

상관계수를 이용한 혼합물실험의 경제적 성분선별 방법

김정숙¹ · 변재현^{1*} · 최경미²

¹경상대학교 산업시스템공학부, 공학연구원/²홍익대학교 교양과

An Economic Screening Method for Mixture Experiments Using Correlation Coefficients

Jeong-Suk Kim¹ · Jai-Hyun Byun¹ · Kyungmee Choi²

¹Department of Industrial and Systems Engineering and Engineering Research Institute
Gyeongsang National University, Jinju 660-701

²Department of Mathematics, Hongik University at Jochiwon, 339-701

A mixture experiment is a special type of the response surface experiment in which the factors are the components of a mixture, and the response variable is a function of the proportions of each component. When a new mixture product is developed, there are a large number of components, and the first objective of the experiment is to identify the ones that are most important by doing a screening experiment. We propose a method of screening mixture components using the correlation coefficients when t-tests cannot identify significant components.

Keywords: Mixture Experiments, Mixture Screening Designs, Component Effects, Correlation Coefficients, Grouping of the Components

1. 서론

실험계획은 제품 개발이나 공정 개선을 효율적으로 수행하는데 유용한 실험수행 방법인데, 실험계획에서 고려하는 어떤 인자의 수준은 다른 인자들의 수준과는 독립적으로 정해진다. 하지만 반도체의 식각공정(etching process)에서 식각에 쓰이는 용액의 최적 혼합비율을 찾는 등의 혼합물 실험에서는, 전체 용액 중 한 성분이 차지하는 비율이 커질수록 다른 성분들의 비율은 작아지기 때문에 각 성분의 수준은 다른 성분의 수준과 독립적으로 정해질 수 없다. 혼합물 실험에서는 혼합물의 성분 비율이 인자가 되고, 각 성분 비율의 함수가 반응변수이며, 이 때 실험의 목표는 반응변수를 최적화하는 혼합비율을 찾는 것이다(Cornell, 2002; Myers and Montgomery, 2002). 이러한 혼합물 실험에서는 성분비율의 합이 항상 일정하므로 실험영역은 심플렉스 형태이고, 실험점은 심플렉스의 각 꼭지

점, 선분 또는 면의 중심, 심플렉스 영역의 중심 등에서 선정된다. 선정된 실험점에서 실험을 실시하여 데이터를 얻은 후 분석을 통해 유의한 모형을 적합하고, 반응변수를 최적화하는 성분비율의 조건을 도출한다.

보통 혼합물 생산 공정을 위한 실험에서 제품의 특성에 영향을 주는 성분이 3~4가지 정도이면 적은 수의 실험만으로도 최적 혼합비율을 비교적 쉽게 찾을 수 있다. 하지만 고려해야 하는 성분의 수가 많을 때에는 성분들의 최적 혼합비율을 구하는 것이 쉽지 않다. 이런 경우에는 적은 수의 실험 수로도 중요한 성분을 경제적으로 선별할 수 있는 실험계획을 이용하여 실험을 실시하고, 분석을 통해 유의한 성분을 우선 선별한 후 최적 혼합비율을 찾기 위한 추가적인 실험을 실시하는 것이 바람직하다.

제품의 특성에 유의한 영향을 주는 성분을 선별하기 위하여 Cornell(2002)은 각 성분 효과의 유의성을 파악하기 위한 t-검

이 논문은 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(과제번호: R05-2003-000-10445-0).

* 연락처 : 변재현, 660-701 경남 진주시 가좌동 900 경상대학교 산업시스템공학부, Tel : 055-751-5336, Fax : 055-762-6599

E-mail : jbyun@gnu.ac.kr

2007년 02월 접수; 2007년 05월 수정본 접수; 2007년 06월 게재 확정.

