The Korean Communications
in Statistics Vol. 10 No. 1, 2003
pp. 11-18

Power Comparison of EGLS Test Statistic for Fixed Effects with Arbitrary Distributions[1)]

Jang-Taek Lee[2)]

Abstract

Quite often normality assumptions are not satisfied in practical applications. In this paper, an estimated generalized least squares(EGLS) analysis are considered in two way mixed linear models with arbitrary types of distributions for random effects. We investigate the power performance of EGLS analysis based on Henderson's method III, ML, REML and MINQUE(1). The power performances depend on the imbalance of design, on the actual values of ratio of variance components, and on the skewness and kurtosis parameters of the underlying distributions slightly.

Results of our limited simulation study suggest that the EGLS F-statistics using four estimators and arbitrary distributions produce similar type I error rates and power performance.

Keywords : mixed models, estimated generalized least squares.

1. 서론

혼합모형은 오차항의 구조가 너무 제한적인 일반선형모형을 확장하여 보다 일반적인 분산행렬을 허용함으로써 실험계획과 관측연구에서 매우 유용하게 사용되어진다. 그리고 최근 10년 동안 컴퓨터 하드웨어와 소프트웨어의 눈부신 발전은 혼합모형에 관한 여러 가지 분석을 현실적으로 처리 가능하게 만들었으며, 그 결과 혼합모형의 필요성과 중요성을 사용자들은 더욱 절실하게 느끼게 된다.

하지만 실제로 Lewis, Montgomery와 Myers(2001)가 소개한 내용처럼 여러 분야의 과학, 공학, 생산 및 품질관리 등에서는 실험자들이 비정규 반응변수들을 많이 접하게 되며, 정규분포이론에 입각한 혼합모형의 이론들 중에서 강건성(robustness)이 밝혀지지 않은 분야 등은 비정규분포에 대한 활용 가능성의 타당성을 살펴보아야 할 시기라고 생각되어진다. 본 연구는 이와 같은 관점에서 이장택(1997) 논문의 연장으로 혼합모형에서 고정효과의 요인 수준간에 유의한 차이가 있는지를 랜덤효과와 오차항은 임의의 확률분포를 따른다는 가정을 사용하여 일반최소제곱 F-통계량이

---

1) This work was supported by Dankook University Research Fund in 2001
2) Professor, Division of Information and Computer Science, Dankook University, San 8, HanNam-Dong, YongSan-Gu, Seoul, Korea
E-mail : jtlee@dankook.ac.kr

적당한 지를 검정력의 관점에서 모의실험을 통하여 살펴본다. 모의실험은 실험계획에서 가장 많이 사용되는 혼합모형 중의 하나인 이원혼합모형을 이용하였으며, 이원혼합모형에 포함된 분산성분들을 핸더슨의 방법 III과 최우추정법, 제한최우추정법 그리고 사전추측값이 1인 MINQUE를 사용하여 추정하고, 추정된 분산성분의 값을 이용하여 고정효과에 대한 일반최소제곱분석을 수행하였다.

본 논문의 구성은 2절에서는 혼합모형과 일반최소제곱 F-통계량 및 분산성분추정량에 대하여 간략하게 서술하고, 3절에서는 모의실험에 사용된 가설, 모형의 종류 및 조건, 모의실험의 과정과 4가지 추정량을 이용하여 추정된 일반최소제곱 F-검정의 제1종 오류율과 검정력의 결과를 랜덤효과와 오차항의 확률분포가 대칭분포와 비대칭분포인 경우로 각각 나누어서 서술한다. 끝으로 4절에서는 본 연구의 결론이 주어진다.

2. 일반최소제곱 F-통계량과 분산성분추정량

이 절에서는 혼합모형과 일반최소제곱 F-통계량, 그리고 모의실험에서 사용된 분산성분추정량들의 종류에 대하여 간략하게 소개한다. 여기에 관한 보다 자세한 설명은 Searle(1987)과 Rao와 Kleffe(1988)의 책에서 찾아볼 수 있다.

2.1 혼합모형과 일반최소제곱 F-통계량

선형혼합모형은 다음과 같이 기술된다.

