응용통계연구 (2008)
21(4), 615–620

때때로 풀링의 규칙에 관한 소고

임용빈[1]

[1]이화여자대학교 통계학과

(2008년 5월 접수, 2008년 5월 채택)

요 약

공업실험에서는 주효과의 검정력을 높일 목적으로 교호작용효과에 대한 예비검정의 결과가 유의하지 않은 경우에, 이 효과를 오차항으로 풀링한 후에 주효과의 유의성을 검정하는 때때로 풀링(Sometimes pooling)의 분석 방법을 사용한다. 이 소고에서는 때때로 풀링의 규칙에 관한 참고문헌의 내용을 정리하고, 블록효과가 오차항으로 풀링되었다는 사실이 알려 졌을 때에 독립표본에 의한 모평균의 차이에 대한 95% 신뢰구간의 길이가 쌍체비교에 의한 신뢰구간의 길이보다 짧을 확률을 다양한 예비검정의 유의수준 α_1값과 충분한 검정력이 보장되지 않은 블록의 크기인 $2 \leq n \leq 13$ 범위에서 살펴보았는데, 그 결과는 풀링한 경우에 주효과에 대한 검정력이 높아진다는 사실을 뒷받침하고 있다.

주요용어: 때때로 풀링의 규칙, 예비검정의 유의수준, 독립표본에 의한 모평균의 차이 비교, 쌍체비교.

1. 서론

실험계획법의 역사는 1930년대에 영국의 로담스테드 농업실험소에서 근무한 R. A. Fisher로부터 시작되었다. Fisher는 농작물 재배 실험방법을 개발하면서 실험계획법의 이론을 체계화하였다 (Fisher, 1935). 실험계획법은 세계 제 2차 대전 후에 화학공업을 주축으로 한 공업실험 방법 연구가 G. E. P. Box (1951)를 중심으로 프로세스 산업에 응용되면서 그 영역이 확장되었다. 특히 공업실험에서는 생산라인에서 실험을 실시해야 하는 경우가 많아서 실험의 크기에 제약조건이 따르고, 최소의 실험으로 최대의 정보를 얻는 것이 실용적으로 중요하다. 초기 단계의 실험에서 실험에 고려될 많은 인자들 중에서 반응변수에 영향을 끼치는 핵심적인 소수 인자들을 선별하는 것이 실험의 목적일 경우에는 2수준 일부실시법을 실시한다. 선별된 인자들과 반응변수와의 계량적인 관계에 관한 정보를 얻기 위하여 반복수가 많아야 2, 3회 정도인 이원배치법 혹은 삼원배치법이나 랜덤화 블록 계획법을 실시한다.

2수준 일부 실시법의 실험자료 분석에서 핵심적인 소수인자들을 선별하기 위하여 자유도가 1인 각 효과들의 크기에 대한 반정규확률그림을 그린다. 반정규확률그림에서 개략적으로 원점을 지나는 직선에서 벗어나는 점들에 대응되는 요인효과를 핵심적인 효과로 선별하고, 나머지효과들을 한데 묶어서 오차항으로 풀링한 후에, 분산분석표를 작성한다. 즉, 각각의 요인효과들에 대한 오차항으로의 풀링의 기준은 반정규확률그림에서 대응되는 점이 개략적으로 직선상에 놓여 있는지를 판단하는 것이다.

공업실험에서의 이원배치법 혹은 삼원배치법의 실험자료의 분석은 축차적으로 진행한다. 먼저 고차의 교호작용효과의 유의성을 검정하고, 고차의 교호작용효과가 유의하지 않은 경우에 이 항을 오차항으로

[1](120-750) 서울시 서대문구 대현동 11-1, 이화여자대학교 자연과학대학 통계학과, 교수.
E-mail: yblim@ewha.ac.kr

