The Korean Journal of Applied Statistics (2015) DOI: http://dx.doi.org/10.5351/KJAS.2015.28.6.1093
28(6), 1093–1101

Variance Components of Nested Designs

Jaesung Choi[a,1]

[a]Department of Statistics, Keimyung University

(Received August 5, 2015; Revised October 24, 2015; Accepted December 4, 2015)

Abstract

This paper discusses nested design models when nesting occurs in treatment structure and design structure. Some are fixed and others are random; subsequently, the fixed factors having a nested design structure are assumed to be nested in the random factors. The treatment structure can involve random and fixed effects as well as a design structure that can involve several sizes of experimental units. This shows how to use projections for sums of squares by fitting the model in a stepwise procedure. Expectations of sums of squares are obtained via synthesis. Variance components of the nested design model are estimated by the method of moments.

Keywords: variance components, nested structure, projection, type I sums of squares, synthesis

1. 서론

지분계획하의 실험자료를 분석하기 위한 모형은 처리요인들이 지분관계(nesting)에 있거나 서로 다른 크기의 실험단위들이 지분관계인 가에 따라 모형의 형태가 달라진다. 지분관계의 처리요인들이 단일 크기의 실험단위에 배정될 때 오차항은 하나이나 지분관계의 요인들이 서로 다른 크기의 실험단위를 취할 때 둘 이상의 오차항들이 존재하게 된다. 지분계획(nested design)에서의 지분관계는 실험의 처리구조나 설계구조 또는 둘다에서 발생할 수 있다.

요인들이 지분관계에 있을 때 분석모형은 요인들의 유형에 따라 고정효과, 확률효과 또는 혼합효과모형이 이용될 수 있다. 지분계획과 관련된 논의는 John (1971), Hicks (1973), Milliken과 Johnson (1984), Montgomery (1976) 그리고 Searle 등 (1992) 등에서 다루어지고 있다. Choi (2011, 2012, 2014)는 선형모형에 근거하여 사영(projection)에 의한 분석방법을 다루고 있다. 사영과 행렬에 관한 이론적 배경은 Searle (1971)과 Graybill (1976) 등에서 구체적 논의를 살펴볼 수 있다.

지분관계가 요인들의 처리구조에서 그리고 실험의 설계구조에서 모두 발생하게 될 때 자료의 분석을 위한 모형 및 방법은 단순하지 않게 된다. 실험단위의 반응에 영향을 미치는 요인들이 지분관계를 가질 때 지분관계의 요인들이 서로 다른 크기의 실험단위를 갖는 가에 따라 오차항은 하나 또는 둘 이상이 될 수 있다. 지분관계의 요인들이 모두 고정효과이고 단일 크기의 실험단위가 이용될 때 모수에 대한 추론은 모형내 모수들의 추정가능함수들에 대해서만 추론가능하게 된다. 지분관계의 요인들이 서로 다른 크기의 실험단위를 갖게 되면 추가적인 오차항이 모형에 포함된다.

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

요인들의 지분관계도 지분구조의 요인들이 모두 고정요인인 경우와 그렇지 않은 경우로 구분할 수 있다. 지분구조의 요인들에 확률요인이 있게 되면 지분계획모형으로 확률성분을 포함하는 혼합효과의 모형을 가정하게 된다. 따라서 확률요인이 포함된 지분계획하의 자료분석 모형에는 고정요인들의 고정효과 외에 분산성분을 나타내는 확률효과와 하나의 오차항이 포함되나 서로 다른 크기의 실험단위가 필요한 지분계획의 경우에 둘 이상의 오차항이 존재하게 된다.

본 논문은 실험단위의 반응에 영향을 미치는 요인들이 지분관계에 있고 지분관계에 있는 요인들의 수준에 서로 다른 크기의 실험단위가 요구될 때 사영에 의한 분석방법을 논의하고자 한다. 지분구조의 요인들은 고정요인들로 구성되거나 고정요인과 확률요인들로 구성될 수 있다. 고정요인들로 구성된 실험의 지분계획과 관련된 다양한 분석방법들이 문헌상에서 제공되고 있으나 분석에 있어서 사영과의 연관성에 관한 논의는 많지 않다. Choi (2011, 2012, 2014)는 실험자료의 분산분석에 필요한 변동요인별 제곱합의 계산에 사영을 이용할 때 사영이 다양한 분석모형하에서 어떻게 활용될 수 있는 가에 관해 구체적으로 다루고 있다.

2. 지분계획모형

지분계획에서의 분석모형은 처리를 나타내는 요인들의 수와 요인들 간의 관계 그리고 서로 다른 크기의 실험단위의 수에 따라 다르게 가정된다. 반응에 영향을 주는 요인들로 $A, B(A), C(A)$의 세 요인을 가정한다. 요인 A는 a개 수준을 갖는 확률요인이고 요인 $B(A)$는 요인 A의 내재요인으로 요인 A의 수준 i내 b개 수준을 갖는 고정요인으로 가정한다. 요인 $C(A)$는 요인 A의 수준 i내 c개 수준을 갖는 고정요인이고 요인 $B(A)$의 분할된 실험단위에 배정되는 세구요인(split-plot factor)이라 가정한다. 요인 A의 실험단위는 가장 큰 크기의 실험단위이고 요인 $B(A)$의 실험단위는 요인 A의 실험단위에 내재되어 있으나 요인 $C(A)$의 실험단위는 요인 $B(A)$의 분할된 실험단위로 가정한다. 실험단위의 반응을 y라 두면

$$y_{ijkm} = \mu + \alpha_i + \epsilon_{j(i)} + \beta_{k(i)} + \epsilon_{k(ij)} + \tau_{m(ij)} + (\beta\tau)_{km(ij)} + \epsilon_{m(ijk)} \tag{2.1}$$

