An Approximate Confidence Interval for the Variance of Random Effects of One-Way Analysis of Variance in the Completely Randomized Design

Reham Alamro¹, Abdullah A. M. Al-Shiha¹, Amal Almohisen¹

Correspondence: Department of Statistics and Operations Research, King Saud University, Saudi Arabia

Received: January 23, 2023    Accepted: July 13, 2023    Online Published: July 27, 2023

doi:10.5539/ijisp.v12n4p40    URL: https://doi.org/10.5539/ijisp.v12n4p40

Abstract

We proposed three methods to find an approximate confidence interval for the variance of the random effects for a one-way analysis of the variance model in completely randomized design. We compared the proposed methods with some other methods reported in the literature. Several criteria are used for the empirical comparisons: the mean width of the confidence interval, the variance of the width, and the coverage probability. We use Simulation and Monte-Carlo techniques to perform the comparison study. We use R language to facilitate the simulation procedures. We found that one of the proposed methods was in general superior to the others.

Keywords: random effects model, one-way analysis of variance, variance of the random effects, variance components

Introduction

Experimentation is the essential component of any research activity in all disciplines. It is therefore important for the researcher to have a general knowledge of the design and analysis of experiments. A model is an appropriate way to describe the relationship between variables in the experiment. Therefore, it is important for the researcher to distinguish between different types of experimental design models. There are three main models, which are the fixed effects model, the random effects model, and the mixed effects model. In this paper, we are interested in finding the variance of the random effects in a one-way analysis of variance in a completely randomized design.

The random effects analysis of the variance component model can be traced as far back as the works of the astronomers Airy (1861) and Chauvenet (1863). Many years later, statisticians re-invented the model beginning with Fisher (1925) who introduced the concept of analysis of variance, and Tippett (1931) who clarified the analysis of variance method of variance component estimation (as cited in Khuri and Sahai, 1985).

The one-way random effects analysis of variance (ANOVA) model can be written as:

\[ y_{ij} = \mu + \tau_i + e_{ij} \quad \{i = 1, 2, ..., k \} \quad \{j = 1, 2, ..., n \} \]

where \( y_{ij} \) is the j-th observation made at the i-th level of the factor (treatment), \( \mu \) is the overall mean, \( \tau_i \) is the unknown effect due to the i-th level, \( e_{ij} \) is a random error, \( k \) is the number of treatments, and \( n \) is the replication of each treatment. It is assumed that the errors \( e_{ij} \) are independently and identically distributed normal random variables with mean 0 and variance \( \sigma^2 \), and that the treatment effects \( \tau_i \) are independently and identically distributed normal random variables with mean 0 and variance \( \sigma^2 \tau \), and that \( e_{ij} \) and \( \tau_i \) are independent.

There are several different quantities that may be of interest in the variance component analysis. One quantity that may be of interest is the variance of random effects \( \sigma^2 \tau \), which is a measure of the variability between the population group means. Another quantity is the total variance \( \sigma^2 + \sigma^2 \). Confidence intervals are one of the most important and informative summary results in statistical applications. There is no exact confidence interval for \( \sigma^2 \) (Searle S.R, Casella G, McCulloch C.E (2006) and Montgomery, D. C, (2017)). Therefore, scientists sought to find approximate confidence intervals for \( \sigma^2 \), and we mention here some of them.

Montgomery, D.C, (2017) mentioned that Satterthwaite (1941, 1946) proposed an approximate confidence interval for \( \sigma^2 \) based on a linear combination of mean squares. In addition, he mentioned that Graybill and Wang (1980) proposed a procedure called the modified large-sample method, which can be a very useful alternative to Satterthwaite’s method. Also, Welch (1956) proposed an approximate confidence interval for \( \sigma^2 \) based on a normal approximation instead of
the Satterthwaite procedure. Bross (1950) derived a fiducial interval for \( \sigma^2. \) Tukey (1951) and others have commented adversely on this procedure because the resultant limits fail to satisfy certain boundary properties. Anderson and Bancroft (1952), while discussing some of the available procedures for confidence intervals on \( \sigma^2, \) they proposed a modified version of the Bross procedure which satisfies the boundary conditions as stated in (Sahai and Ojeda, 2004).

A method for establishing a confidence interval for \( \sigma^2 \) has been independently proposed by Tukey (1951) and Williams (1962). The Tukey-Williams method is based on two quadratic forms in normal variables, which are exactly distributed as multiples of \( \chi^2 \)-distributed random variables. A good description of deriving this Williams interval is given in Graybill (1976).