풀링한 후에 저차의 교호작용효과나 주효과의 유의성을 검정한다. Fisher는 실험의 설계 단계에서 채택한 실험의 랜덤화 방법에 따라서 반응변수에 대한 구조모형을 결정하고, 요인별 평균변동의 기대값에 따라서 실험자료를 얻기 전에 이미 실험자료의 분석 방법이 정해지는 절대로 풀링을 하지 않는 분석 방법(Never pooling)을 제시했다. 그런데 공업실험에서는 주어진 실험의 크기에서 의미 있는 결과를 유추하기 위해서 즉, 주효과의 검정력을 높일 목적으로 고차의 교호작용효과가 유의하지 않은 경우에 오차항으로 풀링한 후에 저차의 교호작용효과나 주효과의 유의성을 검정하는 때때로 풀링(Sometimes pooling)하는 분석 방법을 사용한다. 때때로 풀링의 원칙은 Lorenzen과 Anderson (1993)에 잘 소개 되어 있다. 2절에서는 때때로 풀링에 관한 대표적인 참고문헌의 내용을 정리하고, 3절에서는 독립표본에 의한 모평균 차이 비교와 쌍체비교에 대한 효율 비교에 근거한 관점에서 풀링의 규칙을 살펴보고, 4절에서는 때때로 풀링의 원칙을 종합한다.

2. 때때로 풀링이란?

실험계획법의 구조모형에 있는 항들이 유의하지 않다고 결론을 내린 순간에도, 굉장히 소중한 정보를 얻게 된다. 예를 들면, 항의 효과가 고정된(Fixed) 경우에는 이 인자의 수준 값에 관계없이 반응변수의 평균이 동일하고, 랜덤인 경우에는 이 인자의 수준들 간의 산포가 반응변수의 분산에 영향을 주지 않는다는 정보를 얻게 되며, 다음 단계의 실험에서 해당 인자들을 인자로 고려하지 않아도 되기에 실험의 크기를 줄여 준다.

Bozivich 등 (1956)은 혼합모형(Mixed Model)과 랜덤모형(Random Model)인 경우에 분산분석을 활용한 가설검정에서 때때로 풀링에 대한 가장 간단한 경우를 다음과 같이 설명한다. 자유도 n_1, n_2, n_3와 기댓값 σ_1^2, σ_2^2, σ_3^2를 각각 갖는 평균제곱 V_1, V_2, V_3가 주어져 있다고 하자. 여기서 V_3는 처리평균제곱합, V_2는 오차평균제곱합, V_1은 잠재적인 오차평균제곱합을 표시한다. 우리의 관심사는 V_3와 관련된 가설인 $H_0 : \sigma_3^2 = \sigma_2^2$을 검정하는 것이고, 검정통계량은 $F = V_3/V_2$이다. 그런데 $H_0 : \sigma_2^2 = \sigma_1^2$일 가능성이 존재한다고 하자. 만약에 이 가능성이 참일 경우에는 V_2와 V_1를 통합하여 풀링된 오차 평균제곱인 $V = (n_1V_1 + n_2V_2)/(n_1 + n_2)$를 정의하고 검정통계량으로 $F = V_3/V$를 사용하는 것이 처리효과의 유의성에 대한 가설검정의 검정력을 크게 하는 이점이 있게 된다. 따라서 우선 예비검정인 $H_0' : \sigma_1^2 = \sigma_2^2$를 실시하고, 예비검정의 귀무가설이 채택되면 V_2와 V_1를 통합하여 풀링된 오차 평균제곱인 V와 $F = V_3/V$를 계산하여 $H_0 : \sigma_3^2 = \sigma_2^2$에 대한 검정통계량으로 사용하는 것이다. 예비검정의 귀무가설이 기각되면, $F = V_3/V_2$를 검정통계량으로 사용한다.

구조모형에 따라서 V_2, V_1이 다르게 선택된다. 반복이 있는 이원배치법의 예를 들자. 고정인자인 A의 유의성에 관심이 있는 경우에, 인자 B가 랜덤인자인가 고정인자인가에 따라서 평균제곱의 기대값인 EMS가 달라지고, V_2, V_1도 달라진다. 먼저 인자 B가 랜덤인자인 혼합모형의 경우에는 V_3는 인자 A의 평균제곱인 MS_A이고, V_2는 교호작용효과 $A \times B$의 평균제곱인 $\mathrm{MS}_{A\times B}$이며 V_1은 잔차평균제곱인 MS_E이다. 인자 A, B가 모두 고정인자인 고정모형의 경우에는 V_2는 잔차평균제곱인 MS_E이고, 잠재적인 오차평균제곱인 V_1은 $\mathrm{MS}_{A\times B}$가 된다.

