

▣ 응용논문

실험계획법에서 최소 표준화 검출 가능 효과의
크기에 관한 연구⁺

임용빈

이화여자대학교 통계학과

Study on the Size of Minimal Standardized Detectable
Difference in Balanced Design of Experiments

Yong Bin Lim

Dept. of Statistics, Ewha Womans University

Abstract

In balanced design of experiments sample size is determined to detect a certain standardized size of effect with power $1 - \beta$ at the level of significance α . Tables (Marvin, *et al* (1970) and Lorenzen and Anderson(1993)) and charts (Odeh and Fox(1991)) are available to determine the sample size in balanced design of experiments. To simplify those tables and charts simple MATLAB program is used to find the minimal standardized detectable difference Δ when α , β and sample size are given.

+ 이 논문은 1997년도 교육부 기초과학연구비 지원(과제번호 BSRI-97-1415)에 의한 연구결과임.

1. 서론

실험계획법에 관련한 통계상담에서 상담인이 결정해야 할 첫 번째 과제는 실험에 고려된 요인들의 수준들에 대한 랜덤화의 제약조건 등을 고려하여 실험목적에 부합되는 합리적인 실험계획을 제시하는 일이다. 그 다음으로 결정되어야 할 과제는 각각의 처리에서 실험의 반복의 크기, 즉 표본의 크기를 결정하는 것이다. 일반적으로 표본의 크기는 경제성을 고려하여 결정된다. 특히 신제품의 개발과 관련된 공업실험에서 고려해야 할 요인들의 수는 많고, 실험의 1회 실시에 비용이 많이 드는 경우에는 반복 1회, 혹은 가능한 실험조건들의 일부분에서만 실험을 실시하는 일부실시법을 이용하여 중요한 요인들을 선별하는 실험을 계획하고, 실험자료에 대한 분산분석후에 각 효과의 유의확률(p-value)의 크기에 근거하여 유의성 여부를 판단한다. 경험적으로 알려진 사실은(Lorenzen 등(1993)) 유의확률의 값이 큰($\geq .25$) 경우에는 효과가 존재하지 않는다고 판단하고 해당효과를 오차항에 풀링한다. 그런데 공업실험에서 실험의 비용이 크게 문제가 되지 않고 품질특성치에 영향을 주리라 기대되는 두 세 개의 요인들에 관한 자세한 정보를 얻기 위한 실험에서는 반복이 있는 실험을 계획한다. 의학, 농학, 식품영양학 등의 실험이 이에 해당된다. 이러한 실험에서 각각의 실험조건에서의 표본의 크기는 제1종오류의 확률 α , 실용적으로 차이가 있다고 판단되어서 검출하고자 하는 요인효과의 오차에 대한 상대적인 크기, 그 값에서의 제2종오류의 확률 β 에 따라서 결정된다. 즉, 관심이 있는 효과가 유의하지 않다는 결정을 내릴 때에 실험자는 유의확률의 크기에 의존하는 경험적인 판단이 아닌, 제 2종 오류의 크기가 제어된 결정을 내릴 수 있게 된다.

일원배치법에서의 검정력은 비중심 F-분포로부터 계산되고 표본크기는 요인의 수준수, 제1종오류의 확률 α , 제2종오류의 확률 β , 비중심모수의 값에 따라 결정된다. Marvin 등(1970)은 비중심모수의 값 대신에 실용적으로 차이가 있다고 판단되는 수준효과들의 표준화된 범위의 값을 가지고 실험자가 쉽게 표본의 크기를 결정할 수 있는 표를 만들었다. 이 표는 현재까지도 일원배치법의 표본의 크기를 결정하는데 많이 이용되고 있다. 요인배치법(Factorial Experiment)과 분할법(Split Plot Design) 등과 같은 균형 실험계획법에서의 표본크기의 결정은 일원배치법의 경우보다는 복잡하다. 요인배치법의 경우에 Odeh 등(1991)은 α, β 가 주어진 경우에 검출하려는 효과의 최소크기를 만족하는 표본의 크기를 반복적으로 차트(chart)를 이용하여 결정하는 방법을 제시하였다. 일원배치법의 경우에 Lorenzen 등(1993)은 $\alpha=0.05, \beta=0.10$ 인 경우에 F-검정의 분자, 분모 자유도에 대해서 최소검출효과의 크기를 구할 수 있는 Marvin 등(1970)의 표를 활용하여 Odeh 등 (1991)의 차트를 이용한 방법보다 표본의 크기를 쉽게 결정하는 방법이 제시되었다.

