2000 Proceedings of the Autumn Conference, Korean Statistical Society

오차항이 이분산성을 가지는 일원분류 모형에서 일반 F-검정의 유의수준에 관한 고찰

김기환*, 이준영†

요 약

일원분류 모형에서 표준 F-검정을 하기 위해서는 오차항에 대한 등분산성을 가정한다. 그러나 실제로 이러한 가정은 지켜지기 힘들며, 이에 더불어 관찰치가 각 집단별로 일정하지 않고 불균형한 경우에는 F-검정의 유의수준이 지정된 값을 만족시키지 못하며, 따라서 검정력에 관한 분석은 의미가 없게 된다. 본 연구에서는 등분산성이 지켜지지 않고, 자료가 불균형한 경우, 현실적인 상황에서 일반적으로 사용되는 F-검정의 유의수준 유지라는 문제가 어떤 변화를 겪게 되는지를 확인하고, 더 나아가 유의수준을 유지하기 위해서는 어떤 식의 조정이 가능한지를 살펴보았다.

주요용어 : 일원분류모형; 이분산성; 불균형 정도; 유의수준

1. 서론

실험계획법을 적용하는 많은 경우에, 주어진 확률모형의 오차항에 대하여 등분산성을 가정하지만, 실제로는 여러가지 이유로 인하여 이분산성을 갖게되는 경우가 많다. 오차항에 대한 등분산성의 가정이 깨어지는 경우, 그리고 이와 더불어 자료가 불균형한 경우에는 ANOVA 검정은 심각한 영향을 받게된다(Miller, 1986, Section 3.7). 고정효과 모형(fixed-effects model)하에서 이에 관한 연구가 많이 진행되어 왔으나(Brown and Forsythe, 1974; Bishop and Dudewicz, 1978; Levy, 1978a, 1978b; Krutchkoff, 1988; Weerahandi, 1995), 변량효과 모형(random-effect model) 또는 혼합효과 모형(mixed-effect model) 하에서의 연구는 그다지 많지 않다. Rao, Kaplan과 Cochran(1981)은 오차항이 이분산을 갖는 경우 일원분류 변량 모형(one-way random model)에서의 모수에 대한 점추정에 관한 연구를 하였으며, Jeyaratnam과 Othman(1985)는 Satterthewaite의 근사를 이용하여 $H_0 : \sigma_\alpha^2 = 0$에 대한 근사적인 검정법을 제공하였다. Singh 과 Joshi(1989)는 σ_α^2이 음의 값으로 추정될 확률에 대하여 오차항의 이분산성과 자료의 불균형정도(degree of imbalance)의 효과에 관한 연구

*고려대학교 통계연구소 연구원
†고려대학교 의과대학 예방의학교실 연구강사

를 하였다. 또한 Singh(1991)은 $H_0 : \sigma_\alpha^2 = 0$에 관한 F-검정의 검정력에 대해서도 동일한 연구를 실시하였다. Singh의 연구결과에 의하면 크기가 작은 집단이 큰 오차분산을 갖는 경우 이분산성은 검정력을 증가시키는 것으로 나타났으나, 그와는 반대로 크기가 큰 집단이 큰 분산을 갖는 경우는 검정력을 감소시키는 것을 보였다. 한편 Singh은 균형 자료의 경우에, 비록 적은 크기지만, 이분산성에 의한 검정력의 증가가 발생함을 보였다.

과거의 연구들은 주로 F-검정의 검정력에 관한 것이었다. 그러나 오차항이 이분산성을 갖는 경우 일반적으로 F-검정은 유의수준 $1-\alpha$를 유지하지 못하게 된다. 따라서 이런 상황에서 F-검정의 검정력에 관한 논의는 무의미할 수 있다. 본 연구에서는 자료가 불균형하고 오차항이 이분산성을 갖을 때 일원분류 모형에서 σ_α^2에 대한 F-검정이 유의수준 $1-\alpha$ 유지라는 관점에서 어떤 문제점들이 있는지를 살펴보고, 더 나아가 유의수준을 유지하기 위해서는 F-분포 표의 값이 어떻게 변화 되어야 하는지를 살펴보고자 한다.

2. F 검정에서의 유의수준

다음과 같은 일원분류 모형을 생각하자.

