

11.4: Computations with the Repeated-Measures ANOVA Formula

In order to review how to complete computations for repeated-measures ANOVA we will use Data Set 11.1. Suppose these data were taken from a sample of 8 participants whose confidence was measured under three conditions: before taking a class, immediately after taking a class, and again 6 months after the class. We will use these data to practice testing the hypothesis that mean confidence will be different before (Condition 1), immediately after (Condition 2), and 6 months after taking a class (Condition 3).

Data Set 11.1. Academic Confidence at Three Different Times ($n = 8$)

Participant ID	Condition 1	Condition 2	Condition 3
A	6	10	8
B	5	8	8
C	5	9	7
D	3	6	6
E	3	6	6
F	4	6	5
G	5	7	6
H	7	10	10

Preparatory Computations

Start by completing descriptive computations as a foundation. When performing computations for each condition, numeric subscripts are used to clarify the condition to which the data and results correspond. For example, the sample size for Condition 1 will be noted as n_1 .

First, find the sample size for each condition and the overall size of the data set. For repeated measures ANOVA the sample size refers to the number of raw scores rather than the number of participants. Therefore, the size of the samples for each condition are 8 and the overall size of the data set is 24 (because there are 24 raw scores total across the three conditions). Sample size is indicated with the symbol n and overall size of the data set is indicated with the symbol N . The number of conditions, known as k , should also be noted.

Second, find the means for each condition (e.g. \bar{X}_1) and the grand mean (\bar{X}_{grand}). The grand mean is the mean for all raw scores across participants and conditions together.

Third, find the standard deviations for each condition (e.g. s_1). These are not needed when computing an F -value but are useful when reporting results in APA format which is why they should be computed. Note that computing standard deviations within conditions includes computing sums of squares within each condition. Thus, it is most efficient to either reuse the first three steps of your standard deviations computations when computing sums of squares within or to compute standard deviations after computing sums of squares within.

The results of these preparatory computations are shown in the bottom of Table 11.1.

Table 11.1. Preparatory Computations ($n = 8$)

Participant ID	Condition 1	Condition 2	Condition 3
A	6	10	8
B	5	8	8
C	5	9	7
D	3	6	6
E	3	6	6

Participant ID	Condition 1	Condition 2	Condition 3
F	4	6	5
G	5	7	6
H	7	10	10

Table 11.1. Preparatory Computations ($n = 8$)

$\bar{x}_{\text{grand}} = 6.50$	$\bar{x}_1 = 4.75$	$\bar{x}_2 = 7.75$	$\bar{x}_3 = 7.00$
$N = 24$	$n_1 = 8$	$n_2 = 8$	$n_3 = 8$
$k = 3$	$s_1 = 1.39$	$s_2 = 1.75$	$s_3 = 1.60$

Sum of Squares Between

Understanding SS_b

Now that the preparatory computations are complete, we can begin our computations for the ANOVA formula, starting with the sum of squares between. In a repeated-measures ANOVA the SS_b is the sum of squares between conditions whereas in a simple, independent groups ANOVA SS_b is the sum of squares between different groups of participants. The formulas and steps used to calculate the SS_b for each of these are identical.

Despite this, however, they include different forms of error due to the differences in how data were collected in their corresponding research designs. Specifically, when the sum of squares between are calculated in repeated-measures ANOVA, they only include two sources of variation: those due to the condition and random variation and do not include error due to using different samples of participants. Because the only source of error contributing to the numerator is random error (i.e. SS_e), only this source of error needs to be addressed in the denominator when using repeated measure ANOVA. We must keep this in mind because we will need to address it when computing the denominator of the F -formula. Now that we have reviewed what SS_b includes for a repeated-measures design, we can proceed to its computations.

The sum of squares between is computed using the same formula for independent-groups ANOVA and repeated-measures ANOVA:

$$SS_b = \sum n_i \left[(\bar{x}_i - \bar{x}_{\text{grand}})^2 \right]$$

The resulting value includes treatment effects, random error, and differences between different groups of participants for independent-groups ANOVA.

