Songklanakarin J. Sci. Technol.
43 (1), 57-71, Jan. - Feb. 2021

http://www.sjst.psu.ac.th

Original Article

Expected mean squares for model effects in the two-way ANOVA model when sampling from finite populations

Chawanee Suphirat[1], Boonorm Chomtee[2*], and John J. Borkowski[3]

[1] *Department of Statistics, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900 Thailand*

[2] *Department of Statistics, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand*

[3] *Department of Mathematical Science, Montana State University, Bozeman, Montana, 59717 United States of America*

Received: 11 July 2019; Revised: 30 September 2019; Accepted: 7 November 2019

Abstract

The expected mean squares (EMS) for the effects in a two-way ANOVA model are derived when sampling from a finite population, for factors A and B. The A-effects and B-effects are represented by τ_i and β_j, respectively, in the model. Model effects are studied for the random and mixed effects models. Thus, in terms of hypothesis testing, we are interested in the expected mean square formulas when the random effects are sampled from finite populations. For balanced data, the EMS for the A, B and AB interaction effects when the effects are sampled from a finite population are the same as the EMS for an infinite population. For unbalanced data, the EMS when the effects are sampled from a finite population in factors A, B and the AB interaction, the EMS are not the same as for an infinite population because the values that multiply the variance components differ.

Keywords: expected mean square, factorial design, finite population, mixed effect model, variance components

1. Introduction

Analysis of variance (ANOVA) is a statistical technique used in many fields of research, with applications in industry, economics, education, natural and biological sciences, and agriculture. A frequently-used experimental design is the factorial design, which can allow investigating multiple factors that may influence the response variable (y). The data are often analyzed with a multifactor ANOVA model, which can be associated with a fixed, random or mixed effects model (Montgomery, 2013). In the design of experiments, we must consider the suitability of the data related to the proposed statistical analysis and the model effects of interest to ensure appropriate statistical inferences.

For example, suppose a company has 50 machines that make cardboard cartons for canned goods, and they want to understand the variation in strength of the cartons. They choose 10 machines at random from the 50 machines. In addition to variation due to the machines, variation in operators may also influence the strength of the cartons. Thus, the manufacturer also chooses 10 operators at random. Each operator will produce 4 cartons for each machine, with the cardboard feedstock assigned at random to the machine-operator combinations. We now have a two-way factorial treatment structure with both factors being random and completely randomized assignment of treatments to units (Oehlert, 2010).

*Corresponding author
Email address: fsciboc@ku.ac.th

In general, it is assumed that the levels of the factor (e.g., machines and machine operators) are randomly selected from an infinite population. This assumption is violated in this example because the factor levels are selected from finite populations. When the population size is large relative to the sample size for each factor, it may be reasonable to assume the population size infinite (Searle, Casella & McCulloch, 2009). In many cases, however, the population size is not large relative to the sample sizes, and then the finite size can affect the estimation of variance components of random effects.

In introductory statistics courses, the variance of the sample mean is given as $Var(\hat{y}) = \frac{\sigma^2}{n}$ which is appropriate if the random sample was taken from an infinite or extremely large population. To deal with a finite population, we have to adjust variance formulas by $\frac{N-n}{N}$, which is called "the finite population correction (FPC)". In this finite population case, $Var(\hat{y}_{SRS}) = \left(\frac{N-n}{N}\right)\frac{\sigma^2}{n} = \left(1-\frac{n}{N}\right)\frac{\sigma^2}{n}$ will be the variance of the sample mean. The ratio $\frac{n}{N}$ is called " the sampling fraction of the population". If $\frac{n}{N}$ is small, the FPC is close to 1, then the population size has negligible effect on the estimated variance of the sample mean. In practice, it is recommended that the FPC can be ignored when the sampling fraction does not exceed 5% or 10% (Cochran, 1977).

Many researchers were interested in finite population effects for variance component estimation, including Tukey who in 1956 studied the variances of variance component estimation for balanced data under the assumptions of independence and normality. However, if we sampled from finite populations, the finite population correction would be related to the estimation of variances. Next, Cornfield and Tukey (1956) considered the expected values of mean squares in experiments with balanced data. They used a model of sufficient generality and flexibility to define the formulas for crossed and nested classifications. Moreover, for unbalanced data, Tukey (1957) discussed the variance components for one-way classification, while Searle and Henderson (1961) dealt with the two-way classification model with one fixed factor. Subsequently, Hartley (1967) developed a general procedure for directly yielding the numerical values of the coefficients in the formulas of expected mean squares (EMS) with random and mixed models for one-way and two-way classifications with unequal numbers when sampling from an infinite population. It was useful to obtain mathematical formulas for the numerical coefficients used to produce the variance and covariance formulas for expected mean squares. After that, Searle and Fawcett (1970) studied the EMS in variance component models with random effects which are assumed to be sampled from finite populations. They developed a rule for converting expectations under infinite population models to finite population models. In addition, it can be applied to balanced and unbalanced data, and used for nested and cross classifications when it is assumed the set levels for each factor is finite. Accordingly, Simmachan, Borkowski, and Budsaba (2012) determined the EMS of treatments and error for random effects in only the one-way ANOVA model assuming a finite population, but with normal errors.

In this research, we focus on the random effects in the two-way ANOVA model, in other words, the two-factor factorial model. We consider the case where the population of the model effects are sampled from a finite population. We apply the guidelines given in Searle and Fawcett (1970) who assumed the model error is also sampled from a finite population, to models with random effects but adjust the random error to be normally distributed, which will affect expected mean squares.

Finally, we derive the expected mean square formulas for the two-factor factorial with random effects and mixed effects models when the random effects are sampled from a finite population. In this article, Section 2 contains the research methodology. In Section 3, the results of this research are presented. Conclusions are summarized in Section 4.

2. Materials and Methods

In this section, we describe the methodology for finding the expected mean squares in the model effects of the two-factor factorial design.

2.1 The model effects of the two-factor factorial design

A factorial design can be very efficient for studying the effects of two or more factors. In this research only, the two-factor factorial design is studied. If we assume that factors A and B have a large number of population levels such that the number of levels for each factor is assumed to be infinite, then the two-factor factorial model is:

$$y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}\,, \tag{1}$$

where μ is the mean, τ_i and β_j are the i^{th} and j^{th} level effects of factors A and B with $i = 1, 2, ..., a$ and $j = 1, 2, ..., b$, $(\tau\beta)_{ij}$ is the interaction effect of the (i, j) combination, and random error $\varepsilon_{ijk} \backsim N(0, \sigma^2)$. Whether the τ_i and β_j effects in the model are fixed or random effects depends on the research problem.

Consider the case when the numbers of randomly selected factor levels $(a \text{ and } b)$ are small relative to the population sizes $(N_a \text{ and } N_b)$. That is, the sampling fractions $\frac{a}{N_a}$ and $\frac{b}{N_b}$ are small and are related to the finite population corrections (FPCs) $\frac{N_a - a}{N_a}$ and $\frac{N_b - b}{N_b}$. It is assumed that sampling from the finite populations is done without replacement.

For factors A and B, let G_τ and G_β be the finite distributions of the τ_i and β_j effects from populations of effect of sizes N_a and N_b , respectively.

Case 1: Factors A and B are random with a and b sampled levels, respectively. Thus τ_i , β_j , $(\tau\beta)_{ij}$, and ε_{ijk} are random effects. We take a simple random sample)SRS(of τ_i from $G_\tau(0,\sigma_\tau^2)$ and a SRS of β_j from $G_\beta(0,\sigma_\beta^2)$.

Case 2: Factor A is fixed with a levels and factor B is random with b sampled levels. Thus, the τ_i are fixed effects and β_j , $(\tau\beta)_{ij}$, and ε_{ijk} are random effects. We take a SRS of β_j from $G_\beta(0,\sigma_\beta^2)$.

Both Case 1 and 2 differ from Searle and Fawcett (1970) because the random error ε is assumed to be normally distributed (and not finite).

