응용통계연구 제10권 1호
1997년 3월 pp. 61-68

(M, S)-최적인 동반수 h($h \geq 2$)을 갖는 가장 균형된 계획에 관한 연구[1)]

배 종 성[2)]

요 약

본 논문에서는 블록의 크기와 반복수가 동일한 이항계획인 경우 동반수 $h(h\geq 2)$ 을 갖는 가장 균형된 계획은 (M, S) -최적계획임을 보인다. 동반수가 3이상인 경우는 배치계획을 작성하기가 매우 힘들기 때문에 기존의 PBIBD(3)의 배치계획들 중에서 동반수 3을 갖는 가장 균형된 배치계획들만을 조사하였다.

1. 서론

Bose와 Nair(1939)에 의해 불완비 블럭계획의 일반화된 경우로 상반부류수가 2인 부분적으로 균형된 불완비 블럭계획(Partially Balanced Incomplete Block Design ; PBIBD(2))이 소개된 이래 PBIBD(2)는 충분할 정도로 연구가 되었다. 그러나 상반부류수가 3 이상인 PBIBD($h \geq 3$) 는 주어진 모수에 따르는 배치계획을 구성하기가 힘들기 때문에 연구가 많이 이루어 지지 않았다(John,1987,pp.42). 임의의 쌍의 처리가 블럭에서 만나는 횟수를 λ_i라할 때 Takeuchi(1961)가 $\lambda_2=\lambda_1+1$인 그룹 분해 가능 계획(group divisible design)은 E-최적임을 보인 이래 Cheng(1978,1980), Jacroux(1980)등에 의해 동반수가 2인 경우의 연구가 수행되었다. Jarrett(1983)은 PBIBD의 정의와 비슷한 방법으로 동반수 h을 갖는 계획(h- concurrence design)을 정의하고 λ_i가 연속적인 두개의 정수 값을 갖는 정규 그래프 계획 (regular graph design)은 (M, S)-최적임을 보였다. John 과 Michell(1977)은 불완비 블럭계획이 D-최적(혹은 A-최적, 혹은 E-최적)이면 그 계획은 정규 그래프 계획일 것이라고 추측하였다. John 과 Williams(1982)는 (M, S)-최적이면서 A-최적인 계획은 D-최적이고, (M, S)-최적이면서 D-최적인 계획은 A-최적일 것이라고 추측하였다. Sinha와 Shah(1988)는 동반수 3을 갖는 급간과 급내에서 가장 균형된 그룹 계획(3-concurrence most balance intra-and inter group balanced design)은 주어진 모수를 갖는 계획 중에서 E-최적임을 보였지만 일반적인 배치계획 작성 방법을 보여주지 못하고 단지 예를 1개 보였다. Sinha와 Kageyama(1992)는 Sinha와 Shah의 계획 중에서 그룹의 크기가 모두 동일한 경우, 동반수 3을 갖는 가장 균형된

1) 이 논문은 1995년도 전남대학교 학술연구비 지원에 의하여 연구되었음.
2) (500-757) 광주광역시 북구 용봉동 300 전남대학교 통계학과 교수

계획(3-concurrence most balanced design)중에서 $\lambda_2 > \lambda_3 > \lambda_1$인 경우가 E-최적임을 보였고, 배치계획을 구성하는 일반적인 방법을 2가지 제시하고 이에 따라서 14개의 배치계획을 작성하였다. 그러나 14개의 배치계획 중 블럭 수가 30 이하인 것은 단지 2개 밖에 없었다. 따라서 Sinha와 Kageyama가 제시한 배치계획들은 실험계획의 실제응용이라는 측면에서 매우 회의적이다.

