The Korean Journal of Applied Statistics (2017) DOI: https://doi.org/10.5351/KJAS.2017.30.3.335
30(3), 335–344

Projection analysis for split-plot data

Jaesung Choi[a,1]

[a]Department of statistics, Keimyung University

(Received December 5, 2016; Revised April 1, 2017; Accepted April 1, 2017)

Abstract

This paper discusses a method of analyzing data from split-plot experiments by projections. The assumed model for data has two experimental errors due to two different experimental sizes and some random components in treatment effects. Residual random models are constructed to obtain sums of squares due to random effects. Expectations of sums of squares are obtained by Hartley's synthesis. Estimable functions of fixed effects are discussed.

Keywords: split-plot designs, random components, projection, fitting constants method, estimable functions

1. 서론

실험의 성격에 있어 실험단위의 반응에 영향을 미치는 요인들의 수준배치에 서로 다른 크기의 실험단위가 요구되는 실험을 생각해 보기로 한다. 하나의 실험구가 서로 다른 크기의 실험단위들로 구성될 때 크기가 큰 실험단위를 주구(whole-plot) 작은 크기의 실험단위를 세구(sub-plot)라 부르며 주구와 세구의 실험구를 분할구(split-plot)라 부른다. 분할구 유형의 실험계획은 작은 크기의 실험단위가 큰 크기의 실험단위에 내포되어 있는 설계구조를 갖는다. 따라서 적어도 둘 이상의 서로 다른 크기의 실험단위를 갖는 실험상황을 가정하게 된다.

분할구 유형의 실험계획에서 제기되는 주된 관심문제는 처리구조의 처리들에 배정되는 서로 다른 크기의 실험단위들을 확인하고 결과의 자료를 분석하기 위한 타당한 모형을 구축하는 문제로 기술된다. 실험단위의 개별크기와 관련된 변이를 측정하는 변동요인을 확인하는 것은 중요하다. 이들 변동요인은 각각의 서로 다른 오차항들을 계산하기 위하여 이용되고 각 오차항은 추정평균의 표준오차의 추정치를 계산하기 위해 이용된다. 이는 변동요인들이 서로 다른 크기의 실험단위들의 이용으로 부터 발생함을 의미하고 있다. 둘 이상의 서로 다른 크기의 실험단위를 포함하고 있기때문에 모수나 모수 간의 비교를 위한 표준오차 추정값들은 하나 또는 다수의 변동요인들을 포함하게 된다.

분할구 유형의 실험자료에 대한 분석방법과 모형에 관한 논의는 Milliken과 Johnson (2009), Montgomery (1976) 그리고 Searle (1971) 등의 많은 문헌에서 다루어지고 있다. 고정모형, 확률모형 또는 혼합모형의 가정하에 확률효과와 관련된 분산성분의 추론방법은 Searle 등 (1992) 등에서 보여진다. 이와는 달리 실험설계와 관련된 서로 다른 크기의 실험단위들 간의 변이는 서로 다른 특성의 오차성분들로 간주되고 자료변동의 요인으로 간주된다. 처리구조의 처리 또는 요인들의 수준변화에 따른 확률효과

[1]Department of Statistics, Keimyung University, 1095 Dalgubeoldaero, Dalseogu, Daegu 42601, Korea.
E-mail: jschoi@kmu.ac.kr

와 실험단위의 변이를 나타내는 오차는 서로 독립임을 가정하고 있다. 이러한 가정은 사영분석의 기초를 제공하게 된다.

본 논문은 처리에 해당하는 효과들 외에 분할구 유형의 실험설계와 관련된 다수의 오차항을 갖는 선형모형의 가정하에 실험자료의 분석을 위한 분석방법을 논의하고자 한다. 처리구조에 포함된 확률효과가 아닌 실험설계의 구조로부터 발생하는 오차항의 추가로 인해 고정효과의 분석에 영향을 주고 있는 경우이다. 다시말하면, 분석모형이 처리구조내 고정효과와 확률효과 그리고 설계구조와 관련된 다수의 오차항을 포함할 때 사영의 관점에서 자료를 분석하는 방법을 다루고자 한다. 사영에 관한 논의는 Graybill (1983), Johnson과 Wichern (1988) 등에서 보여진다.

관측반응에 대한 선형모형의 가정에서 모형내 모수의 추정방법으로 최소제곱법의 이용이 가능할 때 사영의 관점에서 동일한 분석이 이루어질 수 있다. 본 논문은 분할구 유형의 실험자료를 분석하기 위한 사영의 이용에서 변동요인에 따른 제곱합의 계산과 관련된 모형의 표현식, 잔차제곱합, 이차형식의 기댓값 그리고 추론을 위한 분포의 자유도 계산 등을 논의한다. 사영의 이용은 벡터공간에서 사영의 개념을 이해할 수 있을 뿐만아니라 행렬의 다양한 성질을 활용할 수 있다는 점에서 유용하며 또한 아직 실험계획과 관련된 사영분석의 논의가 그리 활발하지 않다는 점에서 그 의미가 있다고 볼 수 있다.

