Research Journal of Mathematical and Statistical Sciences ISSN 2320-6047
Vol. **7(1),** 7-16, January **(2019)** *Res. J. Mathematical and Statistical Sci.*

Estimation of variance components in one-way random model with completely missing information

Amany Mousa Mohammed, Ahmed Amin El-Sheikh and Alaa Sayed Shehata*
Department of Applied Statistics and Econometrics, Institute of Statistical Studies and Research, Cairo University, Egypt
alaa_sayed_sh@yahoo.com

Available online at: www.isca.in, www.isca.me
Received 10th November 2018, revised 1st January 2019, accepted 10th January 2019

Abstract

In this paper, the estimators of variance components are derived of one-way random model when the problem of missing information exists using combination between Modified Minimum Variance Quadratic Unbiased Estimation (MMIVQUE) and Analysis of Variance (ANOVA) methods that is called MMIV (ANOVA) method.

Keywords: Variance Components, MMIVQUE, MIVQUEI, Missing Information.

Introduction

In regression analysis, independent variables may have missing values. It is likely that information in variance component model is missing. The information in variance component model has the same importance as the independent variables in regression analyses. With missing information, it may make some variance components inestimable.

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. Discarding data which have missing information leads losing useful information[1].

Song and Shulman[2] presented three types of missing information, completely missing information, partially missing information and not at all on any observation. i. Completely missing information: There is no subgroup information for any observation in the main group. ii. Partially missing information: There is some subgroup information for some observations in the main group. iii. There is no subgroup information for any observation is missing[2].

Song and Shulman[2] estimated the variance components when the problem of missing information exists of two-stage unbalanced nested random model. They compared these estimates by Variances and co variances using four procedures of weights.

Saleh and El Sheikh[3] modified analysis of variance method and the combined symmetric sums with analysis of variance method for estimation of variance components of three-stage unbalanced nested random models for data with completely missing information.

By a simulation study, they compared the bias and the mean squares errors of the estimates of variance components of 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.

This paper is presented explicit forms of estimators of variance components of one-way random model when the problem of completely missing information exists using MMIV (ANOVA) method.

The paper is organized as follows: The second section concerns with the Modified MIVQUE method introduced by Subramani[4]. The third section illustrates the proposed estimators of variance components for ANOVA and modified MIVQUE methods for data with complete information for one-way random model. Estimation of variance components for data with completely missing information is considered for one-way random model by the fourth section. The fifth section presents simulation study for one-way random model when part of information is missing.

Assume:

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

Where: Y is an $N \times 1$ vector of observations, N is the sample size. X is a $N \times s$ matrix with known numbers. β is a $s \times 1$-vector of fixed (unknown) parameters, Z_i is a $N \times c_i$ matrix with known constants, $i = 1, \ldots, d$. $(Z_d = I, c_d = N)$, δ_i is a $c_i \times 1$-vector of random variables. $(\delta_d = e)$, $\sigma_i^2 \geq 0, i = 1,2,..,d-1, \sigma_d^2 = \sigma_e^2 > 0$

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) shall be in a compact form as follows:

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

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 dispersion matrix and $\sigma_1^2, \ldots, \sigma_d^2$ are variance components[5].

Subramani's Modification

Rao[5] suggested a method of estimation "MINQUE" Minimum Norm Quadratic Unbiased Estimation that does not require the normality assumption for the estimation of variance components. He[6] proposed a method of estimation that called MIVQUE, Minimum Variance Quadratic Unbiased Estimation. Subramani[4] suggested a modification on MIVQUE method that introduced two modified MIVQUE (MIVQUE I and MIVQUE II) methods to estimate variance components of one-way random model.

Subramani[4] developed estimating variance components based on Rao[5] method. He estimated variance components linear combinations of variance components $\sum_{j=1}^{d} \rho_{ij}\sigma_i$ using quadratic functions $Y^t A_i Y$ [$A_i$is a symmetric matrixand$\rho_{ij} = Tr(A_i V_j)$]. He estimated the variance components under normality assumptions by solving the following:

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

He introduced different formulas of A_i. The formulas of $A_i (i = 1,2, \ldots, d)$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,

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

Where: $(U_i^t V^{-1} U_i)^{-}$ is the generalized inverse of $U_i^t V^{-1} U_i$

For the case (ii), he derived the estimators, their variances and covariance matrix of one-way random model. The resulting method are referred to as MIVQUE I.

The proposed estimators of variance components are derived through replacing A_i in eq. (3) by A_i for case (ii) in equation (4).

So the steps of MIVQUE method: i. Selecting a symmetric matrix A_i, ii. Solving the equation (3), iii. obtain the estimators of MIVQUE method.

