Original Research Article

On Estimating Variance Components of Two-Way Nested Random Model with Missing Information

ABSTRACT:

In this paper, Alternative estimators have been derived for estimating the variance components according to modified Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of two-way nested random model when problem of missing information exists.

Keywords: Variance Components, MINQUE, MIVQUE, Missing Information.

INTRODUCTION

Rao (1970) suggested a method of estimation “MINQUE” that does not require the normality assumption for the estimation of variances. Rao (1971) proposed a method of estimation that called MIVQUE, Minimum Variance Quadratic Unbiased Estimation. Subramani (2012) suggested a modification on the computational aspects of MIVQUE of variance components in mixed linear models. He introduced two modified MIVQUE (MIVQUE I and MIVQUE II). He estimated variance components in unbalanced one-way random model by Modified MIVQUE and compared between MIVQUE I, MIVQUE II, MIVQUE based on different optimality criteria.

The meaning of incomplete (missing) information is different from the meaning of missing values. Missing values related to the losing of the observation while missing information related to the losing of location of the observation. This means that the value of the observation is known which could happen because it may be not recorded or lost for any other reasons. (Saleh and El-sheikh, 2002)

Discarding data with missing information loses useful information, and in some circumstances, this approach cannot separate variance components from one another (i.e. it may make some variance components inestimable)

Missing information has three types completely missing information, partially missing information and not at all on any observation.

1. Completely missing information
   No observation in the main group has subgroup information.
2. Partially missing information
   Some of the observations in the main group have missing subgroup information.
3. Not at all on any observation in the main group is missing.

(Song and Shulman, 1997)

Song and Shulman (1997) estimated the variance components for the data with missing nesting information in the two-stage unbalanced nested random model. They combined sum of squares for the data with missing nesting information with the sum of squares for the data with complete nesting information linearly. Prespecified weights are used for the combination. Different estimates are obtained by using different weights. Variances and covariances of these estimators are derived and used to compare these estimators. Saleh and El Sheikh (2003) modified the analysis of variance method and the combined symmetric sums with the analysis of variance method for estimation of the variance components of three-stage unbalanced nested random models for the data with complete missing nesting information. By a simulation study, they compared the bias and the mean squares errors of the estimates of variance components of the five methods of estimation namely: ANOVA method (Henderson’s method 1), Modified ANOVA method, Combined analysis of means with ANOVA method, Combined symmetric sums method with ANOVA method, Combined symmetric sums method with modified ANOVA method.

The paper is organized as follows: The second section concerns with the Modified MIVQUE (I) method introduced by Subramani (2012). The third section illustrates the proposed estimators of modified MIVQUE I method for complete information and completely missing information in case of unbalanced two-way nested random model. The fourth section illustrates simulation study to compare different methods by relative efficiency criteria.

MINIMUM VARIANCE QUADRATIC UNBIASED ESTIMATION (MIVQUE I)

Assume the model:

$$Y = X\beta + Z_1\delta_1 + Z_2\delta_2 + \cdots + Z_d\delta_d \qquad (1)$$

where Y is an $N \times 1$ vector of observations, N is the sample size
X is a $N \times s$ matrix with known constants,
β is a $s \times 1$- vector of fixed (unknown) parameters,
Z_i is a $N \times c_i$ matrix with known constants, $i = 1, \dots, d$. ($Z_d = I, c_d = N$)
δ_i is a $c_i \times 1$-vector of random variables. ($\delta_d = e$)
Assume that δ_i is random variable with zero mean value and dispersion matrix $\sigma_i^2 I_{c_i}$. Further, δ_i and δ_j are uncorrelated.
Model (1) can be expressed in a compact form as:

$$Y = X\beta + Z\delta \qquad (2)$$

where $Z = (Z_1 : Z_2 : \cdots : Z_d)$ and $\delta = (\delta_1 : \delta_2 : \cdots : \delta_d)$.
$E(Y) = X\beta$ and $D(Y) = V = \sum_{i=1}^{d} \sigma_i^2 V_i$ where $V_i = Z_i Z_i^t$.
$D(Y)$ is called the dispersion matrix and the parameters $\sigma_1^2, \dots, \sigma_d^2$ are the unknown variance components whose values should be estimated. (Subramani, 2012)
Subramani (2012) developed the estimation of variance components based on Rao (1971) approach. Instead of dealing with one linear combination, he decided to estimate a set of linear combinations of variance components $\sum_{j=1}^{d} \rho_{ij}\sigma_i$ through a set of quadratic functions $Y^t A_i Y$ [A_i is a symmetric matrix and $\rho_{ij} = Tr(A_i V_j)$]. He claimed that estimating variance components under normality obtained by solving the following equations:

$$\begin{bmatrix} Tr(A_1V_1) & \cdots & Tr(A_1V_d) \\ \vdots & \ddots & \vdots \\ Tr(A_dV_1) & \cdots & Tr(A_dV_d) \end{bmatrix}_{d*d} \begin{bmatrix} \sigma_1^2 \\ \vdots \\ \sigma_d^2 \end{bmatrix}_{d*1} = \begin{bmatrix} Tr(A_1W) \\ \vdots \\ Tr(A_dW) \end{bmatrix}_{d*1} \qquad (3)$$

He introduced different formulas of A_i to obtain MIVQUE(I). The formulas of A_i $(i = 1,2, \dots, d)$ have the following form:
$A_i = V^{-1}(I - P_{U_i})$ where $P_{U_i} = U_i(U_i^t V^{-1} U_i)^- U_i^t V^{-1}$. U_i has a variety of choices,

$$U_1 = X, U_2 = [X \quad Z_1], U_3 = [X \quad Z_2], \dots, \qquad U_d = [X \quad Z_{d-1}]$$
$$U_1 = X, U_2 = [X \quad Z_1], U_3 = [X \quad Z_1 \quad Z_2], \dots, \quad U_d = [X \quad Z_1 \quad Z_2 \quad \dots \quad Z_{d-1}] \qquad (4)$$
$$U_1 = X, U_2 = [X \quad X \quad Z_1], U_3 = [X \quad X \quad Z_1 \quad Z_2], \dots, \quad U_d = [X \quad X \quad Z_1 \quad \dots \quad Z_{d-1}]$$

where $(U_i^t V^{-1} U_i)^-$ is the generalized inverse of $U_i^t V^{-1} U_i$
For the case (i), he derived the estimators, their variances and covariance matrix in the unbalanced one-way random model. The resulting method are referred to as MIVQUE I.
The proposed estimators of variance components are derived by replacing A_i in eq. (3) by A_i for case (ii) in eq. (4).
So the steps of MIVQUE method: 1- Selecting a symmetric matrix A_i, 2- Solving the equation (3), 3-obtain the estimators of MIVQUE method.

ESTIMATION OF VARIANCE COMPONENTS WITH COMPLETELY MISSING INFORMATION

In this section, the variance components will be estimated in case of completely missing information by MMIVQUE I. Pre-specified weights will be used to combine complete information and missing information.

Consider the two- way nested random model

$$Y_{ijk} = \eta + \gamma_i + \beta_{j(i)} + e_{k(ij)} \qquad (5)$$

$$i = 1,2,\dots,S,\; j = 1,2,\dots,D_i,\; k = 1,2,\dots,n_{ij}$$

where Y_{ijk} is the k^{th} observation at the j^{th}level of factor β within the i^{th} level of factor γ.

η is the general mean.

γ_i , $\beta_{ij}$and e_{ijk} are mutually independent random variables with zero means and variances σ^2_γ , $\sigma^2_\beta$and σ^2_e respectively. The variance components to be estimated are σ^2_γ , σ^2_β and σ^2_e.

When $D_i = D$ for all i and $n_{ij} = n$ for all i and j, each ij cell contains the same number of observations. The data shall be described as balanced data.

So the model (5) can be written in matrix form of balanced data as:

$$Y = X\eta + T\gamma + T\beta + Te \qquad (6)$$

where Y is an $N \times 1$ vector of observations $X = 1_{N*1}, N = SDn$

$$T_1 = \begin{bmatrix} 1_{\sum_{j=1}^{D_1} n_{1j}*1} & 0 & 0 & 0 & 0 \\ 0 & 1_{\sum_{j=1}^{D_2} n_{2j}*1} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1_{\sum_{j=1}^{D_S} n_{Sj}*1} \end{bmatrix}$$

$$T_2 = \begin{bmatrix} \underline{1}_{n_{11}} & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \underline{1}_{n_{1D_1}} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \underline{1}_{n_{SD_S}} \end{bmatrix}$$

, $T_3 = I_N$,

with $E(Y) = X\eta$ and $D(Y) = V = V_1\sigma^2_\gamma + V_2\sigma^2_\beta + V_3\sigma^2_e$