2. 모형의 가정

실험단위의 반응에 영향을 주는 요인들로 고정요인과 확률요인 그리고 다수의 오차항이 포함되어 있을 때 실험자료를 분석하기 위한 모형으로 혼합효과의 선형모형을 가정하게 된다. 실험자료의 분석을 위한 일반적인 혼합모형의 행렬표현식은 다음과 같다.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \sum_{i=1}^{k} \boldsymbol{Q}_i \boldsymbol{\delta}_i + \sum_{j=1}^{r} \boldsymbol{W}_j \boldsymbol{\nu}_j + \boldsymbol{\epsilon}. \tag{2.1}$$

단, $\boldsymbol{y}$는 $n \times 1$인 관측벡터이고 $\boldsymbol{X}$는 원소가 0 또는 1로 구성되며 크기가 $n \times p$인 고정효과벡터의 계수행렬로 계수(rank)가 $q(< p)$이다. $\boldsymbol{\beta}$는 $p \times 1$인 모수벡터이다. $\boldsymbol{Q}_i$ $(i = 1, 2, \ldots, k)$는 확률효과벡터 $\boldsymbol{\delta}_i$의 계수행렬로 0 또는 1로 구성되며 크기가 $n \times c_i$인 완전열계수행렬(full column rank matrix)이다. $\boldsymbol{\delta}_i$는 $N(\boldsymbol{0}, \sigma^2_{\delta_i} \boldsymbol{I}_{c_i})$인 분포를 따르며 $\boldsymbol{\delta}_i$ $(i = 1, 2, \ldots, k)$는 상호 독립이라고 가정한다. $\boldsymbol{W}_j$ $(j = 1, 2, \ldots, r)$는 확률효과벡터 $\boldsymbol{\nu}_j$의 계수행렬로 0 또는 1로 구성되며 크기가 $n \times d_j$인 완전열계수행렬이다. $\boldsymbol{\nu}_j$는 $N(\boldsymbol{0}, \sigma^2_{\nu_j} \boldsymbol{I}_{d_j})$인 분포를 따르며 $\boldsymbol{\nu}_j$ $(j = 1, 2, \ldots, r)$는 상호 독립이라고 가정한다. $\boldsymbol{\epsilon}$은 $n \times 1$인 오차벡터이며 $N(\boldsymbol{0}, \sigma^2_{\epsilon} \boldsymbol{I}_n)$인 분포를 따른다고 가정한다. 오차벡터와 확률효과벡터들은 서로 독립으로 가정한다. 식 (2.1)은 $\boldsymbol{X}\boldsymbol{\beta}$로 주어지는 고정효과부분과 $\boldsymbol{\Sigma}_{i=1}^{k} \boldsymbol{Q}_i \boldsymbol{\delta}_i$인 확률효과부분 그리고 서로 다른 크기의 실험단위로부터의 변이를 나타내는 오차성분 $\boldsymbol{\Sigma}_{j=1}^{r} \boldsymbol{W}_j \boldsymbol{\nu}_j$와 가장 작은 크기의 실험단위의 변이를 나타내는 오차항의 네 성분으로 구성되어 있다. 고정효과들로 구성되는 고정부분(fixed part)과 확률효과들로 구성되는 확률부분(random part)으로 나누면 혼합모형의 분석은 고정부분의 분석과 확률부분의 분석으로 구분되어 진다. 고정부분의 분석은 확률부분의 분석후에 행해지게 된다. 즉, 분산성분들이 추정된 후 고정효과들의 추정가능함수에 관한 신뢰구간추정이나 가설검정에 관한 추론이 가능하기 때문이다. 혼합모형에서 분산성분을 추정하기 위한 방법으로 상수적합법이라 불리우는 Henderson (1953) 방법 III(Henderson's Method III)을 적용하기로 한다. 상수적합법은 다양한 모형의 적합방식을 이용하여 변동요인에 따른 제곱합을 구하고 있기 때문에 사영에 근거한 분석이 가능하게 된다. 분산성분을 얻기위한 모형으로 혼합모형에서 고정효과를 제외한 잔차모형을 구한다. 고정효과의 모형은

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}^*. \tag{2.2}$$

로 표현된다. 단, $\boldsymbol{\epsilon}^* = \sum_{i=1}^{k}\boldsymbol{Q_i}\boldsymbol{\delta_i} + \sum_{j=1}^{r}\boldsymbol{W}_j\boldsymbol{\nu}_j + \boldsymbol{\epsilon}$이다. $\mathrm{Var}(\boldsymbol{\epsilon}^*) = \boldsymbol{\Sigma}$라 두면 $\boldsymbol{\Sigma}$는

$$\boldsymbol{\Sigma} = \sum_{i=1}^{k}\sigma_{\delta_i}^2\boldsymbol{Q_i}\boldsymbol{Q_i'} + \sum_{j=1}^{r}\sigma_{\nu_j}^2\boldsymbol{W_j}\boldsymbol{W_j'} + \sigma_\epsilon^2\boldsymbol{I}_n \tag{2.3}$$

이다. 고정효과모형 또는 고정부분의 식 (2.2)에 사영을 이용하여 고정효과벡터 $\boldsymbol{\beta}$를 추정한 후 잔차벡터를 구한다. 잔차벡터를 $\boldsymbol{r}$이라 두자. Moore-Penrose의 일반화된 역행렬을 이용한 정규방정식의 해벡터 $\hat{\boldsymbol{\beta}}$은 $\hat{\boldsymbol{\beta}} = \boldsymbol{X}^-\boldsymbol{y}$로 구해지므로

$$\begin{aligned}\boldsymbol{r} &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{y} \\ &= \sum_{i=1}^{k}(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_i\boldsymbol{\delta}_i + \sum_{j=1}^{r}(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_j\boldsymbol{\nu}_j + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon}\end{aligned} \tag{2.4}$$

이다. 식 (2.4)의 잔차모형은 혼합모형의 고정효과부분인 $\boldsymbol{X}\boldsymbol{\beta}$에 종속되지 않는 확률모형이다. 확률모형의 분산성분을 추정하는 방법으로 상수적합법으로 불리우는 Henderson 방법 III을 적용하기로 한다. Henderson 방법 III은 혼합모형에서 모형의 적합을 이용하여 분산성분을 구하는 방법을 제공하므로 모형의 적합방식으로 부터 유도되는 사영행렬을 이용한 사영분석이 가능하게 된다. 고정효과의 사영분석에 관한 논의는 Choi (2012)에서 보여진다.