이다. 여기서 μ는 전체 평균이고 α_i는 A의 수준 i의 확률효과이고 $i = 1, 2, \ldots, a$이다. α_i에 대한 확률분포로 $N(0, \sigma_A^2)$를 가정하며 a개의 확률효과들은 독립으로 가정한다. $\epsilon_{j(i)}$는 확률요인 A의 수준 i가 행해진 실험단위의 오차항으로 $N(0, \sigma_1^2)$인 분포를 따른다고 가정한다. r개의 오차 $j = 1, 2, \ldots, r$는 상호독립으로 간주된다. $\beta_{k(i)}$는 요인 B의 수준 k가 요인 A의 수준 i에서 지분효과(nested effect)를 나타낸다. 요인 B의 수준들이 요인 A의 수준에 내재되어 있으므로 요인 A와 요인 B 간의 교호작용은 성립하지 않는다. 요인 B의 b개 수준 $k = 1, 2, \ldots, b$는 고정효과들로 간주된다. $\epsilon_{k(ij)}$은 요인 A의 수준 i가 행해진 실험단위 $j(i)$에 내재된 실험단위에서 요인 B의 수준 k에서의 오차항을 나타낸다. 두 요인의 수준은 지분관계이고 개별수준의 처리가 행해진 실험단위도 지분관계임을 가정하고 있다. b개의 오차 $k = 1, 2, \ldots, b$는 서로 독립이고 $N(0, \sigma_2^2)$인 분포를 따른다고 가정한다. $\tau_{m(ij)}$는 요인 C의 수준 m이 요인 A의 수준 i와 요인 B의 수준 k에 내재되어 있을 때의 지분효과를 나타낸다. 요인 C의 c개 수준 $m = 1, 2, \ldots, c$는 고정효과이다. $(\beta\tau)_{km(ij)}$는 두 요인 B와 C의 교호작용을 나타내는 고정효과이다. $\epsilon_{m(ijk)}$는 오차항으로 $N(0, \sigma_\epsilon^2)$인 분포를 따르며 오차들은 서로 독립이라고 가정한다. $\epsilon_{j(i)}$, $\epsilon_{k(ij)}$ 그리고 $\epsilon_{m(ijk)}$는 모두 독립인 확률변수들이다. 모형을 행렬표현식으로 나타내면

$$\boldsymbol{y} = \boldsymbol{j}\mu + \boldsymbol{X}_A\boldsymbol{\alpha} + \boldsymbol{X}_1\boldsymbol{\epsilon}_1 + \boldsymbol{X}_{B(A)}\boldsymbol{\beta} + \boldsymbol{X}_2\boldsymbol{\epsilon}_2 + \boldsymbol{X}_{C(A)}\boldsymbol{\tau} + \boldsymbol{X}_{BC(A)}(\boldsymbol{\beta\tau}) + \boldsymbol{\epsilon} \tag{2.2}$$

이다. 단, $\boldsymbol{y}$는 크기 $n \times 1$인 관측벡터이고 모평균 μ의 계수벡터 $\boldsymbol{j}$는 n개의 원소가 모두 1인 열벡터이다. $n = a \times r \times b \times c$이다. $\boldsymbol{X}_A$는 크기가 $n \times a$인 0과 1로 구성되는 계수행렬이다. $\boldsymbol{\alpha}$는 확률효과벡터이고 크기가 $a \times 1$인 열벡터이다. $\boldsymbol{X}_1$은 크기가 $n \times ar$이고 요인 A의 실험단위와 관련된 계수행렬이