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

여기서 $\underline{y}$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, $\underline{u}$는 랜덤효과로 불리는 $q \times 1$ 열벡터, 그리고 $\underline{e}$는 $n \times 1$ 오차벡터이다. 그리고 다음 식(2.2)가 성립한다고 가정한다.

$$\underline{u} \sim N(\underline{0}, G) \quad , \quad \underline{e} \sim N(\underline{0}, \sigma_e^2 I_n) \quad , \quad cov(\underline{u}, \underline{e}') = 0. \tag{2.2}$$

그러면 $\underline{y}$는 평균이 $X\underline{\beta}$이고 분산행렬이 $V = \sigma_e^2 I_n + ZGZ'$인 다변량 정규벡터이다.

만일 식(2.1)에서 벡터 $\underline{y}$의 분산행렬 V가 $V = \sigma_e^2 I_n$이고, 가설의 모양이 $H_0 : K'\underline{\beta} = \underline{m}$, (단 K'은 알려진 $r \times p$ 행렬로 계수는 r이며, $\underline{m}$은 알려진 $r \times 1$ 벡터) 라고 할 때 우리가 사용하는 표준최소제곱 F-통계량은 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 일반화 역행렬이다. 그리고 분산행렬 V가 알려져 있으면, 가설 $H_0 : K'\underline{\beta} = \underline{m}$에 대하여 다음과 같은 일반최소제곱 F-통계량을 얻게 된다.

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

식(2.4)에서 $\underline{b}_g = (X'V^{-1}X)^{-}X'V^{-1}\underline{y}$는 $\underline{\beta}$의 일반최소제곱추정량이며, $K'X\beta$의 최소선형불편추정량(BLUE)는 다음 식(2.5)와 같다.

$$\mathrm{BLUE}(K'X\beta) = K'X(X'V^{-1}X)^{-}X'V^{-1}\underline{y} \tag{2.5}$$

한편 대부분의 사용자는 분산행렬 V의 구조를 알 수 없고, 주어진 데이터로부터 추정하여야 한다. 이 경우 분산행렬 V는 분산성분의 함수이므로 분산성분이 추정되어지면, 보편화된 방법으로 추정치의 값을 모수의 값으로 대체하여 고정인자에 관한 일반최소제곱분석을 수행하게 되는데, 우리는 이것을 경험적 일반최소제곱분석(EGLS)이라고 한다. 하지만 EGLS추정량의 가장 큰 단점은 유한표본성질이 일반적으로 알려져 있지 않다는 점이다. 왜냐하면 추정량 $\underline{b}_g$은 자료벡터 $\underline{y}$의 복잡한 함수이기 때문에 유한표본성질을 유도하는 것이 매우 어려운 일이기 때문이다. 이런 이유로 통계적 추론을 위해서는 EGLS 추정량의 점근적 성질을 주로 이용하게 된다.

2.2 분산성분추정량

혼합모형의 분산성분을 추정하기 위한 방법들은 크게 세 가지 클래스로 나뉘어 진다. 첫째는 불편추정량의 원리에 입각한 추정량들로 1980년대까지 매우 보편적으로 사용되던 헨더슨방법 III을 포함한 분산분석에 기반한 추정량들이 포함된다. 분산분석에 기반한 추정량들은 이차형식이 그 기대값과 같다고 하거나 분산분석표에서 제곱평균이 평균제곱기대값과 같다고 두고 구한 추정량으로 정의된다. 둘째는 우도원리에 입각한 추정량들로 대표적인 추정량이 최우추정량(MLE)과 제한최우추정량(REMLE)을 들 수 있는데 불편추정량과는 달리 혼합모형의 랜덤효과와 오차항이 특정분포를 따른다는 가정이 필요하게 된다. 이 방법들은 확실히 불편추정량들보다 통계적 최적성질들을 많이 가지고 있다. 셋째는 베이지안 방법에 입각한 추정량들이다. 베이지안 방법은 모든 모수들이 랜덤으로 간주되어진다.

본 연구에서는 보편적으로 가장 많이 사용되는 추정량들인 헨더슨의 방법 III 추정량, 최소노옴이차불편추정량(MINQUE), 우도원리에 입각한 MLE와 REMLE를 사용한다.

3. 모의실험

이 절에서는 모의실험에 사용된 모형, 실험계획의 종류, 모의실험의 과정 및 그 결과를 설명한다.