3. 확률효과의 분산성분모형

식 (2.4)의 잔차모형은 처리구조내 확률요인에 따른 분산성분을 구하기 위한 확률모형으로 이용된다. 확률효과모형의 분산성분에 대한 논의는 Choi (2011)에서 보여진다. 처리구조내 고정요인과 확률요인들의 주효과와 교호작용에 따른 k개 분산성분에 대한 잔차확률모형식은

$$\begin{aligned}\boldsymbol{r} &= \sum_{i=1}^{k}(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_i\boldsymbol{\delta}_i + \boldsymbol{\epsilon}_\delta \\ &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_1\boldsymbol{\delta}_1 + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_2\boldsymbol{\delta}_2 + \cdots + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_k\boldsymbol{\delta}_k + \boldsymbol{\epsilon}_\delta\end{aligned} \tag{3.1}$$

로 표현된다. 단, $\boldsymbol{\epsilon}_\delta = \sum_{j=1}^{r}(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_j\boldsymbol{\nu}_j + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon}$이다. $k = 1$이고 $r = 1$일 때 식 (3.1)은 처리구조내 단일의 확률성분을 갖는 분할구 실험에서의 확률모형으로 주어진다. 잔차벡터를 $\boldsymbol{r}_1$이라 두면 분할구의 주구(whole plot)와 세구(subplot)의 오차항들을 포함하는 잔차모형식은

$$\begin{aligned}\boldsymbol{r}_1 &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{y} \\ &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_1\boldsymbol{\delta}_1 + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_1\boldsymbol{\nu}_1 + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon}\end{aligned} \tag{3.2}$$

로 표현된다. 식 (3.2)에서 $\boldsymbol{\nu}$는 주구 간의 변이를 나타내는 오차벡터이고 $\boldsymbol{\epsilon}$은 세구 간의 변이를 나타내는 오차벡터이다. 이들 오차벡터들은 처리구조에 포함된 확률요인의 수준효과벡터와는 독립인 것으로 간주한다. 이때 확률효과벡터 $\boldsymbol{\delta}_1$의 분산성분 $\sigma_{\delta_1}^2$을 구하기 위한 모형식은

$$\begin{aligned}\boldsymbol{r}_1 &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_1\boldsymbol{\delta}_1 + \boldsymbol{\epsilon}_{\delta_1} \\ &= \boldsymbol{X}_Q\boldsymbol{\delta_1} + \boldsymbol{\epsilon}_{\delta_1}\end{aligned} \tag{3.3}$$

이다. 단, $\boldsymbol{X}_Q = (I - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_1$이고 $\boldsymbol{\epsilon}_{\delta_1} = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_1\boldsymbol{\nu}_1 + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon}$이다.

4. 분할구의 분산성분모형

분할구의 실험계획으로부터 발생하는 변동요인으로 주구의 변이와 세구의 변이를 예상할 수 있다. 처리 구조에 포함되는 처리요인들과는 별도로 서로 다른 크기의 실험단위를 나타내는 주구와 세구는 변동을 나타내는 요인들로 간주되고 실험단위들 간의 변이는 주구오차와 세구오차로 가정된다. 주구요인의 분산성분을 구하기 위한 모형은 식 (3.3)으로 부터 확률효과벡터 $\boldsymbol{\delta}_1$을 적합시킨 잔차벡터를 $\boldsymbol{r}_\delta$라 두면

$$\boldsymbol{r}_\delta = (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)\boldsymbol{r}_1 \tag{4.1}$$

이다. $\boldsymbol{r}_\delta$를 이용한 주구요인의 확률모형은

$$\begin{aligned}
\boldsymbol{r}_\delta &= (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)\boldsymbol{r}_1 \\
&= (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{y} \\
&= (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Q}_1\boldsymbol{\delta}_1 + \boldsymbol{W}_1\boldsymbol{\nu}_1 + \boldsymbol{\epsilon}) \\
&= (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_1\boldsymbol{\nu}_1 + \boldsymbol{\epsilon}_{\nu_1} \\
&= \boldsymbol{X}_W\boldsymbol{\nu}_1 + \boldsymbol{\epsilon}_{\nu_1}
\end{aligned} \tag{4.2}$$

이다. 단, $\boldsymbol{X}_W = (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{W}_1$이고 $\boldsymbol{\epsilon}_{\nu_1} = (\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon}$이다. 주구의 확률벡터에 의존하지 않는 잔차벡터를 $\boldsymbol{r}_{\nu_1}$이라 둘 때 $\boldsymbol{r}_{\nu_1}$은 세구의 분산성분을 구하기 위한 확률모형을 제공한다. 주구의 확률효과에 종속되지 않는 잔차벡터 $\boldsymbol{r}_{\nu_1}$을 이용한 세구의 분산성분에 대한 확률모형은

$$\begin{aligned}
\boldsymbol{r}_{\nu_1} &= (\boldsymbol{I} - \boldsymbol{X}_W\boldsymbol{X}_W^-)\boldsymbol{r}_\delta \\
&= (\boldsymbol{I} - \boldsymbol{X}_W\boldsymbol{X}_W^-)(\boldsymbol{X}_W\boldsymbol{\nu}_1 + \boldsymbol{\epsilon}_{\nu_1}) \\
&= (\boldsymbol{I} - \boldsymbol{X}_W\boldsymbol{X}_W^-)(\boldsymbol{I} - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{\epsilon} \\
&= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)\boldsymbol{\epsilon} \\
&= \boldsymbol{X}_\epsilon\boldsymbol{\epsilon}
\end{aligned} \tag{4.3}$$

이다. 단, $\boldsymbol{X}_\epsilon = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)$이다.