2.2 The expected mean squares assuming a finite population

Derivation of the expected mean squares when sampling is from a finite population will differ from the expected mean squares when sampling from an infinite population. Gaylor and Hartwell (1969) assumed that the mean of each population is zero, so that the population variance can be defined as follows:

For factor A: $\sum_{i=1}^{N_a}\tau_i = 0$ and $\sigma_\tau^2 = \frac{1}{N_a-1}\sum_{i=1}^{N_a}\tau_i^2$. (2)

Consequently,

$$\left(\sum_{i=1}^{N_a}\tau_i\right)^2 = \sum_{i=1}^{N_a}\tau_i^2 + \sum_{i\neq i^*}^{N_a}\sum^{N_a}\tau_i\tau_{i^*} = 0 \text{ and}$$

$$\sum_{i=1}^{N_a}\tau_i^2 = \left(N_a-1\right)\sigma_\tau^2 .$$

Thus,

$$\sum_{i\neq i^*}^{N_a}\sum^{N_a}\tau_i\tau_{i^*} = -\left(N_a-1\right)\sigma_\tau^2 . \quad (3)$$

For factor B : $\sum_{j=1}^{N_b}\beta_j = 0$ and $\sigma_\beta^2 = \frac{1}{N_b-1}\sum_{j=1}^{N_b}\beta_j^2$. (4)

Consequently,

$$\left(\sum_{j=1}^{N_b}\beta_j\right)^2 = \sum_{j=1}^{N_b}\beta_j^2 + \sum_{j\neq j^*}^{N_b}\sum^{N_b}\beta_j\beta_{j^*} = 0 \text{ and}$$

$$\sum_{j=1}^{N_b}\beta_j^2 = \left(N_b-1\right)\sigma_\beta^2 .$$

Thus,

$$\sum_{j\neq j^*}^{N_b}\sum^{N_b}\beta_j\beta_{j^*} = -\left(N_b-1\right)\sigma_\beta^2 . \quad (5)$$

If τ_i is a randomly sampled effect then using equation (2) we get

$$E\left(\tau_i\right) = \frac{1}{N_a}\sum_{i=1}^{N_a}\tau_i = 0 \text{ and}$$

$$Var\left(\tau_i\right) = E\left(\tau_i^2\right) - \left[E\left(\tau_i\right)\right]^2 = \frac{1}{N_a}\sum_{i=1}^{N_a}\tau_i^2 = \frac{N_a-1}{N_a}\sigma_\tau^2 . \quad (6)$$

For two sampled values τ_i and τ_{i^*} by (3) we get

$$Cov\left(\tau_i,\tau_{i^*}\right) = E\left(\tau_i\tau_{i^*}\right) = \frac{1}{N_a(N_a-1)}\sum_{i\neq i^*}^{N_a}\sum^{N_a}\tau_i\tau_{i^*} = -\frac{1}{N_a}\sigma_\tau^2 . \quad (7)$$

If β_j is a randomly sampled effect then using equation (4) we get

$$E\left(\beta_j\right) = \frac{1}{N_b}\sum_{j=1}^{N_b}\beta_j = 0 \text{ and}$$

$$Var\left(\beta_j\right) = E\left(\beta_j^2\right) - \left[E\left(\beta_j\right)\right]^2 = \frac{1}{N_b}\sum_{j=1}^{N_b}\beta_j^2 = \frac{N_b-1}{N_b}\sigma_\beta^2 . \quad (8)$$

For two sampled values β_j and β_{j^*} by (5) we get

$$Cov\left(\beta_j,\beta_{j^*}\right) = E\left(\beta_j\beta_{j^*}\right) = \frac{1}{N_b(N_b-1)}\sum_{j\neq j^*}^{N_b}\sum^{N_b}\beta_j\beta_{j^*} = -\frac{1}{N_b}\sigma_\beta^2 . \quad (9)$$

Furthermore, the population of interaction effects is defined in the same way:

$\sum_{i=1}^{N_a}\sum_{j=1}^{N_b}(\tau\beta)_{ij} = 0$ and the variance is defined as

$$\sigma^2_{\tau\beta} \;=\; \frac{\sum_{i=1}^{N_a}\sum_{j=1}^{N_b}(\tau\beta)^2_{ij}}{(N_{ab}-1)}, \quad N_{ab} = N_a \times N_b . \tag{10}$$

Extensions of the procedures in equations (6) to (10) were derived which lead to

$$Cov\left[(\tau\beta)_{ij},(\tau\beta)_{i^*j^*}\right] \;=\; E\left[(\tau\beta)_{ij}(\tau\beta)_{i^*j^*}\right]$$

$$= \begin{cases} \dfrac{(N_{ab}-1)}{N_{ab}}\sigma^2_{\tau\beta} & ,i = i^* \text{ and } j = j^* \\ -\dfrac{1}{N_{ab}}\sigma^2_{\tau\beta} & \text{,otherwise} \end{cases} \tag{11}$$

where, $(\tau\beta)_{ij}$ and $(\tau\beta)_{i^*j^*}$ are two sampled values of the interaction effects.

For infinite populations, the values of (2) and (3) are σ^2_τ and 0, respectively, and the values of (4) and (5) are σ^2_β and 0, respectively. There will be changes in the expected values of the mean square in finite population models. The expected mean squares are linear functions of the variance components. The coefficients were determined for finite population models, and the expected values of mean squares will not be the same as those assuming an infinite population model.

2.3 Quadratic form for deriving the expected mean square

Suppose $\mathbf{y}$ is a vector of random variables, $\mathbf{A}$ is a symmetric matrix of real numbers, $\boldsymbol{\mu}$ is the vector of means, and $\mathbf{V}$ is the covariance matrix for the random vector $\mathbf{y}$. Each mean square can be written as a quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ in $\mathbf{y}$. From the properties of quadratic forms, the expected value is

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) \;=\; tr(\mathbf{AV}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}, \tag{12}$$

where the sums of each row in matrix **A** are zero; i.e. $\mathbf{A1} = \mathbf{0}$ and **1** is a vector of ones. Moreover, in random effects models $\boldsymbol{\mu} = \mu\mathbf{1}$, and $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{A1}\mu = \mathbf{0}$ in (12). Consequently, the expected values of mean squares can be written as:

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) \;=\; tr(\mathbf{AV}).$$

Because matrix $\mathbf{A}$ is the same in the cases of sampling from finite and infinite populations of random effects, the results can be applied to models with either finite or infinite populations of random effects. That is,

$$E_\infty(\mathbf{y}'\mathbf{A}\mathbf{y}) \;=\; tr(\mathbf{AV}_\infty) \text{ and } E_F(\mathbf{y}'\mathbf{A}\mathbf{y}) \;=\; tr(\mathbf{AV}_F), \tag{13}$$

where $\mathbf{V}_\infty$ and $\mathbf{V}_\mathrm{F}$ represent the covariance matrix when sampling from infinite and finite populations, respectively. From (13), the only difference between the expected mean squares is in choice of $\mathbf{V}_\mathrm{F}$ and $\mathbf{V}_\infty$. Thus, the difference between the corresponding forms of the covariance matrices is determined by looking at the way $\mathbf{V}_\infty$ gets altered to become $\mathbf{V}_\mathrm{F}$.

2.3.1 The matrices of the quadratic forms

For each projection matrix $\mathbf{P}$ on a component subspace, there is an appropriate decomposition to a matrix $\mathbf{C}$ such that $\mathbf{C}'\mathbf{C} = \mathbf{P}$ and $\mathbf{CC}' = \mathbf{I}$, and $\mathbf{I}$ is an identity matrix of order equal to the dimension of the projection space (Clarke, 2008). Suitable choices for $\mathbf{C}$ are as follows:

$$\mathbf{C}_\mathrm{M} = \frac{1}{\sqrt{ab}}\mathbf{1}'_{ab} \otimes \frac{1}{\sqrt{n}} \;=\; \frac{1}{\sqrt{abn}}\mathbf{1}'_{abn},$$

$$\mathbf{C}_\mathrm{A} = \frac{1}{\sqrt{b}}\left[\mathbf{I}_a \otimes \mathbf{1}'_b\right] \otimes \frac{1}{\sqrt{n}}\mathbf{1}'_n = \frac{1}{\sqrt{bn}}\left[\mathbf{I}_a \otimes \mathbf{1}'_{bn}\right],$$

$$\mathbf{C}_\mathrm{B} = \frac{1}{\sqrt{a}}\mathbf{1}'_a \otimes \mathbf{I}_b \otimes \frac{1}{\sqrt{n}}\mathbf{1}'_n = \frac{1}{\sqrt{an}}\left[\mathbf{1}'_a \otimes \mathbf{I}_b \otimes \mathbf{1}'_n\right],$$

$$\mathbf{C}_\mathrm{AB} = \mathbf{I}_a \otimes \mathbf{I}_b \otimes \frac{1}{\sqrt{n}}\mathbf{1}'_n = \frac{1}{\sqrt{n}}\left[\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{1}'_n\right],$$

where, the dimension of these matrices are $1 \times abn$ for $\mathbf{C}_\mathrm{M}$, $a \times abn$ for $\mathbf{C}_\mathrm{A}$, $b \times abn$ for $\mathbf{C}_\mathrm{B}$, $ab \times abn$ for $\mathbf{C}_\mathrm{AB}$, $\mathbf{I}_a$ and $\mathbf{I}_b$ are the identity matrices, and $\otimes$ denotes a Kronecker product.

