1.147	1.128
	ML	0.769	0.790	0.786	0.780	0.788	0.807	0.755	0.796	0.796
	MINQE	0.379	0.235	0.258	0.296	0.354	0.493	0.408	0.486	0.602
P7	ANOVA	0.339	0.495	0.795	1.044	1.606	2.044	1.366	2.062	2.302
	MIVQUE(0)	0.286	0.575	1.153	1.575	2.635	3.344	2.176	3.377	3.767
	REML	1.501	1.431	1.418	1.299	1.759	1.151	1.100	2.157	2.913
	ML	0.777	0.781	0.836	0.818	0.943	0.839	0.771	0.930	1.025
	MINQE	0.175	0.130	0.164	0.208	0.276	0.446	0.352	0.494	0.585
P8	ANOVA	0.998	1.050	0.991	1.114	0.956	1.038	1.078	1.078	1.067
	MIVQUE(0)	1.096	1.255	1.120	1.418	1.076	1.229	1.287	1.337	1.339
	REML	1.019	1.024	1.022	1.014	1.011	1.006	1.003	1.002	1.119
	ML	0.723	0.800	0.811	0.838	0.809	0.825	0.812	0.823	0.887
	MINQE	0.781	0.414	0.382	0.443	0.427	0.523	0.538	0.631	0.579
P9	ANOVA	0.635	0.916	1.040	1.137	1.152	1.175	1.333	1.277	1.281
	MIVQUE(0)	0.574	1.114	1.350	1.506	1.520	1.489	1.785	1.721	1.742
	REML	1.440	1.276	1.260	1.108	1.107	1.066	1.043	1.033	1.159
	ML	0.903	0.917	0.947	0.891	0.869	0.858	0.849	0.844	0.879
	MINQE	0.543	0.317	0.307	0.380	0.397	0.480	0.560	0.638	0.627
P11	ANOVA	0.399	0.616	0.878	1.898	1.486	1.923	2.441	2.899	2.674
	MIVQUE(0)	0.421	0.866	1.417	3.405	2.597	3.403	4.454	5.254	4.813
	REML	1.672	1.541	1.440	1.357	1.220	1.134	1.367	2.780	3.272
	ML	0.980	0.975	0.975	1.024	0.935	0.920	0.957	1.247	1.170
	MINQE	0.246	0.141	0.149	0.294	0.319	0.456	0.597	0.701	0.760
P12	ANOVA	0.843	1.097	1.236	1.327	1.323	1.373	1.366	1.362	1.371
	MIVQUE(0)	0.954	1.328	1.521	1.677	1.674	1.749	1.737	1.702	1.753
	REML	1.216	1.158	1.083	1.065	1.055	1.032	1.021	1.011	1.005
	ML	0.410	0.479	0.510	0.508	0.526	0.513	0.523	0.497	0.506
	MINQE	0.312	0.284	0.294	0.301	0.331	0.335	0.363	0.345	0.372
P13	ANOVA	0.841	1.139	1.289	1.310	1.324	1.165	1.303	1.267	1.281
	MIVQUE(0)	0.921	1.386	1.610	1.622	1.641	1.400	1.609	1.582	1.587
	REML	1.188	1.106	1.071	1.039	1.030	1.017	1.011	1.006	1.001
	ML	0.437	0.490	0.513	0.515	0.496	0.518	0.514	0.530	0.529
	MINQE	0.378	0.323	0.335	0.349	0.341	0.366	0.379	0.408	0.421

1. MINQE 추정량의 효율성은 ρ의 값이 매우 작을 때 그 효율성은 떨어지며, ρ=0.2 근처에서 가장 좋은 효율성을 가지나 ρ의 값이 증가할수록 효율성은 떨어져서 ρ=0.9 근처가 되면, 거의 ML 추정량과 효율성이 비슷해진다.
2. ANOVA 추정량과 MIVQUE(0) 추정량의 효율성은 ρ의 값이 증가할수록 효율성이 떨어지나, 그 정도는 MIVQUE(0) 추정량이 훨씬 심하다. ML 추정량과 REML 추정량의 효율성은 다른 추정량에 비해 ρ의 값의 변화에 따른 뚜렷한 경향을 보이지 않는다.