정 방법과 유의성 검정의 결과 유의한 효과가 나타나지 않을 때 이용할 수 있는 공분산 이용 방법을 제시하였다. 반면, Myers and Montgomery(2002)는 Response trace plot을 이용하여 유의한 효과를 선별하는 방법을 제시하였다. 그러나 성분 효과의 공분산이나 Response trace plot을 이용하는 방법에서는 성분을 선별하는 기준이 명확하지 않아서 중요한 성분을 선별하는데 어려움이 따른다. 따라서 명확한 기준을 통하여 중요한 성분을 체계적으로 선별하는 방법이 필요하다. 본 논문에서는 성분 선별을 위한 t-검정 결과 중요한 성분을 선별할 수 없을 때, 성분 효과 간 상관계수를 이용하여 체계적으로 주요 성분들을 선별해내는 방법을 제시하고자 한다. 기존 실험계획과는 달리 본 연구에서는 상관관계가 강한 성분들을 그루핑함으로써 추후 실험을 할 때에는 동일 그룹에 있는 성분 하나만을 대표적으로 고려하는 방식의 선별방법을 이용하고자 한다.

제 2절에는 혼합물 실험의 성분 선별을 위한 실험계획의 종류와 실험 데이터를 얻은 후 각 성분의 요인효과를 이용하는 선별방법에 대한 기존 연구를 소개한다. 본 연구의 주제인 상관계수를 이용하여 성분을 그루핑함으로써 성분을 선별하는 절차는 제 3절에 기술하였다. 제 4절에는 예제를 통하여 이러한 선별 절차를 예시하고, 제 5절에 본 연구의 결론을 기술한다.

2. 혼합물 선별 실험계획의 종류와 성분 선별 방법

2.1 혼합물 선별 실험계획의 종류

제품을 이루는 성분의 수가 많은 경우, 제품의 특성에 중요한 영향을 주는 성분을 경제적으로 선별하기 위해서는 실험점의 수를 최소화하면서 각 성분의 효과를 구할 수 있는 실험계획법을 이용해야 한다. 혼합물 선별 실험계획은 크게 두 가지 경우로 나누어 볼 수 있다. 첫째는 성분비율에 제약이 없거나 제약이 있더라도 실험가능 영역이 심플렉스 형태를 이룰 때이다. 이 경우에는 심플렉스 선별계획법을 이용한다(Snee and Marquardt, 1976). 심플렉스 선별계획법은 실험가능영역이 심플렉스이거나 심플렉스 영역 내에 작은 심플렉스 형태이며, 실험점은 꼭지점, 모서리점, 내부점, 중심점으로 이루어진다.

Table 1. Simplex screening design points

Name	Number	Composition
Vertices	q	$x_i = 1, x_j = 0, j \neq i$
End Effects	q	$x_i = 0, x_j = \frac{1}{q-1}, j \neq i$
Interior	q	$x_i = \frac{q+1}{2q}, x_j = \frac{1}{2q}, j \neq i$
Centroid	1	$x_i = \frac{1}{q}, i = 1, 2, \dots, q$

q개의 성분으로 이루어진 혼합물의 경우, 심플렉스 선별계획법을 이용하면 <Table 1>과 같이 3q+1개의 실험점이 생긴다.

두 번째는 성분비율에 제약이 있고 실험가능 영역이 심플렉스 형태를 이루지 않을 때이다. 이 경우에는 XVERT(Extreme Vertices Design)와 XVERT를 수정한 MXMSDs (Modified XVERT Mixture Screening Designs)를 이용할 수 있다(Snee and Marquardt, 1974; Piepel, 1990). XVERT는 제한된 영역의 꼭지점을 이용하여 실험을 계획하는 것으로, 일반적인 선별 실험에 이용되는 일부 실시법 또는 Plackett-Burman 계획법을 혼합물 실험에 적용한 것이다. 성분의 상한에서 하한을 뺀 값에 순위를 매긴 후 가장 작은 값을 가지는 성분부터 순서대로 나열하여 X_1, X_2, \dots, X_q 로 다시 정의한다. X_1, \dots, X_{q-1} 을 대상으로 그들의 상한과 하한을 수준으로 삼아서 2^{q-1} 개의 수준조합으로 실험계획을 생성하고, q번째 성분인 X_q 의 수준은 모든 성분 비율의 합이 1이 되도록 계산된 값을 입력한다. X_q 열에 입력된 값이 X_q 의 상한 또는 하한을 벗어난 경우에는 X_1, \dots, X_{q-1} 의 값을 조절하도록 한다.

