NIPES Journal of Science and Technology Research 2(4) 2020 pp.115 - 122 pISSN-2682-5821, eISSN-2682-5821

NIPES

Journal of Science and Technology Research

Journal homepage: www.nipesjournals.org.ng

An Investigation of ANOVA Properties of Variance Components into the Variance Component Estimator of Treatment Effect of Sudoku Square Design's Models

[1]Shehu, A. and [2]Dauran, N. S.

[1]Department of Mathematical Sciences, Federal University Dutsin-Ma, Katsina State, Nigeria

[2]Department of Mathematics Usmanu Danfodiyo University, Sokoto, Nigeria

Article Info

Received 2 Oct. 2020
Revised 20 Oct. 2020
Accepted 21 Oct. 2020
Available online 31 Dec. 2020

Keywords:
Sudoku square design, Variance components, Estimator and ANOVA properties

https://doi.org/10.37933/nipes/2.4.2020.13

https://nipesjournals.org.ng

Abstract

In this research, variance component of Sudoku square design models of balanced random effect without interaction were considered. The derivation of variance component estimator of treatment effect of Sudoku square models was carried out using ANOVA method, the estimator was subjected to all the properties of ANOVA method and the result showed that the variance component estimator of treatment effect satisfies all the properties.

1. Introduction

Variance component has been an important tool in estimations of sampling design, quality control procedures or calculating heritability as well as genetic correlation. However, variance components reveal the contributions (in terms variance) of each of the effects in the models. Though, there are several methods of deriving variance components estimators [1]. The estimators frequently used have been that, from analysis of variance (ANOVA), which are obtained by equating observed and expected mean squares from analysis of variance [2, 3]. The ANOVA estimators have properties, especially, when balanced data are used (that is, if the equal number of observations in corresponding rows, or column, or rowblock or sub-square) which are unbiased and has a least variance in an array of unbiased estimators. However, for the unbalanced data which is not the case in this research, all of these properties failed with an exception of unbiasedness [4].

In estimation of variance components, random effects and mixed linear models are useful, because, these models have different sources of variability that variance components can be obtained. Unlike the fixed effect linear model that has its error term, random and has an error variance component σ^2. Variance components of unbalanced data is possible, research have it that a method which doesn't depend on balanced data and which could give rise to a unique criterion for the determination of variance components e.g Henderson method.

The advantage ANOVA technique has over other methods of estimating variance component estimators is that data need not be normal. Other usefulness of the method, the estimator still provides unbiased estimates, when compared with maximum likelihood whose data must be normally distributed.

Henderson [4] used ANOVA technique to derive the variance components for unbalanced data, he equates a series of quadratic forms to their expected values. He considered more than one factor from a mixed model, modified the method and proposed three method of estimating variance components. [5], [6, 7] and [8] studied the variance components of unbalanced data, the study was carried out under the influence of assumption that the data is normal to obtaining the estimator for sampling variances. [9], [10] and [11] gave some advantages of ANOVA method over other methods in the case of special designs. Graybill and Worthan [12] revealed ANOVA estimations of variance components of random effect linear model of balanced data and when the data is normally distributed. Searle [1] used variance components estimation of one factor random effect model to investigate the properties of ANOVA method, the estimator satisfies the properties for balanced data. Searle [1] also investigated the properties of ANOVA method for unbalanced data using ANOVA estimator, the result revealed that only unbiasedness property is satisfied and all other properties failed. Several authors had worked on Sudoku square models in many areas ranging from construction to analysis, for example [13, 14, 15]. Subramani [16] discussed the orthogonal Sudoku squares, [15] proposed a simple method of constructing Samurai Sudoku designs and orthogonal (Graeco) Samurai Sudoku design. Danbaba [17] explained combined analysis of Sudoku square design with same treatment. Shehu and Danbaba [18] discussed variance components of models of Sudoku square design of odd order.

From the literatures reviewed, none has revealed investigation on the properties of ANOVA method of variance component estimator of treatment effect of Sudoku square design models.

This study derives variance components estimator for the treatment effect only for Sudoku square models presented by [13]. The estimator is subjected to investigation under the properties of ANOVA method. This study considered the Sudoku square design model of balanced data and as well as odd order.

2. Methodology

2.1 ANOVA Method of Estimation of Variance Components

The ANOVA method of estimating variance component is simply equating the expected mean square to their observed value in the analysis of variance table and solving for the variance components [1]. Mean square of effects of the Sudoku models I-IV can be obtained as sum of squares divided by degrees of freedom. The means, variances and co-variances of the models are instrumental to the evaluation of mean squares of all random factors. This gave an indication that expectation of effects of each factor is zero, has constant variance and expectation of cross products of effects gives a zero covariance.