3.1 모의실험계획

모의실험에서는 혼합모형의 단순화를 위하여 고정인자가 하나이면서 교호작용이 없는 이원가법모형 및 이원지분모형을 고려하였다. 다양한 왜도와 첨도에 대한 영향을 살펴보기 위하여 사용된 랜덤효과와 오차항에 대한 분포가정으로 대칭분포인 경우에는 첨도가 -1.2인 균등분포, 첨도가 0인 정규분포, 첨도가 9인 오염정규분포를 이용하였는데, 이 경우 첨도가 9인 오염정규분포의 확률밀도함수는 다음과 같이 주어진다.

$$f(x) = \frac{w_1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{x^2}{2\sigma_1^2}\right) + \frac{w_2}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{x^2}{2\sigma_2^2}\right).$$

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 오염정규확

률변수의 분산이다. 또한 비대칭분포인 경우에는 모수가 (β, θ)인 감마분포의 왜도가 $2/\sqrt{\beta}$이고 첨도가 $6/\beta$인 사실을 이용하여 다양한 왜도와 첨도를 갖는 확률분포를 생성하였다. 이원가법모형에서 모의실험에 사용되는 자료의 불균형도를 정확하게 비교하기 위하여 Khuri(1987)가 제안한 측도를 사용하였는데, 상수 k를 식(3.1)과 같이 정의하면,

$$k = c/n_{..} \ , \quad c = \sum_{i}^{a}\sum_{j}^{b} \frac{(n_{ij} - n_{..}/ab)^2}{n_{..}/ab} . \tag{3.1}$$

불균형도의 측도 $\Phi(D)$는 $\Phi(D) = 1/(1+k)$로 정의된다. 이 경우, 각 셀 당 반복수가 모두 동일한 균형자료이면 $\Phi(D)$의 값은 1이 되며, 자료의 불균형도가 심할수록 $\Phi(D)$의 값은 0에 가까워진다. 이상을 고려하여 선택된 모의실험계획은 다음과 같다.

<표 1> 모의실험에 사용된 모형의 종류와 조건

모형의 종류		고려된 조건	사용된 조건의 값
모형1	$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$. $i = 1, 2, \cdots, a;\ j = 1, 2, \cdots, b$ $k = 1, 2, \cdots, n_{ij}$	총관측치의 개수	$n = 30$
		(a, b) 의 값	(a, b) = (3, 5) , (a, b) = (5, 3)
		확률분포	일양분포, 정규분포, 오염정규분포, 감마분포
		분산성분비율	$\sigma_\beta^2/\sigma_e^2 = 0.25, 1, 4;\ \sigma_e^2 = 1$
		독립성 가정	β_j , e_{ijk} : 서로 독립
		불균형도	$\Phi(D) = 0.5$; 0.97
모형2	$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + e_{k(ij)}$, $i = 1, 2, \cdots, a;\ j = 1, 2, \cdots, b_i$ $k = 1, 2, \cdots, n_{ij}$	총관측치의 개수	$n = 30$
		a의 값	$a = 3$
		확률분포	일양분포, 정규분포, 오염정규분포, 감마분포
		분산성분비율	$\sigma_\beta^2/\sigma_e^2 = 0.25, 4;\ \sigma_e^2 = 1$
		독립성 가정	$\beta_{j(i)}$, $e_{k(ij)}$: 서로 독립
		불균형도	N[1], N[2], N[3], N[4]

불균형도가 $\Phi(D) = 0.97$인 실험계획은 거의 균형적인 자료이며, 불균형도가 $\Phi(D) = 0.5$인 자료는 불균형도가 매우 심한 자료이다. 또한 같은 $\Phi(D)$의 값이 존재하지 않는 경우에는 가장 비슷한 $\Phi(D)$의 값을 제공하는 실험계획을 선택하였다. 모형2는 불균형도를 나타내기 위하여 b_i와 n_{ij}를 동시에 고려한 모형 4가지를 사용하였으며, 기술된 실험계획은 $(b_1, b_2, b_3 | n_{11}, n_{12}, \cdots, n_{3b_3})$을 의미한다. 이 경우 N[1], N[2]는 N[3], N[4]와 비교하여 b_i에 대한 불균형도가 약하며, N[1], N[3]는 N[2], N[4]에 비하여 n_{ij}에 대한 불균형도가 약한 실험계획이다.

$$\begin{aligned}
N[1] &= (2,\ 3,\ 4 \mid 3,\ 4,\ 3,\ 4,\ 4,\ 3,\ 3,\ 2,\ 4),\\
N[2] &= (2,\ 3,\ 4 \mid 8,\ 1,\ 1,\ 9,\ 1,\ 1,\ 1,\ 1,\ 7),\\
N[3] &= (1,\ 3,\ 5 \mid 3,\ 4,\ 3,\ 4,\ 4,\ 3,\ 3,\ 2,\ 4),\\
N[4] &= (1,\ 3,\ 5 \mid 8,\ 1,\ 1,\ 9,\ 1,\ 1,\ 1,\ 1,\ 7).
\end{aligned}$$

다음 <표 2>는 모의실험에 사용된 가설의 모양 및 조건을 설명한다.