다. $\boldsymbol{\epsilon}_1$은 크기가 $ar \times 1$이며 요인 A의 실험단위와 관련된 오차벡터이다. $\boldsymbol{X}_{B(A)}$는 $n \times b$인 계수행렬이다. $\boldsymbol{\beta}$는 고정효과벡터이고 크기가 $b \times 1$인 열벡터이다. $\boldsymbol{X}_2$는 요인 A의 지분요인 $B(A)$의 실험오차벡터 $\boldsymbol{\epsilon}_2$와 관련된 계수행렬이며 크기는 $n \times arb$이다. $\boldsymbol{\epsilon}_2$는 지분요인 $B(A)$의 수준들이 행해진 실험단위에서 측정되는 오차벡터이고 크기는 $arb \times 1$이다. $\boldsymbol{X}_C$는 요인 $C(A)$의 고정효과벡터 $\boldsymbol{\tau}$와 관련된 계수행렬로 크기는 $n \times c$이다. $\boldsymbol{X}_{BC}$는 요인 A내 지분의 두 요인 간의 교호작용과 관련된 계수행렬로 크기 $n \times bc$이고 $\boldsymbol{\beta\tau}$는 요인 $B(A)$와 $C(A)$의 교호작용을 나타내는 크기 $bc \times 1$인 고정효과벡터이다. $\boldsymbol{\epsilon}$은 $arbc \times 1$인 오차벡터이다. 요인 A의 확률효과벡터와 실험단위의 크기에 따른 오차벡터들의 분포로 다변량 정규분포를 가정한다. 즉, $\boldsymbol{\alpha}$는 $N(\mathbf{0}, \sigma_A^2 \boldsymbol{I}_a)$, $\boldsymbol{\epsilon}_1$은 $N(\mathbf{0}, \sigma_1^2 \boldsymbol{I}_{ar})$, $\boldsymbol{\epsilon}_2$는 $N(\mathbf{0}, \sigma_2^2 \boldsymbol{I}_{arb})$ 그리고 $\boldsymbol{\epsilon}$의 분포로 $N(\mathbf{0}, \sigma_\epsilon^2 \boldsymbol{I}_n)$을 가정한다. 식 (2.2)의 행렬모형식에 포함된 모수의 총수는 $bc+4$개이고 고정효과를 나타내는 모수들과 분산성분들로 구분된다. 고정효과의 모수들은 모평균 μ, $b-1$개의 요인 $B(A)$의 효과, $c-1$개의 요인 $C(A)$의 효과 그리고 $(b-1)(c-1)$개의 교호작용 $BC(A)$ 효과들로 구성된다. 4개의 분산성분들은 σ_A^2, σ_1^2, σ_2^2 그리고 σ_ϵ^2이다.

3. 분산성분의 추정

지분계획에 따른 모형식 (2.1)은 실험단위의 반응에 영향을 미치는 요인들의 고정효과와 확률효과로 주어지며 실험단위의 크기에 따른 다수의 오차항들로 구성되어 있다. 따라서 식 (2.2)는 혼합효과모형의 행렬표현식임을 알 수 있다. 혼합효과모형의 분석은 고정효과가 적합된 잔차의 분석모형으로 시작한다. 잔차의 분석모형은 고정효과가 적합되어 제외되었기 때문에 더이상 고정효과에 종속되지 않는 확률효과와 오차항 만을 포함하는 확률모형이다. 식 (2.2)의 고정효과와 관련된 계수행렬을 $\boldsymbol{X}_f$라 두면 $\boldsymbol{X}_f = (\boldsymbol{j}, \boldsymbol{X}_{B(A)}, \boldsymbol{X}_{C(A)}, \boldsymbol{X}_{BC(A)})$이다. 잔차의 확률모형을 얻기 위한 모형식 (2.2)로부터

$$\boldsymbol{y} = \boldsymbol{X}_f \boldsymbol{\beta}_f + \boldsymbol{\delta} \tag{3.1}$$

를 얻는다. 단, $\boldsymbol{\beta}_f = (\mu, \boldsymbol{\beta}, \boldsymbol{\tau}, (\boldsymbol{\beta\tau}))'$이다. 여기서

$$\boldsymbol{\Sigma} = \mathrm{Var}(\boldsymbol{\delta}) = \sigma_A^2 \boldsymbol{X}_A \boldsymbol{X}_A' + \sigma_1^2 \boldsymbol{X}_1 \boldsymbol{X}_1' + \sigma_2^2 \boldsymbol{X}_2 \boldsymbol{X}_2' + \sigma_\epsilon^2 \boldsymbol{I} \tag{3.2}$$

이다. 최소제곱법에 의한 $\boldsymbol{\beta}_f$의 추정량은 $\hat{\boldsymbol{\beta}}_f = (\boldsymbol{X}_f' \boldsymbol{X}_f)^- \boldsymbol{X}_f' \boldsymbol{y}$이다. 식 (3.1)에서 $\boldsymbol{y}$의 추정벡터를 $\boldsymbol{y}_f$라 두고 잔차벡터를 $\boldsymbol{r}_f$라 두면 $\boldsymbol{r}_f = \boldsymbol{y} - \hat{\boldsymbol{y}}_f$이므로 $\boldsymbol{r}_f = (\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-) \boldsymbol{y}$로 구해진다. 잔차벡터 $\boldsymbol{r}_f$에 대한 확률모형은

$$\boldsymbol{r}_f = \left(\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-\right) \boldsymbol{X}_A \boldsymbol{\alpha} + \left(\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-\right) \boldsymbol{X}_1 \boldsymbol{\epsilon}_1 + \left(\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-\right) \boldsymbol{X}_2 \boldsymbol{\epsilon}_2 + \left(\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-\right) \boldsymbol{\epsilon} \tag{3.3}$$

로 주어진다. 잔차벡터의 확률모형에서 변동요인에 따른 제곱합과 제곱합의 기댓값에서 요구되는 분산성분의 계수를 구하기 위한 방법으로 Hartley (1967)의 합성법(synthesis)을 이용한다. Hartley의 합성법은 다수의 확률효과나 오차항을 포함하는 어떤 모형에서도 제곱합의 기댓값 계산에 이용가능하다. 변동요인에 따른 제곱합의 계산에 상수적합법(fitting constants method) 또는 Henderson의 방법3 (1953)에 의한 제1종 제곱합을 적용해 보기로 한다. 변동요인 A에 따른 제곱합을 계산하기 위해 적합시킬 모형은

$$\boldsymbol{r}_f = \left(\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-\right) \boldsymbol{X}_A \boldsymbol{\alpha} + \boldsymbol{\epsilon}_f \tag{3.4}$$