5. 확률요인의 사영제곱합

사영에 근거한 분산성분을 추정하기 위한 방법으로 적률법을 이용할 때 확률요인의 변동을 나타내는 제곱합은 상수적합법에 의해 계산된다. 적률법을 적용하기 위해 변동요인에 따른 제곱합과 제곱합의 기댓값이 요구된다. 상수적합법은 변동요인에 따른 제곱합의 계산을 위해 비교되는 모형들의 적합에서 구해지는 축소제곱합(reduction in sums of squares)을 이용한다. 반면에 사영제곱합은 분산성분을 얻기 위한 잔차확률모형의 모형행렬을 이용하여 변동요인의 제곱합을 구한다는 점에서 차이가 있다. 확률요인의 분산성분 $\sigma^2_{\delta_1}$을 구하기 위한 모형식 (3.3)으로부터 사영제곱합은 $\boldsymbol{\delta}_1$의 계수행렬 $\boldsymbol{X}_Q$로의 사영에 의해 구해진다. 사영은 $\boldsymbol{X}_Q\boldsymbol{X}_Q^-\boldsymbol{r}_1$이므로 사영제곱합은 $\boldsymbol{r}_1'\boldsymbol{X}_Q\boldsymbol{X}_Q^-\boldsymbol{r}_1$으로 구해진다. 주구오차에 따른 분산성분의 사영제곱합은 식 (4.2)의 확률벡터 $\boldsymbol{\nu}_1$의 계수행렬 $\boldsymbol{X}_W$로의 사영을 이용하여 구한다. $\boldsymbol{X}_W$로의 사영은 $\boldsymbol{X}_W\boldsymbol{X}_W^-\boldsymbol{r}_\delta$이고 사영제곱합은 $\boldsymbol{r}_\delta'\boldsymbol{X}_W\boldsymbol{X}_W^-\boldsymbol{r}_\delta$로 구해진다. 식 (4.3)을 이용하여 오차벡터의 분산성분 σ^2_ϵ의 사영제곱합을 구할 수 있다. $\boldsymbol{\epsilon}$에 따른 사영제곱합은 $\boldsymbol{r}_{\nu_1}'\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-\boldsymbol{r}_{\nu_1}$으로 구해진다. 확률요인들의 변동량을 사영제곱합으로 구한 후 이들 제곱합의 기댓값을 구해 적률법으로 분산성분들을 추

정하게 된다. 확률요인의 제곱합의 기댓값은 분산성분들의 선형함수로 주어지고 분산성분들의 계수는 Hartley (1967)의 합성법을 이용하여 구한다. 적률법에 의해 분산성분을 추정하기 위한 방정식들은

$$\begin{aligned}\boldsymbol{r}_1'\boldsymbol{X}_Q\boldsymbol{X}_Q^-\boldsymbol{r}_1 &= k_{1\delta}\sigma_\delta^2 + k_{2\delta}\sigma_\nu^2 + k_{3\delta}\sigma_\epsilon^2\\ \boldsymbol{r}_\delta'\boldsymbol{X}_W\boldsymbol{X}_W^-\boldsymbol{r}_\delta &= k_{1\nu}\sigma_\delta^2 + k_{2\nu}\sigma_\nu^2 + k_{3\nu}\sigma_\epsilon^2\\ \boldsymbol{r}_{\nu_1}'\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-\boldsymbol{r}_{\nu_1} &= k_{1\epsilon}\sigma_\delta^2 + k_{2\epsilon}\sigma_\nu^2 + k_{3\epsilon}\sigma_\epsilon^2\end{aligned} \tag{5.1}$$

로 주어진다. $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Q\delta} + \boldsymbol{W\nu} + \boldsymbol{\epsilon}$으로 표현될 때 $\boldsymbol{y}$의 $\text{Var}(\boldsymbol{y})$을 $\boldsymbol{\Sigma}$라 두자.

$$\boldsymbol{\Sigma} = \sigma_\delta^2\boldsymbol{QQ}' + \sigma_\nu^2\boldsymbol{WW}' + \boldsymbol{\sigma}_\epsilon^2 \tag{5.2}$$

이다. 식 (5.1)의 확률벡터 $\boldsymbol{\delta}$에 따른 제곱합 $\boldsymbol{r}_1'\boldsymbol{X}_Q\boldsymbol{X}_Q^-\boldsymbol{r}_1$의 기댓값을 E_Q라 두면

$$\begin{aligned}E_Q &= E(\boldsymbol{r}_1'\boldsymbol{X}_Q\boldsymbol{X}_Q^-\boldsymbol{r}_1)\\ &= \text{tr}\left[(\boldsymbol{I}-\boldsymbol{XX}^-)\boldsymbol{X}_Q\boldsymbol{X}_Q^-(\boldsymbol{I}-\boldsymbol{XX}^-)\left(\sigma_\delta^2\boldsymbol{QQ}' + \sigma_\nu^2\boldsymbol{WW}' + \sigma_\epsilon^2\boldsymbol{I}\right)\right]\\ &= \sigma_\delta^2\text{tr}(\boldsymbol{Q}'\boldsymbol{Q}_P\boldsymbol{Q}) + \sigma_\nu^2\text{tr}(\boldsymbol{W}'\boldsymbol{Q}_P\boldsymbol{W}) + \sigma_\epsilon^2\text{tr}(\boldsymbol{Q}_P)\\ &= k_{1\delta}\sigma_\delta^2 + k_{2\delta}\sigma_\nu^2 + k_{3\delta}\sigma_\epsilon^2\end{aligned} \tag{5.3}$$