Estimation of Variance Components in One-Way Random Model

Consider the model
$$Y_{ij} = \tau + a_i + e_{ij} \quad (5)$$
$i = 1,2, \ldots, r, j = 1,2, \ldots, n_i$
Where: Y_{ij} is the j^{th} observation in the i^{th} level. τ is an unknown parameter.

$a_i \sim N(0, \sigma_a^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$, The variance components are σ_a^2 and σ_e^2.

The model (5) can be in matrix form of unbalanced data as follows:
$$Y = X\tau + Z_1 a + Z_2 e \quad (6)$$

Y is an $N \times 1$ vector of observations, with $E(Y) = X\tau$ and $D(Y) = V = V_1\sigma_a^2 + V_2\sigma_e^2$,

$X = \underline{1}_N$,
$$Z_1 = \begin{bmatrix} \underline{1}_{n_1} & 0 & \ldots & 0 \\ 0 & \underline{1}_{n_2} & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & \underline{1}_{n_r} \end{bmatrix} = \sum_{i=1}^{+r} \underline{1}_{n_i}$$
$$Z_2 = \begin{bmatrix} \underline{I}_{n_1} & 0 & \ldots & 0 \\ 0 & \underline{I}_{n_2} & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & \underline{I}_{n_r} \end{bmatrix} = \sum_{i=1}^{+r} \underline{I}_{n_i}$$

$N = \sum_{i=1}^{r} n_i$, $V_i = Z_i Z_i^t$,$a^t = (a_1, a_2, \ldots, a_r)$
and $e^t = (e_{11}, e_{12}, \ldots, e_{1n_1}, \ldots, e_{r1}, e_{r2}, \ldots, e_{rn_r})$.
$\Sigma +$ is used to denote the direct sum.

Let $V^* = \alpha_1 V_1 + \alpha_0 V_2$ where α_1 and α_0 are a priori values for σ_a^2 and σ_e^2 respectively. So the inverse of $V^{*(-1)}$ is:
$$V^{*(-1)} = \frac{1}{\alpha_0}\left[I - \alpha_0 \sum_{i=1}^{+r} \frac{g_i}{n_i} J_{n_i}\right]$$
Where: $g_i = \frac{n_i}{\alpha_0 + n_i\alpha_1}$

$$V_1 = Z_1 Z_1^t = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & J_r \end{bmatrix} = \sum_{i=1}^{+r} J_{n_i}$$
$V_2 = I_N$
Where: $J_i$denote $n_i \times n_i$ matrix consisting of 1's [4].

In this section, the variance components will be estimated by the following methods: i. ANOVA method, ii. MMIVQUE I method.

Analysis of Variance (ANOVA): Lemma (1): The estimators for σ_a^2, σ_e^2 in one-way random model are:

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r} \frac{Y_{i.}^2}{n_i}}{\sum_{i=1}^{r} n_i - r}$$

$$\hat{\sigma}_a^2 = \frac{\sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i} - \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=1}^{r} n_i} - (r-1)\hat{\sigma}_e^2}{(r-1)c} \quad (7)$$

Where $c = \frac{\left(\sum_{i=1}^{r} n_i\right)^2 - \sum_{i=1}^{r} n_i^2}{(r-1)\sum_{i=1}^{r} n_i}$

Proof: According to steps of ANOVA method, the sums of squares equal to their expectations. For model (6), the equations of sum of squares take the following form:

$$SSB = E(SSB) \quad (8)$$

$$SSE = E(SSE) \quad (9)$$

Where, $SSB = \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i} - \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=1}^{r} n_i}$

$$E(SSB) = (r-1)(\sigma_e^2 + c\sigma_a^2)$$

$$SSE = \sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i}$$

$$E(SSE) = \left(\sum_{i=1}^{r} n_i - r\right)\sigma_e^2$$

By solving the eq. (8) and eq. (9), the estimators for σ_a^2, σ_e^2 are

$$\hat{\sigma}_e^2 = \frac{SSE}{\sum_{i=1}^{r'} n_i - r} = \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i}\right]}{\sum_{i=1}^{r} n_i - r}$$

$$\hat{\sigma}_a^2 = \frac{SSB - (r-1)\hat{\sigma}_e^2}{(r-1)c} = \frac{\sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i} - \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=1}^{r} n_i} - (r-1)\hat{\sigma}_e^2}{(r-1)c}$$

Modified Minimum Variance Quadratic Unbiased Estimation: Lemma (2): for model (6), the estimators for σ_a^2, σ_e^2 in unbalanced one-way random model are:

$$\hat{\sigma}_e^2 = \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i}\right]}{N-r}$$

$$\hat{\sigma}_a^2 = \frac{\left[Q_1 - Tr(A_1V_2)\hat{\sigma}_e^2\right]}{Tr(A_1V_1)} \quad (10)$$