Numeric Computation and Visualization Software)로 전자공학, 경영학(특히, Finance), 통계학 등에서 최근 활발하게 이용되는 소프트웨어이다. MATLAB에 관심이 있는 독자들은 MATLAB User's Guide(1995)를 참조하기 바란다. 이 논문의 목적은 비중심 F-분포를 계산하기 위한 알고리즘인 Norton(1983)과 Narula 등(1986)의 결과를 이용하여 최소검출효과의 크기를 구하는 프로그램을 MATLAB으로 작성하는 것이다. 이 프로그램은 Marvin 등(1970)과 Lorenzen 등(1993)에 주어진 표를 일반화하고 Odeh 등 (1991)에 주어진 약 110쪽의 챕터를 대체하여, 요인배치법의 경우에 표본의 크기를 쉽게 결정하는 것을 도와준다. 랜덤요인(Random Factor)을 포함하는 균형된 실험계획(Balanced Design of Experiments)에서의 표본크기의 결정에도 이 프로그램이 활용될 수 있어서 이에 관한 연구도 진행되고 있고 연구결과도 곧 발표될 예정이다. 2절에서는 용어와 알고리즘을 소개하고, 3절에서는 Odeh 등(1991)에 소개된 예를 들어서 MATLAB 프로그램을 이용하여 표본의 크기를 결정하는 방법을 설명한다. MATLAB 프로그램은 부록에 첨가한다.

2. 용어와 알고리즘

균형된 실험계획법 모형에서 모수효과(Fixed Effect) A의 유의성 검정을 생각하자. F-검정의 분모에 오는 요인효과를 X로 표시하고 각각의 자유도를 ν_1, ν_2 라 하자. A의 수준효과들의 산포의 크기를 $\Phi(A) = \sum_i A_i^2 / \nu_1$, X의 평균제곱의 기대값 $E(MS_X)$ 을 σ_X^2 , A의 평균제곱의 기대값 $E(MS_A)$ 에서 $\Phi(A)$ 의 계수인 A의 각 수준에서의 반복수를 c로 표시한다. 일반적으로 c와 ν_2 는 각 실험조건에서의 표본의 크기인 n에 종속한다. A의 수준효과들의 산포의 크기 $\Phi(A)$ 에서의 검정력은 자유도 ν_1, ν_2 비중심 모수의 값이

$$\lambda = \frac{c\nu_1\Phi(A)}{2\sigma_X^2}$$

인 비중심 F-분포의 분포함수에 의해 결정된다.(Tiku(1967), Marvin 등(1970)) 자유도 ν_1, ν_2 인 F-분포의 꼬리의 확률이 α 가 되는 값을 $F_\alpha(\nu_1, \nu_2)$ 로 표시할 때, 유의수준 α 의 기각역은 $\{F: F > F_\alpha(\nu_1, \nu_2)\}$ 이다. 실용적으로 차이가 있다고 판단되어 실험자가 검출하고자 하는 A의 수준효과의 산포의 크기를 $\Phi^*(A)$ 라 하고, 이때의 비중심 모수의 값을 λ^* 라 할 때에, λ^* 에서 기각역에 속할 확률인 검정력이 적어도 $1 - \beta$ 가 되도록 표본의 크기 n이 결정되어야 한다. 이를 위해서는 비중심 F-분포의 분포함수