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij} \tag{1}$$

$i = 1, 2, \cdots, k;\ j = 1, 2, \cdots, n_i$. 여기서 α_i와 ϵ_i 는 각각 독립적으로 평균이 0이고 분산이 σ_α^2, $\sigma_i^2 (i = 1, 2, \cdots, k)$를 따른다고 가정한다. 즉 오차항의 분산은 각 고정된 i, 즉 각 집단 내에서는 같고, 서로 다른 i, 즉 집단간에는 서로 다른 분산을 갖게된다. 이 경우 모형 (1)은 다음과 같이 쓸 수 있다.

$$\mathbf{y} = \mu \mathbf{1}_N + \left(\bigoplus_{i=1}^{k} \mathbf{1}_{n_i} \right) \boldsymbol{\alpha} + \boldsymbol{\epsilon}, \tag{2}$$

여기서 $N = \sum_{i=1}^{k} n_i$, $\mathbf{1}_{n_i}$는 크기가 n_i, $i = 1, 2, \cdots, k$인 열벡터이며, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_k)'$이고, $\mathbf{y}$와 $\boldsymbol{\epsilon}$은 각기 관찰치와 오차로 구성된 벡터이다. 관찰치 벡터 $\mathbf{y}$의 분산-공분산 행렬을 $\boldsymbol{\Sigma}$라 할때 $\boldsymbol{\Sigma}$는 다음과 같이 표현된다.

$$\boldsymbol{\Sigma} = \sigma_\alpha^2 \bigoplus_{i=1}^{k} \mathbf{J}_{n_i} + \bigoplus_{i=1}^{k} (\sigma_i^2 \mathbf{I}_{n_i}). \tag{3}$$

여기서 $\mathbf{J}_{n_i}$는 크기가 $n_i \times n_i$이고 모든 원소가 1인 행렬이다. 귀무가설 $H_0 : \sigma_\alpha^2 = 0$에 대한 F-검정의 검정 통계량은

$$F = \left(\frac{N-k}{k-1} \right) \frac{\mathbf{y}'\mathbf{Q}\mathbf{y}}{\mathbf{y}'\mathbf{R}\mathbf{y}} \tag{4}$$

이고, $\mathbf{Q}$와 $\mathbf{R}$은 다음과 같이 정의된다.

$$\mathbf{Q} = \bigoplus_{i=1}^{k}\left(\frac{1}{n_i}\mathbf{J}_{n_i}\right) - \frac{1}{N}\mathbf{J}_N, \qquad \mathbf{R} = \mathbf{I}_N - \bigoplus_{i=1}^{k}\left(\frac{1}{n_i}\mathbf{J}_{n_i}\right).$$

식 (4)는 σ_i^2들이 모두 같다는 가정하에, 귀무가설하에서 자유도 $k-1$과 $N-k$를 갖는 F-분포에 따르게 된다.

식 (4)를 이용한 F-검정의 유의수준은 다음과 같이 정의되며,

$$P[F \leq F_{\alpha,k-1,N-k} | \sigma_\alpha^2 = 0] = 1-\alpha \tag{5}$$

이때 $F_{\alpha,k-1,N-k}$는 자유도 $k-1$, $N-k$를 갖는 F-분포의 상위 $\alpha\%$ 백분위 값이다. 식 (5)는

$$\begin{aligned} &P\left[\frac{N-k}{k-1}\frac{(\mathbf{y}'\mathbf{Q}\mathbf{y})}{(\mathbf{y}'\mathbf{R}\mathbf{y})} \leq F_{\alpha,k-1,N-k}|\sigma_\alpha^2 = 0\right] \\ &= P[\mathbf{y}'(\mathbf{Q} - h_\alpha\mathbf{R})\mathbf{y}|\sigma_\alpha^2 = 0] \end{aligned} \tag{6}$$

로 표현될 수 있으며, 여기서 $h_\alpha = \frac{k-1}{N-k}F_{\alpha,k-1,N-k}$이다. 모형에 의하여 벡터 $\mathbf{y}$는 정규분포를 따르므로 이차형식 $\mathbf{y}'(\mathbf{Q}-h_\alpha\mathbf{R})\mathbf{y}$는 다음과 같은 분포를 따르게 된다.

$$\mathbf{y}'(\mathbf{Q}-h_\alpha\mathbf{R})\mathbf{y} = \sum_{l=1}^{r}\lambda_l\chi_{\nu_l}^2 \tag{7}$$