본 논문에서는 정규 그래프 계획을 일반화하여 단지 λ_i가 연속적인 $h(h \geq 2)$개의 정수를 갖는 동반수 h인 가장 균형된 계획은 (M,S)-최적임을 보이 고자 한다. John이 지적한 바와 같이 동반수가 3이상인 경우는 배치계획 구성이 매우 힘들기 때문에 기존의 PBIBD(3)들 중에서 동반수 3을 갖는 가장 균형된 배치계획들을 찾아 표로써 제시하였다. 따라서 이번 연구에서는 블럭의 크기와 반복 수가 동일한 이항계획만으로 제한하고 응용이라는 면에서 블럭 수나 처리 수가 30 이하인 배치계획들로 제한하였다.

2. 최적 계획

D(v, b, k, r)을 처리수 v, 블럭수 b, 블럭크기 k, 반복수 r을 갖는 블럭의 크기와 반복수가 동일한 이항계획(proper equi-replicate binary design)을 나타낸다고 하자. D (v, b, k, r)의 처리효과를 추정하기 위한 축소된 정규 방정식은

$$C = \gamma I - (k)^{-1} NN'$$

이고, 여기서 I는 크기 $v*v$인 단위행렬, N는 크기 $v*b$ 인 블럭 결합행렬, N'는 N의 전치행렬이다. 대칭이고 비음정치인 C 행렬의 열의 합은 0이다. 이제 D(v, b, k, r)의 정준효율을 나타내는 행렬

$$A = I - (r.k)^{-1} NN' \tag{2.1}$$

의 0이 아닌 특성근을 e_1, e_2, $\cdots$, e_{v-1}이라면

$$\sum e_i = trace(A) = v(k-1)/k$$

이고

$$\sum e_i^2 = trace(A^2) = v - 2vr/k + (vr^2 + \sum\sum \lambda_{ij}^2)/k^2 \tag{2.2}$$

이다. 여기서 λ_{ij}는 i번 처리와 j번 처리를 포함하는 블럭 수를 나타낸다. e_i를 정준효율 인자(canonical efficiency factor)이라 하고 e_i에 기초를 둔 여러가지 최적기준이 연구 되었다. 아래의 θ_i를 최대로 하는 기준에 따라서

A-최적 : $\theta_1 = (v-1)/\sum e_i^{-1}$, 정준효율 인자의 조화평균을 최대로 하는 즉 모든 처리 쌍 비교들의 분산의 평균을 최소로 하는 기준이다.

E-최적 : $\theta_2 = \min e_i$, 정준효율 인자가 최소인것이 최대로하는 즉 임의의 처리 대비 분산을 최소로 하는 기준이다.

D-최적 : $\theta_3 = (\prod e_i)^{1/(v-1)}$, 정준효율 인자의 기하평균을 최대로하는 즉 일반화된 분산을 최소로 하는 기준이다.

A, D, E-최적 외에 다른 유용한 최적기준이 Shah(1960), Ecclecton과 Hedayet(1974)에 의해 소개된 2 단계로 구성된 (M, S)－최적이다. (M, S)－최적은 정준효율 인자의 합을 최대로 하는 즉 $\sum e_i$를 최대로 하는 M-최적 부분과, e_i의 평균 $\bar{e} = v(k-1)/k(v-1)$가 주어졌을 때 $\sum e_i^2$을 최소로 택하는 S-최적 부분 즉 정준효율 인자의 분산을 최소로 하는 최적기준이다. Jarrett에 의하면 반복수가 동일한 이항계획은 M-최적조건을 만족하고 S-최적 기준으로 $\theta_4 = \bar{e}^2(v-1)^{1/2}/\sum e_i^2$를 사용했다.

3. 동반수 $h(h \geq 2)$을 갖는 가장 균형된 계획

우리는 Sinha와 Shah가 정의한 가장 균형된 계획의 정의를 임의의 동반수 h로 확장하여 다음과 같이 정의한다.

정의 1. 정규 그래프 계획을 λ_h로 연장한 계획 즉 λ_i가 연속적인 $h(h \geq 2)$개의 정수 값을 갖는 계획을 동반수가 h인 가장 균형된 계획이라 한다.