Moriguti (1954) and Bulmer (1957) independently developed the same confidence limits for \( \sigma^2. \) Boardman (1974) showed that these two methods were identical, and Bulmer (1957) showed that the method gives very accurate approximations. Scheffe (1959) gave general formulas for obtaining approximate confidence limits for variance components based on Bulmer’s method.

Sahai and Ojeda (2004) stated that Howe (1974) proposed a general procedure for constructing confidence intervals on a difference between two expected mean squares. An approximate procedure that seems to provide a shorter interval and has better coverage probability is given by Ting et al. (1990), and Burdick and Graybill (1992). Also, approximate procedures are given in Graybill (1961) and Searle (1971).


Since the variance component model was first formally introduced in the 1930’s, it has been used to model experiments in many disciplines, including Astronomy, Agriculture, Animal Breeding, Medicine, Engineering, Education, and other fields. In all these disciplines, researchers need both point and interval estimates of the variance components to take decisions or test hypotheses.

**Background**

The error mean of squares (\( M_{SE} \)) and the treatment mean of squares (\( M_{STr} \)) can be converted into chi-square variables by multiplying each one of them by corresponding degree of freedom and then dividing by the corresponding expected mean square, that is: \( \frac{(N-k)M_{SE}}{\sigma^2} \sim \chi^2_{(N-k)} \), where \( N = nk \), and \( \frac{(k-1)M_{STr}}{\sigma^2+n\sigma_t^2} \sim \chi^2_{(k-1)}. \) Therefore, it can be easily finding a confidence interval for the variance component \( \sigma^2 \), where the exact 100(1- \( \alpha \))% confidence interval for \( \sigma^2 \) is:

\[
\frac{(N-k)M_{SE}}{\chi^2_{(\alpha/2,N-k)}} \leq \sigma^2 \leq \frac{(N-k)M_{SE}}{\chi^2_{(1-(\alpha/2),N-k)}}
\]

where \( \chi^2_{(\alpha/2,N-k)} \) and \( \chi^2_{(1-(\alpha/2),N-k)} \) are the \( \frac{\alpha}{2} \)-th and \( (1-\frac{\alpha}{2}) \)-th percentiles of \( \chi^2_{(N-k)} \), respectively.

Regarding the variance component \( \sigma_t^2 \), it is known that an unbiased point estimator of it is:

\[
\hat{\sigma}_t^2 = \frac{M_{STr} - M_{SE}}{n}
\]

The distribution of \( \hat{\sigma}_t^2 \) is a linear combination of two chi-square random variables, say:

\[
\frac{\sigma^2 + n\sigma_t^2}{n(k-1)} \chi^2_{(k-1)} = \frac{\sigma^2}{n(N-k)} \chi^2_{(N-k)}
\]

Unfortunately, there is no exact formula (closed form) for the distribution of this linear combination (Sahai, H. & Ojeda, 2004).
M. M. (2004); Searle, S. R., Casella G, McCulloch, C. E. (2006) and Montgomery, D.C, (2017). Therefore, an exact confidence interval for $\sigma^2$ cannot be found. Different approximate confidence intervals have been suggested in the literature. In this paper, we suggested new approximate confidence intervals for $\sigma^2$.

**Methodology**

We compared the proposed three methods for finding approximate confidence intervals for $\sigma^2$ with some other methods reported in the literature. Several criteria are used for the empirical comparisons, which are the mean and the variance of the width of the confidence interval and the coverage probability. We use Simulation and Monte-Carlo techniques to perform the comparison study. We use R language to facilitate the simulation procedures.

**The Proposed Methods**

It is known that an unbiased estimator of the variance component $\sigma^2$ is $MSE$, and an unbiased estimator of the variance component $\sigma^2$ is $\hat{\sigma}^2 = (MST_r - MSE)/n$. Also, it is known that $\frac{(N-k)MSE}{\sigma^2}$ $\sim \chi^2_{(N-k)}$, and an exact 100 $(1-\alpha)$% confidence interval (C.I.) for $\sigma^2$ is $L < \sigma^2 < U$, where:

$$L = \frac{(N-k)MSE}{\chi^2_{(N-k)}}$$

and

$$U = \frac{(N-k)MSE}{\chi^2_{1-\alpha}(N-k)}$$

Also, it is known that $\frac{(k-1)MST_r}{\sigma^2 + n\sigma^2}$ $\sim \chi^2_{(k-1)}$, $F = \frac{MSE}{\frac{\sigma^2}{\frac{k}{n} + \frac{\hat{\sigma}^2}{\sigma^2}}}$ $\sim F_{(N-k,k-1)}$, and an exact 100 $(1-\alpha)$% confidence interval for $\sigma^2$ is:

$$\frac{\sigma^2}{n} \left[ \frac{MST_r}{MSE} F_{1-\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right] < \sigma^2 < \frac{\sigma^2}{n} \left[ \frac{MST_r}{MSE} F_{\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right]$$