표 4. ρ의 변화에 따른 σ_a^2의 추정량들의 MSE에 대한 효율성

추 정 량	ρ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
ANOVA	0.810	0.911	1.036	1.135	1.227	1.289	1.371	1.454	1.512
REML	1.228	1.195	1.151	1.116	1.088	1.078	1.058	1.178	1.358
MIVQUE(0)	0.903	1.107	1.340	1.521	1.700	1.813	1.971	2.137	2.249
ML	0.604	0.675	0.708	0.717	0.724	0.718	0.711	0.733	0.761
MINQE	0.418	0.263	0.274	0.323	0.394	0.470	0.546	0.650	0.738

3.2 $\hat{\sigma}_e^2$의 MSE와 편의들

표 5는 추정된 σ_e^2에 대한 MSE의 상대적 효율성의 평균값과 표준편차를 보여 주며, 우리는 다음과 같은 결론들을 얻을 수 있다. 따라서, 오차항의 분포가 정규분포를 따른다는 가정을 떠나서도, 우리가 주로 사용하는 ANOVA 추정량을 사용하는 것이 바람직한 선택으로 생각되어진다.

1. 6가지의 σ_e^2의 추정량들은 σ_a^2의 추정량들과는 달리 MIVQUE(0)를 제외하고, MSE 판정기준 아래에서 비슷한 효율성을 제공한다. MINQE는 K_e의 값이 증가할수록 효율성은 증가하나, 그 효율성은 처리효과의 MINQE보다 훨씬 떨어진다.
2. 표에서 제시되지는 않았지만, MINQE 추정량은 다른 추정량들에 비해 전반적으로, 편의가 크다. 나머지 σ_e^2의 추정량들은 편의 판정기준 아래서 비슷한 효율성을 제공한다. ML 추정량은 비록 편의추정량이지만, K_e=0가 아닌경우에도 그 편의는 심각할 정도로 크지는 않다. 반면에, MIVQUE(0)는 불편추정량이지만, ρ의 값이 증가할수록 편의는 증가한다.
3. 급내상관계수 ρ값의 변화에 대한 σ_e^2에 대한 각 추정량들의 MSE에 대한 효율성도 뚜렷한 변화가 없는 까닭에 표로 제시되지는 않았지만, 추정된 각 추정량들의 효율성을 비교하여 보면, ANOVA 추정량은 ρ의 값이 증가할수록 효율성이 증가하며, MIVQUE(0) 추정량은 변량인자들이 정규분포를 따르는 경우와 마찬가지로, ρ의 값이 증가할수록 효율성이 매우 심하게 떨어진다. 실제로 ρ=0.9인 경우, MIVQUE(0)는 MIVQUE(A)에 비해 MSE가 27.024 배나 크다. 또한 ML 추정량과 REML 추정량의 효율성은 다른 추정량들에 비해서 σ_a^2의 경

우와 마찬가지로 ρ의 값의 변화에 따른 뚜렷한 경향을 보이지 않으며, MINQE 추정량인 경우에도 σ_a^2의 경우와 달리 ρ=0.9인 경우를 제외하고 (이 경우, MINQE는 MIVQUE(A)에 비해 MSE가 1.344배 커진다), ρ의 값에 크게 종속되지 않는다.

4. 사전추측값의 선택

이 절에서는 MINQE의 사전추측값를 결정하는 문제를 고려해 보기로 하겠다. 일반적으로 MINQUE 이론에 나오는 r은 σ_a^2/σ_e^2의 사전추측값이다. 하지만 다행스럽게도 특정의 사전추측값의 선택이 MSE 에 영향을 민감하게 미치지는 않으나(Hess 1979), 우리는 이 MINQE의 사전추측값를 잘 결정하면 MSE를 보다 줄일 수 있다. 사실 MINQE의 MSE는 σ_a^2/σ_e^2의 이차함수로 표시되고, 자세한 표현(CW 1987 ; Ahrens et al. 1981)은 굉장히 복잡하여서, 이론적으로는 MSE가 최소로 되는 사전추측값를 구할 수 있으나, 모수 σ_a^2/σ_e^2를 우리는 알 수가 없기 때문에 실질적으로 의미가 없다. 따라서, 우리는 모의실험을 통하여 다음과 같은 경험적 규칙을 얻었다.

알고리즘:

단계 1. $n^* = \sum_{i=1}^{k}(1/n_i)$ 를 구한다.

단계 2. $n^* < 1$ 이면, 사전추측값 $r = 1$을 선택하고,
$1 \le n^* < 2$ 이면, 사전추측값 $r = n^*$를 선택하고,
$n^* \ge 2$ 이면, 사전추측값 $r = 2$를 선택한다.