블럭의 크기와 반복수가 동일한 이항계획 중에서 λ_i의 갯수와 위치에 관계없이 임의의 동반수 h을 갖는 가장 균형된 계획이 (M, S)－최적임을 보이려한다.

정리 1. D(v, b, k, r)중 동반수가 $h(h \geq 2)$인 가장 균형된 계획은 (M, S)－최적 계획이다.

증명. (M, S)－최적 계획임을 보이려면 M-최적과 S-최적을 만족함을 보여야된다.

1) M-최적; $\sum_{i=1}^{v-1} e_i$가 최대임을 보여야 한다. 그런데 D(v, b, k, r)은 Jarrett에 의해서 $\sum e_i$는 최대가 됨이 보여졌다.

2) S-최적; $\sum e_i^2$이 최소임을 보여야한다. (2.2) 식에서 $\sum \sum_{i \neq j} \lambda_{ij}^2$ 값이 최소이면 $\sum e_i^2$ 값

도 최소가 된다. λ_{ij}는 크기 $v*v$인 NN'의 원소이다. NN'의 특성으로부터 λ_{ij}는 행에따라 원소의 위치가 달라질 뿐 원소들의 구조(형태)는 같다는 사실로 부터

$$\sum\sum_{i\neq j}\lambda_{ij}^2 = v\sum_{i=1}^{h} n_i\,\lambda_i^2,$$

임을 쉽게 알수 있다. 여기서 n_i는 λ_i의 갯수이다. λ_i의 값을 올림차순으로 $\mu+a_1$, $\mu+a_2$, $\mu+a_3$, $\cdots$, $\mu+a_h$ (단, a_i는 양의 정수, μ는 λ_i의 최소값)이라 하자. 그러면

$$\sum\sum_{i\neq j}\lambda_{ij}^2 = v\sum_{i=1}^{h} n_i(\mu+a_i)^2$$

이고 $a_1=0$, $a_i=a_{i-1}+1$ $(i=2,\cdots,h)$인 경우 최소가 된다. 즉 λ_i의 값이 순서에 무관하게 연속되는 h 개의 정수이면 $\sum e_i^2$은 최소가 된다.

위의 (정리 1)에서 h=2이면 정규 그래프 계획이고 정규 그래프 계획은 (M,S)-최적계획임을 쉽게 알수 있다.

4. 동반수가 3인 가장 균형된 계획의 배치계획.

(정리 1)을 효율적으로 사용하기 위해서는 배치계획이 준비되어야 한다. 그러나 동반수가 4 이상인 경우는 배치계획을 작성하기가 매우 힘들고, 알려진 배치계획도 거의 없기 때문에 대안으로 기존의 PBIBD(3)에서 동반수가 3인 가장 균형된 계획 중 $v\leq 30$이거나 $b\leq 30$인 배치계획만을 찾아서 (표 1)을 작성하였다. (표 1)의 참고난에서 '쌍대'는 주어진 계획의 쌍대계획도 (M,S)-최적계획이 된다는 의미다. D(v, b, k, r)가 (M,S)−최적계획이면 (정의 3)의 D^+(v, b, k^*, r^*)도 (M,S)−최적이 된다(Jacroux,1978). 따라서 (표 1)의 10개 배치계획의 여 계획도 (M,S)−최적계획이된다.

정의 2. N을 $D(v, b, k, r)$의 결합행렬이라면, N'를 결합행렬로 하는 $D^*(v^*=b, b^*=v, k^*=r, r^*=k)$를 $D(v,b,k,r)$의 쌍대 계획(dual design)이라한다. 물론 $D(v,b,k,r)$와 $D^*(v^*,b^*,k^*,r^*)$는 전혀 다른 계획이다.