The resulting value includes only treatment effects and random error for repeated-measures ANOVA.

For this reason, the denominator of the F -formula for independent-samples ANOVA accounts for random error and sample differences using SS_w and the denominator for repeated measures ANOVA only accounts for random error using SS_e .

Computing SS_b

The numerator in the repeated-measures ANOVA is computed using the sum of squares between (SS_b). The formula for computing SS_b is as follows:

$$SS_b = \sum n_i \left[(\bar{x}_i - \bar{x}_{\text{grand}})^2 \right]$$

The subscript i indicates that the computations will be carried out for each condition (i.e. subgroup of data). Therefore, this formula requires computations for each condition separately which will then be summed to get the overall SS_b . The formula can be translated into simpler language and computed for each condition as follows:

$$SS_{b_{\text{condition}}} = n_{\text{condition}} \left[(\bar{x}_{\text{condition}} - \bar{x}_{\text{grand}})^2 \right]$$

The pieces needed are as follows:

$n_{\text{condition}}$: The sample size for the condition

Note

This should be the same for all conditions because the same sample is being used in each condition

$\bar{x}_{\text{condition}}$: The mean of raw scores for a specific condition such as \bar{x}_1

\bar{x}_{grand} : The mean when all data for all conditions are treated as one, grand, group

The sample sizes, condition means, and grand mean can be found in Table 11.1 and can be entered into the formula to find the sum of squares between for each condition. Once the SS_b has been computed for each condition, they are summed together to get the overall SS_b for the test as indicated in the SS_b formula.

The subscript i stands in for the names of all conditions being tested such that the computations should each be computed for Condition 1, then Condition 2, then Condition 3, and so on until computations for all groups have been completed. Once the computations have been performed for each condition they can be summed to get the overall SS_b for the test.

Computing SS_b is the first major step of using the repeated-measures ANOVA formula and it includes several sub-steps. Let's walk through how to compute SS_b for Data Set 11.1 one sub step at a time:

1. Find the total SS_b across all groups:
 - a. Subtract the grand mean (\bar{x}_{grand}) from the mean for Condition 1 (\bar{x}_1) to get the deviation for Condition 1.
 - b. Square the deviation for Condition 1 from step 1a.
 - c. Repeat steps 1a and 1b for each condition to get the squared deviations for each condition.
 - d. Sum the squared deviations for all conditions.
 - e. Multiply the sum of squared deviation (which is the result of step 1d) by the sample size to find the sum of squares between (SS_b).

These computations for steps 1a through 1e are shown in the bottom of Table 11.2. The final result is $SS_b = 39.00$.

Table 11.1. Preparatory Computations ($n = 8$)

Participant ID	Condition 1	Condition 2	Condition 3
A	6	10	8
B	5	8	8
C	5	9	7
D	3	6	6
E	3	6	6
F	4	6	5
G	5	7	6
H	7	10	10
$\bar{x}_{\text{grand}} = 6.50$	$\bar{x}_1 = 4.75$	$\bar{x}_2 = 7.75$	$\bar{x}_3 = 7.00$
	$n_1 = 8$	$n_2 = 8$	$n_3 = 8$
$SS_b = 8[(4.75 - 6.50)^2 + (7.75 - 6.50)^2 + (7.00 - 6.50)^2]$ $SS_b = 8[3.0625 + 1.5625 + 0.25]$ $SS_b = 8[4.875]$ $SS_b = 39.00$			

Sum of Squares Within

Understanding SS_w

When the sum of squares within is computed, it contains the total estimate of pre-existing differences among participants (known as participant variance; SS_p) and unexplained differences (known as error variance; SS_e). However, the repeated-measures ANOVA requires the use of SS_e alone for its denominator (unlike simple ANOVA which uses SS_w for its denominator). However, SS_e cannot be directly computed. Therefore, SS_w must be computed and then partitioned into its two sources of variation in order to find SS_e . To say it another way, we compute the SS_w first so that we can then parse the SS_e from it which is needed for the denominator of the F -formula.