2.2. Properties of ANOVA Estimators

The following are considered properties of ANOVA estimators namely:

(i) Sampling Distribution
(ii) Unbiased
(iii) Minimum variance
(iv) Negative estimate
(v) Sampling variance [1,3]

2.3 Sudoku Square Design Models

Subramani and Ponnuswamy [13] constructed Sudoku square designs of order m^2, and obtained four models of different forms. These different forms were refered to as Sudoku model of Types I, II, III and IV models.

Type I Model

$$Y_{ij(k,l,p,q)} = \mu + \alpha_i + \beta_j + \tau_k + C_p + \gamma_l + s_q + e_{i,j(k,l,p,q)} \quad \begin{cases} i = 1 \cdots m \\ j = 1 \cdots m \\ k = 1 \cdots m^2 \\ l = 1 \cdots m^2 \\ p = 1 \cdots m^2 \\ q = 1 \cdots m^2 \end{cases} \qquad (1)$$

Where

μ = Grand mean
α_i = ith Row block effect
β_j = jth Column block effect
τ_k = kth treatment effect
C_p = pth Column effect
γ_l = lth Row effect
s_q = qth square effect
$e_{i,j(k,l,p,q,r)}$ are normally distributed having mean zero and variance σ^2.

Type II Model

$$Y_{ij(k,l,p,q)} = \mu + \alpha_i + \beta_j + \tau_k + \gamma(\alpha)_{l(i)} + C(\beta)_{p(j)} + s_q + e_{i,j(k,l,p,q)} \begin{cases} i = 1 \cdots m \\ j = 1 \cdots m \\ k = 1 \cdots m^2 \\ l = 1 \cdots m \\ p = 1 \cdots m \\ q = 1 \cdots m^2 \end{cases} \qquad (2)$$

Where

$\gamma(\alpha)_{l(i)}$ = lth Row effect nested in ith row block effect
$c(\beta)_{p(j)}$= pth Column effect nested in jth column block effect
s_q = qth square effect
$e_{i,j(k,l,p,q,r)}$ are normally distributed having mean zero and variance σ^2.

Type III Model

$$Y_{ij(k,l,p,q,r)} = \mu + \alpha_i + \beta_j + \tau_k + \gamma_l + c_p + s(\alpha)_{q(i)} + \pi(\beta)_{r(j)} + e_{i,j(k,l,p,q,r)}$$

$$\begin{cases} i = 1 \cdots m \\ j = 1 \cdots m \\ k = 1 \cdots m^2 \\ l = 1 \cdots m^2 \\ p = 1 \cdots m^2 \\ q = 1 \cdots m \\ r = 1 \cdots m \end{cases} \quad (3)$$

Where

$s(\alpha)_{q(i)}$ = qth horizontal square effect nested in ith Row block effect

$\pi(\beta)_{r(j)}$ = rth vertical square effect nested in the jth column block effect

s_q =qth square effect

$e_{i,j(k,l,p,q,r)}$ are normally distribute having mean zero and variance σ^2.

Type IV Model

$$Y_{ij(k,l,p,q,r)} = \mu + \alpha_i + \beta_j + \tau_k + \gamma(\alpha)_{l(i)} + c(\beta)_{p(j)} + s(\alpha)_{q(i)} + \pi(\beta)_{r(j)} + e_{i,j(k,l,p,q,r)}$$

$$\begin{cases} i = 1 \cdots m \\ j = 1 \cdots m \\ k = 1 \cdots m^2 \\ l = 1 \cdots m \\ p = 1 \cdots m \\ q = 1 \cdots m \\ r = 1 \cdots m \end{cases} \quad (4)$$

Where

$\gamma(\alpha)_{l(i)}$ = lth Row effect nested in ith row block effect

$c(\beta)_{p(j)}$= pth Column effect nested in jth column block effect

$s(\alpha)_{q(i)}$ = qth Horizontal square effect nested in ith Row block effect

$\pi(\beta)_{r(j)} = rth$ vertical square effect nested in the jth column block effect

$e_{i,j(k,l,p,q,r)}$ are normally distribute having mean zero and variance σ^2.