MXMSDs는 XVERT를 수정한 선별 계획법으로, q번째 성분의 비율이 상한 또는 하한을 벗어날 때 X_1, X_2, \dots, X_{q-1} 의 값을 계산하는 방법이 XVERT와 다르다. XVERT에서는 q번째 성분 비율을 계산하기 위해 나머지 (q-1)개의 성분을 모두 고려하지만, MXMSDs에서는 우선 X_{q-1} 열에 입력되어 있는 값을 조절하여 제약조건이 만족되도록 X_{q-1} 의 수준 값을 정한다. 만일 X_{q-1} 의 값을 조절하여도 제약조건이 만족될 수 없으면 그 다음 성분인 X_{q-2} 의 값을 조절한다. 이러한 방법을 순차적으로 적용하여 제약조건이 만족되도록 한다. 위와 같은 실험 계획에 의해 생성된 실험점에서 실험을 실시하고, 데이터를 분석하여 반응변수에 유의한 영향을 주는 성분을 선별한다.

성분 선별을 위한 혼합물 모형은 식 (1)과 같이 Scheffe의 1차 혼합물 모형을 이용하며, 각 성분의 효과는 실험으로 얻은 데이터를 이용하여 성분비율의 변화에 따른 반응값의 변화량으로 구할 수 있다(Cornell, 2002; Myers and Montgomery, 2002). 성분 i의 효과는 x_i 축을 따라 i번째 성분비율이 1인 경우의 반응값과 0인 경우의 반응값의 차이로 정의한다. x_i 축은 성분 i의 꼭지점에서 대응하는 선분 혹은 면의 중심에 수직으로 그은 직선이다. <Figure 1>은 q=3 인 경우에 각 성분의 축을 나타낸다.

$$E(y) = \beta_1x_1 + \beta_2x_2 + \dots + \beta_qx_q \tag{1}$$

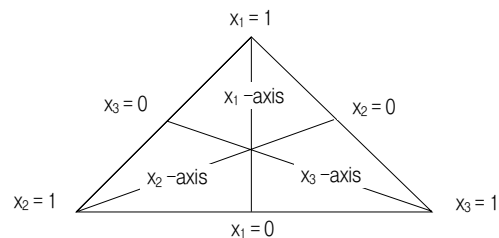


Figure 1. q=3, x_i axes, $i=1, 2, 3$

성분 i 의 효과는 식 (2)와 식 (3)으로 나타낼 수 있다. 여기서 b_i 는 β_i 의 추정치이다.

$$E_i = \left(b_i - (q-1)^{-1} \sum_{j \neq i}^q b_j \right) \quad (2)$$

$$E_i = R_i \left(b_i - (q-1)^{-1} \sum_{j \neq i}^q b_j \right) \quad (3)$$

성분비율에 제약이 없는 경우에는 식 (2)를 이용하고, 성분 비율에 제약이 있을 때에는 식 (3)을 이용한다. 여기서, 식 (3)의 R_i 는 성분 x_i 의 상한 U_i 와 하한 L_i 의 차이인 $U_i - L_i$ 를 말한다. 일반적으로 혼합물 실험에서는 성분비율에 제약이 있는 경우가 많으므로 성분의 효과를 구하는 데에는 식 (3)이 주로 이용된다.

2.2 주요성분의 기준 선별 방법

혼합물 실험 데이터를 이용하여 변수를 선별하는 기존 방법은 3가지가 있다(Kim and Byun, 2006). 우선 각 성분의 효과를 구한 다음, 각 성분비율의 변화에 따른 반응값의 경향을 그래프를 통하여 시각적으로 파악할 수 있는 Response trace plot을 이용하는 방법이 있다. 이 방법은 각 성분 값의 이동에 따른 반응값의 변화를 직선 또는 곡선으로 나타낸 것으로서 그래프에 나타난 선의 변화의 정도를 이용하여 성분의 유의성을 판단하는 것이다. Response trace plot을 이용하면 성분비율의 변화에 따른 반응값을 그래프로 나타내어 실험자가 유의한 성분을 시각적으로 쉽게 선별할 수 있는 장점이 있으나, 성분을 선별하는 기준이 실험자의 주관적 판단에 의존하게 되는 단점이 있다.