<표 2> 모의실험에 사용된 가설의 모양 및 조건

수준조합	가설의 모양	사용된 가설		실제 사용된 α_i의 값
(3,5)	$\lambda'\alpha = \alpha_1 - 2\alpha_2 + \alpha_3$	제1종오류율	$\lambda'\alpha = 0$	$\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1$
		검정력	$\lambda'\alpha = 1$	$\alpha_1 = 2, \alpha_2 = 1, \alpha_3 = 1$
(5,3)	$\lambda'\alpha = \alpha_1 - 2\alpha_2 + \alpha_3 + \alpha_4 - \alpha_5$	제1종오류율	$\lambda'\alpha = 0$	$\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 1$ $\alpha_4 = 0, \alpha_5 = 0$
		검정력	$\lambda'\alpha = 1$	$\alpha_1 = 2, \alpha_2 = 1, \alpha_3 = 1$ $\alpha_4 = 0, \alpha_5 = 0$

3.2 모의실험의 과정

난수생성은 통계패키지 SAS/IML을 이용하였는데, 균등분포 난수는 RANUNI, 정규분포와 오염정규분포 난수는 RANNOR, 그리고 감마분포 난수는 RANGAM를 사용하여 생성되었다. 또한 REMLE와 MLE를 구할 경우에 두 추정량 모두 수렴조건을 만족하기 위한 최대반복횟수는 30번으로 제한하였고, k번째와 k+1번째 반복이 $|\hat{\theta}_{k+1} - \hat{\theta}_k| < 10^{-8}$을 만족할 때, k번째에서 추정량이 수렴하였다고 정의하였으며, 두 추정량이 수렴되지 않는 경우에는 제1종 오류와 검정력을 구하는 경우에 제외하였다. 그리고 MINQUE를 구하는 경우에 가장 많이 사용하는 사전추측값 1을 사용하였다. 두 가지 모형에 있어서 고려된 여러 가지 경우에 대하여 각각 3000개의 자료를 여러 요인들 사이의 교호작용 유무를 확인하기 위하여 2번 생성하였다.

한편 가설검정의 경험적 검정력은 각 모형에 있어서 대립가설을 이용하여 확률난수를 생성하고 분산성분을 4가지 방법을 이용하여 추정한 뒤 EGLS 검정통계량을 계산한 후에 검정통계량 값과 유의수준을 이용하여 채택여부를 결정하여 모든 시행이 끝난 뒤 경험적 검정력을 총기각수/총반복수로 구하였다. 제1종 오류율은 귀무가설을 이용하여 확률난수를 생성한 후 같은 절차로 구하였다.

3.3 EGLS의 경험적 검정력

3.3.1 대칭분포인 경우

모의실험을 통하여 구한 경험적 검정력을 관측치로 하여 검정력에 영향을 미치는 인자로 수준의 수, 자료의 불균형도, 랜덤효과와 오차항의 첨도, 분산성분비율의 값, 분산성분추정량의 종류를 사용하여 다원배치모형을 고려하여 분산분석을 통한 각 인자들의 유의성 검정을 하였다. 그 결과 뚜렷이 드러나는 교호작용은 존재하지 않았으며, 다음 <표 3>은 각 요인들에 대한 유의수준 0.05와 0.10을 이용한 제1종 오류율과 검정력의 값을 정리한 결과이다. <표 3>을 통하여 알 수 있는 중요한 결론들은 다음과 같다.

(a) 이원가법혼합모형

1. 랜덤효과와 오차항이 대칭분포를 따르는 경우에는 고려된 여러 가지 요인과 거의 무관하게 EGLS은 경험적 제1종 오류율의 값이 실제 유의수준의 값에 가까운 유의수준 α인 검정법인 것을 확인할 수 있다.
2. 자료의 개수가 고정되어 있는 경우에 수준의 수가 증가하면 제1종 오류율에는 영향을 크게 미치지 않으나 각 셀 당 반복수가 적어지기 때문에 검정력은 떨어진다.

