두 개의 처리 비교를 위한 실용적인 실험 계획 전략과 분석*

임용빈*†

* 이화여자대학교 통계학과

Practical Experimental Design Strategy and Analysis for the Comparison of Two Treatments

Yong Bin Lim*†

* Department of Statistics, Ewha Womans University

Key Words : Comparison of Two Treatments, Split Plot Designs, Standardized Detectable Differences

Abstract

We consider practical experimental design strategies and analysis to find out whether a modified method give better results than the standard method. The most practical design strategy is for experimenter to make r successive runs under the current standard method and then, change the standard method to a modified method to make another r successive runs under a modified method. To test a statistically significant difference between the population mean of the standard method and a modified method, additional recent data for sufficient number of consecutive responses under the standard method is needed to construct external reference distribution (Box, et al., 1968). Upon those informations unavailable, the practical design strategy is to run the experiment by split plot designs. In this paper, two types of split plot designs are proposed and how to determine efficiently the number of repetition within a given method and replication of those two methods are discussed based on results of the level of significance $\alpha=0.05$ and the power being at least 0.9 at the detectable difference of $\mu_2-\mu_1=1.5\times\sigma$.

1. 서 론

두 가지 제조 방법 중에 어느 방법이 수율을 더 향상시킬 수 있는가를 실험을 통해서 규명하는 작업은 가장 초보적인 실험 설계로, 두 가지의 처리 효과를 상호 비교하는 것이다. 식스 시그마 프로젝트 팀의 과제 수행 중에서 흔히 접하게 되는 문제의 하나는 팀이 제안한 개선 혹은 수정된 방법인 신공법이 기존의 표준공법보다 실제로 품질 면에서 우월한지, 즉 좋은 결과를 주는 지를 알아보는 것으로, 이 범주에 해당된다. 이 경우에 표준공법을 사용하든지 신공법을 사용하든지 '공법에 관계없이 모집단의 평균은 동일하다'가 실험자가 기각하고 싶은 귀무가설이 된다. 품질 특성치가 인장강도, 수율, 두께, 길이 등과 같이 계량치인 경우에 엔지니어들은 흔히 실험 자료가 어떻게 얻어졌는가에 관한 적절한 검토 없이, t-검정에 의하여 분석을 하고, 기존 방법의 변경 여부에 관한 결정을 내린다. t-검정에 의한 분석 결과가 타당성을 얻기 위해서는 실험 자료들이 서로 독립이어야 하고, 이를 위해서는 실험이 랜덤하게 실시되어야 한다. 예를 들어서 각각의 방법에 의해서 10회씩 실험을 실시하는 경우에, 검은색 카드 10장, 빨강색 카드

† 교신저자 yblim@ewha.ac.kr

※ 본연구는 한국과학재단 특정기초연구 과제번호 : R01-2003-000-10220-0에 의해서 지원된 연구결과입니다.

10장을 잘 섞어서 카드를 차례대로 뒤집어서 검은 색이 나오면 기존의 표준공법에 의해서 실험을 실시하고 빨강색이 나오면 제안된 신공법에 의해서 실험을 실시하여 20개의 실험 자료를 얻는다. 다른 색의 카드가 연속해서 나오면 실험조건을 바꾸어야 하기에, 실험을 랜덤하게 실시하면, 공정조건을 자주 바꾸어 가면서 실험을 실시해야 한다. 그런데 현업에서 공정조건을 바꾸기 위해서는 생산을 잠시 멈추어야 하기에 반도체나 LCD 등 첨단 산업에서는 원료 탱크나 램프 교체 등 생산라인의 예방정비(preventive maintenance) 시에나 가능하며, 공정조건을 바꾸기 위해서는 다음 예방 정비 시기까지 기다려야 하기에 공정조건 변경에 시간이 오래 걸리고, 공정조건을 자주 바꿀 수 없는 어려움이 따르게 된다. 따라서 실용적인 실험 설계 전략은 공정조건을 자주 바꾸지 않으면서, 얻어진 실험 자료를 가지고, 두 모집단의 평균에 차이가 있는지를 알아보는 것이다. 이 논문의 목적은 실용적인 실험 설계 전략 및 분석 방법에 관해서 살펴보는 것이다. 가장 실용적인 실험 설계는 표준공법으로 진행되고 있는 공정을 신공법으로 바꾼 후에 연속적으로 생산해서 실험 자료를 얻는 것이다. 이 실험 자료로부터 평균 수율의 차이가 있는 지를 검정하기 위해서는 추가적으로 표준공법으로 연속해서 생산된 제품들의 수율에 관한 정보를 필요로 한다. 즉, 올바른 분석을 위해서는 실험을 편리하게 실시한 것에 대한 대가로 추가적인 정보인 표준공법으로 연속해서 생산된 제품들의 수율에 관한 자료를 필요로 하는 것이다. 2장에서는 Box 등(1968)에서 연구된 이 실험 자료에 관한 분석 방법인 외부 분포를 이용한 두 모평균의 비교방법을 간략하게 소개한다. 3장에서는 공정조건을 적게 바꾸는 실험 설계인 분할법 Ⅰ에 의한 축차적인 실험 설계를 소개하고, 엔지니어들의 가장 커다란 관심사의 하나인 주어진 공정조건에서의 효율적인 반복수 r^*의 결정에 관하여 토의한다. 4장에서는 두 개의 공정조건을 2*rep개의 일차단위에 랜덤하게 배치한 다음에, 일차단위에 랜덤하게 배치된 공정조건을 고정시키고 실험을 r회 반복해서 실시하는 분할법 Ⅱ에 의한 실험 설계와 효율적인 반복수 r^*의 결정에 관하여 토의하고, 5장에서는 논문의 결과를 요약한다.