Where,

$$Q_1 = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \alpha_1\sum_{i=1}^{r}\frac{g_i}{n_i}Y_{i.}^2\right] - g\left[\sum_{i=1}^{r}\frac{g_i}{n_i}Y_{i.}\right]^2$$

$$Tr(A_1V_2) = \frac{1}{\alpha_0}\left[N - \alpha_1\sum_{i=1}^{r} g_i\right] - g\sum_{i=1}^{r}\frac{g_i^2}{n_i}$$

$$Tr(A_1V_1) = \sum_{i=1}^{r} g_i - g\sum_{i=1}^{r} g_i^2$$

$$g_i = \frac{n_i}{\alpha_0 + n_i\alpha_1},\ g = \frac{1}{\sum_{i=1}^{r} g_i}$$

Proof: The MIVQUE I estimators are derived by selecting A_i for case (ii) in equation (4).

For model (6), A_i, i = 1, 2 takes the following form:

$$A_1 = V^{*(-1)} - V^{*(-1)}U_1\left(U_1^tV^{*(-1)}U_1\right)^{-}U_1^tV^{*(-1)}$$

$$A_2 = V^{*(-1)} - V^{*(-1)}U_2\left(U_2^tV^{*(-1)}U_2\right)^{-}U_2^tV^{*(-1)}$$

Where:

$U_1 = X, U_2 = [X \quad Z_1]$

So

$$A_1 = \frac{1}{\alpha_0}\left[I - \alpha_1\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i}\right] - \frac{g}{\alpha_0^2}\ G$$

$$A_2 = \frac{1}{\alpha_0}\left[I - \alpha_1\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i}\right] - \sum_{i=1}^{+r}\frac{g_i}{n_i^2}J_{n_i}$$

Where:

$$G_U = \left[XX^t - \alpha_1\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i}XX^t - \alpha_1XX^t\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i} + \alpha_1^2\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i}XX^t\sum_{i=1}^{+r}\frac{g_i}{n_i}J_{n_i}\right]$$

So the equations are

$$\begin{bmatrix} Tr(A_1V_1) & Tr(A_1V_2) \\ Tr(A_2V_1) & Tr(A_2V_2) \end{bmatrix}_{2\times 2}\begin{bmatrix}\sigma_a^2 \\ \sigma_e^2\end{bmatrix}_{2\times 1} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}_{2\times 1}$$

Where: $Q_i = Tr(A_iW), W = YY^t, i = 1,2$

$$Q_1 = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \alpha_1\sum_{i=1}^{r}\frac{g_i}{n_i}Y_{i.}^2\right] - \frac{g}{\alpha_0^2}\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij} - \alpha_1\sum_{i=1}^{r} g_i Y_{i.}\right]^2$$

$$Q_2 = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i}\right]$$

Where: $g = \frac{1}{\sum_{i=1}^{r} g_i}$

and $Tr(A_1V_1) = \sum_{i=1}^{r} g_i - g\sum_{i=1}^{r} g_i^2$

$$Tr(A_1V_2) = \frac{1}{\alpha_0}\left[N - \alpha_1\sum_{i=1}^{r} g_i\right] - g\sum_{i=1}^{r}\frac{g_i^2}{n_i}$$

$$Tr(A_2V_1) = 0$$

$$Tr(A_2V_2) = \frac{N-r}{\alpha_0}$$

Since

$$Tr(A_1V_1)\sigma_a^2 + Tr(A_1V_2)\sigma_e^2 = Q_{1U} \quad (11)$$

$$Tr(A_2V_2)\sigma_e^2 = Q_{2U} \quad (12)$$

By solving the equation (11) and equation (12), the MIVQUE I estimators for σ_a^2, σ_e^2 are

$$\acute{\sigma}_e^2 = \frac{Q_2}{Tr(A_2V_2)} = \frac{\left[\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r}\frac{Y_{i.}^2}{n_i}\right]}{N-r}$$

$$\acute{\sigma}_a^2 = \frac{[Q_1 - Tr(A_1V_2)\acute{\sigma}_e^2]}{Tr(A_1V_1)}$$

Estimation of variance components using data with missing information

In this section, the variance components will be estimated in case of data with missing information by the following methods: i. ANOVA method, ii. MMIV (ANOVA) method: combined MMIVQUE I method with ANOVA method.

Assume that the total number of the main group will be: $r = r' + r''$.

Where: r' be the numbers of treatments for data with complete information and r'' be the numbers of treatments for data with completely missing information.