예비검정인 $H_0' : \sigma_1^2 = \sigma_2^2$의 유의수준을 α_1, 예비검정이 기각된 경우에 $H_0 : \sigma_3^2 = \sigma_2^2$의 유의수준을 α_2라 하고 예비검정이 채택된 경우에 $H_0 : \sigma_3^2 = \sigma_2^2$의 유의수준을 α_3라 하자. 때때로 풀링에서 분석자의 선택사항은 유의수준 α_1, α_2, α_3들의 값을 결정하는 것이다. 주된 관심사인 $H_0 : \sigma_3^2 = \sigma_2^2$의 유의수준을 α로 정한 경우에는, $\alpha_2 = \alpha_3 = \alpha$로 정하는 것이 때때로 풀링에 의해서 실험자료를 분석할 때에 $H_0 : \sigma_3^2 = \sigma_2^2$의 유의수준을 α로 제어하기가 쉽다. Bozivich 등 (1956)은 관심사인 $H_0 : \sigma_3^2 = \sigma_2^2$에 대한 유의수준이 α인 가설검정에서 Fisher의 분석방법인 절대 풀링 안하기와 때때로 풀링 방법에 대

한 검정력의 비교를 통해서 때때로 풀링에서 예비검정인 $H_0' : \sigma_1^2 = \sigma_2^2$의 유의수준 α_1값을 추천하였다. Bozivich 등 (1956)은 $\theta_{21} = \sigma_2^2/\sigma_1^2$로 정의하였을 때에, $1 \leq \theta_{21} \leq 2$인 경우에 때때로 풀링을 추천하였는데, $n_3 < n_2$이거나 $n_1 < 5n_2$인 경우에는 예비검정의 유의수준으로 $\alpha_1 = 0.25$를 사용하고 $n_3 \geq n_2$ & $n_1 \geq 5n_2$인 경우에는 $H_0 : \sigma_3^2 = \sigma_2^2$의 가설 검정의 유의수준을 α로 제어하기 위해서 $\alpha_1 = 0.50$을 사용하기를 추천하고, $\theta_{21} > 2$인 경우에는 때때로 풀링의 이점이 별로 없음을 보였다.

Mead 등 (1975)은 고정모형(Fixed Model)의 경우에 $n_1 > 2n_2$인 경우에는 예비검정의 유의수준으로 α_1=0.25를 사용하고 $n_1 \simeq n_2$인 경우에는 $H_0 : \sigma_3^2 = \sigma_2^2$의 가설 검정의 유의수준을 α로 제어하기 위해서 α_1=0.50을 사용하기를 추천하고, $n_1 < n_2$이고 n_2가 적당히 큰 경우에는 때때로 풀링의 이점이 별로 없음을 보였다.

Lorenzen과 Anderson (1993)은 공업실험에서 실용적으로 적용하는 때때로 풀링의 규칙으로 우선 가장 고차 교호작용효과의 오차항으로 풀링 여부를 검토한다. 고차 교호작용효과의 p-값(p-value)이 0.25 이상이면, 유의하지 않다고 결론내리고, 해당 평균제곱의 기대값인 EMS에서 대응되는 수준효과의 산포의 크기인 σ^2이나 Φ값을 0으로 대체하여 제거하는 것인데, 자세한 내용은 아래와 같다.

- 모형에 있는 항(요인효과)이 유의 수준 0.25에서 유의하지 않으면 즉, p-값(p-value)이 0.25 이상이면 관련된 항은 모형과 평균제곱의 기대값인 EMS에서 제거될 후보가 된다.
- 저차의 교호작용효과나 주효과는 이 항들을 포함하는 고차의 교호작용효과가 유의 수준 0.25에서 유의한 경우에는 모형에서 제거되지 않는다.
- 특정 항(요인효과)이 모형에서 제거되면, 해당 평균제곱의 기대값인 EMS에서 대응되는 수준효과의 산포의 크기인 σ^2이나 Φ값을 0으로 대체하여 제거한다.
- 평균제곱의 EMS가 동일한 항들을 묶어서(pooling) 새로운 항을 구성하는데, 새로운 항의 제곱합과 자유도는 각각 묶인(pooled) 항들의 제곱합들과 자유도들의 합이다. 새로운 항의 평균제곱을 구한 후에, 계산된 평균제곱 값들과 EMS를 적용하여 처리효과에 대한 가설 검정을 실시한다.