두 번째 방법은 회귀계수를 이용하여 t-검정에 의하여 효과의 유의성을 판단하는 것이다(Cornell, 2002). 이 방법에서는 우선 t-검정을 실시하여 각 성분의 t-통계량과 p-값을 얻고, p-값이 유의수준 보다 작은 성분이 반응값에 유의한 영향을 주는 성분으로 선별된다. 이 방법을 이용하여 적절한 수의 변수가 선별되면 좋으나 유의한 성분이 나타나지 않을 때에는 문제가 생긴다. 특히 선별실험처럼 실험의 크기에 비하여 고려하는 성분의 수가 많으면 t-검정 결과 유의한 성분이 쉽게 선별되지 않을 가능성이 크다.

세 번째로는 성분 효과의 공분산을 이용하는 방법이 있는데, 이것은 t-검정 결과 거의 모든 성분의 p-값이 유의수준보다 커서 유의한 성분을 선별할 수 없는 경우에 유의한 성분을 선별하기 위하여 사용된다. 이 방법은 Cornell(2002)이 제안했는데, 성분 효과의 분산-공분산 행렬에서 절대값이 큰 공분산 값에 대응하는 두 성분을 선택한 후, 두 성분을 묶어서 하나의 성분으로 만들고, 이 성분과 나머지 성분들을 이용하여 다시 t-검정을 실시하는 것이다. 어떤 두 성분의 공분산 절대값이 크다는 것은 그 두 성분이 반응변수에 미치는 영향을 구분할 수 없

다는 것을 의미하므로, 설계행렬에서 이 두 성분의 혼합비를 합쳐서 새로운 설계행렬을 구하고 검정을 순차적으로 실시하면 유의한 성분을 선별할 수 있다. 이 방법은 상관관계가 큰 두 성분을 결합함으로써 성분 효과의 분산과 공분산이 작아지도록 하여 유의한 효과를 찾을 수 있게 하는 장점을 갖고 있으나, 최종 결과를 선정하기 위한 기준이 없고, 이렇게 성분을 합하는 절차를 어느 시점까지 진행할 것인지에 대한 정지 규칙(stopping rule)이 없다는 단점이 있다.

3. 성분 효과의 상관계수를 이용한 성분 선별

효과 유의성 검정을 통하여 유의한 성분을 선별할 수 없을 때 이용하는 공분산 이용방법은 성분을 선별하는 기준이 명확하지 않다. 본 논문에서는 표준화된 통계량으로서 정확한 통계적 검정을 할 수 있는 상관계수(correlation coefficient)를 이용하여 성분변수를 그루핑함으로써 성분을 선별하는 기준을 제시하고자 한다. 이를 위하여 데이터 수에 따른 상관계수의 기각역을 이용하면 두 성분 효과의 상관관계의 유의성을 파악할 수 있다.

상관계수를 이용하여 유의한 성분변수들로 그루핑하는 절차는 다음과 같다. 먼저 성분 효과의 분산-공분산 행렬을 구하고 다음의 식 (4)에 대입하여 두 성분 효과의 상관계수를 구한다.

$$r_{ij} = \frac{\widehat{Cov}(E_i, E_j)}{S_{E_i} \cdot S_{E_j}}, \quad i \neq j \quad (4)$$

계산된 상관계수를 식 (5)에 대입하여 t-통계량을 구하고, 식 (6)을 이용하여 t-분포표에서 데이터 수에 따른 기각역을 찾는다.