, $V_i = T_i T_i^t$.

$$V_1 = T_1 T_1^t = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & J_S \end{bmatrix} = \sum_{i=1}^{+S} J_i$$

where J_i denote $\sum_{j=1}^{D_i} n_{ij} \times \sum_{j=1}^{D_i} n_{ij}$ matrix consisting of $1'$s.

and

$$V_2 = T_2 T_2^t = \begin{bmatrix} K_{n_{11}} & 0 & \dots & 0 \\ 0 & K_{n_{12}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_{n_{SD_S}} \end{bmatrix} = K$$

where $K_{n_{ij}}$ denote $n_{ij} \times n_{ij}$ matrix consisting of $1'$s.

$$V^{-1} = \frac{1}{\sigma^2_e} I - \sum_{i=1}^{+S} B_i - \sum_{i=1}^{+S} \frac{\sigma^2_\gamma}{1 + \sigma^2_\gamma \sum_{j=1}^{D_i} \frac{n_{ij}}{(\sigma^2_e + n_{ij}\sigma^2_\beta)}} C_i$$

where

$$B_i = \begin{bmatrix} \frac{\sigma^2_\beta}{\sigma^2_e(\sigma^2_e + n_{i1}\sigma^2_\beta)} K_{n_{i1}\times n_{i1}} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & & \dots & \vdots \\ \vdots & \vdots & & \\ 0 & 0 & \dots & \frac{\sigma^2_\beta}{\sigma^2_e(\sigma^2_e + n_{iD_i}\sigma^2_\beta)} K_{n_{iD_i}\times n_{iD_i}} \end{bmatrix}$$

and

$$C_i=\begin{bmatrix}\frac{1}{(\sigma_e^2+n_{i1}\sigma_\beta^2)}\\ \vdots\\ \frac{1}{(\sigma_e^2+n_{iD_i}\sigma_\beta^2)}\end{bmatrix}\begin{bmatrix}\frac{1}{(\sigma_e^2+n_{i1}\sigma_\beta^2)} & \cdots & \frac{1}{(\sigma_e^2+n_{iD_i}\sigma_\beta^2)}\end{bmatrix}$$

Let S' and $S''=S-S'$ be the numbers of main groups with no missing and missing subgroup nesting information respectively. We assume that all $D_i's$ and $n_{ij}'s$ in Model (6) are known.

The variance components will be estimated for balanced two–way nested random model using MMIVQUEI.

Let $V^*=\alpha_1V_1+\alpha_2V_2+\alpha_0V_3$ where α_1 , α_2 and α_0 are a priori values for σ_γ^2 , σ_β^2 and σ_e^2 respectively.

$$V^{*(-1)}=\frac{1}{\alpha_0}I-\sum_{i=1}^{+S}B_i-\sum_{i=1}^{+S}\frac{\alpha_1}{1+\alpha_1\sum_{j=1}^{D_i}\frac{n_{ij}}{(\alpha_0+n_{ij}\alpha_2)}}C_i$$

Steps of estimation:

Estimation of variance components for S' groups. (data with complete information)

Estimation of variance components for S'' groups. (data with missing information)

Pre-specified weights will be used to combine data with complete information and missing information.

If there is a notation with a single prime (double primes) is defined as the quantity with complete information (with missing information).

Lemma (1): The estimators of variance components in balanced two-way nested random model with completely missing information are:

$$\hat{\sigma}_e^2=\frac{Q_3'}{Tr(A_3',V_3')}$$

$$\hat{\sigma}_\beta^2=\frac{w_2Q_2'+(1-w_2)Q_2''-[w_2Tr(A_2'V_3')+(1-w_2)Tr(A_2''V_3'')]\hat{\sigma}_e^2}{[w_2Tr(A_2'V_2')+(1-w_2)Tr(A_2''V_2'')]}$$

$$\hat{\sigma}_\gamma^2=\frac{w_1Q_1'+(1-w_1)Q_1''-P_2\hat{\sigma}_e^2-P_1\hat{\sigma}_\beta^2}{[w_1Tr(A_1'V_1')+(1-w_1)Tr(A_1''V_1'')]} \quad (7)$$

Proof:

Part (1): data with complete information:

For model (6), the matrix $A_i', i=1,2,3$ for data with complete information is:

$$A_{1U}'=V_{(c)}^{*(-1)}-h'\left[V_{(c)}^{*(-1)}X'X'^tV_{(c)}^{*(-1)}\right]$$

$$A_{2U}'=V_{(c)}^{*(-1)}-\left[\sum_{i=1}^{+S'}\frac{1}{h_i\left(1+\alpha_1\sum_{j=1}^{D_i}\frac{n_{ij}}{\alpha_0+n_{ij}\alpha_2}\right)^2}C_i\right]$$

$$A_{3U}'=V_{(c)}^{*(-1)}-\left[\sum_{i=1}^{+S'}F_i-\sum_{i=1}^{+S'}\frac{\alpha_1}{\left(1+\alpha_1\sum_{j=1}^{D_i}\frac{n_{ij}}{\alpha_0+n_{ij}\alpha_2}\right)}C_i\right]$$

where

$$F_i=\begin{bmatrix}\frac{1}{n_{i1}(\sigma_e^2+n_{i1}\sigma_\beta^2)}K_{n_{i1}\times n_{i1}} & 0 & & \cdots & 0\\ 0 & 0 & & \cdots & 0\\ & & & \cdots & \vdots\\ \vdots & \vdots & & \frac{1}{n_{iD_i}(\sigma_e^2+n_{iD_i}\sigma_\beta^2)}K_{n_{iD_i}\times n_{iD_i}} & \\ 0 & 0 & \cdots & & \end{bmatrix}$$

$$V_{(c)}^{*(-1)}=\frac{1}{\alpha_0}I_{N'*N'}-\sum_{i=1}^{+S'}B_i-\sum_{i=1}^{+S'}\frac{\alpha_1}{1+\alpha_1\sum_{j=1}^{D_i}\frac{n_{ij}}{(\alpha_0+n_{ij}\alpha_2)}}C_i$$

$$h' = \frac{1}{\sum_{i=1}^{S'} h_i}$$

The normal equations are:

$$\begin{bmatrix} \mathrm{Tr}(A_1'V_1') & \mathrm{Tr}(A_1'V_2') & \mathrm{Tr}(A_1'V_3') \\ \mathrm{Tr}(A_2'V_1') & \mathrm{Tr}(A_2'V_2') & \mathrm{Tr}(A_2'V_3') \\ \mathrm{Tr}(A_3'V_1') & \mathrm{Tr}(A_3'V_2') & \mathrm{Tr}(A_3'V_3') \end{bmatrix} \begin{bmatrix} \sigma_\gamma^2 \\ \sigma_\beta^2 \\ \sigma_e^2 \end{bmatrix} = \begin{bmatrix} Q_1' \\ Q_2' \\ Q_3' \end{bmatrix}$$

where $Q_i' = \mathrm{Tr}(A_i'W), W' = Y'Y'^t, i = 1,2,3$

$$Q_1' = \left[\frac{1}{\alpha_0}\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{k=1}^{n} Y_{ijk}^2 - \sum_{i=1}^{S}\sum_{j=1}^{D_i}\frac{\alpha_2}{\alpha_0 a_{ij}} Y_{ij.}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{\alpha_1}{a_{ij}^2 b_i} Y_{ij.}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i}\frac{\alpha_1}{a_{ij} b_i a_{il}} Y_{ij.} Y_{il.}\right]$$

$$- h'\left[\frac{1}{\alpha_0}\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{k=1}^{n} Y_{ijk} - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}\alpha_2}{\alpha_0 a_{ij}} Y_{ij.} - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i}\frac{\alpha_1 n_{ij}}{b_i}\left(\frac{1}{a_{ij}^2} Y_{ij.} + \frac{1}{a_{ij} a_{il}} Y_{il.}\right)\right]^2$$

$$Q_2' = \left[\frac{1}{\alpha_0}\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{k=1}^{n} Y_{ijk}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{\alpha_2}{\alpha_0 a_{ij}} Y_{ij.}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{\alpha_1}{a_{ij}^2 b_i} Y_{ij.}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i}\frac{\alpha_1}{a_{ij} b_i a_{il}} Y_{ij.} Y_{il.}\right]$$

$$- \left[\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{1}{h_i a_{ij}^2 b_i^2} Y_{ij.}^2 + \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i}\frac{1}{h_i a_{ij} b_i^2 a_{il}} Y_{ij.} Y_{il.}\right]$$

$$Q_3' = \frac{1}{\alpha_0}\left[\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\sum_{k=1}^{n} Y_{ijk}^2 - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{1}{n_{ij}} Y_{ij.}^2\right]$$