3. 독립표본에 의한 모평균 차이 비교와 쌍체비교의 효율성에 근거한 풀링의 규칙 비교

김태민과 김상부 (2006)은 두 모집단의 평균에 대한 비교 방법인 쌍체비교와 독립표본에 의한 모평균 차이 비교의 효율성을 모평균 차이에 대한 신뢰구간의 길이 비교를 통해서 통계적으로 고찰하였다. 간단성을 위하여 독립표본에 의한 모평균 차이 비교를 독립비교라 하자. 쌍체비교는 랜덤화 블록 계획법에서 처리의 수가 2인 경우의 모평균의 차이를 비교하는 방법이다. 독립비교는 랜덤화 블록 계획법에서 블록효과를 오차항에 풀링한 경우의 모평균의 차이를 비교하는 방법과 동일하다. R을 독립비교와 쌍체비교에 의한 모평균의 차이에 대한 신뢰구간의 길이의 비라고 정의하자. R이 1보다 작으면, 독립비교에 의한 모평균의 차이에 대한 신뢰구간의 길이가 쌍체비교에 의한 신뢰구간의 길이보다 짧게 되어서, 독립비교의 정밀도가 높음을 의미한다. 블록의 수를 n이라 할 때에 R은

$$R = \frac{t_{\frac{\alpha}{2};2(n-1)}}{t_{\frac{\alpha}{2};(n-1)}}\sqrt{\frac{1+F_B}{2}} \tag{3.1}$$

(김태민과 김상부, 2006)와 같이 표현된다. 여기서 F_B는 블록효과에 대한 검정통계량으로 MS_B, MS_E를 블록과 오차에 대한 평균제곱이라 표시할 때에 $F_B = \mathrm{MS}_B/\mathrm{MS}_E$이다. 식 (3.1)에서 R의 값이 1이 되는 F_B의 값을 f_B로 정의하면

$$f_B = \frac{2F_{\alpha;1,n-1}}{F_{\alpha;1,2(n-1)}} - 1$$

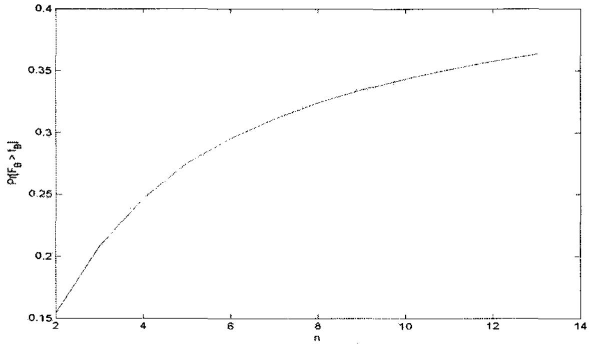

그림 **3.1.** 블록의 크기의 범위인 $2 \le n \le 13$에서 $\text{Pr}[F_B \ge f_B]$의 그래프

이다. 식 (3.1)에서 R은 F_B의 증가함수이기 때문에 $\{R < 1\}$사상은 $\{F_B < f_B\}$사상과 동치이다.

2절에서 소개한 바와 같이 Bozivich 등 (1956)과 Mead 등 (1975)은 예비검정의 유의수준인 α_1값으로 일반적으로 0.25를 추천하고, 특수한 경우에는 0.5를 추천하였다. Lorenzen과 Anderson (1993)은 간편성을 적용하여 α_1값으로 0.25를 추천하였다. 이제, 김태민과 김상부 (2006)의 연구결과를 때때로 풀링의 관점에서 접근해 보자. 예비검정의 유의수준 α_1에서 블록효과를 오차항으로 풀링이 결정된 경우에 독립비교의 정밀도가 쌍체비교 보다 높을 확률은 $\Pr[F_B < f_B \,|\, F_B \le F_{\alpha_1;n-1,n-1}]$이다. 이 확률을 다양한 $0.20 \le \alpha_1 \le 0.50$에서 구하고, 이 확률에 대한 예비검정의 유의수준 α_1값의 영향력을 살펴보려고 한다.