2.4 The sum of squares of effects in the models

$$G = \sum_{i=1}^{m^2}\sum_{j=1}^{m^2} Y_{i,j} \quad and \; N = m^4 \quad (5)$$

$$TSS = \sum_{i=1}^{m^2}\sum_{j=1}^{m^2} Y_{i,j}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (6)$$

$$SStreatmt = \sum_{k=1}^{m^2} m^2\overline{T_k}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (7)$$

$$SSrowblock = \sum_{i=1}^{m} m^3\overline{RB_\iota}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (8)$$

$$SScolblock = \sum_{j=1}^{m} m^3\overline{CB_j}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (9)$$

$$SSrow = \sum_{l=1}^{m^2} m^2\bar{R_l}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (10)$$

$$SScol = \sum_{p=1}^{m^2} m^2\overline{C_p}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (11)$$

$$SSsquare = \sum_{q=1}^{m^2} m^2\bar{s_q}^2 - m^4\,\bar{Y}_{\cdots}^2 \quad (13)$$

$$SSrowwithinRb = \sum_{l=1}^{m^2} m^2\bar{R_l}^2 - \sum_{i=1}^{m} m^3\overline{RB_\iota}^2 \quad (14)$$

$$SScolwithinCb = \sum_{p=1}^{m^2} m^2\bar{c_p}^2 - \sum_{j=1}^{m} m^3\overline{CB_j}^2 \quad (15)$$

$SShwrowblock = \sum_{i=1}^{m}\sum_{q=1}^{m} m^2 \bar{s}_{q(i)}^2 - \sum_{i=1}^{m} m^3 \overline{RB_i}^2$ (16)
$SSvwrowblock = \sum_{j=1}^{m}\sum_{r=1}^{m} m^2 \bar{\delta}_{r(j)}^2 - \sum_{j=1}^{m} m^3 \overline{cB}_j^2$ (17)
See [13]

2.5 Expected Values of Some of The Terms

In this study, the following expected values of sum of squares will be of importance in the derivation of variance components estimator for the treatment effect.

$E[e_{i,j(k,l,p,q)}] = 0$ (18)
$cov(\overline{T_k} e_{i,j(k,l,p,q)}) = cov(T_k e_{i,j(k,l,p,q)}) = 0$ (19)
$var[e_{i,j(k,l,p,q)}] = \sigma^2$ *(20)*
$var[T_k] = E[T_k{}^2] - (E[T_k])^2$ (21)
$E[T_k] = \mu$ (22)

3. Results

3.1 Variance Component Estimator of Treatment Effect

The value of $E\left[\overline{T_k}^2\right]$ is obtained by modifying Equation (21) which is given as follows

$E\left[\overline{T_k}^2\right] = var[\overline{T_k}] + (E[\overline{T_k}])^2$ (23)

$var[\overline{T_k}] = \frac{\sigma_\alpha^2}{m} + \frac{\sigma_\beta^2}{m} + \sigma_\tau^2 + \frac{\sigma_\gamma^2}{m^2} + \frac{\sigma_C^2}{m^2} + \frac{\sigma_s^2}{m^2} + \frac{\sigma^2}{m^4}$ (24)

$E[\overline{T_k}] = \mu$ (25)

Subtitute Equations (24) and (25) into (23), we have

$E\left[\overline{T_k}^2\right] = \frac{\sigma_\alpha^2}{m} + \frac{\sigma_\beta^2}{m} + \sigma_\tau^2 + \frac{\sigma_\gamma^2}{m^2} + \frac{\sigma_C^2}{m^2} + \frac{\sigma_s^2}{m^2} + \frac{\sigma^2}{m^2} + \mu^2$ (26)

$E[SStreatmt] = E\left[\sum_{k=1}^{m^2} m^2 \overline{T_k}^2 - m^4 \bar{Y}_{\cdots}^2\right]$
$= m^2 \sum_{k=1}^{m^2} E[\overline{T_k}^2] - m^4 E[\bar{Y}_{\cdots}^2]$

$E[\bar{Y}_{\cdots}^2] = [\frac{\sigma_\alpha^2}{m} + \frac{\sigma_\beta^2}{m} + \frac{\sigma_\tau^2}{m^2} + \frac{\sigma_\gamma^2}{m^2} + \frac{\sigma_C^2}{m^2} + \frac{\sigma_s^2}{m^2} + \frac{\sigma^2}{m^4} + \mu^2]$ (27)

See [18]