를 계산하고, 그 다음에 제2종오류의 확률인 $F_\alpha(\nu_1, \nu_2)$ 에서의 비중심 분포함수의 값이 β 를 만족하는 비중심 모수의 값 f_λ 를 구한다. $\Phi^*(A)$, σ_X^2 와 ν_1 은 표본의 크기 n 에 종속되지 않고 c 가 표본의 크기 n 의 증가함수여서 비중심 모수의 값 λ 는 n 의 증가함수이다. $F_\alpha(\nu_1, \nu_2)$ 에서의 비중심 분포함수의 값이 비중심 모수의 감소함수임을 주시할 때에

$$f_\lambda \leq \frac{c\nu_1 \Phi^*(A)}{2\sigma_X^2}$$

를 만족하는 최초의 n 이 표본의 크기이고, $\Phi^*(A)$ 에서 제2종오류의 확률이 처음으로 β 보다 작게 된다. 따라서 표본의 크기 n 을 결정하기 위해서, 먼저 검출하고자 하는 표준화된 효과 A 의 크기

$$\Delta^* = \sqrt{\frac{\Phi^*(A)}{\sigma_X^2}}$$

를 구하고, n 을 증가시키면서 최초로

$$\sqrt{\frac{2f_\lambda}{c\nu_1}} \leq \Delta^*$$

를 만족하는 n 을 찾는다. 특히 주효과에 관심이 있을 때에, 실험자는 검출하고자 하는 효과의 크기를 Δ^* 보다는 표준화된 각 수준에서의 평균의 범위인

$$\Delta_1^* = \frac{(\max A_i - \min A_i)}{\sigma_X}$$

로 표현하기가 쉽다고 생각한다(Marvin 등(1970), Lorenzen 등(1993)). 이 때, Δ^* 와 Δ_1^* 의 관계는

$$\Delta_1^* \leq \Delta^* \sqrt{2\nu_1}$$

이고 등호는 $\max A_i$ 와 $\min A_i$ 를 제외한 모든 다른 수준에서 $A_i = (\max A_i + \min A_i)/2$ 인 경우에 성립한다.

비중심 F-분포함수를 계산하는 알고리즘과 관련된 참고문헌은 Tiku(1967, 1972), Marvin 등(1970), Norton(1983) 등이다. Norton(1983)의 알고리즘은 다음과 같이 요약된다. 비중심 베타분포의 분포함수는

$$B(x:a,b;\lambda) = e^{-\lambda} \sum_{k=0}^{\infty} (\lambda^k / k!) B(x:a+k,b),$$

여기서 $0 \leq x \leq 1$, $a > 0$, $b > 0$ 이고 $B(x:a+k,b)$ 는 모두 $a+k$, b 인 베타분포의 분포함수이다. 비중심 베타분포의 분포함수 $B(x:a,b;\lambda)$ 는

$$B_m(x:a,b;\lambda) = e^{-\lambda} \sum_{k=0}^m (\lambda^k / k!) B(x:a+k,b),$$

로 근사될 수 있다. 여기서 m 은 오차의 크기

$$e_m(x:a,b;\lambda) = B(x:a,b;\lambda) - B_m(x:a,b;\lambda)$$

이 충분히 작아지도록 ($< 10^{-8}$) 결정한다. 비중심 F-분포의 분포함수는

$$F(w:\nu_1, \nu_2; \lambda) = B(x:a,b;\lambda),$$

식에 의해서 비중심 베타분포의 분포함수로부터 계산된다. 여기서 $a = .5\nu_1$, $b = .5\nu_2$ 이고 $x = \nu_1 w / (\nu_1 w + \nu_2)$ 이다. Norton(1983)의 알고리즘은 부록에 첨가된 비중심 F-분포의 누적분포함수를 계산하는 MATLAB function subroutine인 noncentf.m에 이용되었다.

3. MATLAB 프로그램과 적용 예제

예3.3에서 유의수준 $\alpha = 0.05$, 뇌손상 주효과의 표준화된 산포의 크기 $\Delta^* = 1$ 에서의 검정력이 적어도 0.8 이상이기 위한 표본의 크기의 결정을 위한 MATLAB 프로그램이 부록에 첨가되어 있다. data.m file에서는 요인 A의 수준수 l1=3, 요인 B의 수준수 l2=2, 제1종오류의 확률 alpha=0.05, 제 2종오류의 확률 beta=0.2, 뇌손상 주효과의 표준화된 산포의 크기 Deltao=1을 입력한다. MATLAB 프로그램인 size.m file에서는 data를 읽고, function noncentf를 이용하여 $F_\alpha(\nu_1, \nu_2)$ 에서의 비중심 F-분포의 누적