2. 외부 참고 분포를 이용한 두 모평균의 비교

가장 실용적인 실험 설계는 표준공법으로 생산되고 있는 공정을 신공법으로 바꾼 후에 연속적으로 생산해서 실험 자료를 얻는 것이다. 즉, 기준 시점에서부터 표준공법으로 연속해서 10개의 제품을 생산한 직후에 공정조건을 신공법으로 바꾸고, 신공법에 의하여 연속해서 10개의 제품을 생산하여 수율을 측정하는 것이다. 이제 검정통계량인 신공법에 의한 표본 평균 수율과 표준공법에 의한 표본 평균 수율의 차이를 구한다. 이 차이가 아주 큰 값이 나오면, 신공법과 표준공법의 모평균 수율에 차이가 있다고 결론을 내릴 수 있을 것이다. 문제는 표본 평균 수율의 차이가 어느 정도로 큰 값이 나와야 하는가를 결정하는 일이다. 이를 위해서는 유사한 환경에서 얻어진 표본 평균 수율의 차이가 '공법에 관계없이 모집단의 평균 수율은 동일하다'라는 귀무가설이 참인 경우에도 우연히 나올 수 있는 값인지를 판단할 수 있어야 한다. 따라서 귀무가설이 참인 경우에 검정통계량의 분포를 알아야 하고, 이를 참고분포(reference distribution)라 한다(Box 등, 1968). 귀무가설이 참인 경우는 표준공법과 신공법에 의한 수율의 분포가 동일하게 되고, 표준공법으로 연속해서 생산된 제품들의 수율에 관한 최근 자료가 보관되어 있는 경우에 이로부터 외부 참고분포(external reference distribution)를 구할 수 있다. 실험자료가 아닌 추가적인 정보인 외부의 자료로부터 구해진 참고분포라 하여 외부의 이름을 붙인다. 외부 참고분포(external reference distribution)는 데이터베이스에 저장된 표준공법으로 연속해서 생산된 제품들의 수율 자료로부터 구할 수 있다. 먼저 연속해서 생산된 10개 제품의 표본평균을 구한 후에, 인접한 표본평균의 차이를 구한 값들로 외부 참고분포가 구성된다. 이제 검정통계량인 신공법과 표준공법에 의한 표본평균 수율의 차이 값의 위치를 외부 참고분포로부터 확인해야 한다. 외부 참고분포로부터 검정통계량의 값 이상이 나올 확률을 계산하면, 이 값이 가설 검정에서의 p값(p-value)에 대응된다. p값이 아주 작은 값이 나오면, 귀무가설이 참인 경우에 드물게 일어나는 큰 값의 차이가 관측되었기에, 귀무가설을 기각하고, 신공법과 표준공법의 평균 수율의

차이를 결론짓게 된다. 외부 참고분포를 이용한 두 모평균의 비교의 예제는 Box 등(1968)이나 임용빈 등(1998)에 상세하게 설명되어 있으니, 참조하기를 바란다.