단, $\boldsymbol{Q}_P = (\boldsymbol{I}-\boldsymbol{XX}^-)\boldsymbol{X}_Q\boldsymbol{X}_Q^-(\boldsymbol{I}-\boldsymbol{XX}^-)$이다. $k_{1\delta} = \text{tr}(\boldsymbol{Q}'\boldsymbol{Q}_P\boldsymbol{Q})$, $k_{2\delta} = \text{tr}(\boldsymbol{W}'\boldsymbol{Q}_P\boldsymbol{W})$이고 $k_{3\delta} = \text{tr}(\boldsymbol{Q}_P)$를 나타낸다. 식 (5.1)의 확률벡터 $\boldsymbol{\nu}$에 따른 제곱합 $\boldsymbol{r}_\delta'\boldsymbol{X}_W\boldsymbol{X}_W^-\boldsymbol{r}_\delta$의 기댓값을 E_W라 두면

$$\begin{aligned}E_W &= E(\boldsymbol{r}_\delta'\boldsymbol{X}_W\boldsymbol{X}_W^-\boldsymbol{r}_\delta)\\ &= \text{tr}\left[(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)\boldsymbol{X}_W\boldsymbol{X}_W^-(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)(\sigma_\delta^2\boldsymbol{QQ}' + \sigma_\nu^2\boldsymbol{WW}' + \sigma_\epsilon^2\boldsymbol{I})\right]\\ &= \sigma_\delta^2\text{tr}(\boldsymbol{Q}'\boldsymbol{W}_P\boldsymbol{Q}) + \sigma_\nu^2\text{tr}(\boldsymbol{W}'\boldsymbol{W}_P\boldsymbol{W}) + \sigma_\epsilon^2\text{tr}(\boldsymbol{W}_P)\\ &= k_{1\nu}\sigma_\delta^2 + k_{2\nu}\sigma_\nu^2 + k_{3\nu}\sigma_\epsilon^2\end{aligned} \tag{5.4}$$

단, $\boldsymbol{W}_P = (\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)\boldsymbol{X}_Q\boldsymbol{X}_Q^-(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^-)$이다. $k_{1\nu} = \text{tr}(\boldsymbol{Q}'\boldsymbol{W}_P\boldsymbol{Q})$, $k_{2\nu} = \text{tr}(\boldsymbol{W}'\boldsymbol{W}_P\boldsymbol{W})$이고 $k_{3\nu} = \text{tr}(\boldsymbol{W}_P)$를 나타낸다. 식 (5.1)의 확률벡터 $\boldsymbol{\epsilon}$에 따른 제곱합 $\boldsymbol{r}_{\nu_1}'\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-\boldsymbol{r}_{\nu_1}$의 기댓값을 E_R라 두면

$$\begin{aligned}E_R &= E(\boldsymbol{r}_{\nu_1}'\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-\boldsymbol{r}_{\nu_1})\\ &= \text{tr}\left[(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)\right.\\ &\quad \left.\times\left(\sigma_\delta^2\boldsymbol{QQ}' + \sigma_\nu^2\boldsymbol{WW}' + \sigma_\epsilon^2\boldsymbol{I}\right)\right]\\ &= \sigma_\delta^2\text{tr}(\boldsymbol{Q}'\boldsymbol{R}_P\boldsymbol{Q}) + \sigma_\nu^2\text{tr}(\boldsymbol{W}'\boldsymbol{R}_P\boldsymbol{W}) + \sigma_\epsilon^2\text{tr}(\boldsymbol{R}_P)\\ &= k_{1\epsilon}\sigma_\delta^2 + k_{2\epsilon}\sigma_\nu^2 + k_{3\epsilon}\sigma_\epsilon^2\end{aligned} \tag{5.5}$$

단, $\boldsymbol{R}_P = (\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)\boldsymbol{X}_Q\boldsymbol{X}_Q^-(\boldsymbol{I}-\boldsymbol{XX}^- - \boldsymbol{X}_Q\boldsymbol{X}_Q^- - \boldsymbol{X}_W\boldsymbol{X}_W^-)$이다. $k_{1\epsilon} = \text{tr}(\boldsymbol{Q}'\boldsymbol{R}_P\boldsymbol{Q})$, $k_{2\epsilon} = \text{tr}(\boldsymbol{W}'\boldsymbol{R}_P\boldsymbol{W})$이고 $k_{3\epsilon} = \text{tr}(\boldsymbol{R}_P)$를 나타낸다.

6. 고정효과의 추정가능함수

식 (2.2)의 모수벡터 $\boldsymbol{\beta}$의 한 함수 $\boldsymbol{a}'\boldsymbol{\beta}$가 추정가능함수이면 $\boldsymbol{\beta}$의 추정량 $\hat{\boldsymbol{\beta}}$을 이용하여 $\boldsymbol{a}'\boldsymbol{\beta}$에 대한 추론을 행할 수 있다. 단, $\boldsymbol{a}$는 $p\times 1$인 상수 벡터이다. $\boldsymbol{a}'\boldsymbol{\beta}$의 추정가능성은

$$\boldsymbol{a} = \boldsymbol{X}'\boldsymbol{X}\boldsymbol{r} \tag{6.1}$$

Table 7.1. Data for tensile strength of paper

Block	Method	Temperature			
		200	225	250	275
1	1	30	35	37	36
1	2	34	41	38	42
1	3	29	26	33	36
2	1	28	32	40	41
2	2	31	36	42	40
2	3	31	30	32	40
3	1	31	37	41	40
3	2	35	40	39	44
3	3	32	34	39	45