확률을 계산한다. function flambda는 $F_\alpha(\nu_1, \nu_2)$ 에서의 비중심 F-분포의 누적 확률에서 β 를 뺀 값을 비중심모수 λ 의 함수로 정의해서 MATLAB root finding subroutine인 fzero를 이용하여 $F_\alpha(\nu_1, \nu_2)$ 에서의 비중심 F-분포의 누적 확률인 제 2종오류의 확률이 β 가 되는 f_1 을 찾는다. 그 다음에 주어진 표본의 크기 n 에서의 검출 가능한 최소 표준화된 산포의 크기 Δ 가 목표치 Δ^* 보다 작게 되는 최초의 표본의 크기 n 를 구한다. MATLAB 프로그램을 실행하기 위해서는 MATLAB Path에 지정된 directory에 size.m, data.m, noncentf.m, flambda.m을 저장하고 MATLAB Command Window에서 size를 실행하면 검출 가능한 최소 표준화 산포의 크기 Δ 가 Δ^* 보다 작게 되는 최초의 표본의 크기 $n=4$ 와 Δ 값인 Delta=0.8454가 출력된다.

주어진 표본의 크기 n 에서 Marvin 등(1970)에 주어진 표와 Odeh 등(1991)에 주어진 차트들로부터 얻어지는 표준화된 최소 검출효과의 크기는 부록에 첨가된 MATLAB 프로그램의 약간의 수정을 통해서 얻어질 수 있다.

일원배치법과 이원배치법의 경우에 Odeh 등(1991)에 소개된 예를 들어서 MATLAB 프로그램을 적용하여 표본의 크기를 결정해 보자.

예3.1 표준양궁시험은 30, 40, 50 야드에서 각각 6발을 발사하여 순위를 결정한다. 6회의 양궁강습을 실시하는데, 각 강습마다 발사거리를 달리한다. 3종류의 강습순서를 결정한 후에 동수의 사람들을 각각의 그룹에 랜덤하게 배치한 후에, 해당 순서에 따라서 강습을 실시한 후에 50야드 거리에서의 점수를 측정한다. 이 실험에서 3 그룹간의 평균점수에 차이가 없다는 가설검정을 할 때에, 유의수준 $\alpha=0.025$, 표준화된 산포의 크기 $\Delta^*=\sqrt{1/2}$ 에서 검정력이 적어도 0.7, 즉 $\beta=0.3$ 을 만족하도록 표본의 크기 n 을 구하라.

(풀이) $\nu_1=2$ $\alpha=0.025$, $\beta=0.3$, $\Delta^*=\sqrt{1/2}$ 를 입력한다. 3수준 일원배치법실험이기에 $\nu_2=3(n-1)$, $c=n$ 이다. 따라서 MATLAB 프로그램 size.m을 약간 수정하여 for loops 문장을 사용하여 n 을 점차적으로 증가시키면서 검출 가능한 최소 표준화 산포의 크기 Δ 가 Δ^* 보다 작게 되는 최소의 n 을 찾는다. 프로그램의 실행결과, 최초의 $\Delta=0.6953$, $n=11$ 이어서 표본의 크기는 11이 되고, 이는 Odeh 등(1991)의 예 2.2.2에서 차트를 반복적으로 적용해서 구한 $n=11$ 과 일치한다.

예3.2 4종류의 마그네시움 알약의 분해되는 평균속도에 차이가 있는지에 관심이 있어서 일원배치 완전확률화계획법에 의해서 실험을 실시하려 한다. 과거의 자료에 의하면 $\sigma=0.6$ 정도이고 유의수준 $\alpha=0.05$ 에서, 평균의 범위가 4초인 경우의 검정력

이 적어도 0.9 이기를 원한다. 표본의 크기를 구하라.

(풀이) $\nu_1 = 3$, $\alpha = 0.05$, $\beta = 0.1$, $\Delta_1^* = 4/1.6$ 을 입력한다. 4수준 일원배치법실험이기에 $\nu_2 = 4(n-1)$, $c = n$ 이다. 수정된 MATLAB 프로그램을 실행시킨 결과 검출 가능한 최소 표준화 산포의 크기 Δ_1 가 Δ_1^* 보다 작게 되는 최초의 $\Delta_1 = 2.4007$, $n = 6$ 이어서 표본의 크기는 6이 되고, 이는 Lorenzen 등(1993)의 2.6절에서 구한 $n=6$ 과 일치한다.