이다. 식 (3.4)의 적합에서 $\boldsymbol{X}_\alpha = (\boldsymbol{I} - \boldsymbol{X}_f \boldsymbol{X}_f^-) \boldsymbol{X}_A$라 두면 $\hat{\boldsymbol{\alpha}} = \boldsymbol{X}_\alpha \boldsymbol{X}_\alpha^- \boldsymbol{r}_f$로 구해지고 $\boldsymbol{\epsilon}_f$는 식 (3.3)에서 첫 항을 제외한 나머지 세 항에 $\boldsymbol{X}_\alpha$가 곱해진 항들의 결합으로 주어진다. 확률효과벡터 $\boldsymbol{\alpha}$에 따른 제

곱합은 $\boldsymbol{r}_f$를 $\boldsymbol{X}_\alpha$로의 사영에 의해 구해진다. 즉, 제곱합은 $\boldsymbol{r}_f'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{r}_f$로 구해진다. $\boldsymbol{r}_f$에서 $\boldsymbol{X}_\alpha$로의 사영을 제외한 잔차를 $\boldsymbol{r}_A$라 두면 $\boldsymbol{r}_A = (\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)\boldsymbol{r}_f$이다. $\boldsymbol{\epsilon}_1$에 따른 확률효과를 추정하기 위한 모형은

$$\boldsymbol{r}_A = \left(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{X}_1\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_A \tag{3.5}$$

이다. 단, $\boldsymbol{\epsilon}_A = (\boldsymbol{I}-\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)[(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_2\boldsymbol{\epsilon}_2+(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{\epsilon}]$이다. 식 (3.5)의 모형행렬을 $\boldsymbol{X}_{\boldsymbol{\epsilon}_1}$이라 두면 $\boldsymbol{X}_{\boldsymbol{\epsilon}_1} = (\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_1$으로 정의된다. $\boldsymbol{r}_A$의 추정벡터 $\hat{\boldsymbol{r}}_A$는 $\boldsymbol{r}_A$를 $\boldsymbol{X}_{\boldsymbol{\epsilon}_1}$으로의 사영이므로 $\hat{\boldsymbol{r}}_A = \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-\boldsymbol{r}_A$로 구해진다. 오차벡터 $\boldsymbol{\epsilon}_1$에 따른 제곱합은 $\boldsymbol{r}_A'\boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-\boldsymbol{r}_A$이다. 잔차벡터 $\boldsymbol{r}_{\boldsymbol{\epsilon}_1}$는 $(\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-)\boldsymbol{r}_A$로 주어진다. 잔차벡터 $\boldsymbol{r}_{\epsilon_1}$에 대한 확률모형은

$$\boldsymbol{r}_{\epsilon_1} = \left(\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{X}_2\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_B. \tag{3.6}$$

단, $\boldsymbol{\epsilon}_B = (\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-)(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{\epsilon}$이다. 확률벡터 $\boldsymbol{\epsilon}_2$의 계수행렬을 $\boldsymbol{X}_{\boldsymbol{\epsilon}_2}$라 두면 $\boldsymbol{X}_{\boldsymbol{\epsilon}_2} = (\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-)(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_2$이다. $\boldsymbol{r}_{\boldsymbol{\epsilon}_1}$의 추정벡터 $\hat{\boldsymbol{r}}_{\boldsymbol{\epsilon}_1}$는 $\boldsymbol{X}_{\boldsymbol{\epsilon}_2}\boldsymbol{X}_{\boldsymbol{\epsilon}_2}^-\boldsymbol{r}_{\boldsymbol{\epsilon}_1}$로 구해진다. 따라서 오차벡터 $\boldsymbol{\epsilon}_2$에 따른 제곱합은 $\boldsymbol{r}_{\epsilon_1}$을 $\boldsymbol{X}_{\boldsymbol{\epsilon}_2}$로의 사영으로부터 구해진 사영까지의 거리제곱합을 나타내는 $\boldsymbol{r}_{\boldsymbol{\epsilon}_1}'\boldsymbol{X}_{\boldsymbol{\epsilon}_2}\boldsymbol{X}_{\boldsymbol{\epsilon}_2}^-\boldsymbol{r}_{\boldsymbol{\epsilon}_1}$이다. 오차벡터 $\boldsymbol{\epsilon}$에 따른 제곱합을 Q_ϵ이라 두고 Q_ϵ을 구하기 위한 잔차벡터를 $\boldsymbol{r}_{\boldsymbol{\epsilon}_2}$로 두면 $\boldsymbol{r}_{\boldsymbol{\epsilon}_2} = \boldsymbol{r}_{\epsilon_1} - \hat{\boldsymbol{r}}_{\epsilon_1} = (\boldsymbol{I} - \boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-)\boldsymbol{r}_{\epsilon_1}$이다. 잔차벡터 $\boldsymbol{r}_{\boldsymbol{\epsilon}_2}$에 대한 모형은

$$\boldsymbol{r}_{\epsilon_2} = \left(\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_2}\boldsymbol{X}_{\boldsymbol{\epsilon}_2}^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\right)\left(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{\epsilon} \tag{3.7}$$