실험을 랜덤하게 실시한 경우에 참고분포는 자유도가 18인 t-분포로부터 구할 수 있지만, 실험을 랜덤하게 실시하지 않고, 실용적인 방법인 각각의 공법에 의해서 10회 연속으로 실시하는 경우에는 참고분포를 구하기 위해서 추가적인 정보인 표준공법으로 연속해서 생산된 제품들의 수율에 관한 자료가 필요하다. 따라서 실험의 용이성과 추가적인 정보의 필요라는 상반관계(trade-off)가 존재하는 것이다.

3. 분할법 I에 의한 두 모평균의 비교

이제 다음으로 실용적인 방법으로 분할법(split-plot design) I에 의한 실험의 축차적인 실시를 제안한다. 단계 1에서는 동전을 던져서 표준공법과 신공법의 실시순서를 랜덤하게 정한다. 그다음에 정해진 방법에 따라서 연속해서 실험을 r회 실시한다. 단계 2, 3, …에서도 동일한 방법으로 실험을 실시한다. 이 방법의 장점은 실험이 단계별로 축차적으로 실시될 수 있다는 점이다. 단계 2의 실험 후에는 단계 1과 단계 2의 실험 자료를 분석하고, 분석 결과에 따라서 단계 3의 실험의 실시 여부를 결정한다. 각 공법의 평균 수율의 차이가 없다는 귀무가설의 기각 여부를 판단할 근거를 실험 자료가 충분히 제공하지 못하는 경우에는 단계 3의 실험을 실시한다. 여기에서 우리의 관심사는 각 단계에서의 실험의 반복수 r의 합리적인 결정 방법이다. 여기서 난괴법(랜덤화 블록 설계)은 분할법 I의 반복수인 r값이 1인 경우로, 난괴법은 분할법 I의 특수한 경우가 된다. 분할법 I의 모형은

$$y_{ijk} = \mu + \text{block}_k + \alpha_i + \varepsilon_{1(ik)} + \varepsilon_{2(ijk)},$$
$$i = 1, 2,\ j = 1, \cdots, r,\ k = 1, \cdots, rep \qquad (1)$$

이다. 여기서 주구만의 모형은 난괴법과 동일하게 되고, 주구에 배치된 공법효과 α_i의 유의성을 검토하기 위해서는 주구오차 $\varepsilon_{1(ik)}$와 비교해야 한다. 따라서 F-검정의 분자의 자유도는 1, 분모의 자유도는 (rep-1)이다. 우리의 목표는 제 1종 오류의 확률 α를 0.05로 제어한 상태에서, 실험자가 검출하고 싶은 표준화된 평균 수율의 차이인 $\Delta_1 = \frac{\mu_2 - \mu_1}{\sigma}$에서 제 2종 오류의 확률이 $\beta = 0.1$, 즉 검정력이 0.9를 만족하는 반복수 r과 단계수 rep를 구한 후에, 경제성을 고려하여 효율적인 반복수 r^*과 예상되는 단계의 수 rep^*를 결정하는 것이다. 분할법에서의 검정력은 비중심 F-분포로부터 계산되고, 분할법의 반복수 r과 단계수 rep가 주어진 경우에 $\alpha = 0.05$, $\beta = 0.1$을 만족하는 비중심모수의 값으로부터 검출 가능 표준화된 최소 평균 수율의 차이 값인 Δ_1이 결정된다. 임용빈(1998)에서 소개된 균형된 실험계획에서의 표준화된 최소 검출 효과의 크기 Δ_1를 결정하는 MATLAB 프로그램을 활용하면 Δ_1값을 구할 수 있다. <표 1>에는 반복수와 단계수가 $1 \le r \le 7$, $2 \le rep \le 10$인 범위에서 표준화된 최소 검출 효과의 크기인 Δ_1값이 구해져 있다. <표 1>을 살펴보면, 반복수가 주어진 경우에, 단계의 수가 많아짐에 따라서, Δ_1값이 작아진다. 또한 단계의 수가 주어진 경우에도 반복수가 늘어남에 따라서 Δ_1값이 작아진다. 즉, 실험정보가 많아짐에 따라서 검출의 정도가 높아짐을 <표 1>에서 확인할 수 있다. 식스 시그마 경영에서는 품질 특성치의 평균이 1.5σ만큼 이동된 경우에 대비해서 산포 관리를 하여 불량률을 줄이는 것을 목표로 한다. 예를 들어서 평균 수율의 차이가 1.5σ인 경우, 즉 표준화된 검출 효과의 크기인 Δ_1=1.5가 실험자가 검출하고 싶은 최소 효과의 크기라고 하자. 반복수 r=4인 경우에는 단계의 수 rep=5를 실시한 경우에 Δ_1값이 처음으로 1.5 보다 작은 1.39가 되어서, 제1종 오류의 확률의 크기와 제2종 오류의 확률의 크기인 $\alpha = 0.05$, $\beta = 0.1$를 통제하는 결정을 내릴 수 있게 된다. 한편 Δ_1=1이 실험자가 검출하고 싶은 최소 효과의 크기인 경우에는 rep=8에서 처음으로 Δ_1값이 1보다 작게 되기에, 8단계를 실시해야 한다.