예3.3 쥐의 뇌의 손상이 기억력에 미치는 영향을 연구하는 실험에서, 2가지 방법에 의해서 미로를 벗어나는 훈련을 시킨다. 훈련을 받은 쥐의 3가지 위치에 뇌의 손상을 가하는 수술을 한 후에 미로를 벗어나는데 걸린 시간을 측정한다. 유의수준 $\alpha = 0.05$, 뇌손상 주효과의 표준화된 산포의 크기 $\Delta^* = 1$ 에서의 검정력이 적어도 0.8, 뇌손상과 훈련방법 교호작용효과의 표준화된 산포의 크기 $\Delta^* = \sqrt{1.5}$ 에서의 검정력이 0.9 이상이기 위한 표본의 크기를 구하라.

(풀이) 뇌손상 주효과; $\nu_1 = 2$, $\alpha = 0.05$, $\beta = 0.2$, $\Delta^* = 1$ 를 입력한다. 3x2 이원배치법실험이기에 $\nu_2 = 6(n-1)$, $c = 2n$ 이다. 수정된 MATLAB 프로그램을 실행시킨 결과 검출 가능한 최소 표준화 산포의 크기 Δ 가 Δ^* 보다 작게 되는 최초의 $\Delta = 0.8454$, $n = 4$ 이어서 표본의 크기는 4가 되고, 이는 Odeh 등(1991)의 예 2.3.1에서 쳐트에 의해서 구한 $n=4$ 와 일치한다.

교호작용효과: $\nu_1 = 2$, $\alpha = 0.05$, $\beta = 0.1$, $\Delta^* = \sqrt{1.5}$ 를 입력한다. 3×2 이원배치법실험의 교호작용효과이기에 $\nu_2 = 6(n-1)$, $c = n$ 이다. 수정된 MATLAB 프로그램을 실행시킨 결과 검출 가능한 최소 표준화 산포의 크기 Δ 가 Δ^* 보다 작게 되는 최초의 $\Delta = 1.1986$, $n = 5$ 이어서 표본의 크기는 5가 되고, 이는 Odeh 등(1993)의 예 2.3.1에서 쳐트에 의해서 구한 $n=5$ 와 일치한다.

주효과와 교호작용효과의 검정력을 동시에 만족시켜야 하기에 이 실험의 표본의 크기는 $n=5$ 가 된다.

4. 결론 및 요약

요인배치법에서의 표본크기는 유의수준 α , 검출하고자 하는 효과의 표준화된 산포

크기에서의 검정력 $1 - \beta$ 에 따라서 결정된다. 표본의 크기 n 이 주어진 경우에, Marvin 등(1970)과 Lorenzen 등(1993)에는 $\alpha = 0.05$, $\beta = 0.10$ 에서 검출가능한 효과의 최소 표준화된 산포의 크기를 구할 수 있는 표가 주어지고, Odeh 등(1991)에서는 α , β 가 주어진 경우에 차트를 이용하여 최소 표준화된 산포의 크기가 결정될 수 있다. 이 논문에서는 비중심 F-분포를 계산하기 위한 알고리즘인 Norton(1983)과 Narula 등(1986)의 결과를 이용하여 최소검출효과의 크기를 구하는 프로그램이 MATLAB으로 작성되었고, 이 프로그램은 Marvin 등(1970)과 Lorenzen 등(1993)에 주어진 표를 일반화하고 Odeh 등(1991)에 주어진 약 110쪽의 차트를 대체하여, 요인 배치법에서 표본의 크기를 결정하는 것을 도와준다. 또한 Odeh 등(1991)에 소개된 예를 들어서 MATLAB 프로그램을 이용하여 표본의 크기를 결정하는 방법이 설명되어서 유사한 공업실험에서도 표본의 크기가 쉽게 결정될 수 있을 것이다.