표 6과 표 7은 각각 ML, CW의 MINQE, 그리고 이 절에서 제안된 사전추측값를 사용한 MINQE의 MSE와 편의를 제시한다. 여기서 사용된 분포는 (K_a, K_e)=(0,0)인 경우이며, K_a와 K_e의 값이 변하더라도 비슷한 효율성을 얻을 수 있다. ML의 MSE 와 편의는 SM에 제시되어 있는 값을 그대로 사용하였고, CW의 MINQE의 MSE와 편의는 각각의 N-형식과 ρ의 값에 대하여 10000개의 자료를 생성하여 다시 계산하였는데, CW의 MINQE의 값하고는 매우 근소한 차가 생겼다. 다음은 표 6과 표 7을 통해서 얻을 수 있는 결론이다.

1. 제안된 방법이 전반적으로 CW의 선택보다 상당히 MSE 를 줄여 준다. CW에서는 MINQE의 우수성을 $\sigma_a^2 > \sigma_e^2$ 인 경우라고 하였는 데, 이젠 $\sigma_a^2 > 0.5\sigma_e^2$ 인 경우까지 가능하다고 말을 할 수가 있다. $\sigma_a^2/\sigma_e^2 \le 0.2$인 경우, 아직도 ML에 비해서 효율성이 나아진다고는 할 수 없지만, 제안된 방법이 CW보다는 상당히 큰 효율성을 가져다 준다. 그러나, 제안된 방법의 한 가지 단점은 σ_a^2/σ_e^2의 값이 ∞로 접근할 때 CW보다 효율성이 다소 나빠진다는 것이다.
2. 제안된 방법이 CW의 선택보다 눈에 띌만한 편의를 줄여 주지는 못하며, 따라서MINQE 의 큰 편의값은 이 절에서 제시된 방법으로는 극복할 수가 없다.

표 5. σ_e^2의 추정량들의 MSE에 대한 효율성

Ka	Ke	추정량	평균	표준편차	Ka	Ke	추정량	평균	표준편차
-1.2	-1.2	ANOVA	0.998	0.022	3	-1.2	ANOVA	0.996	0.027
		REML	1.005	0.017			REML	1.016	0.048
		MIVQUE(0)	7.828	15.918			MIVQUE(0)	1.844	24.663
		ML	1.027	0.037			ML	1.034	0.043
		MINQE	1.245	0.357			MINQE	1.407	0.670
-1.2	0	ANOVA	1.010	0.033	3	0	ANOVA	1.010	0.034
		REML	1.004	0.015			REML	1.016	0.078
		MIVQUE(0)	3.889	6.366			MIVQUE(0)	5.277	8.728
		ML	1.018	0.025			ML	1.021	0.027
		MINQE	0.953	0.157			MINQE	1.020	0.244
-1.2	3	ANOVA	1.023	0.042	3	3	ANOVA	1.024	0.041
		REML	1.001	0.010			REML	1.005	0.013
		MIVQUE(0)	2.280	2.834			MIVQUE(0)	2.939	3.981
		ML	1.009	0.013			ML	1.010	0.012
		MINQE	0.828	0.091			MINQE	0.859	0.109
-1.2	9	ANOVA	1.032	0.050	3	9	ANOVA	1.032	0.049
		REML	0.999	0.020			REML	1.022	0.127
		MIVQUE(0)	1.489	1.163			MIVQUE(0)	1.914	2.058
		ML	1.006	0.010			ML	1.015	0.068
		MINQE	0.752	0.066			MINQE	0.772	0.074
0	-1.2	ANOVA	0.999	0.029	9	-1.2	ANOVA	0.993	0.026
		REML	1.008	0.022			REML	1.036	0.130
		MIVQUE(0)	7.932	14.813			MIVQUE(0)	6.569	32.203
		ML	1.028	0.037			ML	1.046	0.080
		MINQE	1.271	0.386			MINQE	1.561	0.981
0	0	ANOVA	1.010	0.031	9	0	ANOVA	1.012	0.030
		REML	1.004	0.016			REML	1.241	2.119
		MIVQUE(0)	4.173	6.599			MIVQUE(0)	6.964	13.780
		ML	1.019	0.023			ML	1.168	1.369
		MINQE	0.987	0.183			MINQE	1.102	0.428
0	3	ANOVA	1.023	0.044	9	3	ANOVA	1.024	0.040
		REML	1.002	0.013			REML	1.019	0.077
		MIVQUE(0)	2.360	2.904			MIVQUE(0)	4.238	6.662
		ML	1.009	0.013			ML	1.014	0.021
		MINQE	0.832	0.091			MINQE	0.894	0.184
0	9	ANOVA	1.032	0.052	9	9	ANOVA	1.045	0.060
		REML	1.000	0.019			REML	1.025	0.193
		MIVQUE(0)	1.632	1.312			MIVQUE(0)	2.462	3.677
		ML	1.005	0.010			ML	1.018	0.115
		MINQE	0.765	0.070			MINQE	0.797	0.107