The sum of squares (SS) in a two-factor factorial design is derived from the following quadratic forms,

The correction term $\mathbf{y}'\mathbf{C}'_\mathbf{M}\mathbf{C}_\mathbf{M}\mathbf{y}$.

The total SS

$$(\mathrm{SS_T}) \;=\; \mathbf{y}'\mathbf{C}'_\mathbf{T}\mathbf{C}_\mathbf{T}\mathbf{y} - \mathbf{y}'\mathbf{C}'_\mathbf{M}\mathbf{C}_\mathbf{M}\mathbf{y}$$

$$= \mathbf{y}'\left[\mathbf{I}_{abn} - \frac{1}{abn}\mathbf{J}_{abn}\right]\mathbf{y}.$$

Note that: $\mathbf{C}_\mathrm{T} = \mathbf{I}_\mathrm{abn}$ then $\mathbf{C}'_\mathrm{T}\mathbf{C}_\mathrm{T} = \mathbf{I}'_\mathrm{abn}\mathbf{I}_\mathrm{abn} \;=\; \mathbf{I}_\mathrm{abn}$.

The SS for A

$$(\mathrm{SS_A}) \;=\; \mathbf{y}'\mathbf{C}'_\mathbf{A}\mathbf{C}_\mathbf{A}\mathbf{y} - \mathbf{y}'\mathbf{C}'_\mathbf{M}\mathbf{C}_\mathbf{M}\mathbf{y}$$

$$= \mathbf{y}'\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right) \otimes \frac{1}{bn}\mathbf{J}_{bn}\right]\mathbf{y}.$$

The SS for B

$$(\mathrm{SS_B}) = \mathbf{y}'\mathbf{C}_\mathbf{B}'\mathbf{C}_\mathbf{B}\mathbf{y} - \mathbf{y}'\mathbf{C}_\mathbf{M}'\mathbf{C}_\mathbf{M}\mathbf{y}$$

$$= \mathbf{y}'\left[\frac{1}{a}\mathbf{J}_a \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n\right]\mathbf{y}.$$

The interaction SS $(\mathrm{SS_{AB}})$

$$= \mathbf{y}'\mathbf{C}_\mathbf{AB}'\mathbf{C}_\mathbf{AB}\mathbf{y} - \mathbf{y}'\mathbf{C}_\mathbf{A}'\mathbf{C}_\mathbf{A}\mathbf{y} - \mathbf{y}'\mathbf{C}_\mathbf{B}'\mathbf{C}_\mathbf{B}\mathbf{y} + \mathbf{y}'\mathbf{C}_\mathbf{M}'\mathbf{C}_\mathbf{M}\mathbf{y}$$

$$= \mathbf{y}'\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right) \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n\right]\mathbf{y},$$

and the error SS $(\mathrm{SS_E}) = \mathrm{SS_T} - \mathrm{SS_A} - \mathrm{SS_B} - \mathrm{SS_{AB}}$

$$= \mathbf{y}'\left[\mathbf{I}_{ab} \otimes \left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\right]\mathbf{y}.$$

From these results, we get the following matrices $\mathbf{A}_i; i = 1,2,3,4$ for the quadratic forms:

$$\mathbf{A}_1 = \left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right) \otimes \frac{1}{bn}\mathbf{J}_{bn} \text{ for } \mathrm{SS_A},$$

$$\mathbf{A}_2 = \frac{1}{a}\mathbf{J}_a \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n \text{ for } \mathrm{SS_B},$$

$$\mathbf{A}_3 = \left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right) \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n \text{ for } \mathrm{SS_{AB}},$$

and $$\mathbf{A}_4 = \mathbf{I}_{ab} \otimes \left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right) \text{ for } \mathrm{SS_E}.$$

2.3.2 Variances and covariance in the random effects model

For the random effects model, the τ_i, β_j, $(\tau\beta)_{ij}$, and ε_{ijk} are random effects sampled from infinite populations (Hocking, 1985). Then:

$$
\begin{aligned}
E(y_{ijk}) &= \mu, \\
Cov\left(y_{ijk}, y_{i^*j^*k^*}\right) &= \sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 && i = i^*, j = j^*, k = k^* \\
&= \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 && i = i^*, j = j^*, k \neq k^* \\
&= \sigma_\tau^2 && i = i^*, j \neq j^* \\
&= \sigma_\beta^2 && i \neq i^*, j = j^* \\
&= 0, && i \neq i^*, j \neq j^*.
\end{aligned}
$$

In this research, the covariance may be different from the infinite case. We take a SRS of τ_i from $G_\tau(0,\sigma_\tau^2)$ and a SRS of β_j from $G_\beta(0,\sigma_\beta^2)$. Then, $Var(\tau_i) = \left(1 - \frac{1}{N_a}\right)\sigma_\tau^2$ and $Cov(\tau_i, \tau_{i^*}) = -\frac{\sigma_\tau^2}{N_a}$, and $Var(\beta_j) = \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2$ and $Cov(\beta_j, \beta_{j^*}) = -\frac{\sigma_\beta^2}{N_b}$. Because G_τ and G_β are finite, $(\tau\beta)_{ij}$ is finite too. That is, $(\tau\beta)_{ij} \sim G_{\tau\beta}(0,\sigma_{\tau\beta}^2)$ then $Var(\tau\beta)_{ij} = \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2$ and $Cov\left[(\tau\beta)_{ij}, (\tau\beta)_{i^*j^*}\right] = -\frac{\sigma_{\tau\beta}^2}{N_{ab}}$ for all i, j, i^* and j^* with $N_{ab} = N_a \times N_b$.

Since τ_i and β_j are selected from finite populations, then the variance and covariance for Case 1 is,

$$
\begin{aligned}
Cov_F\left(y_{ijk}, y_{i^*j^*k^*}\right) &= \sigma^2 + \left(1 - \frac{1}{N_a}\right)\sigma_\tau^2 + \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i = i^*, j = j^*, k = k^* \\
&= \left(1 - \frac{1}{N_a}\right)\sigma_\tau^2 + \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i = i^*, j = j^*, k \neq k^* \\
&= \left(1 - \frac{1}{N_a}\right)\sigma_\tau^2 + \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i = i^*, j \neq j^* \\
&= \left(-\frac{1}{N_a}\right)\sigma_\tau^2 + \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i \neq i^*, j = j^* \\
&= \left(-\frac{1}{N_a}\right)\sigma_\tau^2 + \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2, && i \neq i^*, j \neq j^*.
\end{aligned}
$$

Thus, $\mathbf{V_{F_1}}$ represents $Cov_F\left(y_{ijk}, y_{i^*j^*k^*}\right)$ and has matrix form: $\mathbf{V_{F_1}} = f_1\mathbf{11}' + \mathbf{V_{F_1^*}}$, where $f_1 = \left(-\frac{1}{N_a}\right)\sigma_\tau^2 + \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2$ and $\mathbf{V_{F_1^*}}$ is formed as follows:

$$\begin{aligned}
\mathbf{V_{F_1^*}} &= \sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 &&= \mathbf{V_{11}}, \quad i = i^*, j = j^*, k = k^* \\
&= \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 &&= \mathbf{V_{12}}, \quad i = i^*, j = j^*, k \neq k^* \\
&= \sigma_\tau^2 &&= \mathbf{V_{13}}, \quad i = i^*, j \neq j^* \\
&= \sigma_\beta^2 &&= \mathbf{V_{14}}, \quad i \neq i^*, j = j^* \\
&= 0 &&= \mathbf{V_{15}}, \quad i \neq i^*, j \neq j^*.
\end{aligned} \tag{14}$$

2.3.3 Variances and covariance for the mixed effects model

For the mixed effects model, τ_i are fixed effects and β_j, $(\tau\beta)_{ij}$ and ε_{ijk} are random effects sampled from infinite populations (Hocking, 1985). Then,

$$\begin{aligned}
E(y_{ijk}) &= \mu + \tau_i, \\
Cov\left(y_{ijk}, y_{i^*j^*k^*}\right) &= \sigma^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 && i = i^*, j = j^*, k = k^* \\
&= \sigma_\beta^2 + \sigma_{\tau\beta}^2 && i = i^*, j = j^*, k \neq k^* \\
&= \sigma_\beta^2 && i \neq i^*, j = j^* \\
&= 0, && j \neq j^*.
\end{aligned}$$

The τ_i are fixed effects which correspond to $\sigma_\tau^2 = 0$. We take a SRS of β_j from $G_\beta(0, \sigma_\beta^2)$. Then, $Var\left(\beta_j\right) = \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2$ and $Cov\left(\beta_j, \beta_{j^*}\right) = -\frac{\sigma_\beta^2}{N_b}$.