$$t_0 = \frac{r_{ij}}{\sqrt{\frac{1-r_{ij}^2}{n-2}}} \quad (5)$$

$$|t_0| > t(n-2, 2/\alpha) \quad (6)$$

식 (5)에 의해 계산된 t-통계량의 절대값과 식 (6)에 의한 기각역을 비교하여, 기각역 보다 큰 값을 나타내는 것들 중에서 가장 큰 값에 대응하는 두 개의 성분을 선택한다. 선택된 두 성분을 그루핑하여 하나의 성분으로 만들고, 그루핑된 성분과 나머지 성분으로 다시 성분 효과의 유의성 검정을 실시한다. 이러한 과정을 t-통계량이 기각역보다 작을 때까지 반복하여 수행한다. 최종적으로는 그루핑된 변수의 집합들 중에서 F-value가 가장 크게 되는 결과를 선택한다. 성분 선별 절차를 순서대로 나타내면 <Figure 2>와 같다. 이 방법은 상관관계가 큰 두 성분을 결합함으로써 성분 효과들의 분산과 공분산이 작아지도록 하여 유의한 효과를 찾을 수 있게 하며, 데이터 수

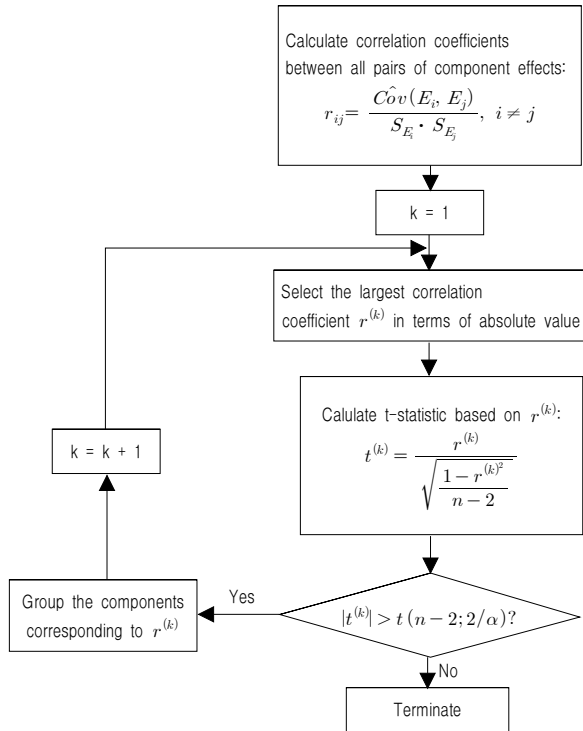


Figure 2. Components grouping procedure

에 따른 기각역을 선별 기준으로 삼음으로써 체계적으로 성분을 선별할 수 있도록 한다.

본 논문에서 제시하는 방법을 이용할 때에 t-검정 결과가 모든 성분에 대하여 항상 유의하게 나오는 것은 아니다. 유의하지 않은 성분은 경제성이나 편리성을 고려하여 일정 수준에 값을 고정하는 것이 좋다. 또한 F-값이 유의하게 나오지 않는 경우가 생기면 이는 성분들 간에 상관관계가 작은 경우이므로 2.2절에서 지적인 일반적 t-검정을 이용하면 된다.

4. 예제

Cornell(2002)은 옥탄가를 결정하는 가솔린 혼합 성분 7가지를 이용하여 1.5mL Pb/gal 모터법 옥탄가를 최적화하기 위한 실험을 실시하였다. 실험에 고려된 성분은 x_1, x_2, \dots, x_7 이고, 총 12개의 실험점에서 실험을 실시하였으며, 성분비율의 상·하한은 다음과 같다(Cornell, 2002).

$$0 \leq x_1 \leq 0.21, \quad 0 \leq x_2 \leq 0.62, \quad 0 \leq x_3 \leq 0.12, \\ 0 \leq x_4 \leq 0.62, \quad 0 \leq x_5 \leq 0.12, \quad 0 \leq x_6 \leq 0.74, \\ 0 \leq x_7 \leq 0.08$$

Response trace plot을 이용하여 분석한 결과를 <Figure 3>에 나타내었다. <Figure 3>을 보면 A(x_1), C(x_3), D(x_4), F(x_6), G(x_7) 직선의 기울기가 상대적으로 크며, B(x_2), E(x_5) 직선의 기울기는 아주 작게 나타남을 알 수 있다.