Variables and coefficients without prime-notation or with prime or double primes symbols will be defined as: i. If there is a symbol without prime then we do not specify the range for i if the variable or coefficient is summed over i. ii. The same symbol with a prime (double primes) is then defined as the same symbol summed over i from, 1 to r' (from $r'+1, \dots, r$, respectively).

Steps of estimation: i. Estimation of variance components for r' treatments (data with complete information), ii. Estimation of variance components for r'' treatments. (data with completely missing information), iii. Prespecified weights will be used to combine data with complete information and completely missing information.

Analysis of Variance method (ANOVA): In this section, the variance components will be estimated for data with complete information and data with completely missing information by ANOVA method. Prespecified weights will be used to combine data with complete information and missing information.

Lemma (3): The estimators of variance components with completely missing information are:

$$\dot{\hat{\sigma}}_e^2 = \frac{\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'} \frac{Y_{i.}^2}{n_i}\right]}{\sum_{i=1}^{r'} n_i - r'}$$

$$\dot{\hat{\sigma}}_a^2 = \frac{wSSB' + (1-w)SSBE'' - s_2\dot{\hat{\sigma}}_e^2}{s_1} \quad (13)$$

$0 \le w \le 1$ (the weight)
$s_1 = w(r' - 1)c' + (1-w)(r'' - 1)c''$
$s_2 = w(r' - 1) + (1-w)(\sum_{i=r'+1}^{r} n_i - 1)$

Proof: Part (1): For data with complete information: According to steps of ANOVA method, the sums of squares equal to their expectations.

For model (6), the equations of sum of squares take the following form:

$$SSB' = E(SSB') \quad (14)$$

$$SSE' = E(SSE') \quad (15)$$

Where,

$$SSB' = \left[\sum_{i=1}^{r'} \frac{Y_{i.}^2}{n_i}\right] - \frac{\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=1}^{r'} n_i}$$

$$E(SSB') = (r' - 1)(\sigma_e^2 + N_c\sigma_a^2)$$

$$N_c = \frac{\left(\sum_{i=1}^{r'} n_i\right)^2 - \sum_{i=r'+1}^{r} n_i^2}{(r'-1)\sum_{i=1}^{r'} n_i}$$

$$SSE' = \sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'} \frac{Y_{i.}^2}{n_i}$$

$$E(SSE') = \left(\sum_{i=1}^{r'} n_i - r'\right)\sigma_e^2$$

Part (2): For data with missing information: For model (6), the sum of squares of the data with completely missing information (where $i = r' + 1, \dots, r$) will take the following form:

$$SSB'' = E(SSB'') \quad (16)$$

Where

$$SSB'' = \left[\sum_{i=r'+1}^{r} \frac{Y_{i.}^2}{n_i}\right] - \frac{\left[\sum_{i=r'+1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=r'+1}^{r} n_i}$$

Then, for data with completely missing information $\sum_{i=r'+1}^{r} Y_{i.}^2$ can not to be computed directly in SSB''. To solve this problem, define $SSBE''$ as a linear combination of SSB'' and SSE''as follows:

$$SSBE'' = SSB'' + SSE''$$

$$= \left[\sum_{i=r'+1}^{r} \frac{Y_{i.}^2}{n_i} - \frac{\left[\sum_{i=r'+1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=r'+1}^{r} n_i}\right] + \left[\sum_{i=r'+1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=r'+1}^{r} \frac{Y_{i.}^2}{n_i}\right]$$

Then,

$$SSBE'' = \sum_{i=r'+1}^{r}\sum_{j=1}^{n_i} Y_{ij}^2 - \frac{\left[\sum_{i=r'+1}^{r}\sum_{j=1}^{n_i} Y_{ij}\right]^2}{\sum_{i=r'+1}^{r} n_i} \quad (17)$$

According to equation (17):

$$E(SSBE'') = (r'' - 1)(\sigma_e^2 + c''\sigma_a^2) + \left(\sum_{i=r'+1}^{r} n_i - r''\right)\sigma_e^2$$
$$= \left(\sum_{i=r'+1}^{r} n_i - 1\right)\sigma_e^2 + c''(r'' - 1)\sigma_a^2 \quad (18)$$

By equating the sums of squares in equation (17) to their expectations in equation (18), then

$$SSBE'' = E(SSBE'') \quad (19)$$

Part (3): combination data with complete information and data with completely missing information: In order to obtain the estimators of variance components (σ_a^2 and σ_e^2), the equations (14, 15, and 19) will be linearly combined into two equations. Pre-specified weights will be used to combine (SSB' and $SSBE''$) and then solve the combined equations as:

$$SSE' = \left(\sum_{i=1}^{r'} n_i - r'\right)\dot{\tilde{\sigma}}_e^2 \quad (20)$$

$$w_1 SSB' + (1-w_1)SSBE'' = s_1\dot{\tilde{\sigma}}_a^2 + s_2\dot{\tilde{\sigma}}_e^2 \quad (21)$$