The $(\tau\beta)_{ij}$ interactions are sampled from $G_{\tau\beta}(0, \sigma_{\tau\beta}^2)$. Then, $Var\,(\tau\beta)_{ij} = \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2$ and $Cov\left[(\tau\beta)_{ij}, (\tau\beta)_{i^*j^*}\right] = -\frac{\sigma_{\tau\beta}^2}{N_{ab}}$ for all i, j, i^* and j^* with $N_{ab} = a \times N_b$.

For this case, the mean is $E_F(y_{ijk}) = \mu + \tau_i$, Therefore , the variance and covariance in Case 2 is,

$$\begin{aligned}
Cov_F\left(y_{ijk}, y_{i^*j^*k^*}\right) &= \sigma^2 + \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i = i^*, j = j^*, k = k^* \\
&= \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(1 - \frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i = i^*, j = j^*, k \neq k^* \\
&= \left(1 - \frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && i \neq i^*, j = j^* \\
&= \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2 && j \neq j^*
\end{aligned}$$

Thus, $\mathbf{V_{F_2}}$ represents $Cov_F\left(y_{ijk}, y_{i^*j^*k^*}\right)$ and has the matrix form $\mathbf{V_{F_2}} = f_2\mathbf{11}' + \mathbf{V_{F_2^*}}$, where $f_2 = \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2$ and $\mathbf{V_{F_2^*}}$ is formed as follows:

$$\begin{aligned}\mathbf{V}_{\mathbf{F}_2^*} &= \sigma^2+\sigma_\beta^2+\sigma_{\tau\beta}^2 = \mathbf{V}_{21},\ i=i^*, j=j^*, k=k^* \\ &= \sigma_\beta^2+\sigma_{\tau\beta}^2 = \mathbf{V}_{22},\ i=i^*, j=j^*, k\neq k^* \\ &= \sigma_\beta^2 = \mathbf{V}_{23},\ i\neq i^*, j=j^* \\ &= 0 = \mathbf{V}_{24},\ j\neq j^*.\end{aligned} \tag{15}$$

3. Results and Discussion

In this section, we present the expected mean squares when sampling from finite distributions of effects by using the quadratic forms and variance-covariance matrices previously mentioned.

In this research, expected mean squares are derived for both balanced and unbalanced data for any finite distribution using matrix notation in linear models. We begin with the matrices which are in the quadratic forms of the two-factor factorial design: $\mathrm{SS_A} = \mathbf{y}'\mathbf{A}_1\mathbf{y}$, $\mathrm{SS_B} = \mathbf{y}'\mathbf{A}_2\mathbf{y}$, $\mathrm{SS_{AB}} = \mathbf{y}'\mathbf{A}_3\mathbf{y}$, and $\mathrm{SS_E} = \mathbf{y}'\mathbf{A}_4\mathbf{y}$ then, we compute the covariance matrices for the random and mixed effects cases. Next, we multiply $\mathbf{A}_\mathbf{i}$ by its covariance matrix in each case to find the expected sum of squares by applying the property of quadratic forms: $\mathrm{E}(\mathbf{y}'\mathbf{A}\mathbf{y}) = \mathrm{tr}(\mathbf{AV}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$. Finally, we divide the result by the degree of freedom of each factor to get the expected mean squares.

3.1 Expected mean squares for balanced data

Case 1: (The random effects model) τ_i, β_j and $(\tau\beta)_{ij}$ are random effects. We now derive the expected sum of squares for the effects in Case 1 by taking the product of a row from matrix $\mathbf{A}$ and the corresponding column in matrix $\mathbf{V}_{\mathbf{F}_1^*}$ given in (14). The product will be the same for each of the abn row and column pairs.

For example: Row 1 of the matrix $\mathbf{A}_1$ is

$$\Big[\underbrace{1-\tfrac{1}{a}\ \ 1-\tfrac{1}{a}\ \cdots\ 1-\tfrac{1}{a}}_{bn}\ \vdots\ \underbrace{-\tfrac{1}{a}\ \ -\tfrac{1}{a}\ \cdots\ -\tfrac{1}{a}}_{bn}\ \vdots \cdots \vdots\ \underbrace{-\tfrac{1}{a}\ \ -\tfrac{1}{a}\ \cdots\ -\tfrac{1}{a}}_{bn}\Big]$$

and column 1 of the variance covariance matrix is

$$\Big[\underbrace{\underbrace{\mathrm{M}_1}_{1}\ \underbrace{\mathrm{M}_2\ \cdots\ \mathrm{M}_2}_{n-1}\ \Big|\ \underbrace{\sigma_\tau^2\ \cdots\ \sigma_\tau^2}_{n(b-1)}}_{bn}\ \Big|\ \underbrace{\sigma_\beta^2\ \cdots\ \sigma_\beta^2\ 0\ \cdots\ 0}_{bn}\ \Big|\cdots\Big|\ \underbrace{\sigma_\beta^2\ \cdots\ \sigma_\beta^2\ 0\ \cdots\ 0}_{bn}\Big]',$$

where $\mathrm{M}_1 = \sigma^2+\sigma_\tau^2+\sigma_\beta^2+\sigma_{\tau\beta}^2$ and $\mathrm{M}_2 = \sigma_\tau^2+\sigma_\beta^2+\sigma_{\tau\beta}^2$.

For factor A, the expected sum of squares is

$$\begin{aligned}E_F(\mathrm{SS_A}) &= E_F(\mathbf{y}'\mathbf{A}_1\mathbf{y}) = tr\left(\mathbf{A}_1\mathbf{V}_{\mathbf{F}_1}\right) \\ &= tr\left(\left[\left(\mathbf{I}_a-\frac{1}{a}\mathbf{J}_a\right)\otimes\frac{1}{bn}\mathbf{J}_{bn}\right]\left[f_1\mathbf{11}'+\mathbf{V}_{\mathbf{F}_1^*}\right]\right);\quad \mathbf{A}_1\mathbf{1}=\mathbf{0} \\ &= abn\left(\frac{1}{abn}\right)\Big[(a-1)\left(\sigma^2+\sigma_\tau^2+\sigma_\beta^2+\sigma_{\tau\beta}^2\right)+ \\ &\quad (a-1)(n-1)\left(\sigma_\tau^2+\sigma_\beta^2+\sigma_{\tau\beta}^2\right)+(a-1)(b-1)n\sigma_\tau^2-(a-1)n\sigma_\beta^2\Big].\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_A}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_A}\right)}{\mathrm{df_A}} = \sigma^2 + bn\sigma_\tau^2 + n\sigma_{\tau\beta}^2$.

For factor B, the expected sum of squares is

$$\begin{aligned} E_{\mathrm{F}}\left(\mathrm{SS_B}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_2\mathbf{y}\right) = tr\left(\mathbf{A}_2\mathbf{V}_{\mathbf{F}_1}\right) \\ &= tr\left(\left[\frac{1}{a}\mathbf{J}_a \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n\right]\left[f_1\mathbf{1}\mathbf{1}' + \mathbf{V}_{\mathbf{F}_1^*}\right]\right);\ \mathbf{A_2 1} = \mathbf{0} \\ &= abn\left(\frac{1}{abn}\right)\Big[(b-1)\left(\sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right) + \\ &\quad (b-1)(n-1)\left(\sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right) - (b-1)n\sigma_\tau^2 + (a-1)(b-1)n\sigma_\beta^2\Big] \\ &= (b-1)\sigma^2 + an(b-1)\sigma_\beta^2 + (b-1)n\sigma_{\tau\beta}^2. \end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_B}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_B}\right)}{\mathrm{df_B}} = \sigma^2 + an\sigma_\beta^2 + n\sigma_{\tau\beta}^2$.

For the AB interaction, the expected sum of squares is

$$\begin{aligned} E_{\mathrm{F}}\left(\mathrm{SS_{AB}}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_3\mathbf{y}\right) = tr\left(\mathbf{A}_3\mathbf{V}_{\mathbf{F}_1}\right) \\ &= tr\left(\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right) \otimes \left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right) \otimes \frac{1}{n}\mathbf{J}_n\right]\left[f_1\mathbf{1}\mathbf{1}' + \mathbf{V}_{\mathbf{F}_1^*}\right]\right);\ \mathbf{A_3 1} = \mathbf{0} \\ &= abn\left(\frac{1}{abn}\right)\Big[(a-1)(b-1)\left(\sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right) + \\ &\quad (a-1)(b-1)(n-1)\left(\sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right) - (a-1)(b-1)n\sigma_\tau^2 - \\ &\quad (a-1)(b-1)n\sigma_\beta^2\Big] \\ &= (a-1)(b-1)\sigma^2 + (a-1)(b-1)n\sigma_{\tau\beta}^2. \end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_{AB}}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_{AB}}\right)}{\mathrm{df_{AB}}} = \sigma^2 + n\sigma_{\tau\beta}^2$,

and for the error, the expected sum of squares is

$$\begin{aligned} E_{\mathrm{F}}\left(\mathrm{SS_E}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_4\mathbf{y}\right) = tr\left(\mathbf{A}_4\mathbf{V}_{\mathbf{F}_1}\right) \\ &= tr\left(\left[\mathbf{I}_{ab} \otimes \left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\right]\left[f_1\mathbf{1}\mathbf{1}' + \mathbf{V}_{\mathbf{F}_1^*}\right]\right);\ \mathbf{A_4 1} = \mathbf{0} \\ &= abn\left(\frac{1}{n}\right)\Big[(n-1)\left(\sigma^2 + \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right) - (n-1)\left(\sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2\right)\Big] \\ &= ab(n-1)\sigma^2. \end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_E}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_E}\right)}{\mathrm{df_E}} = \sigma^2$.