우선 작은 크기의 처리효과와 블록효과에 대한 검정력이 충분하도록 블록의 크기인 n을 결정하자. 제1종 오류의 확률 $\alpha = 0.05$와 제 2종 오류의 확률 $\beta = 0.1$가 주어진 경우에 이연수 등 (1999)의 MATLAB 프로그램을 활용하여 처리효과와 블록효과의 표준화된 최소 검출 가능 효과인 Δ의 크기가 1.5 보다 적게 되는 블록의 크기를 구한 결과 $n = 14$인 경우에 처리효과에 대한 Δ값은 0.9382이고, 블록효과에 대한 Δ값은 1.4766으로 이 실험 계획은 Lorenzen과 Anderson (1993)에 따르면 Small Δ 그룹에 속할 정도로 충분한 검정력을 가지고 있다. 따라서 블록의 크기가 $2 \le n \le 13$ 범위에서는 충분한 검정력이 보장되지 않는다.

독립비교의 정밀도가 높을 확률인 $\Pr[F_B < f_B]$를 처리효과와 블록효과에 대한 충분한 검정력이 보장되지 않는 블록의 크기의 범위인 $2 \le n \le 13$ 경우에 각각 계산하고, 예비검정의 유의수준인 α_1값을 $\Pr[F_B \ge f_B]$이상의 값으로 선택하면, 유의수준 α_1에서 블록효과를 오차항으로 풀링했다는 정보가 주어진 경우에 독립비교의 정밀도가 높을 확률인 $\Pr[F_B < f_B \,|\, F_B \le F_{\alpha_1;n-1,n-1}]$는 1이 된다. 블록의 크기인 n의 값에 따라서 예비검정의 유의수준 α_1값으로 취한 $\Pr[F_B \ge f_B]$이 0.1539에서 0.3634로 곡선효과를 가지고 증가함을 그림 3.1에서 시각적으로 확인할 수 있다.

표 3.1에는 α_1의 값과 n의 값에 따른 $\Pr[F_B < f_B \,|\, F_B \le F_{\alpha_1;n-1,n-1}]$이 주어진다. 표 3.1은 실험자료를 분석한 후에 블록효과의 오차항에 대한 풀링 여부를 위한 추가 정보를 제공한다. 예를 들어 n=8이고, 실험자료를 분석한 결과 블록효과에 대한 p-값이 0.301이라하자. 이 경우에 유의수준 α_1을 0.30으로 하면, 블록효과를 오차항으로 풀링하는 결정을 내리게 되고, 표 3.1에 따르면, 블록효과를 오차항으로 풀링했다는 정보가 주어진 경우에 독립비교의 정밀도가 높을 확률은 0.97이다. 산업실험에서 실용적으로 많이 이용되는 α_1=0.25인 경우에, n이 4이하인 경우에는 독립비교의 정밀도가 높을 확률은 1이지만, n이 증가함에 따라서 감소하여, n=13인 경우에는 0.85가 된다.

표 **3.1.** 유의수준 α_1의 값과 블록의 크기 n에서 $\mathrm{PR}[F_B < f_B \mid F_B \le F_{\alpha_1;n-1,n-1}]$

α_1 \ n	5	6	7	8	9	10	11	12	13
0.25	0.97	0.94	0.92	0.90	0.89	0.88	0.87	0.86	0.85
0.26	0.98	0.95	0.93	0.91	0.90	0.89	0.88	0.87	0.86
0.27	0.99	0.97	0.94	0.93	0.91	0.90	0.89	0.88	0.87
0.28	1.00	0.98	0.96	0.94	0.92	0.91	0.90	0.89	0.88
0.29	1.00	0.99	0.97	0.95	0.94	0.92	0.91	0.90	0.90
0.30	1.00	1.00	0.98	0.97	0.95	0.94	0.93	0.92	0.91
0.31	1.00	1.00	1.00	0.98	0.96	0.95	0.94	0.93	0.92
0.32	1.00	1.00	1.00	0.99	0.98	0.97	0.95	0.94	0.94
0.33	1.00	1.00	1.00	1.00	0.99	0.98	0.97	0.96	0.95
0.34	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.97	0.96
0.35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98
0.36	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
0.37	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