3. 불균형도는 심할수록 제1종 오류율과 검정력이 다소 떨어지나 큰 차이는 없다. 또한 분산성분 비율의 값은 커질수록 제1종 오류율에는 큰 영향을 미치지 못하나 검정력은 많이 떨어진다.

<표 3> 대칭분포인 경우의 제1종 오류율과 검정력

모형의 종류	요인	조건	제1종 오류율		검정력	
			0.05	0.10	0.05	0.10
모형 1	수준의 수	(a,b)=(3,5)	.9451	.8915	.1892	.2787
		(a,b)=(5,3)	.9461	.8940	.1340	.2126
	불균형도	약함	.9459	.8929	.1667	.2517
		강함	.9452	.8927	.1566	.2396
	랜덤효과 (첨도값)	-1.2	.9456	.8933	.1476	.2304
		0	.9461	.8934	.1467	.2283
		9	.9451	.8917	.1906	.2783
	오차항 (첨도값)	-1.2	.9456	.8921	.1485	.2306
		0	.9457	.8936	.1288	.2104
		9	.9455	.8927	.2076	.2960
	분산성분 비율	0.25	.9460	.8938	.2700	.3744
		1	.9457	.8923	.1330	.2154
		4	.9451	.8923	.0819	.1473
	분산성분 추정량	헨더슨(III)	.9453	.8922	.1620	.2461
		REMLE	.9452	.8921	.1621	.2462
		MINQUE(1)	.9453	.8923	.1619	.2460
		MLE	.9466	.8945	.1605	.2443
모형 2	불균형도 (수준)	약함	.9602	.9185	.1664	.2464
		강함	.9596	.9193	.1494	.2257
	불균형도 (관측)	약함	.9594	.9183	.1587	.2400
		강함	.9604	.9194	.1571	.2320
	랜덤효과 (첨도값)	-1.2	.9598	.9194	.1488	.2271
		0	.9595	.9182	.1520	.2310
		9	.9604	.9190	.1729	.2499
	오차항 (첨도값)	-1.2	.9606	.9189	.1049	.1755
		0	.9589	.9179	.1196	.1961
		9	.9602	.9198	.2492	.3364
	분산성분 비율	0.25	.9583	.9154	.2493	.3507
		4	.9615	.9223	.0665	.1214
	분산성분 추정량	헨더슨(III)	.9562	.9162	.1793	.2585
		REMLE	.9646	.9257	.1402	.2152
		MINQUE(1)	.9654	.9273	.1375	.2116
		MLE	.9534	.9062	.1746	.2586

(b) 이원지분혼합모형

1. 수준 및 관측에 대한 불균형도는 모두 제1종 오류율과 검정력에 큰 영향을 미치지 못한다.
2. 랜덤효과 및 오차항의 첨도는 모두 제1종 오류율에 영향을 미치지 못하지만 첨도의 값이 커지면 검정력이 커지며, 분산성분비율의 값이 커지면 제1종 오류율과 검정력이 나빠진다.
3. 분산성분추정량의 선택은 MLE를 사용하는 것이 다른 추정량보다 제1종 오류율과 검정력의 측면에서 좋아짐을 확인할 수 있다.
4. 경험적 제1종 오류율은 고려된 실험계획 모두 유의수준 값을 상회하는 것을 확인할 수 있다.

<표 4> 비대칭분포인 경우의 제1종 오류율과 검정력

모형의 종류	요인	조건	제1종 오류율		검정력	
			0.05	0.10	0.05	0.10
모형 1	수준의 수	(a,b)=(3,5)	.9468	.8956	.1508	.2438
		(a,b)=(5,3)	.9470	.8971	.1077	.1858
	불균형도	약함	.9490	.9005	.1301	.2169
		강함	.9448	.8922	.1284	.2127
	분산성분 비율	0.25	.9460	.8952	.1705	.2691
		4	.9478	.8976	.0880	.1605
	랜덤효과 (왜도값)	2	.9476	.8985	.1282	.2087
		4	.9455	.8950	.1336	.2199
		6	.9476	.8956	.1260	.2158
	분산성분 추정량	헨더슨(III)	.9463	.8957	.1295	.2152
		REMLE	.9462	.8957	.1297	.2155
		MINQUE(1)	.9466	.8960	.1296	.2154
		MLE	.9484	.8981	.1283	.2131
모형 2	불균형도 (수준)	약함	.9611	.9168	.2394	.3210
		강함	.9585	.9138	.1241	.2014
	불균형도 (관측)	약함	.9594	.9133	.1295	.2099
		강함	.9601	.9174	.2339	.3125
	랜덤효과 (왜도값)	2	.9584	.9139	.1292	.2117
		4	.9593	.9137	.1939	.2736
		6	.9616	.9183	.2221	.2982
	분산성분 비율	0.25	.9583	.9153	.2770	.3670
		4	.9612	.9153	.0865	.1554
	분산성분 추정량	헨더슨(III)	.9531	.9037	.1973	.2827
		REMLE	.9651	.9246	.1700	.2453
		MINQUE(1)	.9654	.9255	.1689	.2429
		MLE	.9554	.9075	.1907	.2738