이제 효율적인 실험의 반복수 r^*와 단계의 수 rep^*를 결정하는 방법을 생각해 보자. 각 단계내의 실험의 반복수인 r을 고정시키고, 단계를 점차적으로 늘려가면서 Δ_1값이 최초로 1.5보다 작게 되는 단계의 수를 구한 후에, 이 단계의 수를 가장 작게 하는 실험의 반복수를 구하면, 실험의 반복이 6인 경우와 7인 경우로 단계의 수는 4가 된다. 단계의

<표 1> 분할법 I실험 자료의 분석 시, α=0.05에서 검정력 0.9를 만족하는 표준화된 최소 검출 효과의 크기

반복수 \ 단계수	2	3	4	5	6	7	8	9	10
1	20.96	5.55	3.55	2.78	2.36	2.09	1.90	1.75	1.63
2	14.82	3.92	2.51	1.97	1.67	**1.48**	1.34	1.24	1.15
3	12.10	3.20	2.05	1.61	**1.36**	1.21	1.10	1.01	0.94
4	10.48	2.77	1.77	**1.39**	1.18	1.05	**0.95**	0.88	0.82
5	9.38	2.48	1.59	**1.24**	1.06	0.94	0.85	0.78	0.73
6	8.56	2.27	**1.45**	1.14	0.96	0.85	0.77	0.71	0.67
7	7.92	2.10	**1.34**	1.05	0.89	0.79	0.72	0.66	0.62

수를 줄이는 것이 실험의 경제성(비용) 측면에서 일차적으로 중요하고, 그 다음에는 반복수를 줄여야 하기에, r^*=6이고, 단계의 수 rep*=4를 실시한 경우에 $\alpha=0.05$, $\Delta_1=1.45$에서 $\beta=0.1$되는 통계적인 결정을 내릴 수 있게 된다. 차선책은 r^*=4이고, 단계의 수 rep*=5를 실시한 경우에 $\alpha=0.05$, $\Delta_1=1.39$에서 $\beta=0.1$되는 통계적인 결정을 내릴 수 있게 된다.

4. 분할법 II에 의한 두 모평균의 비교

서론에서 언급한 바와 같이 현업에서 공정조건을 바꾸기 위해서는 생산을 잠시 멈추어야 하기에 반도체나 LCD 등 첨단 산업에서는 원료 탱크나 램프 교체 등 생산라인의 예방 정비(preventive maintenance) 시에나 가능하며, 공정조건을 바꾸기 위해서는 다음 예방 정비 시기 까지 기다려야 하기에 시간이 오래 걸려서, 실험을 축차적으로 실시하기에 기술적인 어려움이 따를 수 있다. 한 번에 실험을 설계해야 하고, 두 개의 공정조건의 변경이 제한적으로 가능하다면, 실험단위들을 일차단위와 이차단위로 나누어서 두 단계로 진행할 수 있다. 첫 번째 단계에서는 먼저 rep개의 반복을 갖는 각 공정조건을 2*rep 개의 일차단위에 랜덤하게 배치한다. 그 다음에는 일차단위에 랜덤하게 배치된 공정조건을 고정시키고 실험을 r회 반복한다. 일차단위내의 실험의 반복에 해당되는 각각의 실험단위를 이차단위라고 한다. 분할법 II의 모형은

$$y_{ijk} = \mu + \alpha_i + \varepsilon_{1k(i)} + \varepsilon_{2(ijk)},\ i=1,2\ ,\ j=1,\cdots,r,\ k=1,\cdots,rep \quad (2)$$

이다. 여기서 일차단위의 모형은 일원배치법의 모형과 동일하게 되어, 일차단위에 배치된 공법효과 α_i의 유의성을 검토하기 위해서는 일차단위 오차 $\varepsilon_{1k(i)}$와 비교해야 하고, 이때의 F-검정의 분자의 자유도는 1, 분모의 자유도는 2*rep-2이다. 축차적으로 실험을 실시할 수 있는 분할법 I에 비해서, F-검정의 분모에 해당되는 일차단위의 오차의 자유도가 커지기에 검정력이 커지게 되고, 이는 <표 2>에 주어진 분할법 II에 의한 실험자료의 분석시에 $\alpha=0.05$에서 검정력 0.9를 만족하는 표준화된 최소검출 효과의 크기 값인 Δ_1이 <표 1>의 분할법 I의 경우 보다 작음에서 확인할 수 있다. 예를 들어서 r=4, rep=4 인 경우에 <표 2>에서는 $\Delta_1=1.38$인 반면에 <표 1>에서는 $\Delta_1=1.77$이다.