The EMS for A, B, AB interaction and error when sampling effects from finite populations are as same as the EMS when sampling from infinite populations (Sahai & Ojeda, 2003). They differ, however, in the values of the variance components σ_τ^2, σ_β^2 and $\sigma_{\tau\beta}^2$ in each population. For Case1, the σ_τ^2, σ_β^2 and $\sigma_{\tau\beta}^2$ represent finite population variance components that include a FPC, while for the infinite case, σ_τ^2, σ_β^2 and $\sigma_{\tau\beta}^2$ represent the variance components that do not get adjusted by a FPC.

Case 2: (The mixed effects model) τ_i are fixed effects and β_j, $(\tau\beta)_{ij}$, and ε_{ijk} are random effects. The variance and covariance of this case depend on $\mathbf{V}_{\mathbf{F}_2^*}$ given in (15). From the properties of quadratic forms, Case 2 is slightly different from Case 1 for $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$ which does not equal zero because $\boldsymbol{\mu} = \mu + \boldsymbol{\tau}$.

Since $\boldsymbol{\mu}=\mathbf{0}$, then $(\boldsymbol{\mu}+\boldsymbol{\tau})'\mathbf{A}_1(\boldsymbol{\mu}+\boldsymbol{\tau}) = \boldsymbol{\tau}'\mathbf{A}_1\boldsymbol{\tau}$; $\boldsymbol{\tau}' = \begin{bmatrix}\tau_1 & \tau_2 & \ldots & \tau_a\end{bmatrix} \otimes \frac{1}{bn}\mathbf{1}'_{bn}$, so

$$\boldsymbol{\tau}'\mathbf{A}_1 = \left[bn\tau_1 - \frac{bn}{a}\sum\tau_i,\ldots,bn\tau_1 - \frac{bn}{a}\sum\tau_i \quad \ldots \quad bn\tau_a - \frac{bn}{a}\sum\tau_i,\ldots,bn\tau_a - \frac{bn}{a}\sum\tau_i\right]_{1\times abn}.$$

Assuming $\sum\tau_i = 0$, then $\boldsymbol{\tau}'\mathbf{A}_1 = bn\left[\tau_1,\ldots,\tau_1 \quad \tau_2,\ldots,\tau_2 \quad \ldots \quad \tau_a,\ldots,\tau_a\right]_{1\times abn}$ and $\boldsymbol{\tau}'\mathbf{A}_1\boldsymbol{\tau} = bn\sum_{i=1}^{a}\tau_i^2$.

For factor A, the expected sum of squares is

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_A}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A_1}\mathbf{y}\right) + \boldsymbol{\mu}'\mathbf{A_1}\boldsymbol{\mu} = tr\left(\mathbf{A_1}\mathbf{V}_{\mathbf{F}_2}\right) + \boldsymbol{\mu}'\mathbf{A_1}\boldsymbol{\mu} \\
&= tr\left(\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right)\otimes\frac{1}{bn}\mathbf{J}_{bn}\right]\left[f_2\mathbf{11}' + \mathbf{V}_{\mathbf{F}_2^*}\right]\right) + \left[\boldsymbol{\mu}+\boldsymbol{\tau}\right]'\mathbf{A}_1\left[\boldsymbol{\mu}+\boldsymbol{\tau}\right] \\
&= tr\left(\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right)\otimes\frac{1}{bn}\mathbf{J}_{bn}\right]\left[f_2\mathbf{11}' + \mathbf{V}_{\mathbf{F}_2^*}\right]\right) + \boldsymbol{\tau}'\mathbf{A}_1\boldsymbol{\tau} \\
&= (a-1)\sigma^2 + n(a-1)\sigma_{\tau\beta}^2 + bn\sum_{\mathrm{i}=1}^{\mathrm{a}}\tau_{\mathrm{i}}^2.
\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_A}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_A}\right)}{\mathrm{df_A}} = \sigma^2 + n\sigma_{\tau\beta}^2 + \dfrac{bn}{a\text{-}1}\sum_{i=1}^{a}\tau_i^2$.

For factor B, the expected sum of squares is

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_B}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_2\mathbf{y}\right) = tr\left(\mathbf{A}_2\mathbf{V}_{\mathbf{F}_2}\right) \\
&= tr\left(\left[\frac{1}{a}\mathbf{J}_a\otimes\left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right)\otimes\frac{1}{n}\mathbf{J}_n\right]\left[f_2\mathbf{11}' + \mathbf{V}_{\mathbf{F}_2^*}\right]\right), \quad \mathbf{A_2 1} = \mathbf{0} \\
&= (b-1)\sigma^2 + an(b-1)\sigma_\beta^2 + n(b-1)\sigma_{\tau\beta}^2.
\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_B}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_B}\right)}{\mathrm{df_B}} = \sigma^2 + an\sigma_\beta^2 + n\sigma_{\tau\beta}^2$.

For the AB interaction, the expected sum of squares is

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_{AB}}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_3\mathbf{y}\right) = tr\left(\mathbf{A}_3\mathbf{V}_{\mathbf{F}_2}\right) \\
&= tr\left(\left[\left(\mathbf{I}_a - \frac{1}{a}\mathbf{J}_a\right)\otimes\left(\mathbf{I}_b - \frac{1}{b}\mathbf{J}_b\right)\otimes\frac{1}{n}\mathbf{J}_n\right]\left[f_2\mathbf{11}' + \mathbf{V}_{\mathbf{F}_2^*}\right]\right); \quad \mathbf{A}_3\mathbf{1} = \mathbf{0} \\
&= (a-1)(b-1)\sigma^2 + (a-1)(b-1)n\sigma_{\tau\beta}^2.
\end{aligned}$$

Then, the expected mean square $E_{\mathrm{F}}\left(\mathrm{MS_{AB}}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_{AB}}\right)}{\mathrm{df_{AB}}} = \sigma^2 + n\sigma_{\tau\beta}^2$,

and for the error, the expected sum of squares is

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_E}\right) &= E_{\mathrm{F}}\left(\mathbf{y}'\mathbf{A}_4\mathbf{y}\right) = tr\left(\mathbf{A}_4\mathbf{V}_{\mathbf{F}_2}\right) \\
&= tr\left(\left[\mathbf{I}_{ab}\otimes\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\right]\left[f_2\mathbf{11}' + \mathbf{V}_{\mathbf{F}_2^*}\right]\right); \quad \mathbf{A_4 1} = \mathbf{0} \\
&= ab(n-1)\sigma^2.
\end{aligned}$$

Then, the expected mean square $E_{\mathrm{F}}\left(\mathrm{MS_E}\right) = \dfrac{E_{\mathrm{F}}\left(\mathrm{SS_E}\right)}{\mathrm{df_E}} = \sigma^2$.

Like in Case 1, the EMS when sampling effects from finite populations in Case 2 are same as these when sampling from infinite populations. The only difference in the inclusion of FPCs is in the σ_β^2 and $\sigma_{\tau\beta}^2$ variance components.

3.2 Expected mean squares for unbalanced data

The effects model for the unbalanced two-factor factorial design is:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad ; \begin{cases} i = 1,2,...,a \\ j = 1,2,...,b \\ k = 1,2,...,n_{ij} \end{cases}, \tag{16}$$

where the difference between equations (16) and (1) is that equation (16) has unequal replication.

The expected mean squares for unbalanced data are found by applying the method for finding expected mean squares for balanced data but modifying matrices $\mathbf{A}_p$; $p = 1,2,3,4$ to account for unequal n_{ij}. Since the variance and covariances for unbalanced data are the same as for balanced data, then the expected value of quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ is still $E(\mathbf{y}'\mathbf{A}\mathbf{y}) = tr(\mathbf{A}\mathbf{V}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

Case 1: Since τ_i and β_j are selected from finite populations, then the variance and covariance in Case 1 can be written in matrix form as $\mathbf{V}_{\mathbf{F_1}} = f_1\mathbf{1}\mathbf{1}' + \mathbf{V}_{\mathbf{F_1^*}}$ where $f_1 = \left(-\frac{1}{N_a}\right)\sigma_\tau^2 + \left(-\frac{1}{N_b}\right)\sigma_\beta^2 + \left(-\frac{1}{N_{ab}}\right)\sigma_{\tau\beta}^2$ and $\mathbf{V}_{\mathbf{F_1^*}}$ is given in (14).