로 부터 벡터 $\boldsymbol{r}$의 존재로 확인하거나

$$\boldsymbol{a} = \boldsymbol{X}^{-}\boldsymbol{X}\boldsymbol{a} \tag{6.2}$$

를 확인함으로서 알 수 있다. $\boldsymbol{X}^{-}\boldsymbol{X}$는 모형행렬 $\boldsymbol{X}$의 행벡터들로 구성되는 벡터공간으로의 사영헬렬을 나타내므로 행공간으로의 사영행렬을 이용하여 추정가능성을 확인할 수 있다는 것을 의미하고 있다. $\boldsymbol{\beta}$의 추정량 $\hat{\boldsymbol{\beta}}$으로

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}\right)^{-}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y} \tag{6.3}$$

를 이용하면 $\boldsymbol{a}'\hat{\boldsymbol{\beta}}$의 분포는 평균과 분산이

$$\begin{aligned} E(\boldsymbol{a}'\hat{\boldsymbol{\beta}}) &= \boldsymbol{a}'\boldsymbol{\beta} \\ \mathrm{Var}(\boldsymbol{a}'\hat{\boldsymbol{\beta}}) &= \boldsymbol{a}'\left(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}\right)^{-}\boldsymbol{a} \end{aligned} \tag{6.4}$$

인 정규분포를 근사적으로 따른다. 표본이 충분히 큰 경우에

$$Z = \frac{\boldsymbol{a}'\hat{\boldsymbol{\beta}} - \boldsymbol{a}'\boldsymbol{\beta}}{\sqrt{\boldsymbol{a}'\left(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}\right)^{-}\boldsymbol{a}}} \tag{6.5}$$

는 $N(0,1)$인 분포를 근사적으로 따르므로 추정가능함수 $\boldsymbol{a}'\boldsymbol{\beta}$의 추론에 이용된다.

7. 자료분석의 예

Table 7.1은 Montgomery (1976)의 분할구실험에 대한 자료 예이다. 제지업자가 제지의 강도에 영향을 미치는 두 요인들의 효과를 조사하기 위해 세 수준의 펄프준비방법과 네 수준의 공정온도에서 실험한 자료이다. 소규모시험설비(pilot plant)는 두 요인들의 수준결합에서 하루에 1회 가능하므로 삼일 간에 걸쳐 3회 반복된 실험자료이다. 반복은 블록으로 간주되고 방법과 온도는 고정요인이나 세 수준의 방법은 큰 실험단위인 주구의 펄프배취에 임의로 배정되고 작은 실험단위인 세구에 각 방법에 의해 생산된 펄프배취를 4개의 표본으로 나누어 임의로 실험온도에 배정하여 생산된 제지의 인장강도(tensile strength)자료를 얻게 된다. 블록 i와 방법 j 그리고 온도 k에서 실험단위의 반응을 y_{ijk}라 두면

$$y_{ijk} = \mu + \delta_i + \beta_j + \nu_{ij} + \gamma_k + (\delta\gamma)_{ik} + (\beta\gamma)_{jk} + \epsilon_{ijk} \tag{7.1}$$

로 표현된다. 단, μ는 평균이고 δ_i는 블록 i의 수준효과이며 β_j는 방법 j의 수준효과 를 나타낸다. 여기서 $i, j = 1, 2, 3$이다. γ_k는 온도 k의 수준효과이고 $(\delta\gamma)_{ik}$와 $(\beta\gamma)_{jk}$는 교호작용을 나타낸다. ν_{ij}는 주구오차이고 ϵ_{ijk}는 세구오차를 나타낸다. 단, $k = 200, 225, 250, 275$이다. 블록효과 δ_i는 $N(0, \sigma_b^2)$, ν_{ij}는 $N(0, \sigma_w^2)$, $(\delta\gamma)_{ik}$는 $N(0, \sigma_{bt}^2)$이고 ϵ_{ijk}는 $N(0, \sigma_\epsilon^2)$을 가정하며 또한 이들 확률효과는 서로 독립이라고 가정하게 된다. 4개의 분산성분 중 두개는 분할구실험에서 실험단위의 크기와 관련된 오차들로 주구오차와 세구오차이다. 분산성분들을 사영에 의해 구하기 위한 첫 단계로 식 (2.1)과 같은 행렬표현의 모형식을 생각한다. 행렬표현식은

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{j}\mu + \boldsymbol{Q}_b\boldsymbol{\delta} + \boldsymbol{X}_m\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{\nu} \\ &\quad + \boldsymbol{X}_t\boldsymbol{\gamma} + \boldsymbol{Q}_{bt}(\boldsymbol{\delta\gamma}) + \boldsymbol{X}_{mt}(\boldsymbol{\beta\gamma}) + \boldsymbol{\epsilon} \end{aligned} \tag{7.2}$$

이다. 식 (7.2)를 고정부분과 확률부분으로 분리하여 표현하면

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{j}\mu + \boldsymbol{X}_m\boldsymbol{\beta} + \boldsymbol{X}_t\boldsymbol{\gamma} + \boldsymbol{X}_{mt}(\boldsymbol{\beta\gamma}) \\ &\quad + \boldsymbol{Q}_b\boldsymbol{\delta} + \boldsymbol{W}\boldsymbol{\nu} + \boldsymbol{Q}_{bt}(\boldsymbol{\delta\gamma}) + \boldsymbol{\epsilon} \end{aligned} \tag{7.3}$$