By solving the equations (20) and (21), the estimators of variance components are as follows:

$$\dot{\tilde{\sigma}}_e^2 = \frac{\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'}\frac{Y_{i.}^2}{n_i}\right]}{\sum_{i=1}^{r'} n_i - r'}$$

$$\dot{\tilde{\sigma}}_a^2 = \frac{wSSB' + (1-w)SSBE'' - s_2\dot{\tilde{\sigma}}_e^2}{s_1}$$

Combined Modified Minimum Variance Quadratic Estimation with ANOVA method (MMIV(ANOVA)): In this section, the variance components will be estimated in case of data with complete information by ANOVA method and data with completely missing information by MMIV(ANOVA) method. Prespecified weights will be used to combine data with complete information and missing information.

Lemma (4): The estimators of variance components of one-way random model when the problem of missing information exists take the form:

$$\dot{\tilde{\sigma}}_e^2 = \frac{\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'}\frac{Y_{i.}^2}{n_i}\right]}{\sum_{i=1}^{r'} n_i - r'}$$

$$\dot{\tilde{\sigma}}_a^2 = \frac{wQ_1' + (1-w)SSBE'' - \tilde{s}_2\dot{\tilde{\sigma}}_e^2}{\tilde{s}_1} \quad (22)$$

Proof: Part (1): For data with complete information: For model (6), the matrixA_i', $i = 1,2$ for data with complete information is:

$$\mathrm{A}_1' = \frac{1}{\alpha_0}\left[I' - \alpha_1\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i}\right] - \frac{g'}{\alpha_0^2}G_U'$$

$$\mathrm{A}_2' = \frac{1}{\alpha_0}\left[I' - \alpha_1\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i}\right] - \sum_{i=1}^{+r'}\frac{g_i}{n_i^2}J_{n_i}$$

Where;

$$G_U' = \Big[X'X'^t - \alpha_1\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i}X'X'^t - \alpha_1 X'X'^t\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i} + \alpha_1^2\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i}X'X'^t\sum_{i=1}^{+r'}\frac{g_i}{n_i}J_{n_i}\Big]$$

X' denote $\sum_{i=1}^{r'} n_i \times 1$ vector consisting of 1's.

The normal equations are:

$$\begin{bmatrix} Tr(A_1'V_1') & Tr(A_1'V_2') \\ Tr(A_2'V_1') & Tr(A_2'V_2') \end{bmatrix}_{2\times 2}\begin{bmatrix}\sigma_a^2 \\ \sigma_e^2\end{bmatrix}_{2\times 1} = \begin{bmatrix} Q_1' \\ Q_2' \end{bmatrix}_{2\times 1}$$

where

$$Tr(\mathrm{A}_1'\mathrm{V}_1') = \sum_{i=1}^{r'} g_i - g'\sum_{i=1}^{r'} g_i^2$$

$$Tr(\mathrm{A}_1'\mathrm{V}_2') = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r'} n_i - \alpha_1\sum_{i=1}^{r'} g_i\right] - g'\sum_{i=1}^{r'}\frac{g_i^2}{n_i}$$

$$Tr(\mathrm{A}_2'\mathrm{V}_1') = 0$$

$$Tr(\mathrm{A}_2'\mathrm{V}_2') = \frac{\sum_{i=1}^{r'} n_i - r'}{\alpha_0}$$

where $g' = \frac{1}{\sum_{i=1}^{r'} g_i}$

$$\mathrm{V}_1' = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & J_{k'} \end{bmatrix}$$

$$\mathrm{V}_2' = I'_{\sum_{i=1}^{r'} n_i}$$

Where V_1'denote $\sum_{i=1}^{r'} n_i \times \sum_{i=1}^{r'} n_i$ matrix

I'denote $\sum_{i=1}^{r'} n_i \times \sum_{i=1}^{r'} n_i$ identity matrix.

$$Q_1' = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \alpha_1\sum_{i=1}^{r'}\frac{g_i}{n_i}Y_{i.}^2\right] - g'\left[\sum_{i=1}^{r'}\frac{g_i}{n_i}Y_{i.}\right]^2$$

$$Q_2' = \frac{1}{\alpha_0}\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'}\frac{Y_{i.}^2}{n_i}\right]$$

Since

$$Tr(A_1'V_1')\sigma_a^2 + Tr(A_1'V_2')\sigma_e^2 = Q_1' \quad (23)$$

$$Tr(A_2'V_2')\sigma_e^2 = Q_2' \quad (24)$$

Part (2): For data with completely missing information
This part is the same part in lemma (3). So the equation (19) for data with completely missing information (where $\mathrm{i} = \mathrm{r}' + 1, \dots, \mathrm{r}$).