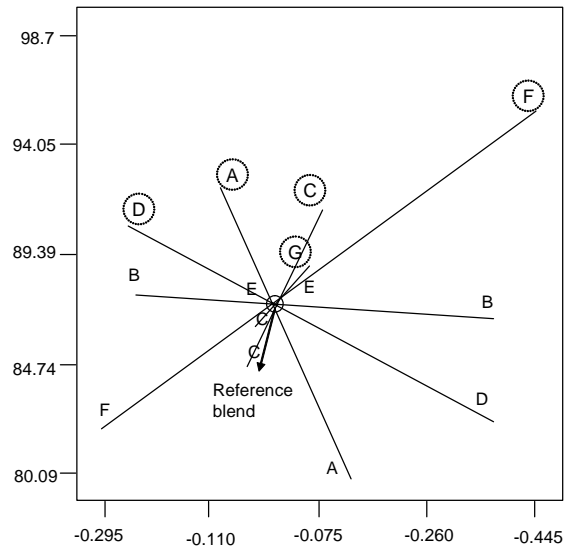


Figure 3. Response trace plot

데이터를 분석한 결과, 각 성분에 대한 효과와 t-검정 결과는 <Table 2>에 나타내었다.

Table 2. Component effects and t-test result(individual components)

Component	Effect	t-value	p-value
x_1	-14.12	-0.25	0.8099
x_2	-4.36	-0.32	0.7594
x_3	6.90	0.14	0.8957
x_4	-10.68	-0.84	0.4385
x_5	-0.59	-0.23	0.8251
x_6	7.21	0.47	0.6563
x_7	2.33	0.28	0.7923

<Table 2>를 보면 p-값이 모두 유의수준 0.05보다 크다. 공분산을 이용하여 변수를 선별한 최종 결과는 <Table 3>에 나타내었다(Kim and Byun, 2006).

Table 3. Final analysis result by covariance method

Component	Effect	t-value	p-value
$(x_1 + x_3)$	24.32	4.13	0.0044
x_2	30.04	6.57	0.0003
$(x_4 + x_6)$	54.06	6.67	0.0003
x_5	7.67	4.95	0.0017
x_7	-21.34	-6.21	0.0004

분석 결과 (x_1+x_3) , (x_4+x_6) , x_5 , x_7 의 p-값이 모두 0.05보다 작으므로 반응변수에 유의한 영향을 주는 성분으로 선별할 수 있다. 여기서 $(x_1 + x_3)$ 은 이들 2개의 성분의 상관관계가 커서

이들이 반응변수에 미치는 영향을 구분하지 않고 종합적으로 본다는 것을 의미한다.

본 논문에서 제시한 성분 효과의 상관계수를 이용하여 유의한 성분을 선별해 보았다. 성분 선별 절차에 따라 성분 효과의 분산-공분산 행렬을 구하였고, 분산과 공분산을 식 (4)에 대입하여 상관계수를 계산하였다. 계산된 상관계수를 식 (5)에 대입하여 t-통계량을 구하고, 식 (6)의 t-분포표를 보고 기각역을 구하였다. 데이터의 수 $n = 12$ 이고 유의수준 $\alpha = 0.05$ 인 경우의 기각역은 2.228이다. 식 (5)에 의해 계산된 t-통계량의 절대값이 기각역보다 큰 것들 중에 가장 큰 값을 가지는 두 성분은 x_1 과 x_3 이며 t-통계량은 -35.2 이다. 이 두 성분을 합하여 하나의 성분 ($x_1 + x_3$)으로 만들고 다시 분석을 실시하였으며, 그 결과를 <Table 4>에 나타내었다.