<표 2> 분할법 II실험 자료의 분석 시, α=0.05에서 검정력 0.9를 만족하는 표준화된 최소 검출 효과의 크기

반복수 \ 일차단위의 처리반복수	2	3	4	5	6	7	8	9	10
4	3.40	1.79	1.38	1.17	1.04	0.94	0.87	0.81	0.77
5	3.04	1.61	1.24	1.05	0.93	0.85	0.78	0.73	0.69
6	2.77	1.47	1.13	0.96	0.85	0.77	0.71	0.66	0.63
7	2.57	1.36	1.05	0.89	0.79	0.71	0.66	0.62	0.58

분할법 Ⅱ에서의 효율적인 실험의 반복수 r^*와 일차단위에 배치된 각 공법의 반복수 rep^*를 3절에서 소개된 분할법 Ⅰ에서와 마찬가지 방법으로 <표 2>에서 구해보면 $r^*=6$이고, $rep^*=3$를 실시한 경우에 $\alpha=0.05$, $\Delta_1=1.47$에서 $\beta=0.1$되는 통계적인 결정을 내릴 수 있게 된다. 차선책은 $r^*=4$이고, $rep^*=4$를 실시한 경우에 $\alpha=0.05$, $\Delta_1=1.38$ 에서 $\beta=0.1$ 되는 통계적인 결정을 내릴 수 있게 된다.

5. 요 약

기존의 표준공법과 신공법간의 품질특성치의 평균에 차이가 있는지를 알아보기 위한 가장 실용적인 실험 설계는 표준공법으로 진행되고 있는 공정을 신공법으로 바꾼 후에 연속적으로 생산해서 실험 자료를 얻는 것이다. 이 실험 자료로부터 평균 수율의 차이가 있는 지를 검정하기 위해서는 표준공법으로 연속해서 생산된 제품들의 수율에 관한 정보를 추가적으로 필요로 한다. 표준공법으로 연속해서 생산된 제품들의 수율에 관한 충분한 최근의 자료가 없으면 분할법에 의한 실험설계가 실용적인 설계이다. 분할법 I은 각 단계의 실험을 축차적으로도 실시할 수 있다는 장점이 있다. 단계 1에서는 표준공법과 신공법의 실시순서를 랜덤하게 정한다. 그다음에 정해진 방법에 따라서 연속해서 실험을 r회 실시한다. 단계 $2,\cdots,rep$에서도 동일한 방법으로 실험을 실시한다. 분할법 Ⅱ는 모든 실험을 한번에 실시하는 설계로, 두 개의 공정조건의 변경이 제한적으로 가능하다면 먼저 rep개의 반복을 갖는 각 공법을 2*rep개의 일차단위에 랜덤하게 배치하고, 일차단위에 랜덤하게 배치된 공정조건을 고정시키고 실험을 r회 반복한다. 분할법에서 일반적으로 공정조건을 변경하는데 비용과 시간이 많이 들어서 rep를 적게 하는 효율적인 실험의 반복수 r^*을 결정하는 것이 엔지니어들의 관심사항이다. 분할법 I이나 분할법 Ⅱ에서의 효율적인 실험의 반복수는 $r^*=6$이고, 차선책은 $r^*=4$이다.

참 고 문 헌

[1] Box, G. E. P., Hunter, W. G., and Hunter, J. S.(1968), *Statistics for Experimenters*, John Wiley & Sons, New York.
[2] The MATHWorks, INC.(1992), *MATLAB User's Guide*, The MATHWorks, INC., Mass
[3] 박성현(2003),「현대실험계획법」, 민영사.
[4] 임용빈(1998), "실험계획법에서 최소 표준화 검출 가능 효과의 크기에 관한 연구",「품질경영학회지」, 26권 4호, pp. 239-249.
[5] 임용빈, 안병진, 성내경(1998),「실험설계의 원리와 응용」, 자유아카데미.