이다. 식 (3.7)에서 오차벡터 $\boldsymbol{\epsilon}$의 모형행렬을 $\boldsymbol{X}_\epsilon$라 두면 $\boldsymbol{X}_\epsilon = (\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_2}\boldsymbol{X}_{\boldsymbol{\epsilon}_2}^-)(\boldsymbol{I} - \boldsymbol{X}_{\boldsymbol{\epsilon}_1}\boldsymbol{X}_{\boldsymbol{\epsilon}_1}^-)(\boldsymbol{I} - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)$이다. 오차벡터 $\boldsymbol{\epsilon}$의 추정벡터 $\hat{\boldsymbol{\epsilon}}$은 $\boldsymbol{X}_{\boldsymbol{\epsilon}}\boldsymbol{X}_{\boldsymbol{\epsilon}}^-\boldsymbol{r}_{\boldsymbol{\epsilon}}$로 구해지고 오차벡터 $\boldsymbol{\epsilon}$에 따른 제곱합은 $\boldsymbol{r}_{\boldsymbol{\epsilon}_2}'\boldsymbol{X}_{\boldsymbol{\epsilon}}\boldsymbol{X}_{\boldsymbol{\epsilon}}^-\boldsymbol{r}_{\boldsymbol{\epsilon}_2}$로 계산된다. 지분계획 모형식 (2.2)에서 고려된 확률효과벡터와 오차벡터에 따른 제1종 제곱합을 사영에 근거하여 계산하는 과정을 다루었다. 사영에 근거한 변동요인들의 변동량 계산을 위한 잔차모형들을 유도하고 잔차모형별 모형행렬로부터 사영을 이용하여 해당하는 변동요인의 변동량을 구할 수 있음을 보여주고 있다. 분산성분들을 적률법으로 추정하기 위해 확률효과와 오차에 따른 제곱합과 제곱합의 기댓값을 구한다. 제곱합의 기댓값을 구하기 위한 방법으로 Hartley의 합성법을 이용하기로 한다. 요인 A의 확률효과 $\boldsymbol{\alpha}$에 따른 제곱합을 Q_A라 두면

$$\begin{aligned}
E(Q_A) &= E\left(\boldsymbol{r}_f'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{r}_f\right) \\
&= E\left(\boldsymbol{y}'\left(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\left(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{y}\right) \\
&= E\left(\boldsymbol{y}'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{y}\right) \\
&= \mathrm{tr}\left(\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{\Sigma}\right) + E(\boldsymbol{y}')\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-E(\boldsymbol{y}) \\
&= \mathrm{tr}\left(\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{\Sigma}\right) + \left(\boldsymbol{\beta}_f'\boldsymbol{X}_f'\right)\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\left(\boldsymbol{X}_f\boldsymbol{\beta}_f\right) \\
&= \mathrm{tr}\left(\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{\Sigma}\right) \\
&= \mathrm{tr}\left[\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\left(\sigma_A^2\boldsymbol{X}_A\boldsymbol{X}_A' + \sigma_1^2\boldsymbol{X}_1\boldsymbol{X}_1' + \sigma_2^2\boldsymbol{X}_2\boldsymbol{X}_2' + \sigma_\epsilon^2\boldsymbol{I}\right)\right] \\
&= \mathrm{tr}(\boldsymbol{X}_A'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{X}_A)\sigma_A^2 + \mathrm{tr}(\boldsymbol{X}_1'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{X}_1)\sigma_1^2 + \mathrm{tr}(\boldsymbol{X}_2'\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-\boldsymbol{X}_2)\sigma_2^2 + \mathrm{tr}(\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)\sigma_\epsilon^2 \\
&= c_{11}\sigma_A^2 + c_{12}\sigma_1^2 + c_{13}\sigma_2^2 + c_{14}\sigma_\epsilon^2
\end{aligned} \tag{3.8}$$

로 구해진다. 단, c_{ij}는 각 분산성분의 계수로 주어지는 행렬의 대각합을 나타낸다. 식 (3.8)의 두 번째 식에서 행렬의 곱 $(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)$은 사영행렬의 성질을 이용할 때 $\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-$로 표현됨을 알 수있다. 즉, $(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_\alpha = (\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)(\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_A = (\boldsymbol{I} - \boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_A$이

므로 $\boldsymbol{X}_\alpha$이다. 행렬 곱 $\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)$의 계산을 위해 $\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-$가 대칭행렬인 성질을 이용한다. 행렬의 곱 $\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)$은 $(\boldsymbol{X}_\alpha^-)'\boldsymbol{X}_\alpha'(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)$로부터 $\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-$임을 알 수 있다. 식에서 $(\boldsymbol{\beta}_f'\boldsymbol{X}_f')\boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-(\boldsymbol{X}_f\boldsymbol{\beta}_f)$의 값은 0이 된다. 왜냐하면, 행렬곱 $\boldsymbol{X}_f'\boldsymbol{X}_\alpha=\boldsymbol{X}_f'(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{X}_A$에서 $\boldsymbol{X}_f'(\boldsymbol{I}-\boldsymbol{X}_f'^-\boldsymbol{X}_f')\boldsymbol{X}_A$가 영행렬로 주어지기 때문이다. 오차벡터 $\boldsymbol{\epsilon}_1$에 따른 제곱합을 Q_1이라 두면

$$\begin{aligned}
E(Q_1) &= E(\boldsymbol{r}_A'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{r}_A) \\
&= E\left(\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)\boldsymbol{y}\right) \\
&= E(\boldsymbol{y}'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{y}) \\
&= \mathrm{tr}(\boldsymbol{X}_A'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{X}_A)\sigma_A^2 + \sum_{i=1}^{2}\mathrm{tr}(\boldsymbol{X}_i'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{X}_i)\sigma_i^2 + \mathrm{tr}(\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-)\sigma_\epsilon^2 \\
&= c_{21}\sigma_A^2 + c_{22}\sigma_1^2 + c_{23}\sigma_2^2 + c_{24}\sigma_\epsilon^2
\end{aligned} \tag{3.9}$$