Table 4. Analysis result 1 by correlation coefficient method

Component	Effect	t-value	p-value
($x_1 + x_3$)	-7.45	-1.36	0.2224
x_2	-4.04	-0.83	0.4404
x_4	-11.32	-1.13	0.3005
x_5	-0.57	-0.42	0.6885
x_6	7.59	0.93	0.3894
x_7	3.35	0.68	0.5208

<Table 4>를 보면 모든 성분의 p-값이 0.05보다 크다. 따라서 다시 분산 - 공분산 행렬을 구하고 두 성분 효과의 상관계수를 이용하여 t-통계량을 계산하였다. 기각역과 t-통계량을 비교해 본 결과 x_4 와 x_7 에 해당하는 t 값이 -26.0으로 가장 컸다. 두 성분을 합하여 하나의 성분으로 만들고 다시 분석을 실시하였다(<Table 5>).

Table 5. Analysis result 2 by correlation coefficient method

Component	Effect	t-value	p-value
($x_1 + x_3$)	-3.78	-4.86	0.0018
x_2	-1.08	-0.97	0.3658
($x_4 + x_7$)	-5.00	-3.88	0.0060
x_5	0.25	0.45	0.6651
x_6	13.57	10.31	0.0001

Table 8. F-values corresponding to the grouping sets

Analysis method	grouping	MS _{model}	MS _{error}	F-value
Correlation coefficient	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	77.38	0.70	110.77
	($x_1 + x_3$), x_2, x_4, x_5, x_6, x_7	92.86	0.59	158.42
	($x_1 + x_3$) $x_2, (x_4 + x_7), x_5, x_6$	155.98	0.55	209.40
	($x_1 + x_3$), $x_2, (x_4 + x_7), (x_5 + x_6)$	153.43	0.94	163.30
	($x_1 + x_3 + x_4 + x_7$), $x_2, (x_5 + x_6)$	229.96	0.87	262.87
Covariance(final result)	($x_1 + x_3$), $x_2, (x_4 + x_6), x_5, x_7$	119.89	2.89	38.73

분석 결과 ($x_1 + x_3$), ($x_4 + x_7$), x_6 의 p-값이 0.05보다 작으므로 반응변수에 유의한 영향을 주는 성분으로 선별할 수 있었다. 여기서 기각역보다 큰 통계량을 가지는 성분이 있는지 확인하기 위하여 t-통계량과 기각역을 비교해 본 결과, 기각역보다 큰 값(-6.38)이 있었으며 그에 대응하는 성분들이 x_5 와 x_6 이다. 따라서 두 성분을 합하여 하나의 성분으로 만들어 다시 분석을 실시하였다(<Table 6>).

Table 6. Analysis result 3 by correlation coefficient method

Component	Effect	t-value	p-value
($x_1 + x_3$)	-3.45	-3.35	0.0101
x_2	-0.43	-0.34	0.7453
($x_4 + x_7$)	-5.18	-4.05	0.0037
($x_5 + x_6$)	16.05	19.47	0.0001

<Table 6>을 보면 ($x_1 + x_3$), ($x_4 + x_7$), ($x_5 + x_6$)의 p-값이 0.05보다 작으므로 유의하다. <Figure 2>의 절차를 따라 기각역과 t-통계량을 비교하였으며, ($x_1 + x_3$)과 ($x_4 + x_7$)에 대응하는 t-통계량의 값이 -4.907로 가장 컸다. 이들을 합하여 성분 ($x_1 + x_3 + x_4 + x_7$)로 만든 후 다시 분석을 실시하여 얻은 t-검정 결과를 <Table 7>에 나타내었다.