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

In order to estimate σ_a^2 and σ_e^2, the equations (19, 23, and 24) will be linearly combined into two equations. Pre-specified weights will be used to combine (Q_1' and $SSBE''$) and then solve the combined equations as:

$$Q_2' = Tr(A_2'V_2')\dot{\tilde{\sigma}}_e^2 \quad (25)$$

$$wQ_1' + (1-w)SSBE'' = \tilde{s}_1\dot{\tilde{\sigma}}_a^2 + \tilde{s}_2\dot{\tilde{\sigma}}_e^2 \quad (26)$$

Where,

$$\tilde{s}_1 = [w(A_1', V_1') + (1-w)N_c(r'' - 1)]$$

$\tilde{s}_2 = \left[w(M'_{1U}, V'_{2U}) + (1-w)\left(\sum_{i=r'+1}^{r} n_i - 1\right)\right]$

By solving the equations (25) and (26), the estimators of variance components are as follows:

$$\dot{\tilde{\sigma}}_e^2 = \frac{\left[\sum_{i=1}^{r'}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{r'}\frac{Y_{i.}^2}{n_i}\right]}{\sum_{i=1}^{r'} n_i - r'}$$

$$\dot{\tilde{\sigma}}_a^2 = \frac{wQ'_{1U} + (1-w)SSBE'' - \tilde{s}_2\dot{\tilde{\sigma}}_e^2}{\tilde{s}_1}$$

Simulation study

In this section, the variance components are estimated for unbalanced one-way random model under normality assumption for data with completely missing information via simulation study by MMIV(ANOVA) and ANOVA methods and to compare the estimates using mean squared error, absolute bias, probability of getting negative estimates and relative efficiency.

A simulation study for unbalanced one-way random model requires a n-pattern, true values for the variance components σ_a^2 and σ_e^2 and initial values α_1 and α_0 for the variance components σ_a^2 and σ_e^2 respectively.

As stated by Swallow and Searle[7], the true values have been chosen such that the ratio of true values is less than or equal or larger than initial values as follows:

$$(1)\frac{\sigma_a^2}{\sigma_e^2} < \frac{\alpha_1}{\alpha_0}, \qquad (2)\frac{\sigma_a^2}{\sigma_e^2} = \frac{\alpha_1}{\alpha_0}, \quad (3)\frac{\sigma_a^2}{\sigma_e^2} > \frac{\alpha_1}{\alpha_0}$$

So it is assumed that the true values $\sigma_e^2 = 1, \sigma_a^2 = 0.1, 0.5, 1, 2, 10$ and initial values $\alpha_0 = 1, \alpha_1 = 1$.

By simulation study, 5000 independent random sample are generated, it is assumed that the sample size is 48, n-pattern $(n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8)$ is (3,3,5,5,7,7,9,9), number of levels $r = 8$, percentage of missing information (25%, 50% ,75%), and the weights $w_1 = 0.1, 0.3, 0.5, 0.7, 0.9$.

According to simulation study for unbalanced one-way random model, a number of conclusions are drawn from the results for all tables of this model which are summarized in the following points: i. In case of unbalanced one-way random model, it does not matter computing the estimates of σ_e^2 for ANOVA and MMIV(ANOVA) methods because they are the same.

25% data with completely missing information: i. The estimates of variance component σ_a^2 for MMIV (ANOVA) method is better than ANOVA method, it has lower mean squared error. Also, the absolute bias of the estimates of variance components σ_a^2 is lower than ANOVA method. ii. There are negligible differences among ANOVA and MMIV (ANOVA) methods with respect to the probability of getting negative estimates. iii. For MMIV (ANOVA) method, mean squared error decreases at high level of weight. iv. For weight = 0.1, mean square error for ANOVA and MMIV (ANOVA) methods are the same and is the largest.

50% data with completely missing information: i. The estimates of variance component σ_a^2 for MMIV (ANOVA) method is better than ANOVA method, it has lower mean squared error. Also, the absolute bias of the estimates of variance components σ_a^2 is lower than ANOVA method. ii. There are negligible differences among ANOVA and MMIV (ANOVA) methods with respect to the probability of getting negative estimates. iii. For MMIV (ANOVA) method, mean squared error decreases at high level of weight. iv. For weight = 0.1, mean squared error for ANOVA and MMIV (ANOVA) methods approaches. Also, absolute bias for ANOVA and MMIV (ANOVA) methods is the same. v. For weight = 0.9, the estimates of variance component σ_a^2 for ANOVA method have larger mean squared error.