이다. 식 (7.3)의 고정부분은 $\boldsymbol{j}\mu+\boldsymbol{X}_m\boldsymbol{\beta}+\boldsymbol{X}_t\boldsymbol{\gamma}+\boldsymbol{X}_{mt}(\boldsymbol{\beta\gamma})$이고 확률부분은 $\boldsymbol{Q}_b\boldsymbol{\delta}+\boldsymbol{W}\boldsymbol{\nu}+\boldsymbol{Q}_{bt}(\boldsymbol{\delta\gamma})+\boldsymbol{\epsilon}$이다. 분산성분을 추정하기 위한 모형은 식 (7.3)의 고정효과부분을 적합시켜 구해지는 잔차에 대한 확률모형이다. 잔차모형식은

$$\begin{aligned} \boldsymbol{r} &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{y} \\ &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)(\boldsymbol{Q}_b\boldsymbol{\delta} + \boldsymbol{W}\boldsymbol{\nu} + \boldsymbol{Q}_{bt}(\boldsymbol{\delta\gamma}) + \boldsymbol{\epsilon}) \end{aligned} \tag{7.4}$$

이다. 단, $\boldsymbol{X} = (\boldsymbol{j}, \boldsymbol{X}_m, \boldsymbol{X}_t, \boldsymbol{X}_{mt})$이다. 식 (7.4)로 부터 확률효과벡터 $\boldsymbol{\delta}$의 분산성분을 추정하기 위한 모형은

$$\begin{aligned} \boldsymbol{r} &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_b\boldsymbol{\delta} + \boldsymbol{\epsilon}_\delta \\ &= \boldsymbol{X}_b\boldsymbol{\delta} + \boldsymbol{\epsilon}_\delta \end{aligned} \tag{7.5}$$

이다. 단, $\boldsymbol{X}_b = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^-)\boldsymbol{Q}_b$이다. $\boldsymbol{\delta}$에 따른 변동량의 계산은 식 (7.5)의 모형행렬 $\boldsymbol{X}_b$로의 사영을 이용하여 구한다. 사영은 $\boldsymbol{X}_b\boldsymbol{X}_b^-\boldsymbol{r}$이다. $\boldsymbol{\delta}$에 따른 변동량을 SS_b라 두면 $\boldsymbol{r}'\boldsymbol{X}_b\boldsymbol{X}_b^-\boldsymbol{r}$로 구해지고 $SS_b = 77.55556$이다. 주구오차벡터 $\boldsymbol{\nu}$의 분산성분을 추정하기 위한 모형은 식 (7.5)의 적합에서 구해진 잔차벡터를 $\boldsymbol{r}_\delta$라 두면

$$\begin{aligned} \boldsymbol{r}_\delta &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_b\boldsymbol{X}_b^-)\boldsymbol{W}\boldsymbol{\nu} + \boldsymbol{\epsilon}_\delta \\ &= \boldsymbol{X}_w\boldsymbol{\nu} + \boldsymbol{\epsilon}_\delta \end{aligned} \tag{7.6}$$

이다. 단, $\boldsymbol{X}_w = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_b\boldsymbol{X}_b^-)\boldsymbol{W}$이다. $\boldsymbol{\delta}$에 따른 변동량은 식 (7.6)의 모형행렬 $\boldsymbol{X}_w$로의 사영을 이용하여 구한다. $\boldsymbol{X}_w$로의 사영은 $\boldsymbol{X}_w\boldsymbol{X}_w^-\boldsymbol{r}_\delta$이다. $\boldsymbol{\delta}$에 따른 변동량을 SS_w라 두면 $\boldsymbol{r}_\delta'\boldsymbol{X}_w\boldsymbol{X}_w^-\boldsymbol{r}_\delta$로 구해지고 $SS_w = 36.27778$이다. 식 (7.6)의 잔차벡터를 $\boldsymbol{r}_\nu$라 둘 때 교호작용의 분산성분 $\sigma_{\delta\gamma}^2$과 관련된 제곱합 SS_{bt}를 구하기 위한 잔차모형은

$$\begin{aligned} \boldsymbol{r}_\nu &= (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_b\boldsymbol{X}_b^- - \boldsymbol{X}_w\boldsymbol{X}_w^-)\boldsymbol{Q}_{bt}(\boldsymbol{\delta\gamma}) + \boldsymbol{\epsilon}_\nu \\ &= \boldsymbol{X}_{bt}(\boldsymbol{\delta\gamma}) + \boldsymbol{\epsilon}_\nu \end{aligned} \tag{7.7}$$

이다. 단, $\boldsymbol{X}_{bt} = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^- - \boldsymbol{X}_b\boldsymbol{X}_b^- - \boldsymbol{X}_w\boldsymbol{X}_w^-)\boldsymbol{Q}_{bt}$이다. $\boldsymbol{X}_{bt}$로의 사영은 $\boldsymbol{X}_{bt}\boldsymbol{X}_{bt}^-\boldsymbol{r}_\nu$이므로 변동량 SS_{bt}는 사영제곱합 $\boldsymbol{r}_\nu'\boldsymbol{X}_{bt}\boldsymbol{X}_{bt}^-\boldsymbol{r}_\nu$에 의해 20.66667로 구해진다. 세구오차에 따른 제곱합을 SS_e로 나타낼 때 $SS_e = 50.83333$로 구해짐을 알 수 있다. 네 개의 분산성분들과 관련된 제곱합이 각 벡터부분공간으로의 사영에 의하여 구해졌기 때문에 적률법에 의한 분산성분을 추정하기 위해 Hartley (1967)의 합성법으로 분산성분의 계수를 추정한다. 식 (7.2)의 분산성분을 구하기 위한 방정식들은

$$\begin{aligned} SS_b &= 24\sigma_\delta^2 + 8\sigma_\nu^2 + 6\sigma_{(\delta\gamma)}^2 + 2\sigma_\epsilon^2, \\ SS_w &= 16\sigma_\nu^2 + 4\sigma_\epsilon^2, \\ SS_{bt} &= 18\sigma_{(\delta\gamma)}^2 + 6\sigma_\epsilon^2, \\ SS_e &= 12\sigma_\epsilon^2 \end{aligned} \tag{7.8}$$