$= m^4\left[\frac{\sigma_\alpha^2}{m} + \frac{\sigma_\beta^2}{m} + \sigma_\tau^2 + \frac{\sigma_\gamma^2}{m^2} + \frac{\sigma_C^2}{m^2} + \frac{\sigma_s^2}{m^2} + \frac{\sigma^2}{m^2} + \mu^2\right] - [m^4[\frac{\sigma_\alpha^2}{m} + \frac{\sigma_\beta^2}{m} + \frac{\sigma_\tau^2}{m^2} + \frac{\sigma_\gamma^2}{m^2} + \frac{\sigma_C^2}{m^2} + \frac{\sigma_s^2}{m^2} + \frac{\sigma^2}{m^4}$
$+\mu^2]]$

$E[SStreatmt] = m^2(m^2-1)\sigma_\tau^2 + (m^2-1)\sigma^2$

$MSStreatmt$ is obtained by dividing $[SStreatmt]$ by degree of freedom of treatments

$E[MSStreatmt] = \frac{E[SStreatmt]}{(m^2-1)}$
$E[MSStreatmt] = m^2\sigma_\tau^2 + \sigma^2$
$\hat{\sigma}_\tau^2 = \frac{1}{m^2}(MSStreatmt - MSSerror)$ (28)

The expectations of mean sum of squares of Equations (5) – (17) and (27) are obtained and substituted into the corresponding models Equations (1) - (4) with which the respective variance components of errors for the four models are obtained.
$\hat{\sigma}^2 = MSSerror$

3.2 Properties of ANOVA Estimator

An investigation into the variance components of treatment effect of Sudoku design through the properties of ANOVA method of estimation is as follows.

(i) Sampling Distribution

With or without normality assumption, $\hat{\sigma}^2 = MSSerror$ is the only estimator that has its sampling distribution in closed form and is Chi-square (χ^2) distributed having degrees of freedom of $MSSerror$. The estimator $\hat{\sigma}_\tau^2 = \frac{1}{(m^2)}(MSStreatmet - MSSerror)$ which is made of two mean squares, each of the mean squares is independent and each is distributed as Chi-square

(ii**) Unbiasedness**

Despite the estimator $\hat{\sigma}_\tau^2 = \frac{1}{(m^2)}(MSStreatmet - MSSerror)$ not in a closed form, it can be shown that is still unbiased.

$$E(\hat{\sigma}_\tau^2) = (m^2)^{-1}E[MSStreatmet - MSSerror]$$

$$= (m^2)^{-1}(m^2)\sigma_\tau^2$$

$$= \sigma_\tau^2$$

(iii) **Minimum Variance**

ANOVA estimator from balanced data, like Sudoku design data are least variance unbiased on the assumption that error terms and other effects are normal [12]. Even without normality assumption ANOVA estimator from balanced data still has the property of minimum variance [19].

(iv) **Negative Estimate**

As shown in property (iii) $\hat{\sigma}_\tau^2 = \frac{1}{(m^2)}(MSStreatmet - MSSerror)$ has a minimum variance, unbiased, it has a greater chance of resulting to a negative result in the evaluation of variance components. Apart from $\hat{\sigma}^2 = MSSerror$ which is always positive, the estimator can result to negative estimates. If for example, $MSSerror > MSStreatmet$ definitely it yields a negative estimate of σ_τ^2. But by definition of variance, it ought not to be negative rather positive. Yet it does occur however, the chance of its occurrence is somehow large in most times.

Naturally, variance components are positive parameter and to give an explanation of a negative value creates a setback, this has been the problem of ANOVA method.

(v) **Sampling Variance**

The expression for sampling variance of treatment effects can be obtained easily especially when data used are balanced normal. This is owing to the fact that analysis of variance mean squares are not dependent of distributional proportion to Chi-squares.

The sampling variance of the treatment effect in Sudoku square Type I mode is given as follows and is similar to three other Sudoku square models.

$$var(\sigma_\tau^2) = var\left(\frac{MSStreatment - MSSerror}{m^2}\right)$$

$$= \frac{2}{m^4}\left[\frac{[E(MSStreatment)]^2}{m^2-1} + \frac{[E(MSSerror)]^2}{(m-1)[(m+1)(m^2-3)-2]}\right]$$

$$= \frac{2}{m^4}\left[\frac{\left(m^2\sigma_\tau^2 + \sigma^2\right)^2}{m^2-1} + \frac{\sigma^4}{(m-1)[(m+1)(m^2-3)-2]}\right]$$

is unbiasedly estimated by

$$= \frac{2}{m^4}\left[\frac{\left(m^2\hat{\sigma}_\tau^2 + \hat{\sigma}^2\right)^2}{m^2+1} + \frac{\hat{\sigma}^4}{(m-1)[(m+1)(m^2-3)-2]+2}\right]$$