Since $\sigma^2$ is unknown, we suggested the following methods for finding confidence intervals:

$$W_1: \frac{MST_r}{n} \left[ \frac{MST_r}{MSE} F_{1-\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right] < \sigma^2 < \frac{MST_r}{n} \left[ \frac{MST_r}{MSE} F_{\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right]$$

$$W_2: \frac{U}{n} \left[ \frac{MST_r}{MSE} F_{1-\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right] < \sigma^2 < \frac{L}{n} \left[ \frac{MST_r}{MSE} F_{\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right]$$

$$W_3: \frac{L}{n} \left[ \frac{MST_r}{MSE} F_{1-\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right] < \sigma^2 < \frac{U}{n} \left[ \frac{MST_r}{MSE} F_{\alpha/2}\left(\frac{N-k}{k-1}\right) - 1 \right]$$

We will investigate these three approximate confidence intervals in a coming section.

**Some Approximate Intervals Found in the Literature**

1. Williams’ confidence interval:

The following approximate interval is found in (Searle S.R., Casella G, McCulloch C.E. (2006), and Sahai, H. & Ojeda, M. M. (2004)):

$$\frac{(k-1)(MST_r - MSE) \ast F_{(k-1,N-k; \frac{\alpha}{k})}}{n\chi^2_{(k-1; \frac{\alpha}{k})}} \leq \sigma^2 \leq \frac{(k-1)(MST_r - MSE) \ast F_{(k-1,N-k; 1-\frac{\alpha}{k})}}{n\chi^2_{(k-1; 1-\frac{\alpha}{k})}}$$

2. Wald’s confidence interval:

The following approximate interval is found in (Kraemer, Kari. (2012)):
3. SAS confidence interval:

The following approximate interval is found in (Kraemer, Kari. (2012)) and used by SAS:

\[
\text{SAS: } \frac{u \hat{\sigma}_i^2}{\chi^2(u; \frac{1-\alpha}{2})} \leq \sigma_i^2 \leq \frac{u \hat{\sigma}_i^2}{\chi^2(u; \frac{1+\alpha}{2})}
\]

where:

\[
u = \frac{(MST_r - MSE)^2}{(\text{MST}_r^2 \cdot \text{MSE})^{1/2}} \cdot \frac{(k-1)}{(N-k)}
\]
Figure 1. Mean and Variance for the Six Intervals of the First Case When $K=5$ and $\alpha=0.01$ (left), 0.05 (right)

It is known that the lower of the $\alpha$, the wider of the interval, and this is what we notice in this Figure. As for the average width of the intervals, we notice that at $\alpha = 0.01$, $W_3$ and $W_1$ have the smallest width average for small values of $n$. For large values of $n$, Wald interval has the smallest width average (but we recall that the Wald interval in this case is at $\alpha$ between 0.03-0.04, so the comparison here is unfair with other methods. So, it can be said in general that $W_3$ is the best for this case. For $\alpha =0.05$, $W_3$ and $W_1$ have the smallest width average for all values of $n$. The same remarks are said for the variance.
Figure 2. Mean and Variance for the Six Intervals of the First Case When K=10 and α=0.01 (left), 0.05 (right)

Here we notice that for all $\alpha$ values, $W_1$ is the best for small values of n and then $W_3$ competes it strongly with the remaining n values. So, we notice that $W_3$ is the best in general. The same remarks are said for the variance.
Here, we notice that the best interval is $W_3$ at $\alpha = 0.01$ and at $\alpha = 0.05$, then $W_1$ comes next and then the Williams interval and Wald come last in the average width. The same applies to the variance of the width of the intervals.