5. 결론

일원변량모형의 분산성분추정에 관한 연구들은 지금까지 변량인자들의 분포들이 정규분포라는 가정아래에서 행하여 졌는데, 이 논문에서는 변량인자들의 분포에 대한 임의성을 가정하고 MINQE의 MSE와 편의의 판정기준아래서 효율성을 알아보았다. MINQE는 변량인자의 분포의 첨도가 커질수록 MSE의 효율성이 증가한다. 그리고, σ_a^2를 추정할 때, MINQE의 사전추측값을 잘 선택함으로서, 효율성이 증가될 수 있는데, 사전추측값은 주관적으로 선택하는 값이지만, 각

수준의 관측치의 함수로서, 1에서 2까지의 값을 선택하는 것이 바람직하다. 점추정을 하는데 어떠한 판정기준을 사용하는가 하는 문제는 확실히 주관적이다. MINQE는 변량인자들의 분포가 정규분포를 따른다는 가정을 벗어나서도 편의는 다른 추정량에 비해 크다. 하지만, 편의라는 기준하나만 고려하는 것은 대단히 위험한 일이다. 편의보다는 CW의 논의처럼 추정된 모수로부터 멀리 떨어져 있을 확률을 줄이는 것이 더 의미가 있지 않을 까? MINQE는 이러한 관점에서 살펴보면 확실히 우수한 추정량이다.

표 6. 세 가지 추정량의 추정된 σ_a^2의 MSE

N-형식	추정량	σ_a^2/σ_e^2					
		0.1	0.2	0.5	1.0	2.0	5.0
P2	ML	4.700	2.500	1.302	0.986	0.759	0.647
	MINQE(1.5)	4.882	1.527	0.597	0.514	0.512	0.521
	MINQE(1.31)	4.100	1.307	0.563	0.509	0.514	0.527
P4	ML	1.900	1.000	0.591	0.455	0.386	0.331
	MINQE(2)	3.285	0.989	0.396	0.300	0.290	0.287
	MINQE(1.35)	2.500	0.753	0.338	0.284	0.289	0.294
P5	ML	2.100	1.200	0.730	0.540	0.433	0.356
	MINQE(2.5)	7.570	1.917	0.507	0.311	0.291	0.294
	MINQE(2)	5.980	1.501	0.427	0.294	0.292	0.301
P7	ML	4.900	2.400	1.160	0.736	0.519	0.389
	MINQE(4)	21.555	5.229	1.012	0.407	0.302	0.289
	MINQE(2)	10.074	2.282	0.472	0.288	0.301	0.331
P9	ML	1.500	0.900	0.514	0.374	0.294	0.251
	MINQE(4)	10.286	2.701	0.528	0.262	0.208	0.202
	MINQE(2)	5.432	1.350	0.293	0.205	0.208	0.224
P11	ML	4.300	2.100	0.953	0.570	0.391	0.279
	MINQE(7)	37.298	9.107	1.542	0.479	0.240	0.201
	MINQE(2)	11.252	2.369	0.355	0.200	0.232	0.278

본 연구에서는 논의되지 않았지만, 일원변량모형에서의 MINQE의 MSE 판정기준아래서의 최적성은 이원변량모형에서도 성립될 것으로 간주되며, 또한 MINQE의 효율성을 높이기 위해서 가장 많이 쓰이는 유클리드 노옴대신 가중노옴, 삼각노옴, L_1노옴, 체비세프 노옴, 스펙트럴 노옴등을 도입하는 것도 매우 흥미로운 일이 될 것으로 생각되어진다. 이 주제들에 대해서도 향후 논의되어 질 것이다.