로 주어지고 분산성분의 추정값들은

$$\hat\sigma_\delta^2 = 2.5416667, \quad \hat\sigma_\nu^2 = 1.2083333, \quad \hat\sigma_{(\delta\gamma)}^2 = -0.2638889, \quad \hat\sigma_\epsilon^2 = 4.2361111 \tag{7.9}$$

로 구해진다. $\hat\sigma_{(\delta\gamma)}^2$이 음수이므로 분산성분은 0으로 추정된다. 식 (2.3)으로부터 $\hat\Sigma$은

$$\hat\Sigma = \hat\sigma_\delta^2\boldsymbol{Q}_b\boldsymbol{Q}_b' + \hat\sigma_\nu^2\boldsymbol{W}\boldsymbol{W}' + \hat\sigma_{(\delta\gamma)}^2\boldsymbol{Q}_{bt}\boldsymbol{Q}_{bt}' + \hat\sigma_\epsilon^2\boldsymbol{I} \tag{7.10}$$

이다. 식 (6.2)의 조건을 만족시키는 $\boldsymbol{a}$로

$$\begin{aligned} \boldsymbol{a} = (&0.4, -0.4, 0.4, 0.4, 0.1, 0.1, 0.1, 0.1, -0.1, -0.1, \\ &-0.1, -0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1) \end{aligned} \tag{7.11}$$

라 두면 $\boldsymbol{a}'\hat{\boldsymbol{\beta}} = 14.7$로 구해진다. $\widehat{\mathrm{Var}}(\boldsymbol{a}'\hat{\boldsymbol{\beta}}) = 0.1694444$로 추정된다.

8. 결론

본 논문은 분할구의 실험자료 분석에서 실험단위의 크기에 따른 추가적인 오차항 외에 처리들의 효과에 확률효과가 포함될 때 혼합모형의 가정하에서 사영분석하는 방법을 논의하였다. 사영분석의 결과는 기존의 분산분석에 따른 결과와 동일함을 보였다. 다시말하면, 사영분석 또한 자료의 분산분석에 이용될 수 있는 한 방법임을 입증하고 있다. 사영을 이용할 때, 변동요인에 따른 변동량의 계산을 위한 모형구축과 모형적합을 구체적으로 논의하였다. 잔차벡터에 대한 확률모형의 모형적합을 통해 분산성분의 추정을 위한 변동요인의 변동량을 사영으로 구할 수 있음을 나타내었다. 즉 변동량은 변동요인과 관련된 벡터부분공간으로의 사영으로 구할 수 있음을 다루었다. 또한 변동요인의 제곱합에 대한 기댓값 계산에 Hartley의 합성법을 논의하였다. 그리고 고정효과의 추정가능함수에 대한 추정방법으로 가중최소제곱법을 이용하였고 추정량의 분산에 대한 추정치의 계산과정도 구체적으로 다루었다.

References

Choi, J. S. (2011). Variance components in one-factor random model by projections, *Journal of the Korean Data & Information Science Society*, **22**, 381−387.

Choi, J. S. (2012). Type II analysis by projections, *Journal of the Korean Data & Information Science Society*, **23**, 155−1163.

Graybill, F. A. (1983). *Matrices with Applications in Statistics*, Wadsworth Publishing Company, Belmont.

Hartley, H. O. (1967). Expectations, variances and covariances of ANOVA mean squares by "synthesis", *Biometrics*, **23**, 105–114.

Henderson, C. R. (1953). Estimation of variance and covariance components, *Biometrics*, **9**, 226–252.

Johnson, R. A. and Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis* (2nd ed), Prentice Hall, Inc., Englewood Cliffs.

Milliken, G. A. and Johnson, D. E. (2009). *Analysis of Messy Data*, Van Nostrand Reinhold, New York.

Montgomery, C. D. (1976). *Design and Analysis of Experiments*, John Wiley and Sons, New York.

Searle, S. R. (1971). *Linear Models*, John Wiley and Sons, New York.

Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance Components*, John Wiley and Sons, New York.

분할구자료의 사영분석

최재성[a,1]

[a]계명대학교 통계학과

(2016년 12월 5일 접수, 2017년 4월 1일 수정, 2017년 4월 1일 채택)

요 약

본 논문은 분할구실험으로 부터 주어진 자료분석을 위해 사영을 이용하는 방법을 다루고 있다. 분할구 실험의 특성으로 서로 다른 크기의 실험단위를 나타내는 오차항과 처리에 포함된 확률효과가 존재할 때 이들 분산성분의 추정에 사영을 이용하여 구하는 방법을 제시하고 있다. 분산성분 추정을 위해 잔차벡터에 대한 확률모형의 구축을 다루고 있다. 고정효과를 제외한 확률효과에 따른 제곱합의 계산을 위해 상수적합법이 적용되고 있다. 적률법에 의한 분산성분 추정을 위해 변동량의 기댓값 계산에 합성법을 이용한다. 고정효과들의 선형함수로 주어지는 추정가능함수에 관한 추정을 다루고 있다.

주요용어: 분할구실험, 분산성분, 사영, 상수적합법, 추정가능함수.

[1](42601) 대구광역시 달서구 달구벌대로 1095, 계명대학교 통계학과. E-mail: jschoi@kmu.ac.kr