부록 MATLAB 프로그램

size.m file:

```
% MATLAB main program for the sample size
% in two way ANOVA model, eg. 3.3.
% l1: levels of fixed factor A
% l2: levels of dixed factor B

% testing significance of effect A; F=MS_A / MS_X
% n: replication of experiment
% nul: df for numerator, effect A
% nu2: df for denominator corresponding to term X
%      : function of rep in general
% c : coef. of preceding the Phi(A) term, a dispersion measure
%      of the effect of levels of A, in EMS
%      :function of rep in general

% alpha: probability of type I error
% Delta: mimimal detectable difference, sqrt(Phi(A))/sqrt(EMS(X))
% Deltao: target mimimal detectable difference, sqrt(Phi(A))/sqrt(EMS(X))
% Delta1: minimal detectable difference, range(A_i)/sqrt(EMS(X))
% Deltalo: target minimal detectable difference, range(A_i)/sqrt(EMS(X))
% beta: probability of type II error at Deltao
```

```
% input l1, l2, alpha, beta, Deltao
data

nu1=l1-1

for n=2:100
    nu2=l1*l2*(n-1);
    c=l2*n;

    % fc: (1-alpha) percentile of F(nu1,nu2) distn
    % x(1): lower bound of noncentrality parameter, fl
    % x(2): upper bound of noncentrality parameter, fl

    fc=finv(1-alpha,nu1,nu2);

    % Find an upper bound of noncentrality parameter
    ul=1;
    while flambda(ul,fc,nu1,nu2,beta)>0,
        ul=ul+1;
    end
    x(1)=0;
    x(2)=ul;
    % flambda(x(1),fc,nu1,nu2,beta)
    % flambda(x(2),fc,nu1,nu2,beta)

    % Find the exact value of noncentrality parameter fl
    % where the probability of type II error is beta
    fl=fzero('flambda',x,[],[],fc,nu1,nu2,beta);

    % Find the minimal detectable diff., Delta
    Delta=sqrt(2*fl/(c*nu1));
    Delta1=Delta*sqrt(2*nu1);
    if Delta <= Deltao, break, end
end
n
Delta
```

data.m file:

```
l1=3;
l2=2;
alpha=0.05;
beta=0.2;
Deltao=1;
```

noncentf.m file:

```
function out=noncentf(fc,nu1,nu2,fl)
% cdf of non-central F distribution of df nu1, nu2 & non-central. par. fl
a=.5*nu1;
b=.5*nu2;
x=a*fc/(a*fc+b);
tem=betacdf(x,a,b);
r=fl;
k=1;
p=betacdf(x,a+k,b);
while r*p*min(1,exp(-fl)+fl/(k+1)) > 1.e-8
    p=betacdf(x,a+k,b);
    tem=tem+r*p;
    r=r*fl/(k+1);
    k=k+1;
end
out=exp(-fl)*tem;
```

flambda.m file

```
function out=flambda(fl,fc,nu1,nu2,beta)
% cdf of non-central F distn - beta as a function of non-central. par. fl
out=noncentf(fc,nu1,nu2,fl)-beta;
```

참고문헌

- [1] Lorenzen, T.J. and Anderson, V.L.(1993) *Design of Experiments: A No-Name Approach*, Marcel Dekker Inc, New York.

- [2] Marvin, A., Kastenbaum, A. and Hoel, D.G.(1970) "Sample size requirements: one-way analysis of variance," *Biometrika*, Vol. 57, No. 2, pp. 421-430.
- [3] Narula S.C. and Weistroffer, H.R.(1986) "Computation of probability and non-centrality parameter of a non-central F-distribution," *Commun. Statist.-Simula.*, Vol. 15, No. 3, pp. 871-878.
- [4] Norton, V.(1983) "A simple algorithm for computing the non-central F distribution," *Appl. Statist.*, Vol. 32, No. 1, pp. 84-85.
- [5] Odeh, R.E. and Fox, M.(1991) *Sample Size Choice: Charts for Experiments with Linear Models, second edition*, Marcel Dekker Inc, New York.
- [6] The MATHWorks, INC.(1992) *MATLAB User's Guide*, The MathWorks, Inc., Mass.
- [7] Tiku, M.L.(1967) "Tables of the power of the F-Test," *Journal of the American Statistical Association*, Vol. 62, pp. 525-539.
- [8] Tiku, M.L.(1972) "More tables of the power of the F-Test," *Journal of the American Statistical Association*, Vol. 67, pp. 709-710.