and

$$\mathrm{Tr}(A_1'V_1') = \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}}{a_{ij} b_i} - h'\sum_{i=1}^{S'}(h_i)^2$$

$$\mathrm{Tr}(A_1'\tilde{V}_2') = \left[\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}[a_{ij}b_i - n_{ij}\alpha_1]}{a_{ij}^2 b_i}\right] - h'\sum_{i=1}^{S'}\sum_{i=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i} n_{ij}\left[\frac{n_{ij}(b_i a_{ij} - n_{ij}\alpha_1)}{a_{ij}^3 b_i^2} - \frac{n_{il}^2\alpha_1}{a_{ij} b_i^2 a_{il}^2}\right]$$

$$\mathrm{Tr}(A_1'V_3') = \sum_{i=1}^{S'}\sum_{j=1}^{D_i} n_{ij}\left[\frac{1}{\alpha_0} - \frac{\alpha_2}{\alpha_0 a_{ij}} - \frac{\alpha_1}{b_i a_{ij}^2}\right] - h'\sum_{i=1}^{S'}\sum_{i=1}^{D_i}\sum_{\substack{l=1 \\ j\neq l}}^{D_i} n_{ij}\left[\frac{(b_i a_{ij} - n_{ij}\alpha_1)}{a_{ij}^3 b_i^2} - \frac{n_{il}\alpha_1}{a_{ij} b_i^2 a_{il}^2}\right]$$

$$\mathrm{Tr}(A_2'V_2') = \left[\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}[a_{ij}b_i - n_{ij}\alpha_1]}{a_{ij}^2 b_i}\right] - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}^2}{a_{ij}^2 b_i^2 h_i}$$

$$\mathrm{Tr}(A_2'V_3') = \sum_{i=1}^{S'}\sum_{j=1}^{D_i} n_{ij}\left[\frac{1}{\alpha_0} - \frac{\alpha_2}{\alpha_0 a_{ij}} - \frac{\alpha_1}{b_i a_{ij}^2}\right] - \sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{n_{ij}}{a_{ij}^2 b_i^2 h_i}$$

$$\mathrm{Tr}(A_3'V_3') = \sum_{i=1}^{S'}\sum_{j=1}^{D_i} n_{ij}\left[\frac{1}{\alpha_0} - \frac{\alpha_2}{\alpha_0 a_{ij}} - \frac{\alpha_1}{b_i a_{ij}^2}\right] - \left[\sum_{i=1}^{S'}\sum_{j=1}^{D_i}\frac{[a_{ij}b_i - n_{ij}\alpha_1]}{a_{ij}^2 b_i}\right]$$

Since

$\mathrm{Tr}(A_1'V_1')\sigma_\gamma^2 + \mathrm{Tr}(A_1'V_2')\sigma_\beta^2 + \mathrm{Tr}(A_1'V_3')\sigma_e^2 = Q_1'$ (8)

$\mathrm{Tr}(A_2'V_2')\sigma_\beta^2 + \mathrm{Tr}(A_2'V_3')\sigma_e^2 = Q_2'$ (9)

$\mathrm{Tr}(A_3'V_3')\sigma_e^2 = Q_3'$ (10)

By solving eq.(8), (9) and (10), the estimators of variance components for data with complete information can be given as follows:

$$\dddot{\sigma}_e^2 = \frac{Q'_3}{Tr(A'_3 V'_3)}$$

$$\dddot{\sigma}_\beta^2 = \frac{Q'_2 - Tr(A'_2 V'_3)\dddot{\sigma}_e^2}{Tr(A'_2 V'_2)}$$

$$\dddot{\sigma}_\gamma^2 = \frac{Q'_1 - Tr(A'_1 V'_3)\dddot{\sigma}_e^2 - Tr(A'_1 V'_2)\dddot{\sigma}_\beta^2}{Tr(A'_1 V'_1)}$$

Part (2): data with missing information by MMIVQUE I(0):