We now show how to find expected mean squares of Case 1. The components that are used to determine $E_F(SS_A)$, $E_F(SS_B)$, $E_F(SS_{AB})$ and $E_F(SS_E)$ in the unbalanced data are shown in Table 1.

Table 1. The components used in determination of $E_F(SS_A)$, $E_F(SS_B)$, $E_F(SS_{AB})$, and $E_F(SS_E)$ for Case 1 for unbalanced data. $\mathbf{V}_{\mathbf{F_1^*}}(c)$ is the component of $\mathbf{V}_{\mathbf{F_1^*}}$ for Case *c*.

Case	Freq	The first row of Matrix $\mathbf{A}_p$				$\mathbf{V}_{\mathbf{F_1^*}}(c)$
		$\mathbf{A}_1$	$\mathbf{A}_2$	$\mathbf{A}_3$	$\mathbf{A}_4$	
$c=1;\ i=i^*, j=j^*, k=k^*$	1	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	$\left(1-\frac{1}{n_{11}}\right)$	$\mathbf{V}_{11}$
$c=2;\ i=i^*, j=j^*, k\neq k^*$	$n_{11}-1$	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	$\left(-\frac{1}{n_{11}}\right)$	$\mathbf{V}_{12}$
$c=3;\ i=i^*, j\neq j^*$	$n_{1\cdot}-n_{11}$	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot1}}\left(-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(-\frac{1}{b}\right)$	0	$\mathbf{V}_{13}$
$c=4;\ i\neq i^*, j=j^*$	$\sum_{i=2}^{a} n_{i1} = n_{\cdot1}-n_{11}$	$\frac{1}{n_{1\cdot}}\left(-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	0	$\mathbf{V}_{14}$
$c=5;\ i\neq i^*, j\neq j^*$	$\sum_{i=2}^{a}(n_{i\cdot}-n_{i1}) = (n_{\cdot\cdot}-n_{1\cdot})-(n_{\cdot1}-n_{11})$	$\frac{1}{n_{1\cdot}}\left(-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot1}}\left(-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(-\frac{1}{a}\right)\left(-\frac{1}{b}\right)$	0	$\mathbf{V}_{15} = 0$

Note that "*c*" is the case of variance component in $\mathbf{V}_{\mathbf{F_1^*}}$ from equation (14) where $c = 1, 2, \ldots, 5$.

In Table 1, the first column presents the possible cases of the (i, j) treatment combination for the k^{th} replication. The second column shows the frequency of each *c* which is in the first row of the matrix $\mathbf{A}_p$; $p = 1,2,3,4$. The $\mathbf{V}_{\mathbf{F_1^*}}$ components in the variance and covariance matrix are in the last column. Note that: $N = \sum_{i=1}^{a} n_{i\cdot} = \sum_{j=1}^{b} n_{\cdot j}$ and define

$C_1 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}$ and $C_2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{\cdot j}}$.

The EMS can be modified by multiplication of each column in this table as follows:

For factor A,
$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_A}\right) &= tr\left(\mathbf{A}_1\mathbf{V}_{\mathbf{F}_1^*}\right) \\
&= \sum_{c=1}^{5}\left[\left(\mathrm{Freq}_c\cdot\mathbf{A}_1\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_1^*}(c)\right)\cdot\mathrm{Dimension}\left(\mathbf{A}_1\right)\right] \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}}{n_{i\cdot}}\left[\left(1-\frac{1}{a}\right)\mathbf{V}_{11}+\left(n_{ij}-1\right)\left(1-\frac{1}{a}\right)\mathbf{V}_{12}+\left(n_{i\cdot}-n_{ij}\right)\left(1-\frac{1}{a}\right)\mathbf{V}_{13}+\left(n_{\cdot j}-n_{ij}\right)\left(-\frac{1}{a}\right)\mathbf{V}_{14}\right] \\
&= \left(a-1\right)\mathbf{V}_{11}+\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}-a\right)\left(1-\frac{1}{a}\right)\mathbf{V}_{12}+ \\
&\quad \left(\sum_{i=1}^{a}n_{i\cdot}-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}\right)\left(1-\frac{1}{a}\right)\mathbf{V}_{13}+\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}}-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}\right)\left(-\frac{1}{a}\right)\mathbf{V}_{14} \\
&= \left(a-1\right)\left[\mathbf{V}_{11}+\left(\frac{C_1}{a}-1\right)\mathbf{V}_{12}+\frac{1}{a}\left(N-C_1\right)\mathbf{V}_{13}+\frac{1}{a\left(a-1\right)}\left(C_1-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}}\right)\mathbf{V}_{14}\right].
\end{aligned}$$

Then, the expected mean square is

$$E_{\mathrm{F}}\left(\mathrm{MS_A}\right) = \mathbf{V}_{11}+\left(\frac{C_1}{a}-1\right)\mathbf{V}_{12}+\left(\frac{1}{a}\right)\left(N-C_1\right)\mathbf{V}_{13}+\frac{1}{a\left(a-1\right)}\left(C_1-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}}\right)\mathbf{V}_{14}.$$

For factor B,
$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_B}\right) &= tr\left(\mathbf{A}_2\mathbf{V}_{\mathbf{F}_1^*}\right) \\
&= \sum_{c=1}^{5}\left[\left(\mathrm{Freq}_c\cdot\mathbf{A}_2\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_1^*}(c)\right)\cdot\mathrm{Dimension}\left(\mathbf{A}_2\right)\right] \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}}{n_{\cdot j}}\left[\left(1-\frac{1}{b}\right)\mathbf{V}_{11}+\left(n_{ij}-1\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{12}+\left(n_{i\cdot}-n_{ij}\right)\left(-\frac{1}{b}\right)\mathbf{V}_{13}+\left(n_{\cdot j}-n_{ij}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{14}\right] \\
&= \left(b-1\right)\mathbf{V}_{11}+\left(C_2-b\right)\frac{\left(b-1\right)}{b}\mathbf{V}_{12}+\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{i\cdot}}{n_{\cdot j}}-C_2\right)\left(-\frac{1}{b}\right)\mathbf{V}_{13}+\left(N-C_2\right)\frac{\left(b-1\right)}{b}\mathbf{V}_{14} \\
&= \left(b-1\right)\left[\mathbf{V}_{11}+\left(\frac{C_2}{b}-1\right)\mathbf{V}_{12}+\frac{1}{b\left(b-1\right)}\left(C_2-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{i\cdot}}{n_{\cdot j}}\right)\mathbf{V}_{13}+\frac{1}{b}\left(N-C_2\right)\mathbf{V}_{14}\right].
\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS_B}\right) = \mathbf{V}_{11}+\left(\frac{C_2}{b}-1\right)\mathbf{V}_{12}+\frac{1}{b\left(b-1\right)}\left(C_2-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{\cdot j}}\right)\mathbf{V}_{13}+\left(\frac{1}{b}\right)\left(N-C_2\right)\mathbf{V}_{14}$.

For the AB interaction,
$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS_{AB}}\right) &= tr\left(\mathbf{A}_3\mathbf{V}_{\mathbf{F}_1^*}\right) \\
&= \sum_{c=1}^{5}\left[\left(\mathrm{Freq}_c\cdot\mathbf{A}_3\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_1^*}(c)\right)\cdot\mathrm{Dimension}\left(\mathbf{A}_3\right)\right] \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\left[\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{11}+\left(n_{ij}-1\right)\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{12}+\right. \\
&\quad \left.\left(n_{i\cdot}-n_{ij}\right)\left(1-\frac{1}{a}\right)\left(-\frac{1}{b}\right)\mathbf{V}_{13}+\left(n_{\cdot j}-n_{ij}\right)\left(-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{14}\right]
\end{aligned}$$

$$= (a-1)(b-1)\mathbf{V}_{11} + (a-1)(b-1)\frac{(N-ab)}{ab}\mathbf{V}_{12} - (a-1)(b-1)\frac{N}{ab}\mathbf{V}_{13} - (a-1)(b-1)\frac{N}{ab}\mathbf{V}_{14}$$

$$= (a-1)(b-1)\left[\mathbf{V}_{11} + \left(\frac{N}{ab}-1\right)\mathbf{V}_{12} - \frac{N}{ab}\mathbf{V}_{13} - \frac{N}{ab}\mathbf{V}_{14}\right].$$

Then, the expected mean square $E_{\mathrm{F}}(\mathrm{MS}_{\mathrm{AB}}) = \mathbf{V}_{11} + \left(\frac{N}{ab}-1\right)\mathbf{V}_{12} - \frac{N}{ab}\mathbf{V}_{13} - \frac{N}{ab}\mathbf{V}_{14}$,

and for the error, $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{E}}) = tr\left(\mathbf{A}_4\mathbf{V}_{\mathbf{F}_1^*}\right)$

$$= \sum_{c=1}^{5}\left[\left(\mathrm{Freq}_c \cdot \mathbf{A}_4\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_1^*}(c)\right)\cdot \mathrm{Dimension}(\mathbf{A}_4)\right]$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\left[n_{ij}\left(1-\frac{1}{n_{ij}}\right)\mathbf{V}_{11} - (n_{ij}-1)\mathbf{V}_{12}\right]$$

$$= N\mathbf{V}_{11} - ab\mathbf{V}_{11} - N\mathbf{V}_{12} + ab\mathbf{V}_{12}$$

$$= (N-ab)(\mathbf{V}_{11} - \mathbf{V}_{12}).$$

Then, the expected mean square $E_{\mathrm{F}}(\mathrm{MS}_{\mathrm{E}}) = \mathbf{V}_{11} - \mathbf{V}_{12}$.