75% data with completely missing information: i. The estimates of variance component σ_a^2 for MMIV (ANOVA) method is better than ANOVA method, it has lower mean squared error. Also, the absolute bias of the estimates of variance components σ_a^2 is lower than ANOVA method. ii. There are negligible differences among ANOVA and MMIV (ANOVA) methods with respect to the probability of getting negative estimates. iii. For weight = 0.9, the estimates of MMIV (ANOVA) method have smaller mean squared error. iv. Mean squared error of estimates for ANOVA and MMIV (ANOVA) methods approaches at high percentage for data with completely missing information.

Conclusion

In this paper, explicit forms of the estimators are presented for estimating the variance components according to Analysis of Variance (ANOVA) and Modified Minimum Variance Quadratic Unbiased Estimation (MMIVQUE) methods of one-way random model in case of data with complete information and data with completely missing information. The relative performance of the proposed estimators is evaluated via simulation study based on different criteria such as mean squared error, absolute bias, the probability of getting negative estimates. So a number of conclusions are drawn from the results which are summarized in the following points:

The estimates of variance component σ_a^2 for MMIV (ANOVA) method is better than ANOVA method, it has lower mean squared error. Also, the absolute bias of the estimates of variance components σ_a^2 is lower than ANOVA method.

There are negligible differences among ANOVA and MMIV (ANOVA) methods with respect to the probability of getting negative estimates. Mean squared error of estimates for ANOVA and MMIV (ANOVA) methods approaches at high level of completely missing information.

Table-1: Comparison of ANOVA and MMIV (ANOVA) estimates for unbalanced one-way random model-25% data with completely missing information based on MSE, Absolute Bias and Probability of getting negative estimates.

$\frac{\sigma_a^2}{\sigma_e^2}$	w_1	ANOVA			MMIV(ANOVA)		
		MSE	Absolute Bias	Prob. negative estimates	MSE	Absolute Bias	Prob. negative estimates
0.1	0.1	0.81	0.71	0.51	0.8	0.71	0.5
	0.3	0.64	0.63	0.5	0.62	0.62	0.51
	0.5	0.54	0.58	0.5	0.45	0.53	0.51
	0.7	0.44	0.53	0.51	0.25	0.39	0.5
	0.9	0.51	0.56	0.54	0.08	0.22	0.51
0.5	0.1	1.68	1.03	0.51	1.68	1.03	0.51
	0.3	1.48	0.97	0.51	1.43	0.96	0.52
	0.5	1.24	0.89	0.5	1.06	0.83	0.51
	0.7	1.04	0.82	0.5	0.66	0.66	0.5
	0.9	1.16	0.88	0.55	0.38	0.54	0.51
1	0.1	3.69	1.55	0.51	3.68	1.56	0.51
	0.3	3.1	1.43	0.52	3	1.41	0.52
	0.5	2.66	1.32	0.5	2.38	1.26	0.5
	0.7	2.43	1.28	0.5	1.75	1.11	0.5
	0.9	2.6	1.35	0.54	1.22	1	0.5
2	0.1	9.73	2.55	0.5	9.71	2.55	0.5
	0.3	8.9	2.46	0.51	8.7	2.44	0.51
	0.5	7.9	2.34	0.5	7.3	2.27	0.51
	0.7	7.23	2.26	0.5	5.68	2.07	0.5
	0.9	7.75	2.39	0.55	4.54	2	0.51
10	0.1	177.7	11.14	0.5	177.4	11.15	0.5
	0.3	165.1	10.91	0.5	163.05	10.9	0.5
	0.5	158.1	10.8	0.52	148.26	10.65	0.53
	0.7	145.2	10.55	0.51	123.95	10.14	0.5
	0.9	148.1	10.77	0.54	107.09	10.02	0.52

Table-2: Comparison of ANOVA and MMIV(ANOVA) estimates for unbalanced one-way random model-50% data with completely missing information based on MSE, Absolute Bias and Probability of getting negative estimates.