For model (6), the matrix $A''_{iU}, i = 1,2,3$ for data with complete information is:

$$A''_1 = V_{(m)}^{*(-1)} - h''\left[V_{(m)}^{*(-1)} X'' X''^t V_{(m)}^{*(-1)}\right]$$

$$A''_2 = V_{(m)}^{*(-1)} - \left[\sum_{i=S'+1}^{+S} \frac{1}{h_i\left(1 + \alpha_1 \sum_{j=1}^{D_i} \frac{n_{ij}}{\alpha_0 + n_{ij}\alpha_2}\right)^2} C_i\right]$$

$$A''_3 = V_{(m)}^{*(-1)} - \left[\sum_{i=S'+1}^{+S} F_i - \sum_{i=S'+1}^{+S} \frac{\alpha_1}{\left(1 + \alpha_1 \sum_{j=1}^{D_i} \frac{n_{ij}}{\alpha_0 + n_{ij}\alpha_2}\right)} C_i\right]$$

where

$$V_{(m)}^{*(-1)} = \frac{1}{\alpha_0} I_{N''*N''} - \sum_{i=S'+1}^{+S} B_i - \sum_{i=S'+1}^{+S} \frac{\alpha_1}{1 + \alpha_1 \sum_{j=1}^{D_i} \frac{n_{ij}}{(\alpha_0 + n_{ij}\alpha_2)}} C_i$$

$$h'' = \frac{1}{\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} n_{ij}}$$

The normal equations are:

$$\begin{bmatrix} Tr(A''_1 V''_1) & Tr(A''_1 V''_2) & Tr(A''_1 V''_3) \\ Tr(A''_2 V''_1) & Tr(A''_2 V''_2) & Tr(A''_2 V''_3) \\ Tr(A''_3 V''_1) & Tr(A''_3 V''_2) & Tr(A''_3 V''_3) \end{bmatrix} \begin{bmatrix} \sigma_\gamma^2 \\ \sigma_\beta^2 \\ \sigma_e^2 \end{bmatrix} = \begin{bmatrix} Q''_1 \\ Q''_2 \\ Q''_3 \end{bmatrix}$$

where

$$Q''_1 = \left[\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} \sum_{k=1}^{n_{ij}} Y_{ijk}^2\right] - h''\left[\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} \sum_{k=1}^{n_{ij}} Y_{ijk}\right]^2$$

$$Q''_2 = \left[\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} \sum_{k=1}^{n_{ij}} Y_{ijk}^2\right] - \left[\sum_{i=S'+1}^{S} \frac{1}{n_{i.}} Y_{i..}^2\right]$$

$$Q''_3 = \left[\sum_{i=1}^{S''} \sum_{j=1}^{D_i} \sum_{k=1}^{n_{ij}} Y_{ijk}^2 - \sum_{i=1}^{S''} \sum_{j=1}^{D} \frac{1}{n_{ij}} Y_{ij.}^2\right]$$

and

$$Tr(A''_1 V''_1) = \sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} n_{ij} - h'' \sum_{i=S'+1}^{S} \left(\sum_{j=1}^{D_i} n_{ij}\right)^2$$

$$Tr(A''_1 V''_2) = \left[\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} n_{ij}\right] - h'' \sum_{i=S'+1}^{S} \sum_{i=1}^{D_i} n_{ij}^2$$

$$Tr(A''_1 V''_3) = \sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} n_{ij} - h'' \sum_{i=S'+1}^{S} \sum_{i=1}^{D_i} n_{ij}$$

$$Tr(A''_2 V''_2) = \left[\sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} n_{ij}\right] - \sum_{i=S'+1}^{S} \sum_{j=1}^{D_i} \frac{n_{ij}^2}{\sum_{j=1}^{D_i} n_{ij}}$$

$$\mathrm{Tr}(A_2''V_3'') = \sum_{i=S'+1}^{S}\sum_{j=1}^{D_i} n_{ij} - \sum_{i=S'+1}^{S}\sum_{j=1}^{D_i}\frac{n_{ij}}{\sum_{j=1}^{D_i} n_{ij}}$$

$$\mathrm{Tr}(A_3''V_3'') = \sum_{i=S'+1}^{S}\sum_{j=1}^{D_i} n_{ij} - \left[\sum_{i=S'+1}^{S} D_i\right]$$

Since

$$\mathrm{Tr}(A_1''V_1'')\sigma_a^2 + \mathrm{Tr}(A_1''V_2'')\sigma_\beta^2 + \mathrm{Tr}(A_1''V_3'')\sigma_e^2 = Q_1'' \quad (11)$$

$$\mathrm{Tr}(A_2''V_2'')\sigma_\beta^2 + \mathrm{Tr}(A_2''V_3'')\sigma_e^2 = Q_2'' \quad (12)$$

$$\mathrm{Tr}(A_3''V_3'')\sigma_e^2 = Q_3'' \quad (13)$$

Part (3): Combination the data with complete information and completely missing information:

Pre-specified weights will be used to combine (Q_1'&Q_1'') and (Q_2'&Q_2'') as:

$$\mathrm{Tr}(A_3'V_3')\hat{\sigma}_e^2 = Q_3' \quad (14)$$

$$\begin{aligned} w_2Q_2' + (1-w_2)Q_2'' &= w_2\left[\mathrm{Tr}(A_2'V_2')\hat{\sigma}_\beta^2 + \mathrm{Tr}(A_2'V_3')\hat{\sigma}_e^2\right] \\ &+ (1-w_2)\left[\mathrm{Tr}(A_2''V_2'')\hat{\sigma}_\beta^2 + \mathrm{Tr}(A_2'V_3'')\hat{\sigma}_e^2\right] \quad (15) \end{aligned}$$

$$\begin{aligned} w_1Q_1' + (1-w_1)Q_1'' &= w_1\left[\mathrm{Tr}(A_1'V_1')\hat{\sigma}_\gamma^2 + \mathrm{Tr}(A_1'V_2')\hat{\sigma}_\beta^2 + \mathrm{Tr}(A_1'V_3')\hat{\sigma}_e^2\right] \\ &+ (1-w_1)\left[\mathrm{Tr}(A_1''V_1'')\hat{\sigma}_\gamma^2 + \mathrm{Tr}(A_1''V_2'')\hat{\sigma}_\beta^2 \right. \\ &\left. + \mathrm{Tr}(A_1''V_3'')\hat{\sigma}_e^2\right] \quad (16) \end{aligned}$$

By solving eq.(14), (15) and (16), the estimators of variance components are:

$$\hat{\sigma}_e^2 = \frac{Q_3'}{\mathrm{Tr}(A_3'V_3')}$$

$$\hat{\sigma}_\beta^2 = \frac{w_2Q_2' + (1-w_2)Q_2'' - [w_2\mathrm{Tr}(A_2'V_3') + (1-w_2)\mathrm{Tr}(A_2''V_3'')]\hat{\sigma}_e^2}{[w_2\mathrm{Tr}(A_2'V_2') + (1-w_2)\mathrm{Tr}(A_2''V_2'')]}$$

$$\dot{\sigma}_\gamma^2 = \frac{w_1Q_1' + (1-w_1)Q_1'' - P_2\hat{\sigma}_e^2 - P_1\hat{\sigma}_\beta^2}{[w_1\mathrm{Tr}(A_1'V_1') + (1-w_1)\mathrm{Tr}(A_1''V_1'')]}$$

where

$$P_1 = [w_1\mathrm{Tr}(A_1'V_2') + (1-w_1)\mathrm{Tr}(A_1''V_2'')],$$
$$P_2 = [w_1\mathrm{Tr}(A_1'V_3') + (1-w_1)\mathrm{Tr}(A_1''V_3'')]$$

SIMULATION STUDY OF TWO –WAY NESTED RANDOM MODEL

In this section, it is used R program to estimate the variance components in unbalanced two-way random nested model and to compare the estimators of modified methods by using Mean Square Error (MSE) and Relative Efficiency in case of completely missing information through a simulation study with samples of size $N = 60$, number of groups $S = 8$, number of subgroups $D_1 = 2, D_2 = 3, D_3 = 2, D_4 = 3, D_5 = 2, D_6 = 3, D_7 = 2, D_8 = 3$, the true values $\sigma_e^2 = 1, \sigma_\beta^2 = \sigma_\gamma^2 = 0.1, 0.5, 2$, initial values $\alpha_0 = 1, \alpha_1 = 0.1, 0.5, 2, \alpha_2 = 0.1, 0.5, 2$.

Percentage of missing information levels 25%, 50% and 75%, the weights $(w_1, w_2) = (0.5, 0.5), (0.78, 0.89), (0.5, 0.91)$, and sample of each subgroup $n_{11} = 3, n_{12} = 5, n_{21} = 3, n_{22} = 2, n_{23} = 2, n_{31} = 3, n_{32} = 4, n_{41} = 5, n_{42} = 4, n_{43} = 2, n_{51} = 3, n_{52} = 2, n_{61} = 2, n_{62} = 4, n_{63} = 2, n_{71} = 3, n_{72} = 2, n_{81} = 3, n_{82} = 2, n_{83} = 4$. The chain is run for 5000 iterations.