정의 3. N을 D(v, b, k, r)의 결합행렬이라면, $J-N$을 결합행렬로 하는 D^+(v, b, $k^*=v-k$, $r^*=b-k$)를 여 계획(complement design)이라 한다. 여기서 J는 크기 $v\times b$ 인 원소가 1인 행렬이다.

표 1 . 동반수 3을 갖는 가장 균형된 계획

배치번호 \ 모수	v	b	k	r	λ_1	λ_2	λ_3	원 전	참 고
1	8	4	4	2	2	1	0	Rao(1956)의 원형격자 계획	
2	18	6	6	2	2	1	0	〃	
3	8	8	3	3	2	0	1	Raghavarao(1960)의 HGD계획	쌍대
4	8	8	5	5	4	2	3	〃	쌍대
5	12	4	6	2	2	1	0	Rao(1956)의 그룹 분해가능계획을 3 반복후의 쌍대계획	
6	27	6	9	2	2	1	0	〃	
7	8	10	4	5	1	3	2	Sinha와 Kageyma(1992)	
8	12	12	4	4	0	1	2	참고문헌[14] pp. 39	
9	12	16	3	4	1	2	0	Rao(1956)의 유사요인계획	
10	16	12	4	3	1	2	0	〃	

(표 1)의 이용방법을 예를 들어 설명하자.

처리하여야 할 처리(변수)수가 8일 때, 블록수는 8, 블럭의 크기는 3 ,반복 수 3이라는 제약조건 아래서 실험을 부득이 수행해야 하는 경우 물론 $D(8,8,3,3)$을 만족하는 BIBD가 존재하면 이 계획이 최적계획이된다. 그러나 BIBD가 존재하지 아니하는 경우 PBIBD 중에서 조건을 만족하는 배치계획이 알려져 있는 계획을 찾아야 한다. 이제 D(8,8,3,3) 및 $\lambda_1=2$, $\lambda_2=0$, $\lambda_3=1$을 만족하는 배치계획을 PBIBD(3)인 Raghavarao(1960)의 HGD계획에서 찾을 수 있었다. 이 배치계획이 (표 1)의 배치번호 3이고 $(M,S)-$최적계획이다. Raghavarao(1960)에서 배치계획을 찾아보니 다음과 같았다. 배종성(1992)을 참조 바람.

배치계획 : (1 2 5) (3 5 6) (1 2 6) (4 5 6) (3 4 7) (1 7 8) (3 4 8) (2 7 8)

이 배치계획의 NN'는 아래와 같다.

$$NN'=\begin{bmatrix}3&2&0&0&1&1&1&1\\2&3&0&0&1&1&1&1\\0&0&3&2&1&1&1&1\\0&0&2&3&1&1&1&1\\1&1&1&1&3&2&0&0\\1&1&1&1&2&3&0&0\\1&1&1&1&0&0&3&2\\1&1&1&1&0&0&2&3\end{bmatrix}$$

우리는 (정의 2)와 (정의 3)으로 부터 이 계획의 쌍대계획은 자신으로 환원되고, 여 계획은 (표 1)의 배치번호 4인 계획이되어 쌍대 및 여 계획도 (M, S) - 최적이 됨을 쉽게 알수 있다.

4. 논의

정준효율 인자 e_i가 모두 같은 BIBD인 경우 위의 4가지 최적기준을 만족함을 우리는 잘 알고 있다. 따라서 PBIBD인 경우는 λ_i의 값들이 최대한 비슷한 경우가 4가지 최적기준에 부합하리라는 것을 추측할 수 있다. (정리 1)은 이러한 추측을 뒷받침 해준다. 4가지 최적 기준 중 어떠한 최적기준이 가장 좋은 기준인가를 한마디로 이야기하기는 어렵다. Jarrett은 균형된 계획이 아니면 4가지 최적기준은 조금씩 다른 결과를 줄수도 있다라고 했다. 또한 A, D, E-최적의 약점(disadvantage)은 (2.1)식에서 A행렬의 특성근이나 역수를 계산해야 하나, (M, S) - 최적은 (2.2)식에서와 같이 단지 NN' 행렬의 원소만 제곱해서 더하는 계산상의 간단함을 잇점으로 생각하였다. Sinha 와 Kageyama의 E-최적은 $\lambda_2 > \lambda_3 > \lambda_1$이라는 조건때문에 사용할 수 있는 배치계획의 수가 극히 제한될 수 밖에 없다. 동반수 $h(h \geq 2)$를 갖는 가장균형된 계획과 A, D, E-최적과의 관계 및 배치계획을 작성하는 방법에 대해서 연구가 이루어 져야할 것이다.