$\frac{\sigma_a^2}{\sigma_e^2}$	w_1	ANOVA			MMIV(ANOVA)		
		MSE	Absolute Bias	Prob. negative estimates	MSE	Absolute Bias	Prob. negative estimates
0.1	0.1	0.5	0.55	0.48	0.49	0.55	0.48
	0.3	0.46	0.53	0.47	0.45	0.52	0.47
	0.5	0.44	0.51	0.46	0.4	0.49	0.47
	0.7	0.44	0.51	0.47	0.34	0.45	0.47
	0.9	0.52	0.55	0.51	0.19	0.34	0.47
0.5	0.1	1.07	0.8	0.47	1.07	0.8	0.47
	0.3	1.08	0.81	0.47	1.06	0.8	0.48
	0.5	1.05	0.79	0.47	0.98	0.77	0.47
	0.7	1.1	0.8	0.48	0.92	0.74	0.48
	0.9	1.22	0.87	0.51	0.59	0.6	0.47
1	0.1	2.47	1.23	0.46	2.46	1.23	0.46
	0.3	2.59	1.26	0.47	2.54	1.25	0.48
	0.5	2.46	1.22	0.47	2.34	1.2	0.47
	0.7	2.39	1.22	0.47	2.07	1.14	0.47
	0.9	2.65	1.32	0.5	1.54	1.03	0.46
2	0.1	7.69	2.23	0.48	7.66	2.23	0.48
	0.3	7.33	2.19	0.47	7.24	2.18	0.47
	0.5	7.02	2.13	0.46	6.76	2.11	0.46
	0.7	7	2.16	0.47	6.34	2.09	0.47
	0.9	7.89	2.36	0.51	5.38	2.04	0.48
10	0.1	148.8	10.4	0.48	148.5	10.4	0.48
	0.3	145.11	10.27	0.48	143.7	10.25	0.48
	0.5	145.2	10.3	0.48	142.06	10.28	0.48
	0.7	139.61	10.11	0.45	130.66	9.99	0.45
	0.9	154.58	10.81	0.52	119.39	10.08	0.49

Table-3: Comparison of ANOVA and MMIV(ANOVA)estimates for unbalanced one-way random model-75% data with completely missing information based on MSE, Absolute Bias and Probability of getting negative estimates.

$\frac{\sigma_a^2}{\sigma_e^2}$	w_1	ANOVA			MMIV(ANOVA)		
		MSE	Absolute Bias	Prob. negative estimates	MSE	Absolute Bias	Prob. negative estimates
0.1	0.1	0.95	0.73	0.43	0.95	0.73	0.43
	0.3	0.91	0.71	0.4	0.91	0.71	0.4
	0.5	0.95	0.73	0.43	0.93	0.72	0.43
	0.7	0.96	0.73	0.42	0.91	0.71	0.42
	0.9	0.98	0.73	0.43	0.75	0.64	0.42
0.5	0.1	2.08	1.03	0.42	2.08	1.03	0.42
	0.3	1.93	0.99	0.41	1.92	0.99	0.41
	0.5	1.92	1	0.41	1.89	1	0.41
	0.7	1.97	1	0.42	1.87	0.98	0.43
	0.9	2.06	1.03	0.44	1.63	0.91	0.42
1	0.1	4.11	1.43	0.41	4.11	1.43	0.41
	0.3	4.07	1.43	0.41	4.05	1.43	0.41
	0.5	4.08	1.44	0.42	4	1.43	0.42
	0.7	3.91	1.41	0.42	3.76	1.39	0.42
	0.9	4.53	1.52	0.45	3.74	1.37	0.43
2	0.1	11.14	2.37	0.41	11.13	2.37	0.41
	0.3	11.37	2.42	0.42	11.31	2.42	0.43
	0.5	10.16	2.31	0.42	10.03	2.3	0.42
	0.7	11.08	2.4	0.42	10.69	2.36	0.42
	0.9	11.42	2.46	0.43	9.64	2.25	0.42
10	0.1	189.72	10.41	0.42	189.56	10.41	0.42
	0.3	194.85	10.62	0.42	194.1	10.61	0.42
	0.5	185.94	10.36	0.42	184.11	10.34	0.42
	0.7	188.73	10.49	0.42	184.1	10.43	0.42
	0.9	184.02	10.53	0.42	163.03	10.05	0.42

References

1. Saleh A.A. and El-Sheikh A.A. (2002). 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 Researche. I.S.S.R. Cairo University, Egypt, 1-15.
2. Song R. and Shulman S.A. (1997). Variance components in the two-way nested model with incomplete nesting information. *Technometrics*, 39(1), 71-80.
3. Saleh A.A. and El-Sheikh A.A. (2003). 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. I.S.S.R. Cairo University, Egypt, 13-16.
4. Subramani J. (2012). On Modified Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components in Mixed Linear Models. *Model assisted statistics and applications*, 7(3), 179-200.
5. Rao C.R. (1970). Estimation of heteroscedastic variances in linear models. *Journal of the American Statistical Association*, 65(329), 161-172.
6. Rao C.R. (1971). Minimum Variance Quadratic Unbiased Estimation of Variance Components. *Journal of Multivariate Analysis*, 1(4), 445-456.
7. Swallow W.H. and Searle S.R. (1978). Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components. *Technometrics*, 20(3), 265-272.