For σ_β^2, we compare MMIV(MIV(0)) and ANOVA by using Mean Square Error (MSE) and Relative Efficiency (MSE of MMIV(MIV(0))'s estimator/ MSE of ANOVA's estimator) which the estimator of σ_β^2 is the same for methods MMIV(ANOVA)and MMIV(MIV(0)).

For σ_γ^2, we compare ANOVA, MMIV(MIV(0)) and MMIV(ANOVA) by using Mean Square Error (MSE) and Relative Efficiency (MSE of any estimator/ MSE of MMIV(MIV(0))'s estimator).

Simulation Results:

In case of unbalanced two-way random nested model, the estimates of σ_e^2 are the same for methods (ANOVA, MMIV(MIV(0)) and MMIV(ANOVA)) in all cases.

In case of unbalanced two-way random nested model, the estimates of σ_β^2 are the same for methods (MMIV(ANOVA) and MMIV(MIV(0))) in all cases.

In 25% missing information, in most cases, MMIV(MIV(0))'s estimator of σ_β^2 is the best, it has minimum mean square error and relative efficiency is less than one.

In 25% missing information, in most cases, ANOVA's estimator of σ_γ^2 is the best, it has minimum mean square error and relative efficiency is less than MMIV(MIV(0)).

In 25% missing information, in most cases, modified MIV(ANOVA)'s estimator of σ_γ^2 is better than MMIV(MIV(0)), relative efficiency is greater than one.

In 50% missing information, in most cases, estimator of σ_γ^2 of ANOVA and MIV(ANOVA) methods are the best, two methods have minimum mean square error and relative efficiency is less than MMIV(MIV(0)).

In 75% missing information, in most cases, estimator of σ_γ^2 of ANOVA and MIV(ANOVA) methods are the best, two methods have minimum mean square error and relative efficiency is less than MMIV(MIV(0)).

When $\alpha_2 = 0.5 > \alpha_1 = 0.1$, It is better to estimate the variance components by using MMIV(MIV(0)). But $\alpha_2 = 0.5 < \alpha_1 = 0.9$, It is better to estimate the variance components by using ANOVA.

When the true values of σ_β^2 increases, It is better to estimate the variance components by using MMIV(MIV(0)).

In balanced two-way random nested model, MMIV(ANOVA) and ANOVA methods have negative estimates.

CONCLUSION

The aim of this thesis was to evaluate the performance of the proposed estimators relative to ANOVA's estimator via simulation studies. Two criteria such as mean square error and relative efficiency are used to show the performance of the estimators under the study.

From simulation study, we estimated the variance components by using MMIV(MIV(0)), MMIV(ANOVA) and ANOVA methods under normality assumption and compared the estimators for unbalanced two-way random nested model.

In unbalanced two-way random nested model, It is better to estimate variance component by MIV(ANOVA) and ANOVA methods. But two methods have negative estimates.

REFERENCES:

1. Rao C.R., Estimation of heteroscedastic variances in linear models, J. Amer. Stat. Assoc.,1970, 65, 112–115.
2. Rao C.R., Minimum Variance Quadratic Unbiased Estimation of Variance Component, Journal of Multivariate Analysis, 1971, 1, 445-456.
3. Sahai, H. and Ojeda, M.M., Analysis of Variance for Random Models, Unbalanced Data, Birkhauser, Boston, Basel, Berlin,2005.
4. Saleh, A.A. and El-Sheikh, A.A., Estimation of Variance Components for Three-Stage Unbalanced Nested Random Mixed Models with Incomplete Information, Proceeding of the 37th Annual Conference On Statistics, Computer Sciences and Operation Research,2002, 1-15, I.S.S.R, Cairo University, Egypt.
5. Saleh, A.A. and El-Sheikh, A.A., Estimation of Variance Components for Three-Stage Unbalanced Nested Random Mixed Models with Complete Missing Information (C.M.I), Proceeding of the 38th Annual Conference On Statistics, Computer Sciences and Operation Research,2003, 13-16, I.S.S.R, Cairo University, Egypt.
6. Song, R. and Shulman, S.A., Variance Components in the Two-Way Nested Model with Incomplete Nesting Information, Technometrics, 1997, 39(1), 71-80.

7. Subramani, J., On Modified Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components in Mixed Linear Models, model assisted statistics and applications, 2012,7, 179-200.