이다. 식 (3.9)의 두 번째 식에서 행렬의 곱 $(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^-)$은 식 (3.8)의 설명에서와 같이 사영행렬의 성질을 이용할 때 $\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-$로 표현됨을 알 수있다. 오차벡터 $\boldsymbol{\epsilon}_2$에 따른 제곱합을 Q_2라 두면

$$\begin{aligned}
E(Q_2) &= E(\boldsymbol{r}_{\epsilon_1}'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\boldsymbol{r}_{\epsilon_1}) \\
&= E\left[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^- - \boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-)\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^- \times (\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^- - \boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-)\boldsymbol{y}\right] \\
&= E(\boldsymbol{y}'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}{}^-\boldsymbol{y}) \\
&= \mathrm{tr}(\boldsymbol{X}_A'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\boldsymbol{X}_A)\sigma_A^2 + \sum_{i=1}^{2}\mathrm{tr}(\boldsymbol{X}_i'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\boldsymbol{X}_i)\sigma_i^2 + \mathrm{tr}(\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-)\sigma_\epsilon^2 \\
&= c_{31}\sigma_A^2 + c_{32}\sigma_1^2 + c_{33}\sigma_2^2 + c_{34}\sigma_\epsilon^2
\end{aligned} \tag{3.10}$$

이다. 오차벡터 $\boldsymbol{\epsilon}$에 따른 제곱합을 Q_ϵ라 두면

$$\begin{aligned}
E(Q_\epsilon) &= E\left(\boldsymbol{r}_{\epsilon_2}'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\boldsymbol{r}_{\epsilon_2}\right) \\
&= E\left[\boldsymbol{y}'(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^- - \boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^- - \boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-)\boldsymbol{X}_\epsilon\boldsymbol{X}_\epsilon^-\right. \\
&\quad \left.\times\left(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^- - \boldsymbol{X}_\alpha\boldsymbol{X}_\alpha^- - \boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^- - \boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\right)\boldsymbol{y}\right] \\
&= E\left(\boldsymbol{y}'\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\boldsymbol{y}\right) \\
&= \mathrm{tr}\left(\boldsymbol{X}_{\epsilon_2}\boldsymbol{X}_{\epsilon_2}^-\right)\sigma_\epsilon^2 \\
&= c_{44}\sigma_\epsilon^2
\end{aligned} \tag{3.11}$$

이다. 한 실험의 실험계획에서 처리구조와 설계구조가 모두 지분구조를 갖는 지분계획의 경우에 고정효과들의 추정가능함수와 확률효과 및 오차성분들의 분산성분을 추정하는 것이 필요로 하게 된다. 자료벡터의 행렬표현으로 부터 벡터공간의 사영의 관점에서 자료분석이 행해질 때 행렬의 다양한 성질과 특성들을 이용할 수 있는 이점이 있다. 적률법으로 분산성분을 구하기 위한 방정식은

$$\begin{aligned}
\boldsymbol{Q}_A &= c_{11}\sigma_A^2 + c_{12}\sigma_1^2 + c_{13}\sigma_2^2 + c_{14}\sigma_\epsilon^2, \\
\boldsymbol{Q}_1 &= c_{21}\sigma_A^2 + c_{22}\sigma_1^2 + c_{23}\sigma_2^2 + c_{24}\sigma_\epsilon^2, \\
\boldsymbol{Q}_2 &= c_{31}\sigma_A^2 + c_{32}\sigma_1^2 + c_{33}\sigma_2^2 + c_{34}\sigma_\epsilon^2, \\
\boldsymbol{Q}_\epsilon &= c_{44}\sigma_\epsilon^2
\end{aligned} \tag{3.12}$$

Table 4.1. Nested design data for the comfort study

Temperature	Sex	Chamber_1	Chamber_2	Chamber_3
$65F$	Male	5, 4	5, 4	4, 3
	Female	1, 2	5, 5	1, 3
Temperature	Sex	Chamber_4	Chamber_5	Chamber_6
$70F$	Male	8, 8	6, 3	5, 7
	Female	10, 7	8, 8	8, 8
Temperature	Sex	Chamber_7	Chamber_8	Chamber_9
$75F$	Male	12, 8	8, 7	6, 6
	Female	11, 13	8, 8	6, 7

이다. 식 (3.2)의 연립방정식으로부터 분산성분들의 해벡터 $\boldsymbol{\sigma}^2$를 구하게 된다. $\boldsymbol{\sigma}^2 = (\sigma_A^2, \sigma_1^2, \sigma_2^2, \sigma_\epsilon^2)'$ 이다. 식 (3.12)에서 c_{ij} $(i, j = 1, 2, 3, 4)$는 $\boldsymbol{y}$의 이차형식으로 주어지는 제곱합 Q_i에서 분산성분들의 계수로 주어지는 대각합을 나타낸다. 분산성분의 추정에 적률법을 비롯한 최대우도법과 MINQUE방법 등의 여러 방법이 가능하나 불균형자료의 경우 적률법에 의한 계산이 다른 방법들에 비해 단순하다는 이점이 있다.