Figure 3. Mean and Variance for the Six Intervals of the First Case When K=15 and $\alpha=0.01$ (left), 0.05 (right)
Here, it is good that the SAS interval enter in the comparisons. But $W_3$ is still the best, then comes the SAS interval and then $W_1$ at $\alpha = 0.01$. The case is different at $\alpha = 0.05$, where we notice that $W_1$ is the best, then $W_3$ competes with the SAS interval, then comes the Williams interval, then $W_2$ and finally the Wald interval. Also, we notice that $W_2$ approaches good intervals at large values of $n$. As for the variance, at $\alpha = 0.01$, we notice that SAS is the best except for small $n$ values. For $\alpha = 0.05$, The variances for all intervals are very similar except for Wald interval.
Figure 5. Mean and Variance for the Six Intervals of the First Case When K=25 and α=0.01 (left), 0.05 (right)

Here, we notice that as the values of k increase, the average width of the intervals approaches each other very significantly, and $W_3$ is still the best in general. For the variance, SAS interval is almost the best then $W_3$ rivals it.
Here, we notice that $W_3$ is the best then $W_1$, then SAS and Williams intervals are compete. As for variance, SAS is best for small values of $n$ and $W_3$ is best for large values of $n$.

Conclusion of the first case:

Regarding the average width of the interval, and in general, $W_3$ is the best, then $W_1$ and then the Williams interval. If the SAS interval is presented in the comparison, it competes strongly with $W_1$ but it is faulted for its excessive sensitivity when $n$ and $k$ are small. In addition, $W_2$ and Wald interval approach the other intervals if $k$ and $n$ are increased. As for the variance, it takes the same behavior as the average, but when SAS interval entering in comparison, it is the best for small and moderate $n$ and $W_3$ is the best for large $n$.

There is a problem that many researchers have had in finding SAS interval, such as Kraemer, Kari, (2012) which is the sensitivity of the SAS interval to small values of $n$ and $k$. By examining the program, we found that some values of $q_{1sas}$ and $q_{2sas}$ are equal to zero, which are values in the denominator of $sasll$ and $sasul$, which results in “NAN”. So, if there are some values of “NAN” then the variance of the length of the intervals will of course be a missing value (NA).

The Second Case

In this case we would like to see what happen if we change the variance for random effects. For this case, the standard deviation of the random effects is $\sigma_\tau = 2$, and the standard deviation of the errors is $\sigma = 1$. The graphs for this case are Figure 7 to Figure 12 which follow:
Figure 7. Mean and Variance for the Six Intervals of the Second Case When K=5 and α=0.01 (left), 0.05 (right)

In this case, we notice that $W_3$ is the best followed by $W_4$ then William’s interval for all $\alpha$ values. In addition, we exclude the Wald interval from the comparison because $\alpha$ in the left part of the graph is between 0.03-0.04 not at 0.01, but at $\alpha=0.05$, we notice that it approaches the other intervals strongly when the value of n increases. As for the variance of the width of the intervals, we find that $W_3$ is the least and therefore it is the best, followed by $W_4$. 
In general, we notice that $W_3$ is the best for all $\alpha$ values, then the comparison is difficult between $W_1$ and SAS interval, then William’s interval. The same applies to variance as well.
Figure 9. Mean and Variance for the Six Intervals of the Second Case When K=15 and α=0.01 (left), 0.05 (right) In general, $W_3$ is still the best of all $\alpha$ values for the mean and the variance.
Figure 10. Mean and Variance for the Six Intervals of the Second Case When K=20 and α=0.01 (left), 0.05 (right)

For all α values, comparison between $W_3$ and $W_1$ in terms of preference is difficult, followed by the Williams and SAS intervals. For the variance, we notice the same between $W_1$ and $W_3$, then comes the Williams and SAS intervals.
Figure 11. Mean and Variance for the Six Intervals of the Second Case When \( K = 25 \) and \( \alpha = 0.01 \) (left), 0.05 (right).

In general, \( W_3 \) is the best than \( W_1 \) for all \( \alpha \) values and for both mean and variance.
Figure 12. Mean and Variance for the Six Intervals of the Second Case When K=30 and α=0.01 (left), 0.05 (right)

In general, $W_3$ is the best for all $\alpha$ values for both mean and variance.

Conclusion of the second case:

For this case ($\sigma_\tau=2$, $\sigma=1$), we notice that $W_3$ is still the best at all $\alpha$ values, then $W_1$, and then SAS interval which competes with the Williams interval. As for the variance, it takes the same behavior as the mean. The variance of $W_3$ is the least, and $W_1$ competes with it, and then SAS interval which competes with the Williams interval.