참 고 문 헌

[1] 배종성 (1992). Hierarchial Group Divisial Design의 분석에 관한 연구, 「응용통계연구」, 제5권 제1호, 1-8.

[2] Bose, R. C. and Nair, K. R. (1939). PBIB designs, Sankhya, 4, 337-372.

[3] Cheng, C, S. (1978). Opmility of certain asymmetrical experimental designs, *The Annals of Statistics*, 6, 1239 -1261.

[4] ——————(1980). On the E-optimality of some block designs, *Journal of the Royal Statistcal Society*, B. 42, 199-204.

[5] Eccleston, J. A. and Hedayat, A. (1974). On the theory of connected designs: characterization and optimality, *The Annals of Statistics*, 2, 1238-1255.

[6] Jacroux, M. A. (1978). On the properties of proper (M.S) optimal block designs, *The Annals of Statistics*, Vol 6, No 6, 1302-1309.

[7] Jarrett, R. G. (1983). Definitions and properties for m-concurrence designs, *Journal of the Royal Statistical Society, Series B*, Vol. 45, No 1, 1-10.

[8] John, J. A. (1987). *Cyclic designs*, Chapman and Hall, New York.

[9] John, J. A. and Williams, E. R. (1982). Conjectures for optimal block designs, *Journal of the Royal Statistical Society B,* Vol. 44, No. 2, 221-225.

[10] John, J. a. and Mitchell, T. J. (1977). Optimal incomplete block designs, *Journal of the Royal Statisticl Society B,* Vol 39, 39-44.

[11] Raghavarao, D. (1960). A generalization of group divisible designs, *The Annals of Mathematical Statistics,* 31, 756-771.

[12] Rao, C. R. (1956). A general class of quasi factorial and related designs, *Sankhya,* 17, 165-174.

[13] Shah, K. R. (1960). Optimality criteria for incomplete block designs, *The Annals of Mathematical Statistics,* 31, 791-794.

[14] Shah, K. R. and Sinha, B. K. (1989). *Theory of optimal designs*, Springer-Verlag, New York.

[15] Sinha, B. K. and Shah, K. R. (1988). Optimality aspects of 3-concurrence most balanced designs, *Journal of Statistical Planning and Inference,* 20, 229-236.

[16] Sinha, K. and Kageyama, S. (1992). Constructions of some E-optimal 3-concurrence most balanced designs, *Journal of Statistical Planning and Inference,* 31, 127-132.

[17] Takeuchi, K. (1961). On the optimality of certain of PBIB disigns, Report on Statistics and applied research, *University of Japen Sevene Engrs,* 8, 140-145.

A study of (M, S)-optimal $h(h \geq 2)$-concorrance most balanced design.[3)]

Jongsung Bae[4)]

Abstract

In this paper, We show that the proper equi-replicate binary designs with $h(h \geq 2)$-concurrence most balanced designs are (M, S)-optimal. It is very difficult to construct with more than 3-concurrence designs, so we choose 3-concurrence most balanced designs out of the known PBIBD(3).

3) This paper was supportded by the NON DIRECTED RESEARCH FUND, Chunnam University Research Foundation, 1995.

4) Professor, Department of statistics, Chonnam National University, Kwangju, 500-757, Korea.