

CONFIDENCE INTERVAL APPROXIMATION FOR TREATMENT VARIANCE IN RANDOM EFFECTS MODELS

M. E. NJA

(Received 19, December 2008; Revision Accepted 12, March 2009)

ABSTRACT

In a random effects model with a single factor, variation is partitioned into two as residual error variance and treatment variance. While a confidence interval can be imposed on the residual error variance, it is not possible to construct an exact confidence interval for the treatment variance. This is because the treatment variance is distributed as a linear combination of two chi-square variables, an expression which does not have a closed form. Approximate procedures have been provided in the literature. This paper proposes a new approximate procedure for the construction of a confidence interval for the treatment variance of a random effects model having a single factor. The new procedure uses the chi-square relationship between the residual error variance and its sample estimate.

KEYWORDS: Confidence interval, random effects model, residual error variance, treatment variance, treatment mean square, error mean square.

1.0 INTRODUCTION

The random effects model is employed in an experimental situation in which levels of a factor are randomly selected from a large number of possible levels. In this paper, we consider the random effects model with a single factor. Since the interest in this model is to test hypotheses about the treatment variance (Montgomery, 1978), σ_t^2 , the construction of approximate confidence interval for σ_t^2 is the focus of this work. The linear statistical model for the single factor random effects model is given as

$$y_{ij} = \mu + \tau_i + e_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}$$

The variance of any observation $V(y_{ij})$ is given as

$$V(y_{ij}) = V(\tau_i) + V(e_{ij}) = \sigma_\tau^2 + \sigma_e^2$$

Based on the relationship

$$\hat{\sigma}_e^2 = MSE \sim \frac{\sigma_e^2}{k_{MSE}} \chi^2(k_{MSE})$$

where k_{MSE} is the degree of freedom associated with MSE, an exact confidence interval has been constructed for σ_e^2

with $\frac{SSE}{\chi_{K_{SSE}, U}^2}$ and $\frac{SSE}{\chi_{K_{SSE}, L}^2}$ as lower and upper limits respectively (Searle, 1971).

Montgomery (1978) also derived the confidence interval for σ_e^2 , obtaining

$$\frac{(N-a)MSE}{\chi_{1-\alpha/2, N-a}^2} \text{ and } \frac{(N-a)MSE}{\chi_{\alpha/2, N-a}^2} \text{ as lower and upper limits respectively.}$$

The random variable $(N-a)MSE/\sigma_e^2$ is distributed as χ_{N-a}^2 and

$(a - 1)MSA / (\sigma_e^2 + n\sigma_\tau^2)$ is distributed as χ_{a-1}^2 . Thus the probability distribution of $\hat{\sigma}_\tau^2$ is

$$\frac{\sigma_e^2 + n\sigma_\tau^2}{N - n} \chi_{a-1}^2 - \frac{\sigma_e^2}{n(N - a)} \chi_{N-a}^2$$

This expression does not have a closed-form. Hence the exact confidence interval cannot be obtained. Williams (1962) derived an approximate confidence interval for σ_τ^2 using the expression $\sigma_e^2 F / (n\sigma_\tau^2 + \sigma_e^2) \sim F_{(a-1, N-a)}$ and the distribution of SSE / σ_e^2

Graybill (1961) and Scheffé (1959) constructed the intervals for $\sigma_e^2 / (\sigma_\tau^2 + \sigma_e^2)$ and $\sigma_\tau^2 / \sigma_e^2$ respectively.

In this paper another approximate interval for the treatment variance σ_τ^2 is proposed.

Montgomery (1978) derived an exact confidence interval for the residual error variance, σ_e^2 , as follows:

For normally and independently distributed observations, $(N - a)MSE / \sigma_e^2$ is distributed as χ_{N-a}^2 . Hence

$$P\left(\chi_{L, N-a}^2 \leq \frac{(N - a)MSE}{\sigma_e^2} \leq \chi_{U, N-a}^2\right) = 1 - \alpha$$

so that $100(1 - \alpha)\%$ confidence interval for σ_e^2 is $\frac{(N - a)MSE}{\chi_{U, N-a}^2} \leq \sigma_e^2 \leq \frac{(N - a)MSE}{\chi_{L, N-a}^2}$

$\chi_{L, N-a}^2$ is the lower critical value of the distribution with N-a degrees of freedom,

where $\chi_{U, N-a}^2$ is the upper critical value, N is the total number of observations a is the number of treatments and MSE is the error mean square.

2.0 Approximate confidence interval for σ_τ^2

Williams (1962) derived an approximate confidence interval for σ_τ^2 in which he obtained the lower and upper

limits as $\frac{SSA(1 - F_U / F)}{n\chi_{U, a-1}^2}$ and $\frac{SSA(1 - F_L / F)}{n\chi_{L, a-1}^2}$ respectively.

F_L and F_U are the lower and upper critical values of F respectively;

a = number of treatments

n = sample size

$$F_L = F_{1-\alpha}(v_1, v_2) = \frac{1}{F_\alpha(v_2, v_1)} \quad F_U = F_\alpha(v_2, v_1)$$

$\chi_{U, a-1}^2$ and $\chi_{L, a-1}^2$ are the lower and upper points of a $(1 - \alpha)\%$ region of the χ^2 distribution such that

$$pr\{\chi_{L, a-1}^2 \leq \chi_{a-1}^2 \leq \chi_{U, a-1}^2\} = 1 - \alpha, \quad \chi_{a-1, L}^2 = \chi_{\alpha/2, a-1}^2, \quad \chi_{a-1, U}^2 = \chi_{1-(\alpha/2), a-1}^2$$

$$F = MSA / MSE, \quad pr\{F_{1-\alpha}(v_1, v_2) \leq F_{[(a-1), (N-a)]} \leq F_\alpha(v_1, v_2)\}$$

MSA = treatment mean square, and MSE = error mean square.