Table 7. Final analysis result by correlation coefficient method

Component	Effect	t-value	p-value
($x_1 + x_3 + x_4 + x_7$)	-9.26	-13.58	0.0001
x_2	-3.09	-3.44	0.0073
($x_5 + x_6$)	15.35	17.40	0.0001

<Table 7>을 보면 성분 ($x_1 + x_3 + x_4 + x_7$), x_2 그리고 ($x_5 + x_6$)의 p-값이 모두 유의수준 0.05보다 작으므로 이들 성분변수의 그룹을 반응변수에 유의한 영향을 주는 성분으로 선별할 수 있다. <Table 8>은 상관계수를 이용한 분석결과를 회귀모형의 F-value 측면에서 정리한 것인데 가장 바람직한 그루핑 결과는 F-value가 가장 높게 나오는 ($x_1 + x_3 + x_4 + x_7$), x_2 , ($x_5 + x_6$)이다.

<Table 8>을 보면 또한 본 논문에서 제시한 상관계수를 이용하여 구한 그루핑 결과의 F-value는 262.87로서 공분산을 이용한 경우의 38.73보다 아주 큰 값을 알 수 있다.

상관계수를 이용한 그루핑에 의한 성분 선별 방법을 본 예제에 적용한 최종 선별 결과는 $(x_1 + x_3 + x_4 + x_7)$, x_2 , $(x_5 + x_6)$ 이다. 7개의 성분변수를 대상으로 12회 실험을 실시하고, 데이터를 분석하여 상관관계를 이용하여 그루핑을 한 결과 3개의 그룹으로 선별되었다. 추후 최적조건을 찾기 위한 실험에서는 각 그룹에서 실험의 용이성 등을 고려하여 성분을 하나씩 선별하여(예를 들면, x_1 , x_2 , x_5) 이들을 대상으로 자세한 혼합물 실험을 실시하면 될 것이다.

5. 결론

본 논문은 제품을 구성하는 성분의 수가 많을 경우 적은 실험횟수로서 제품의 특성에 유의한 영향을 주는 성분을 선별하기 위한 방법을 연구하였다. 본 논문에서는 특히 성분 선별을 위한 t-검정 결과 중요한 성분을 선별할 수 없을 때, 성분 효과 간 상관계수를 이용하여 체계적으로 주요 성분들을 선별해내는 방법을 제시하였다. 요인배치 또는 부분요인배치 실험계획과는 달리 본 연구에서는 상관관계가 강한 성분들을 그루핑함으로써 추후 실험을 할 때에는 동일 그룹에 있는 성분 하나만을 대표적으로 고려하는 방식의 선별방법을 이용하였다.

성분 효과 간 상관관계가 큰 경우의 혼합물실험 성분선별을

위해 기존에 개발된 공분산 이용방법은 선별 기준이 명확하지 않아서 유의한 성분을 선별하는데 어려움이 따른다. 그러나 본 연구에서 제시하는 상관계수를 이용한 검정방법은 성분 선별방법의 표준화된 통계량을 이용함으로써 성분변수를 그루핑하여 선별할 수 있는 명확한 기준을 제시하고 있다. 본 연구의 결과는 혼합물 제품을 개발하는 연구자 또는 혼합물 제품을 생산하는 현장의 실무자들이 적은 실험횟수로 경제적으로 중요한 성분을 선별하는 데에 도움이 될 것으로 기대한다.

참고문헌

- Cornell, J. A. (2002), *Experiments with Mixtures, Designs, Models, and the Analysis of Mixture Data, 3rd Edition*, Wiley, New York.
- Kim, J. -S. and Byun, J. -H. (2006), A Study on Implementation Guideline of Screening Mixture Design and Analysis for the Development of New Mixture Products, *IE Interfaces*, **16**(2), 117-123.
- Myers, R. H. and Montgomery, D. C. (2002), *Response Surface Methodology, Process and Product Optimization Using Designed Experiments, 2nd Edition*, Wiley, New York.
- Piepel, G. F. (1990), Screening Designs for Constrained Mixture Experiments Derived from Classical Screening Designs, *Journal of Quality Technology*, **22**(1), 23-33.
- Snee, R. D. and Marquardt, D. W. (1974), Extreme Vertices Designs for Linear Mixture Models, *Technometrics*, **16**(3), 399-408.
- Snee, R. D. and Marquardt, D. W. (1976), Screening Concept and Designs for Experiments with Mixtures, *Technometrics*, **18**(1), 19-29.