표 7. 세 가지 추정량의 추정된 σ_a^2의 편의

N-형식	추정량	σ_a^2/σ_e^2					
		0.1	0.2	0.5	1.0	2.0	5.0
P2	ML	-0.021	-0.073	-0.202	-0.389	-0.771	-1.790
	MINQE(1.5)	0.094	0.038	-0.134	-0.425	-0.993	-2.715
	MINQE(1.31)	0.081	0.023	-0.155	-0.456	-1.046	-2.828
P4	ML	-0.009	-0.043	-0.110	-0.205	-0.366	-0.859
	MINQE(2)	0.117	0.084	-0.009	-0.170	-0.496	-1.466
	MINQE(1.35)	0.097	0.058	-0.053	-0.244	-0.629	-1.776
P5	ML	-0.018	-0.046	-0.112	-0.213	-0.388	-0.895
	MINQE(2.5)	0.190	0.151	0.045	-0.139	-0.512	-1.618
	MINQE(2)	0.166	0.124	0.007	-0.194	-0.600	-1.807
P7	ML	-0.013	-0.044	-0.119	-0.207	-0.395	-0.838
	MINQE(4)	0.350	0.309	0.202	0.013	-0.371	-1.508
	MINQE(2)	0.230	0.178	0.031	-0.222	-0.730	-2.238
P9	ML	-0.015	-0.040	-0.083	-0.141	-0.273	-0.598
	MINQE(4)	0.253	0.229	0.151	0.020	-0.259	-1.037
	MINQE(2)	0.180	0.145	0.037	-0.143	-0.519	-1.597
P11	ML	-0.007	-0.033	-0.090	-0.141	-0.259	-0.587
	MINQE(7)	0.515	0.487	0.408	0.267	-0.031	-0.827
	MINQE(2)	0.273	0.222	0.071	-0.186	-0.709	-2.224

참고 문헌

[1] Ahrens, H., Kleffe, J., and Tenzler, R. (1981), "Mean Square Error Comparison for MINQUE, ANOVA and Two Alternative Estimators Under the Unbalanced One-Way Random Model," *Biomedical Journal*, 23, 323 - 342.

[2] Chaloner, K. (1987), "A Bayesian Approach to the Estimation of Variance Components for the Unbalanced One Way Random Model," *Technometrics*, 29(3), 323 - 337.

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

[4] Hess, J. L. (1979), "Sensitivity of MINQUE with Respect to A priori Weights," *Biometrics*, 35, 645 - 649.

[5] Lee, J. T. and Lee, K. S. (1990), "A Comparison on Non-Negative Estimators for Ratios of Variance Components," *Proceedings of COMSTAT'90*, 303 - 308.

[6] Palmer, J. L. and Broemeling, L. D. (1990), "A Comparison of Bayes and Maximum Likelihood Estimation of the Intraclass Correlation Coefficient," *Communications in Statistics- Theory and Methodology*, 19(3), 953 - 975.

[7] Rao, C. R. (1971), "Minimum Variance Quadratic Unbiased Estimation of Variance Components," *Journal of Multivariate Analysis*, 1, 445 - 456.

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

[9] Rao, P. S. R. S. and Chaubey, Y. P. (1978), "Three Modifications of the Principle of MINQUE," *Communications in Statistics- Theory and Methodology*, 7(8), 767 -

778.
[10] Searle, S. R. (1971), *Linear Models,* John Wiley, New York.
[11] Searle, S. R. (1987), *Linear Models for Unbalanced Data,* John Wiley & Sons, New York.
[12] Swallow, W. H., and Monahan, J. F. (1984), "Monte Carlo Comparison of ANOVA, MINQUE, REML, and ML Estimators of Variance Components," *Technometrics,* **26**, 47 - 57.
[13] Swallow, W. H. and Searle, S. R. (1978), "Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components," *Technometrics,* **20**, 265 - 272.
[14] Westfall, P. H. (1987), "A Comparison of Variance Components Estimates for Arbitrary Underlying Distributions," *Journal of the American Statistical Association,* **82**, 866 - 874.

Efficiency of MINQE for Arbitrary Underlying Distribution Under One Way Random Effects Model[1)]

Jang Taek Lee[2)]

Abstract

The estimations of variance components for the unbalanced one way random effects model when the underlying distributions are not necessarily normal are considered. ANOVA, REML, ML, MIVQUE, and MINQE estimators are compared with respect to their mean squared errors and biases through a simulation study. Explicit, computable expressions with no matrix inversion necessary are given for these estimators. An efficient rule to provide a prior guess of MINQE is given. Our results indicate that the efficiency of MINQE is excellent for arbitrary underlying distribution in the sense of MSE even in the presence of nontrivial bias. Also, MINQE is a worthwhile improvement over other estimators when kurtosis of underlying distributions become large. 1

---

1) Research supported by the Korea Science and Engineering Foundation, 1992. (KOSEF 921-0100-008-1)
2) (140-714) Department of Computer Science and Statistics, Dankook University, Seoul