Consequently, the EMS of the unbalance in Case 1 is summarized in Table 2.

Table 2. Expectations of mean squares when sampling a finite population for the unbalanced Case 1.

Effect	Component	EMS (Finite) for unbalanced case
τ_i	σ_τ^2	$\mathbf{V}_{11} + \left(\frac{C_1}{a}-1\right)\mathbf{V}_{12} + \left(\frac{1}{a}\right)(N-C_1)\mathbf{V}_{13} + \frac{1}{a(a-1)}\left(C_1 - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{.j}}{n_{i.}}\right)\mathbf{V}_{14}$
β_j	σ_β^2	$\mathbf{V}_{11} + \left(\frac{C_2}{b}-1\right)\mathbf{V}_{12} + \frac{1}{b(b-1)}\left(C_2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{i.}}{n_{.j}}\right)\mathbf{V}_{13} + \left(\frac{1}{b}\right)(N-C_2)\mathbf{V}_{14}$
$(\tau\beta)_{ij}$	$\sigma_{\tau\beta}^2$	$\mathbf{V}_{11} + \left(\frac{N}{ab}-1\right)\mathbf{V}_{12} - \frac{N}{ab}\mathbf{V}_{13} - \frac{N}{ab}\mathbf{V}_{14}$
ε_{ijk}	σ^2	$\mathbf{V}_{11} - \mathbf{V}_{12}$

Case 2: Since τ_i are fixed effects and β_j are selected from the finite population then the variance and covariance of this case can be written in matrix form as $\mathbf{V}_{\mathbf{F}_2} = f_2\mathbf{1}\mathbf{1}' + \mathbf{V}_{\mathbf{F}_2^*}$ where $f_2 = \left(-\frac{\sigma_\beta^2}{N_b}\right) + \left(-\frac{\sigma_{\tau\beta}^2}{N_{ab}}\right)$ and $\mathbf{V}_{\mathbf{F}_2^*}$ is given in (15).

The components that are used to determine $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{A}})$, $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{B}})$, $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{AB}})$, and $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{E}})$ in the unbalanced data are summarized in Table 3.

The method of using Table 1 for Case 1 is now applied to Table 3 for Case 2.

For factor A, $E_{\mathrm{F}}(\mathrm{SS}_{\mathrm{A}}) = tr\left(\mathbf{A}_1\mathbf{V}_{\mathbf{F}_2^*}\right) + \boldsymbol{\mu}'\mathbf{A}_1\boldsymbol{\mu}$

$$= \sum_{c=1}^{4}\left[\left(\mathrm{Freq}_c \cdot \mathbf{A}_1\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_2^*}(c)\right)\cdot \mathrm{Dimension}(\mathbf{A}_1)\right] + \boldsymbol{\mu}'\mathbf{A}_1\boldsymbol{\mu}$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}}{n_{i.}}\left[\left(1-\frac{1}{a}\right)\mathbf{V}_{21} + (n_{ij}-1)\left(1-\frac{1}{a}\right)\mathbf{V}_{22} + (n_{.j}-n_{ij})\left(-\frac{1}{a}\right)\mathbf{V}_{23}\right] + n_{i.}\sum_{i=1}^{a}\tau_i^2$$

$$= (a-1)\mathbf{V}_{21} + \left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}} - a\right)\left(1-\frac{1}{a}\right)\mathbf{V}_{22} + \left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}} - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}\right)\left(-\frac{1}{a}\right)\mathbf{V}_{23} + n_{i\cdot}\sum_{i=1}^{a}\tau_i^2$$

$$= (a-1)\left[\mathbf{V}_{21} + \left(\frac{C_1}{a}-1\right)\mathbf{V}_{22} + \frac{1}{a(a-1)}\left(C_1 - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}}\right)\mathbf{V}_{23} + \frac{n_{i\cdot}}{a-1}\sum_{i=1}^{a}\tau_i^2\right].$$

Then, the expected mean square is $E_{\mathrm{F}}(\mathrm{MS_A}) = \mathbf{V}_{21} + \left(\frac{C_1}{a}-1\right)\mathbf{V}_{22} - \frac{1}{a(a-1)}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}} - C_1\right)\mathbf{V}_{23} + \frac{n_{i\cdot}}{a-1}\sum_{i=1}^{a}\tau_i^2$.

Note that $C_1 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{i\cdot}}$.

Table 3. The components used in determination of $E_{\mathrm{F}}(\mathrm{SS_A})$, $E_{\mathrm{F}}(\mathrm{SS_B})$, $E_{\mathrm{F}}(\mathrm{SS_{AB}})$, and $E_{\mathrm{F}}(\mathrm{SS_E})$ for Case 2 with unbalanced data. $\mathbf{V}_{\mathbf{F}_2^*}(c)$ is the component of $\mathbf{V}_{\mathbf{F}_2^*}$ for Case *c*.

Case	Freq	The first row of Matrix $\mathbf{A}_p$				$\mathbf{V}_{\mathbf{F}_2^*}(c)$
		$\mathbf{A}_1$	$\mathbf{A}_2$	$\mathbf{A}_3$	$\mathbf{A}_4$	
$c=1;\ i=i^*, j=j^*, k=k^*$	1	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot 1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	$\left(1-\frac{1}{n_{11}}\right)$	$\mathbf{V}_{21}$
$c=2;\ i=i^*, j=j^*, k\neq k^*$	$n_{11}-1$	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot 1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	$\left(-\frac{1}{n_{11}}\right)$	$\mathbf{V}_{22}$
$c=3;\ i\neq i^*, j=j^*$	$n_{\cdot 1}-n_{11}$	$\frac{1}{n_{1\cdot}}\left(-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot 1}}\left(1-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)$	0	$\mathbf{V}_{23}$
$c=4;\ j\neq j^*$	$n_{1\cdot}-n_{11}$	$\frac{1}{n_{1\cdot}}\left(1-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot 1}}\left(-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(1-\frac{1}{a}\right)\left(-\frac{1}{b}\right)$	0	$\mathbf{V}_{24}=0$
$c=5;\ i\neq i^*, j\neq j^*$	$\sum_{i=2}^{a}(n_{i\cdot}-n_{i1}) = (n_{\cdot\cdot}-n_{1\cdot})-(n_{\cdot 1}-n_{11})$	$\frac{1}{n_{1\cdot}}\left(-\frac{1}{a}\right)$	$\frac{1}{n_{\cdot 1}}\left(-\frac{1}{b}\right)$	$\frac{1}{n_{11}}\left(-\frac{1}{a}\right)\left(-\frac{1}{b}\right)$	0	$\mathbf{V}_{24}=0$

Note that "*c*" is the case of variance component in $\mathbf{V}_{\mathbf{F}_2^*}$ from equation (15) where $c = 1, 2, \ldots, 4$.