4. 자료의 예

본문에서 논의된 지분계획모형의 축소된 형태이나 실험단위의 지분구조를 나타내는 자료를 이용하여 분산성분을 구하는 과정을 설명해 보기로 한다. Table 4.1은 실험단위의 설계구조에서 지분관계를 갖는 지분계획의 한 형태를 나타내는 Milliken과 Johnson (1984)의 자료이다. 실험의 두 요인은 세 수준의 실내온도와 두 수준의 성별이다. 반응은 고정된 실내온도에서 성별에 따라 평안함을 느끼는 점수(comfort scores)이고 실내온도의 수준은 9개 가능한 실험방 중 임의로 세 방에 배정되고 각 방에 남자 2명 여자 2명이 임의로 배정되어 한 실험방에서 4개의 점수가 구해진다. 실내온도와 성별은 개인별 반응 (y)에 영향을 주는 고정요인들이고 각 실내온도에 임의로 배정된 방과 그 방에 배정된 사람들은 지분관계에 있는 지분구조의 실험단위를 갖는 지분계획으로 간주하고 있다. Table 4.1의 자료분석을 위한 Milliken과 Johnson (1984)의 지분계획모형은

$$y_{ijkm} = \mu_{ik} + c_{j(i)} + p_{m(ijk)} \tag{4.1}$$

이다. 단, μ_{ik}는 온도 i와 성별 k에서 평균반응이고 $c_{j(i)}$는 온도 i가 배정된 방 j의 효과이고 $p_m(ijk)$는 온도 i의 방 j에 배정된 성별 k인 m번째 사람의 효과이다. $c_{j(i)}$들은 독립이고 $N(0, \sigma_c^2)$인 분포를 따른다고 가정한다. $p_m(ijk)$도 독립이고 $N(0, \sigma_p^2)$인 분포를 따른다고 가정한다. $c_{j(i)}$와 $p_m(ijk)$는 독립이라고 가정한다. 식 (4.1)은 실험의 처리구조와 설계구조에서 모두 지분구조를 갖는 경우의 모형식 (2.1)에서

$$y_{ijkm} = \mu + \alpha_i + \epsilon_{j(i)} + \beta_{k(i)} + (\alpha\beta)_{ik(i)} + \epsilon_{m(ijk)} \tag{4.2}$$

로 표현되는 모형이다. 즉, $\mu_{ik} = \mu + \alpha_i + \beta_{k(i)} + (\alpha\beta)_{ik(i)}$로 취급되면 식 (4.1)과 식 (4.2)는 동일함을 알 수 있다. 식 (4.2)의 α_i는 온도의 고정효과를 나타낸다. 식 (4.2)의 행렬표현식은

$$\boldsymbol{y} = \boldsymbol{j}\mu + \boldsymbol{X}_T\boldsymbol{\alpha} + \boldsymbol{X}_1\boldsymbol{\epsilon}_1 + \boldsymbol{X}_S\boldsymbol{\beta} + \boldsymbol{X}_{TS}(\boldsymbol{\alpha\beta}) + \boldsymbol{\epsilon} \tag{4.3}$$

이다. $\boldsymbol{X}_T$는 온도의 고정효과벡터 $\boldsymbol{\alpha}$의 계수행렬이고 $\boldsymbol{X}_S$는 성별효과 $\boldsymbol{\beta}$의 계수행렬이며 $\boldsymbol{X}_{TS}$는 교호작용$(\boldsymbol{\alpha\beta})$의 계수행렬을 나타낸다. 식 (4.3)에서 $\boldsymbol{X}_f = (\boldsymbol{j}, \boldsymbol{X}_T, \boldsymbol{X}_S, \boldsymbol{X}_{TS})$라 두고 $\boldsymbol{r}_f$를 구하면 $\boldsymbol{r}_f =$

$(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-)\boldsymbol{y}$로 주어진다. 잔차벡터 $\boldsymbol{r}_f$에 대한 확률모형은

$$\boldsymbol{r}_f=\left(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{X}_1\boldsymbol{\epsilon}_1+\left(\boldsymbol{I}-\boldsymbol{X}_f\boldsymbol{X}_f^-\right)\boldsymbol{\epsilon} \tag{4.4}$$

이다. 모형식 (4.4)를 이용하여 오차벡터 $\boldsymbol{\epsilon_1}$에 따른 제곱합 $Q_1=66.5$와 $\boldsymbol{\epsilon}$에 따른 제곱합 $Q_\epsilon=39.7$을 구한다. 지분관계의 두 실험단위의 오차에 따른 분산성분 $\sigma^2_{\epsilon_1}$과 σ^2_ϵ을 추정하기 위한 방정식들은

$$\begin{aligned}\boldsymbol{Q}_1&=c_{11}\sigma^2_{\epsilon_1}+c_{12}\sigma^2_\epsilon,\\ \boldsymbol{Q}_\epsilon&=c_{22}\sigma^2_\epsilon\end{aligned} \tag{4.5}$$