3.3.2 비대칭분포인 경우

3.3.1절에서는 EGLS 방법이 랜덤 및 오차항의 확률분포가 대칭분포를 따르는 경우에는 첨도에 영향을 많이 받지 않는다는 사실을 확인할 수 있었다. 따라서 왜도의 영향을 살펴보기 위하여 먼저 오차항은 정규분포를 따른다고 하고, 랜덤효과의 왜도 값을 변화시켰다. <표 4>는 비대칭분포를 사용한 두 가지 모형에 대한 모의실험 결과를 보여주며, 3.3.1절에서 얻은 결론들이 비슷하게 성립하는 것을 확인할 수 있다. 또한 이원가법모형에서는 왜도의 효과가 별로 없으나, 이원지분모형에서는 왜도의 값이 커질수록 제1종 오류율이 나빠짐을 알 수 있다. 이 경우 검정력의 값은 증가하는 데, 이것은 왜도의 값에 기인하는 것이 아니라 첨도의 값에 영향을 받는 것으로 간주되어진다.

<표 5>는 랜덤효과와 오차항이 모두 왜도가 0이 아닌 분포를 따르는 경우에 왜도가 제1종 오류율과 검정력에 미치는 영향을 보여준다. 이 경우 <표 4>에서 확인하였던 것처럼 한 개의 확률변수가 비대칭인 경우보다 왜도의 값이 커지면 커질수록 약간 더 검정력의 값을 크게 한다. 따라서 이원가법모형인 경우에는 유의수준 α를 하회하던 제1종 오류율의 값이 거의 유의수준과 비슷하게 되었고, 이원지분모형인 경우에는 유의수준 값으로부터 점점 더 멀어져 감을 알 수 있다. 나

머지 요인들인 수준의 수, 불균형도, 분산성분비율, 그리고 분산성분추정량의 선택은 앞의 경우하고 해석이 거의 흡사하므로 생략하였다.

<표 5> 랜덤효과의 왜도와 오차항의 왜도가 같은 경우

모형의 종류	요인	조건	제1종 오류율		검정력	
			0.05	0.10	0.05	0.10
모형 1	랜덤효과 (왜도값)	2	.9451	.8925	.1366	.2221
		4	.9502	.8964	.4123	.5163
		6	.9507	.8968	.6561	.7339
모형 2	랜덤효과 (왜도값)	2	.9591	.9140	.1349	.2168
		4	.9637	.9149	.4261	.5277
		6	.9644	.9185	.6667	.7382

4. 결론

본 연구에서는 고정효과의 유의성검정에 랜덤효과와 오차항은 비정규분포를 따른다는 가정을 사용하여 일반최소제곱 F-통계량이 적당한 지를 제1종 오류율과 검정력의 관점에서 모의실험을 통하여 살펴보았다. 그 결과 비록 제한된 모의실험의 결과이지만 고려된 두 가지 혼합모형 모두 제1종 오류인 경우에는 여러 가지 왜도와 첨도의 값의 변화에 큰 영향을 받지 않았다. 따라서 EGLS 검정은 비정규성에 대한 강한 강건성을 갖고 있으며, 실제 유의수준의 값에 가까운 유의수준 α인 검정법인 것을 확인할 수 있었다.

참 고 문 헌

[1] 이장택 (1997). 이원혼합모형에서 고정효과 유의성검정에 대한 검정력 분석, <응용통계연구>, 제10권 1호, 177-187.

[2] Lewis, S. L., Montgomery, D. C., and Myers, R. H. (2001). Examples of Designed Experiments With Nonnormal Responses, *Journal of Quality Technology*, Vol 33(3), 265-278.

[3] Khuri, A. I. (1987). Measures of Imbalance for Unbalanced Models, *Biometrical Journal*, 29, 383-396.

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

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

[2002년 9월 접수, 2003년 1월 채택]