The Third Case

In this case we would like to see what happen if we change the variance for random errors. For this case, the standard deviation of the random effects is $\sigma_\tau=1$, and the standard deviation of the errors is $\sigma=2$. The graphs for this case are Figure 13 to Figure 18 which follow:
Here, we notice that if we increase the variance of random errors, the situation has changed. For all $\alpha$ values, Wald interval is better for small values of $n$ and $W_3$ is better for large values of $n$. As for the variance, it is generally that Wald interval is the best at small $n$ values. At $\alpha = 0.01$, $W_3$ is better for large $n$ values. At $\alpha = 0.05$, it becomes very difficult to compare between $W_3$ and $W_1$ for large $n$ values.
Figure 14. Mean and Variance for the Six Intervals of the Third Case When K=10 and α=0.01 (left), 0.05 (right)

For the mean, in general and at $\alpha =0.01$, $W_1$ is the best then $W_3$. But at $\alpha =0.05$, the competition between them is strong. While for the variance, at $\alpha =0.01$, Wald interval then $W_1$ are better for small values of n, and $W_1$ competes with $W_3$ for large values of n. For $\alpha = 0.05$, $W_2$ is the best for small n values, then $W_1$, $W_3$ and William’s interval are competing at higher n values.
For the average, at $\alpha = 0.01$, $W_1$ and $W_3$ are better in general. Also, at $\alpha = 0.05$, $W_4$ is the best in general, and then $W_3$ competes with the Williams interval. As for the variance, we notice that $W_2$ is the best for only small n values for all $\alpha$ values, then $W_3$, $W_4$ and the Williams interval compete strongly.
Figure 16. Mean and Variance for the Six Intervals of the Third Case When K=20 and α=0.01 (left), 0.05 (right)

Here, we find that at $\alpha = 0.01$, $W_1$ is best for small values of $n$ and $W_3$ is the best for larger values of $n$. At $\alpha = 0.05$, $W_2$ is the best for small values of $n$, but for larger values of $n$, $W_1$ competes with $W_3$. The case is different for the variance, we find that $W_2$ is the best at small $n$ values, followed by $W_1$, $W_3$ and the Williams interval in a different way for all $\alpha$ values.
In general, for the average, we notice that $W_1$ and $W_3$ are the best. As for the variance, $W_2$ is the best for small values of $n$, and for the largest value of $n$, we notice that $W_1$ and $W_3$ are the best.
Here, we notice that the higher the value of k, the closer the average width of the intervals to each other. But, in general, we notice that $W_3$ and $W_1$ are the best. While for variance, $W_2$ remains the best for most values of n.

Conclusion of the third case:
For this case ($\sigma_x = 1, \sigma = 2$), we notice that there is a difference in the behavior of the mean width of the intervals according to the value of k. In the case of too small k, we find that the average width of the Wald interval is better for small values of n and $W_3$ is better for large values of n. For large value of k, and in general, the comparison between $W_3$ and $W_2$ is difficult in terms of preference, and in some cases $W_2$ is better only for small n values. As for the variance, we notice that for too small values of k, Wald is better for small values of n and $W_3$ and $W_1$ are better for large values of n, while for high and moderate values of k, $W_2$ is better at small values of n, and $W_1$ and $W_3$ and the Williams interval are the best at the higher value of n. Finally, for too large value of k, we notice that $W_2$ is better for most values of n.

**General Conclusion**

When comparing our proposed intervals ($W_1$, $W_2$, and $W_3$) with some of the intervals in the literature (Wald, Williams, and SAS), we noticed the following:

1. The higher the value of k, the closer the six intervals are to each other in the mean and variance of the interval's width, and they almost become equal when we increase the number of treatments to 80.

2. In general, $W_3$ is the best among all intervals, although SAS interval competes with it when entering the
comparisons, but it is weakness for not showing its values at small values of n and k, especially since most designs do not take large values of k.

3. In most cases, we found a strong competition between $W_3$ and $W_1$, then Williams competes with SAS interval, but $W_3$ maintains the lead.

4. $W_2$'s interval and Wald's interval approach other intervals at large values of n and k.

5. When we allow an increase in the variance of errors (to 4 for example) then $W_3$ is better for large values of n while Wald and $W_2$ intervals are better for small values of n.

In conclusion, we empirically found that $W_3$ is the best for most cases. This is because of calibration and adjusting the confidence level since they are all approximate methods, and they are somewhat conservative.

Reference


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).