여기서 $\lambda_1, \lambda_2, \cdots, \lambda_r$은 $(\mathbf{Q}-h_\alpha\mathbf{R})\Sigma_0$의 서로 다른 값을 갖는, 0이 아닌 고유치들 이며, 각 λ_i들의 개수는 $\nu_1, \nu_2, \cdots, \nu_r$이다. Σ_0는 귀무가설 하에서의 식 (3)의 값으로, $\sigma_\alpha^2 = 0$인 경우에 해당한다. 이때 $\sum_{l=1}^{r}\nu_l = rank(\mathbf{Q}-h_\alpha\mathbf{R})$을 만족하고, $\chi_{\nu_l}^2$는 자유도가 ν_l인 χ^2-분포에 따르게 된다(Johnson and Kotz, 1970). 그러므로 식 (5)는 다음과 같이 표현할 수 있다.

$$P[F \leq F_{\alpha,k-1,N-k}|\sigma_\alpha^2 = 0] = P\left[\sum_{l=1}^{r}\lambda_l\chi_{\nu_l}^2 \leq 0\right] \tag{8}$$

3. 자료의 불균형도

본 논문에서는 2절에서 언급한 F-검정의 유의수준 유지에 영향을 줄 수 있는 요인으로 오차항의 이분산성과, 자료의 불균형 정도를 고려한다. 일원분류 모형의 경우, 사용된 실험설계를 $D = \{n_1, n_2, \cdots, n_g\}$라 표현하면 여기서 g는 사용되는 그룹을 나타내며, $n_i, i = 1, 2, \cdots, g$는 각 그룹별 관찰치의 수를 나타낸다. 이 경우 자료의 불균형 정도(degree of data unbalancedness)를 수량화하려는 연구가 진행되어 왔으며(Hess, 1979; Ahrens and Pincus, 1981; Khuri, 1987; Lera Marques,1994), 특별히 Ahren과 Pincus는 일원분류 모형

에서의 자료의 불균형 정도를 측정하기 위한 척도(불균형도:measure of imbalance)의 하나로 다음을 제안하였다.

$$\phi = \frac{(\sum_{i=1}^{g} n_i)^2}{g\sum_{i=1}^{g} n_i^2} \tag{9}$$

ϕ는 $\frac{1}{g} \le \phi \le 1$을 만족하며, ϕ 값이 작을 수록 자료의 불균형 정도가 높은 것을 의미하고, 자료가 완전 균형을 이룰 때에는 $\phi = 1$이 된다.

한편 Khuri(1996)는 불균형도 ϕ와 집단 수 g 및 전체 자료 수 N이 주어졌을 때 $\sum_{i=1}^{g} n_i = N$과 $\sum_{i=1}^{g} n_i^2 = \frac{N^2}{g\phi}$의 관계로부터 주어진 불균형도를 만족하는 일원분류 모형의 실험 설계 $D = \{n_1, n_2, \cdots, n_g\}$를 생성하는 방법을 소개하였다. 다음 절에서는 이를 이용하여 집단 수 g가 고정된 경우 N, ϕ가 주어진 상황에서 생성된 D와 서로 다른 σ_i^2을 고려한 상황에서, 2절에서 언급한 오차항이 이분산을 갖는 F-검정의 유의수준 변화를 조사하도록 하겠다.

4. 예

실제 예를 보이기 위해, 본 연구에서는 $g = 5$인 경우, N, ϕ, 그리고 $\sigma_i^2 (i = 1, 2, \cdots, 5)$들 각각이 $N = 25, 50, 75$; $\phi = 0.35, 0.65, 0.95$; $\sigma_i^2 = 0.11, 1.0, 9.0$의 수준을 가지는 design들을 500개 생성하였다.

F-검정의 유의수준을 계산하기 위해서는 다음의 세 가지 단계를 거치게 된다. 첫 단계로 우선 관찰치의 크기 N, 불균형도 ϕ를 정하고 이를 만족시키는 D를 발생시키는 작업이 선행되어야 한다. 이후에 각 D에 상응하는 오차항의 이분산 정의한다. 두 번째 단계에서는 식 (8)을 구하기 위하여 고유치를 구한 후 실제로 확률 값을 구하게 된다. 이 때 실제 확률 값을 구하기 위하여 Davies(1980)의 알고리즘(AS155)을 이용하였다. Davies의 알고리즘은 식 (8)의 고유치에 의한 확률계산을 할 수있도록 되어 있으며, 인터넷을 이용하여 FORTRAN subroutine을 쉽게 구할 수 있다. 마지막 세 번째 단계는 실제로 구해진 확률 값이 $1 - \alpha$를 유지하지 못할 것이 예상되므로 역으로 유의 수준 α를 유지할 수 있는 $F_{\alpha,k-1,N-k}$를 찾아가는 단계이다. 세 가지 단계를 거쳐 계산된 결과의 일부가 표 (1)에 제시되어 있다.