3.0 The proposed σ_τ^2 - confidence interval

In this paper, the proposed approximation for σ_τ^2 - confidence interval is derived as follows:

$$\frac{(a - 1)MSA}{\sigma_e^2 + n\sigma_\tau^2} \sim \chi_{a-1}^2$$

The confidence limits can thus be obtained using the relationship between $(\sigma_e^2 + n\sigma_\tau^2)$ and $\frac{(a - 1)MSA}{\chi_{a-1}^2}$

Thus

$$\chi_{L,a-1}^2 \leq \frac{(a-1)MSA}{\sigma_e^2 + n\sigma_\tau^2} \leq \chi_{U,a-1}^2$$

$$\frac{(a-1)MSA}{\chi_{U,a-1}^2} \leq \sigma_e^2 + n\sigma_\tau^2 \leq \frac{(a-1)MSA}{\chi_{L,a-1}^2}$$

$$\frac{(a-1)MSA - \sigma_e^2}{n\chi_{U,a-1}^2} \leq \sigma_\tau^2 \leq \frac{(a-1)MSA - \sigma_e^2}{n\chi_{L,a-1}^2}$$

The Lower Limit

$$\frac{(a-1)S^2}{\sigma_e^2} \approx \chi^2 \quad (\text{Spiegel, 2006})$$

Hence
$$\frac{(a-1)MSA - \sigma_e^2}{n\chi_{U,a-1}^2} = \frac{(a-1)MSA - \left(\frac{n-1}{a}\hat{S}\right)^2}{n\chi_{U,a-1}^2}$$

$$= \frac{\frac{n}{a}\left(SSA - \frac{(a-1)}{a}\hat{S}^2\right)}{n\chi_{U,a-1}^2} \frac{a}{n}$$

$$= \frac{a\left[\frac{n}{a}SSA - a\frac{(a-1)\hat{S}^2}{a}\right]}{n^2\chi_{U,a-1}^2} = \frac{a\left[\frac{n}{a}SSA - SSE\right]}{n^2\chi_{U,a-1}^2}$$

(Since $(a-1)\hat{S}^2 = SSE$)

$$= \frac{nSSA - aSSE}{n^2\chi_{U,a-1}^2}$$

$$= \frac{SSA - \frac{a}{n}SSE}{n\chi_{U,a-1}^2}$$

The upper limit

Similarly the upper limit is obtained as
$$= \frac{SSA - \frac{a}{n}SSE}{n\chi_{L,a-1}^2}$$

Thus the proposed approximate confidence interval is

$$= \frac{SSA - \frac{a}{n}SSE}{n\chi_{U,a-1}^2} \leq \sigma_\tau^2 \leq \frac{SSA - \frac{a}{n}SSE}{n\chi_{L,a-1}^2}$$

This is comparable with Williams (1962) approximation given as

$$= \frac{SSA - \frac{F_U}{F}SSA}{n\chi_{U,a-1}^2} \leq \sigma_\tau^2 \leq \frac{SSA - \frac{F_L}{F}SSA}{n\chi_{L,a-1}^2}$$

$$P \left(\frac{SSA - \frac{a}{n} SSE}{n\chi_{U,a-1}^2} \leq \sigma_\tau^2 \leq \frac{SSA - \frac{a}{n} SSE}{n\chi_{L,a-1}^2} \right) = 1 - 2\alpha$$

4.0 Numerical Example

The values in an ANOVA table for a single factor random effects model provided by Montgomery (1978) are used to compare the proposed approximation with that of Williams.

The experimental situation has sixteen observations, four treatments (levels of a factor), $\alpha = 0.05$. The table is given below:

Table 1: ANOVA table for a single factor random effect model (Montgomery, 1978)

source of variation	sum of squares	degrees of freedom	mean square	F ₀
Treatment	89.19	3	29.73	15.68
Error	22.75	12	1.90	
Total	111.94	15		

This table is now used to obtain Williams' and the proposed approximate confidence intervals.

Williams approximation

Lower Limit: $\frac{SSA - \frac{F_U}{F} SSA}{n\chi_{U,a-1}^2} = 0.3159$

Upper Limit: $\frac{SSA - \frac{F_L}{F} SSA}{n\chi_{L,a-1}^2} = 15.7207$

The approximate interval is (0.3159, 15.7207)

Proposed approximation

Lower Limit: $\frac{SSA - \frac{a}{n} (SSE)}{n\chi_{U,a-1}^2} = 0.5317$

Upper Limit: $\frac{SSA - \frac{a}{n} SSE}{n\chi_{L,a-1}^2} = 11.7969$

The approximate interval is (0.5317, 11.7969) which is narrower than William's.

5.0 CONCLUSION

An alternative method for obtaining approximate confidence interval for the treatment variance in random effects models is provided. The proposed approximation does not require the lower and upper critical values of the F distribution. This renders it computationally more friendly than Williams' approximation.

REFERENCES

Graybill, F. A., 1961. An introduction to Linear Statistical Models. Vol.1 McGraw-Hill, New York.
 Montgomery. D. C., 1978. Design and Analysis of experiments. Georgia Institute of Technology, John Wiley and Sons.

Scheffé, H., 1959. The Analysis of variance. Wiley, New York.

Searle, S. R., 1971. Linear Models. John Wiley and Sons, New York.

Spiegel, M. R., Schiller, J., Srinivasan, R. A., 2006. Probability and Statistics. 2nd edition. Tata McGraw-Hill publishing company Ltd. New Delhi.

Williams, J. S., 1962. A confidence interval for variance components. *Biometrika*, 49: 278-281.