이다. 식 (4.5)의 계수들은 Hartley의 합성법에 의해 $E(\boldsymbol{Q}_1)$과 $E(\boldsymbol{Q}_\epsilon)$는

$$\begin{aligned}E(Q_1)&=E\left(\boldsymbol{r}_f'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{r}_f\right)\\&=\sigma_1^2\,\mathrm{tr}\left(\boldsymbol{X}_1'\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{X}_1\right)+\sigma_\epsilon^2\,\mathrm{tr}\left(\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\right)\\&=24\sigma^2_{\epsilon_1}+6\sigma^2_\epsilon,\\E(Q_\epsilon)&=E\left[\boldsymbol{r}_f'\left(I-\boldsymbol{X}_f\boldsymbol{X}_f^--\boldsymbol{X}_{\epsilon_1}\boldsymbol{X}_{\epsilon_1}^-\boldsymbol{r}_f\right)\right]\\&=24\sigma^2_\epsilon\end{aligned} \tag{4.6}$$

으로 구해진다. 적률법에 의한 연립방정식의 해는 $\hat{\sigma}^2_{\epsilon_1}=2.36$이고 $\hat{\sigma}^2_\epsilon=1.65$로 구해진다.

5. 결론

본 논문은 지분계획하에서 실험이 행해질 때 실험단위의 반응에 영향을 주는 요인들과 실험단위들이 모두 지분구조를 갖는 경우에 분산성분의 추정방법을 다루고 있다. 실험의 지분계획에서 나타나는 요인들의 지분구조는 다양한 형태로 주어질 수 있으나 본 논문에서는 요인들의 지분구조로 확률요인에 내재된 고정의 교차요인들을 가정하고 있으며 또한 지분구조의 요인들은 지분관계의 상이한 실험단위를 취한다고 가정하고 있다. 따라서 이러한 지분계획하의 실험자료를 분석하기 위한 모형으로 처리구조에서 요인간의 지분구조와 설계구조에서 실험단위 간의 지분구조를 고려한 모형이 논의되었으며 모형에 포함되어 있는 고정효과들과 확률효과들의 분산성분 그리고 이들 간의 교호작용을 구하는 방법을 다루고 있다. 본 논문의 지분계획에 가정된 모형은 지분구조의 실험단위들로 인해 다수의 오차항이 추가된 형태의 혼합모형으로 단순히 두 유형의 처리요인에 대한 일반적인 혼합효과모형과는 차이가 있음을 나타내고 있다. 지분계획모형하의 실험자료를 분석하기 위한 여러 방법이 있으나 벡터공간의 분할을 통한 사영의 관점에서 변동요인의 변동량을 구하는 방법에 관한 논의는 찾아보기가 쉽지 않다. 본 논문은 요인들의 지분구조와 실험단위 간의 지분구조를 갖는 지분계획모형에서 변동요인별 제곱합을 구하기 위한 단계별 잔차모형에서 모형행렬을 정의하고 사영을 이용하여 분석하는 방법을 구체적으로 다루고 있다.

References

Choi, J. S. (2011). Type I analysis by projections, *The Korean Journal of Applied Statistics*, **24**, 373−381.

Choi, J. S. (2012). Type II analysis by projections, *Journal of the Korean Data & Information Science Society*, **23**, 1155−1163.

Choi, J. S. (2014). Projection analysis for two-way variance components, *Journal of the Korean Data & Information Science Society*, **23**, 547−554.

Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Wadsworth, Inc. California.

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

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

Hicks, C. R. (1973). *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, Inc., New York.

John, P. W. M. (1971). *Statistical Design and Analysis of Experiments*, The Macmillan Company, New York.

Milliken, G. A. and Johnson, D. E. (1984). *Analysis of Messy Data*, Van Nostrand Reinhold, New York.

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

Searle, S. R., Casella, G. and McCulloch, C. E. (1971). *Linear Models*, John Wiley and Sons, Inc., New York.

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

지분계획의 분산성분

최재성[a,1]

[a]계명대학교 통계학과

(2015년 8월 5일 접수, 2015년 10월 24일 수정, 2015년 12월 4일 채택)

요 약

본 논문은 요인들의 처리구조와 실험단위들의 설계구조에서 지분이 발생하는 경우의 지분계획모형에서 분산성분을 구하는 방법을 다루고 있다. 지분구조의 고정효과와 확률효과 그리고 실험단위들의 지분구조에 따른 오차성분을 포함하는 지분계획모형을 제안하고 있다. 모형내 확률효과의 분산성분과 다수의 오차항에 따른 분산성분을 추정하는 방법으로 상수적합법을 이용하고 있다. 상수적합법에 의한 제1종 제곱합의 계산은 모형의 단계별 적합에서 주어지는 모형행렬의 사영을 이용하고 구하고 있다. 사영을 이용한 변동요인별 제1종 제곱합의 기댓값 계산에 Hartley의 합성법이 이용된다. 단계별 방법에 의한 모형의 순차적 적합은 모형행렬로의 사영공간을 나타내는 사영행렬의 구조를 파악할 수 있는 이점이 있다.

주요용어: 분산성분, 지분구조, 사영, 제1종 제곱합, 합성법

[1](42601) 대구광역시 달서구 신당동 1000번지, 계명대학교 통계학과. E-mail: jschoi@kmu.ac.kr