참고문헌

Ahrens, H., and Pincus, R. (1981). On Two Measures of Unbalancedness in a One-way Model and Their Relation to Efficiency. *Biometrical Journal, 23,* 227-235.

Bishop, T.A., and Dudewicz, E. J. (1978). Exact Analysis of Variance with Unequal Variances: Test Procedures and Tables. *Technometrics, 20,* 419-430.

Brown, M. B., and Forsythe, A. B. (1974). The ANOVA and Multiple Comparison for Data with Heterogenous Variances. *Biometrics, 30,* 719-724.

Davies, R. B. (1980). The Distribution of a Linear Combination of χ^2 Random Variables. *Applied Statistics, 29,* 323-333.

Hess, J. L. (1979). Sensitivity of MINQUE with respect to A Priori Weights. *Biometrics, 35,* 645-649.

Jeyaratnam, S., and Othman, A. R. (1985). Test of Hypothesis in One-Way Random Effects Model with Unequal Error Variances. *Journal of the American Statistical Association, 69,* 789-794.

Johnson, N. L., and Kotz, S. (1970). *Continuous Univariate Distributions* -2. John Wiley, New York.

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

Khuri, A. I. (1996). A Method for Determining the Effect of Imbalance. *Journal of Statistical Planning and Inference, 55,* 115-129.

Krutchkoff, R. G. (1988). One-Way Fixed Effects Analysis of Variance when the Error Variances may be Unequal. *Journal of Statistical Computation and Simulation, 30,* 259-271.

Lera Marques, L. (1994). Measures of Imbalance for Higher Order Designs. *Biometrical Journal, 36,* 481-490.

Levy, K. J. (1978a). Some Emprical Power Results associated with Welch's Robust Analysis of Variance Technique. *Journal of Statistical Computation and Simulation, 8,* 43-48.

Levy, K. J. (1978b). An Empirical Comparison of the ANOVA F-test with Alternatives which are More Robust against Heterogeneity of Variance. *Journal of Statistical Computation and Simulation, 8,* 49-58.

Miller, R.G. (1986). *Beyond ANOVA.* John Wiley, New York.

Rao, P. S. R. S., Kaplan, J., and Cochran, W. G. (1981). Estimators for the One-Way Random Effects Model with Unequal Error Variances. *Journal of the American Statistical Association, 76,* 89-97.

Singh, B. (1991). Distribution of Variance Ratio in Unbalanced Random Model under Heterogeneity of Error Variances. *Communications in Statistics, Part A - Theory and Methods, 20,* 3021-3028.

Singh, B., and Joshi, D. D. (1989). On the Distribution of Components Sums of Squares in One-Way Unbalanced Random Model under Heterogeneous Error Variances. *Communications in Statistics, Part A - Theory and Methods, 18,* 4137-4144.

Weerahandi, S. S. (1955). ANOVA under Unequal Error Variances. *Biometrics, 51,* 589-599.

표 1: 일원분류 모형에서 오차항의 이분산과 불균형도에 의한 $H_0 : \sigma_\alpha^2 = 0$ 의 F-검정 유의수준의 변화. (일부분)