Similarly

$$var(\hat{\sigma}^2) = var(MSSerror)$$

$$var(\hat{\sigma}^2) = \frac{2[E(MSSerror)]^2}{(m-1)[(m+1)(m^2-3)-2]}$$
$$= \frac{2\sigma^4}{(m-1)[(m+1)(m^2-3)-2]}$$

is estimated unbiasedly by $\frac{2\hat{\sigma}^4}{(m-1)[(m+1)(m^2-3)-2]+2}$

And the covariance of $\hat{\sigma}^2$ with $\hat{\sigma}_\tau^2$ is given as follows

$$Cov(\hat{\sigma}_\tau^2, \hat{\sigma}^2) = \frac{Cov[(MSStreatment-MSS\ error), MSSerror]}{m^2} = -\frac{var(MSSerror)}{m^2}$$
$$= \frac{-2\sigma^4}{m^2(m-1)[(m+1)(m^2-3)-2]}$$

For which an unbiased estimator is $\frac{-2\sigma^4}{m^2[(m-1)[(m+1)(m^2-3)-2]+2]}$

4. Conclusion

In this research, estimator of variance components of treatment effect of Sudoku square design models by Subramanian and Ponnuswamy was obtained using ANOVA method. The Estimator is subjected to investigation, using all the properties of ANOVA method of estimating variance component and thus satisfies all the properties.

Reference

[1] S.R.Searle (1988).Variance Components- Some History and A Summary Account of Estimation methods. Biometrics Unit, Cornell University, Ithaca,NY

[2] S. R. Searle (1971), Linear Models, New York: John Wiley. (1979), "Notes on Variance Component Estimation: A Detailed Account of Maximum Likelihood and Kindred Methodology," Paper No.BU-673-M, Biometrics Unit, Cornell University.

[3] S.R. Searle, G. Casella and C.E. Mcculloch (1992). *Variance components*, John Wiley, New York

[4] C.R.Henderson (1953).Estimation of Variance Components. *Biometrics 9,226-252*

[5] C. Hirotsu (1966). Estimating Variance Components in a Two-Way Layout with unequal numbers of observations. *Reports of Statistical Applications Union of Japanese Scientist and Engineer (JUSE) 13, No. 2, 29-34*

[6] W.R.Blischke (1966).Variances of estimates of variance components in a three-way classification. Biometrics 22,553-565

[7] W.R.Blischke (1968).Variances of moment estimators of variance components in the unbalanced r-way classification. *Biometrics* 24**,** 527-540.

[8] D. A. Rohde and G. M. Tallis (1969).Exact first- and second-order moments of estimates of components of covariance. *Biometrika* **56**,517-525.

[9] R.D. Anderson (1978). Studies on the estimation of variance component Ph.D Thesis, cornell University, Ithaca, New York.

[10] R. L. Anderson and P. P. Crump (1967). Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process. *Technometrics* 9**,** 499-416.

[11] T.R. Bainbridge (1963). Staggered nested design for estimating variance components. American Society for Quality Control Annual Conference Transaction 93-103.

[12] F.A. Graybill and A.M. Worthan (1956). A Note on uniformly best unbiased estimators for variance components. *Journal of the American Statistical Association 51,261-268*

[13] J.Subramani and K.N. Ponnuswamy (2009). Construction and analysis of Sudoku designs, *Model Assisted Statistics and Applications*, 4(4), 287-301.

[14] M.Hui-Dong and X Ru-Gen. (2008). Sudoku Square — a New Design in Field *Experiment, Acta Agron Sin*, 34(9), 1489–1493.

[15] A.Danbaba (2016a). Construction and Analysis of Samurai Sudoku. International *Journal Mathematical, Computational, Physical,Electrical and Computer Engineering*.**10**(4) 126-131.

[16] J. Subramani (2012), Construction of Graeco Sudoku Square Designs of Odd Orders, *Bonfring International Journal of Data Mining*, 2 (2), 37 – 41.

[17] A.Danbaba (2016b). Combined Analysis of Sudoku Square Designs with same treatments *World Academy of* Science and Technology. *International Journal of Mathematical,Computational, Physical,Electrical and Computer Engineering*.**10**(4) 155-159.
[18] A.Shehu and A.Danbaba (2018). Variance Components of Models Sudoku square design. Annals of Computer Science Series. 16th Tome 1st Fasc. – 2018, 106 -113
[19] F.A.Graybill and R.A. Hultquist (1961).Theorems concerning Eisenhart's Model II. *Annals of Mathematical Statistics* 32, 261-269.