4. 결론

공업실험에서는 주어진 실험의 크기에서 주효과의 검정력을 높일 목적으로 고차의 교호작용효과에 대한 예비검정의 결과가 유의하지 않은 경우에, 이 효과를 오차항으로 풀링한 후에 저차의 교호작용효과나 주효과의 유의성을 검정하는 때때로 풀링(Sometimes pooling)이라는 분석 방법을 사용한다. 때때로 풀링의 가장 커다란 관심사는 예비검정의 유의수준인 α_1값을 결정하는 것이다. Bozivich 등 (1956)과 Mead 등 (1975)은 예비검정의 유의수준인 α_1값으로 일반적으로 0.25를 추천하고, 특수한 경우에는 0.5를 추천하였다. Lorenzen과 Anderson (1993)은 간편성을 적용하여 α_1값으로 0.25를 추천하였다. 이 소고에서는 때때로 풀링의 규칙에 관한 참고문헌의 내용을 정리하고, 김태민과 김상부 (2006)이 고려한 두 모집단의 평균에 대한 비교 방법인 쌍체비교와 독립비교에 대한 효율성 비교의 관점에서 예비검정의 유의수준인 α_1값의 영향력을 충분한 검정력이 보장되지 않은 블록의 크기인 $2 \le n \le 13$ 범위에서 살펴보았다. 유의수준 α_1=0.25에서 블록효과를 오차항으로 풀링한 경우의 독립비교의 정밀도가 쌍체비교 보다 높을 확률은 블록의 수 n이 4이하인 경우에는 1이지만, n이 증가함에 따라서 감소하여, n=13인 경우에는 0.85이다. 이 확률 값은 풀링한 경우에 주효과에 대한 검정력이 높아진다는 사실을 뒷받침하고 있다.

참고문헌

김태민, 김상부 (2006). 쌍체비교와 독립비교에 대한 통계적인 고찰, <응용통계연구>, **19**, 231–240.

이연수, 임용빈, 김재주 (1999). 혼합모형에서의 실험의 크기에 관한 연구, <응용통계연구>, **12**, 593–603.

Bozivich, H., Bancroft, T. A. and Hartley, H. O. (1956). Power of analysis of variance test procedures for certain incompletely specified models, I, *The Annals of Mathematical Statistics*, **27**, 1017–1043.

Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions, *Journal of the Royal Statistical Society. Series B*, **13**, 1–45.

Fisher, R. A. (1935). *The Design of Experiments*, Oliver and Boyd, Edinburg.

Lorenzen, T. J. and Anderson, V. L. (1993). *Design of Experiments*, Marcel Dekker, New York.

Mead, R., Bancroft, T. A. and Han, C. (1975). Power of analysis of variance test procedures for incompletely specified fixed models, *The Annals of Statistics*, **3**, 797–808

A Note on Sometimes Pooling Rules

Yong Bin Lim[1]

[1]Dept. of Statistics, Ewha Womans University

(Received May 2008; accepted May 2008)

Abstract

In engineering experiments, 'Sometimes Pooling Rules' to remove insignificant terms from the model has been implemented to increase the power of detecting the small size of main effects when the preliminary test of higher order interaction effects declare to be insignificant. In this note, we review the sometimes pooling rules in the literature and also study the probability of the length of 95% confidence interval of $\mu_1 - \mu_2$ of the comparison of two independent samples being shorter than that of the paired comparison at the various level of significance α_1 of the preliminary test and the insufficient number of blocks n in $[2, 13]$, given the block effects being pooled to the error term. This study supports that the sometimes pooling results in the power improvement of the main effects.

Keywords: Sometimes pooling rules, level of significance of the preliminary test, comparison of two independent samples, paired comparison.

[1]Professor, Dept. of Statistics, Ewha Womans University, 11-1 Daehyun-dong, Seodaemun-gu, Seoul 120-750, Korea. Email: yblim@ewha.ac.kr