N	ϕ	σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_5^2	n_1	n_2	n_3	n_4	n_5	prob	Adj-F	α
25	.35	1.00	0.11	1.00	1.00	1.00	19	2	2	1	1	.0337	2.5223	.05
25	.35	1.00	0.11	1.00	1.00	1.00	18	3	2	1	1	.0410	2.6830	.05
25	.35	1.00	0.11	1.00	1.00	1.00	19	3	1	1	1	.0410	2.6830	.05
25	.35	1.00	9.00	1.00	1.00	1.00	19	2	2	1	1	.2675	8.1571	.05
25	.35	1.00	9.00	1.00	1.00	1.00	18	3	2	1	1	.1964	6.4252	.05
50	.35	1.00	0.11	9.00	1.00	9.00	37	6	5	1	1	.0166	1.8627	.05
50	.35	1.00	0.11	9.00	1.00	9.00	37	7	3	2	1	.0135	1.7677	.05
50	.35	1.00	0.11	9.00	1.00	9.00	37	5	5	2	1	.0166	1.8627	.05
50	.35	1.00	1.00	0.11	1.00	0.11	37	8	2	2	1	.3623	9.8165	.05
50	.35	1.00	1.00	0.11	1.00	0.11	36	10	2	1	1	.3623	9.8165	.05
75	.35	1.00	9.00	1.00	1.00	1.00	55	14	3	2	1	.0557	2.5841	.05
75	.35	1.00	9.00	1.00	1.00	1.00	54	16	2	2	1	.0627	2.6803	.05
75	.35	1.00	9.00	1.00	1.00	1.00	55	9	8	2	1	.0413	2.3694	.05
75	.35	1.00	0.11	1.00	1.00	1.00	55	12	6	1	1	.1306	4.0290	.05
75	.35	1.00	0.11	1.00	1.00	1.00	55	11	4	4	1	.1452	4.2995	.05
25	.65	1.00	1.00	1.00	9.00	0.11	11	7	4	2	1	.4807	12.6571	.05
25	.65	1.00	1.00	1.00	9.00	0.11	12	4	4	3	2	.3077	8.3212	.05
25	.65	1.00	1.00	1.00	9.00	0.11	11	6	5	2	1	.4807	12.6571	.05
25	.65	1.00	1.00	1.00	9.00	0.11	10	9	3	2	1	.4807	12.6571	.05
25	.65	1.00	1.00	1.00	9.00	0.11	11	7	4	2	1	.4807	12.6571	.05
50	.65	1.00	9.00	9.00	9.00	1.00	23	13	6	5	3	.0390	2.3593	.05
50	.65	1.00	9.00	9.00	9.00	1.00	22	14	9	3	2	.0506	2.5908	.05
50	.65	1.00	9.00	9.00	9.00	1.00	21	14	11	3	1	.0656	2.8619	.05
50	.65	1.00	9.00	9.00	0.11	1.00	23	13	7	4	3	.2967	8.8209	.05
50	.65	1.00	9.00	9.00	0.11	1.00	21	14	11	2	2	.4522	14.6041	.05
75	.65	1.00	1.00	1.00	9.00	1.00	35	14	13	12	1	.4188	11.0925	.05
75	.65	1.00	1.00	1.00	9.00	1.00	32	21	16	3	3	.3054	7.9734	.05
75	.65	1.00	1.00	1.00	9.00	1.00	35	18	9	8	5	.2664	7.0280	.05
75	.65	1.00	1.00	1.00	9.00	1.00	32	20	17	5	1	.3818	9.9948	.05
75	.65	1.00	1.00	1.00	9.00	1.00	33	21	13	5	3	.3131	8.1706	.05
25	.95	1.00	9.00	1.00	1.00	1.00	6	6	5	5	3	.0735	3.2972	.05
25	.95	1.00	9.00	1.00	1.00	1.00	6	6	6	4	3	.0735	3.2972	.05
25	.95	1.00	9.00	1.00	1.00	1.00	7	6	4	4	4	.0735	3.2972	.05
25	.95	1.00	0.11	1.00	1.00	1.00	6	6	5	4	4	.0729	3.4535	.05
25	.95	1.00	0.11	1.00	1.00	1.00	7	6	4	4	4	.0729	3.4535	.05
50	.95	1.00	1.00	9.00	1.00	9.00	13	12	9	9	7	.0466	2.5194	.05
50	.95	1.00	1.00	9.00	1.00	9.00	14	10	9	9	8	.0517	2.6085	.05
50	.95	1.00	1.00	9.00	1.00	9.00	13	11	11	8	7	.0574	2.7033	.05
50	.95	1.00	1.00	9.00	1.00	9.00	12	12	10	10	6	.0467	2.5194	.05
50	.95	1.00	1.00	9.00	1.00	9.00	14	10	10	8	8	.0574	2.7033	.05
75	.95	1.00	0.11	1.00	1.00	1.00	20	18	14	12	11	.0669	2.8391	.05
75	.95	1.00	0.11	1.00	1.00	1.00	18	18	16	14	9	.0669	2.8391	.05
75	.95	1.00	0.11	1.00	1.00	1.00	20	17	15	13	10	.0752	2.9978	.05
75	.95	1.00	0.11	1.00	1.00	1.00	18	18	17	13	9	.0669	2.8391	.05
75	.95	1.00	0.11	1.00	1.00	1.00	19	19	14	12	11	.0594	2.6926	.05