For factor B, $E_{\mathrm{F}}(\mathrm{SS_B}) = tr\left(\mathbf{A}_2\mathbf{V}_{\mathbf{F}_2^*}\right)$

$$= \sum_{c=1}^{4}\left[(\mathrm{Freq}_c\cdot\mathbf{A}_2\mathrm{value})\cdot\left(\mathbf{V}_{\mathbf{F}_2^*}(c)\right)\cdot\mathrm{Dimension}(\mathbf{A}_2)\right]$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}}{n_{\cdot j}}\left[\left(1-\frac{1}{b}\right)\mathbf{V}_{21} + (n_{ij}-1)\left(1-\frac{1}{b}\right)\mathbf{V}_{22} + (n_{\cdot j}-n_{ij})\left(1-\frac{1}{b}\right)\mathbf{V}_{23}\right]$$

$$= (b-1)\mathbf{V}_{21} + \left(\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{\cdot j}} - b\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{22} + \left(\sum_{j=1}^{b}n_{\cdot j} - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{\cdot j}}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{23}$$

$$= (b-1)\left[\mathbf{V}_{21} + \left(\frac{C_2}{b}-1\right)\mathbf{V}_{22} + \left(\frac{1}{b}\right)(N-C_2)\mathbf{V}_{23}\right].$$

Then, the expected mean square is $E_{\mathrm{F}}(\mathrm{MS_B}) = \mathbf{V}_{21} + \left(\frac{C_2}{b}-1\right)\mathbf{V}_{22} + \left(\frac{1}{b}\right)(N-C_2)\mathbf{V}_{23}$.

Note that: $N = \sum_{j=1}^{b} n_{\cdot j}$ and $C_2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}^2}{n_{\cdot j}}$.

For the AB interaction,

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS}_{\mathrm{AB}}\right) &= tr\left(\mathbf{A}_3\mathbf{V}_{\mathbf{F}_2^*}\right)\\
&= \sum_{c=1}^{4}\left[\left(\mathrm{Freq}_c\cdot\mathbf{A}_3\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_2^*}(c)\right)\cdot\mathrm{Dimension}\left(\mathbf{A}_3\right)\right]\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\left[\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{21}+\left(n_{ij}-1\right)\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{22}+\left(n_{\cdot j}-n_{ij}\right)\left(-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{23}\right]\\
&= (a-1)(b-1)\mathbf{V}_{21}+\left(N-ab\right)\left(1-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{22}+\left(\sum_{i=1}^{a}\sum_{j=1}^{b}n_{\cdot j}-\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}\right)\left(-\frac{1}{a}\right)\left(1-\frac{1}{b}\right)\mathbf{V}_{23}\\
&= (a-1)(b-1)\left[\mathbf{V}_{21}+\left(\frac{N}{ab}-1\right)\mathbf{V}_{22}-\frac{N}{ab}\mathbf{V}_{23}\right].
\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS}_{\mathrm{AB}}\right) = \mathbf{V}_{21}+\left(\frac{N}{ab}-1\right)\mathbf{V}_{22}-\frac{N}{ab}\mathbf{V}_{23}$

and for the error,

$$\begin{aligned}
E_{\mathrm{F}}\left(\mathrm{SS}_{\mathrm{E}}\right) &= tr\left(\mathbf{A}_4\mathbf{V}_{\mathbf{F}_2^*}\right)\\
&= \sum_{c=1}^{4}\left[\left(\mathrm{Freq}_c\cdot\mathbf{A}_4\mathrm{value}\right)\cdot\left(\mathbf{V}_{\mathbf{F}_2^*}(c)\right)\cdot\mathrm{Dimension}\left(\mathbf{A}_4\right)\right]\\
&= \sum_{\mathrm{i=1}}^{\mathrm{a}}\sum_{\mathrm{j=1}}^{\mathrm{b}}\left[\mathrm{n_{ij}}\left(1-\frac{1}{\mathrm{n_{ij}}}\right)\mathbf{V}_{21}-\left(\mathrm{n_{ij}}-1\right)\mathbf{V}_{22}\right]\\
&= \mathrm{N}\mathbf{V}_{21}-\mathrm{ab}\mathbf{V}_{21}-\mathrm{N}\mathbf{V}_{22}+\mathrm{ab}\mathbf{V}_{22}\\
&= \left(\mathrm{N}-\mathrm{ab}\right)\left(\mathbf{V}_{21}-\mathbf{V}_{22}\right).
\end{aligned}$$

Then, the expected mean square is $E_{\mathrm{F}}\left(\mathrm{MS}_{\mathrm{E}}\right) = \mathbf{V}_{21}-\mathbf{V}_{22}$.

Consequently, the EMS of the unbalanced in Case 2 is summarized in Table 4.

Table 4. Expectations of mean squares when sampling a finite population for the unbalanced Case 2.

Effect	Component	EMS (Finite) for unbalanced case
τ_i	$\sum_{i=1}^{a}\tau_i^2 \big/ a-1$	$\mathbf{V}_{21}+\left(\frac{C_1}{a}-1\right)\mathbf{V}_{22}+\frac{1}{a(a-1)}\left(C_1-\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{n_{ij}n_{\cdot j}}{n_{i\cdot}}\right)\mathbf{V}_{23}+\frac{n_{i\cdot}}{a-1}\sum_{i=1}^{a}\tau_i^2$
β_j	σ_{β}^2	$\mathbf{V}_{21}+\left(\frac{C_2}{b}-1\right)\mathbf{V}_{22}+\left(\frac{1}{b}\right)\left(N-C_2\right)\mathbf{V}_{23}$
$(\tau\beta)_{ij}$	$\sigma_{\tau\beta}^2$	$\mathbf{V}_{21}+\left(\frac{N}{ab}-1\right)\mathbf{V}_{22}-\frac{N}{ab}\mathbf{V}_{23}$
ε_{ijk}	σ^2	$\mathbf{V}_{21}-\mathbf{V}_{22}$

The EMS when sampling from a finite population will be different from the EMS for an infinite population (Sahai & Ojeda, 2004). In particular, the multipliers of the variance components for each model effect are not the same when the random effects are sampled from a finite population.

4. Conclusions

In this research, we have determined the expected mean squares for the random effects in the random and mixed effects models in two-factor factorial model assuming finite populations for A, B, and AB interaction effects. The error, however, is assumed to be $N(0, \sigma^2)$.

For the case of balanced data, the expected mean square formulas for factors A, B, and the AB interaction when the random effects are sampled from finite populations are the same as the infinite population case. However, the primary differences are the values of the variance components. In the infinite case, σ_τ^2, σ_β^2, and $\sigma_{\tau\beta}^2$ are assumed to follow normal distributions, while in the finite case they are the variances of finite populations.

For an unbalanced case, the expected mean square of error for finite population is equal to the expected mean square for an infinite population. For the expected value of mean square in factor A, B, and the AB interaction will not be the same, because they depend on the multiplier values of the population variances. Also, for the infinite case, σ_τ^2, σ_β^2, and $\sigma_{\tau\beta}^2$ are the variances of normally distributed random variables. For the finite case, they are the variances of finite populations.

Acknowledgements

The authors are thankful for the financial support by the Department of Statistics, Faculty of Science, Kasetsart University.

References

Clarke, B. R. (2008). *Linear models: the theory and application of analysis of variance*. New Jersey, NJ: John Wiley and Sons.

Cochran, W. G. (1977). *Sampling techniques.* New York, NY: John Wiley and Sons.

Cornfield, J., & Tukey, J. W. (1956). Average values of mean squares in factorials. *The Annals of Mathematical Statistics*, *27*(4), 907-949.

Gaylor, D. W., & Hartwell, T. D. (1969). Expected mean squares for nested classifications. *Biometrics*, 427-430.

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

Hocking, R. R. (1985). *The analysis of linear models*. California, CA: Brooks/Cole Publishing.

Montgomery, D. C. (2013). *Design and analysis of experiments*. New York, NJ: John Wiley and Sons.

Oehlert, G. W. (2010). *A first course in design and analysis of experiments*. Retrieved from https://hdl.handle.net/11299/168002.

Sahai, H., & Ojeda, M. M. (2003). *Analysis of variance for random models, Volume 1: Balanced data: Theory, methods, applications, and data analysis*. Berlin, Germany: Springer Science and Business Media.

Sahai, H., & Ojeda, M. M. (2004). *Analysis of variance for random models, Volume 2: Unbalanced data: Theory, methods, applications, and data analysis*. Berlin, Germany: Springer Science and Business Media.

Searle, S. R., Casella, G., & McCulloch, C. E. (2009). *Variance components*. New Jersey, NJ: John Wiley and Sons.

Searle, S. R., & Fawcett, R. F. (1970). Expected mean squares in variance components models having finite populations. *Biometrics*, 243-254.

Searle, S. R., & Henderson, C. R. (1961). Computing procedures for estimating components of variance in the two-way classification, mixed model. *Biometrics*, *17* (4), 607-616.

Simmachan, T., Borkowski, J. J., & Budsaba, K. (2012). Expected mean squares for the random effects one-way ANOVA model when sampling from a finite population. *Thailand Statistician*, *10*(1), 121-128.

Tukey, J. W. (1956). Variances of variance components: I. Balanced designs. *The Annals of Mathematical Statistics*, *27*(3), 722-736.

Tukey, J. W. (1957). Variances of variance components: II. The unbalanced single classification. *The Annals of Mathematical Statistics*, *28*(1), 43-56.