The denominator for a repeated-measures ANOVA only needs to include SS_e which is part of SS_w . Therefore, SS_w is calculated so that SS_e can then be parsed out from it in a later step.

Computing SS_w

Computing SS_w is the second major step of using the repeated-measures ANOVA formula and it includes several sub-steps. The formula for computing SS_w is as follows:

$$SS_w = \Sigma \left[\Sigma (x_i - \bar{x}_{\text{condition}})^2 \right]$$

To find SS_w , we must first find the SS_w for each condition and then sum them up. The SS within each condition formula can be written on its own as follows:

$$SS_{w_{\text{condition}}} = \Sigma (x_i - \bar{x}_{\text{condition}})^2$$

The pieces needed to use the formula are as follows:

x_i : Each raw score in the condition

$\bar{x}_{\text{condition}}$: The mean of raw scores for a specific condition such as \bar{x}_1

The raw scores and condition means can be found in Table 11.1 and can be entered into the formula to find the sum of squares within for each condition. $SS_{w_{\text{condition}}}$ is found by finding the deviation for each raw score within the condition, squaring those deviations, and then summing them. Once the SS_w has been computed for each condition, they are summed together to get the overall SS_w for the test as indicated in the SS_w formula.

Each condition SS_w can be organized using subscripts (such as SS_{w1} for the sum of squares within Condition 1 and so on). Therefore, we must find the SS_w for each condition and then add them all together to find the overall SS_w . If we have three conditions, this can be summarized as follows:

$$SS_w = SS_{w1} + SS_{w2} + SS_{w3}$$

This can be expanded to include as many conditions as needed.

Finding SS_w is the second major step in computing a repeated-measures ANOVA. Let's walk through how to compute SS_w for Data Set 11.1, one sub-step at a time:

2. Find the SS_w for each condition then sum them to get the total SS_w :
 - a. Subtract the condition mean (\bar{x}_1) from each raw score in Condition 1 to find each deviation.
 - b. Square each deviation for Condition 1.
 - c. Sum the squared deviations for Condition 1. The result of this step is the SS_{w1} .
 - d. Repeat steps 2a through 2c for each condition until the SS_w is known for all conditions.
 - e. Sum the SS_w from all conditions together to get the overall SS_w .

The computations for steps 2a through 2e are shown in Table 11.3. The final result is $SS_w = 53.00$

Table 11.3. Sum of Squares Within Computations ($n = 8$)

Participant ID	Condition 1	Condition 2	Condition 3
A	$(6 - 4.75)^2 = 1.5625$	$(10 - 7.75)^2 = 5.0625$	$(8 - 7)^2 = 1$

Participant ID	Condition 1	Condition 2	Condition 3
B	$(5 - 4.75)^2 = 0.0625$	$(8 - 7.75)^2 = 0.0625$	$(8 - 7)^2 = 1$
C	$(5 - 4.75)^2 = 0.0625$	$(9 - 7.75)^2 = 1.5625$	$(7 - 7)^2 = 0$
D	$(3 - 4.75)^2 = 3.0625$	$(6 - 7.75)^2 = 3.0625$	$(6 - 7)^2 = 1$
E	$(3 - 4.75)^2 = 3.0625$	$(6 - 7.75)^2 = 3.0625$	$(6 - 7)^2 = 1$
F	$(4 - 4.75)^2 = 0.5625$	$(6 - 7.75)^2 = 3.0625$	$(5 - 7)^2 = 4$
G	$(5 - 4.75)^2 = 0.0625$	$(7 - 7.75)^2 = 0.5625$	$(6 - 7)^2 = 1$
H	$(7 - 4.75)^2 = 5.0625$	$(10 - 7.75)^2 = 5.0625$	$(10 - 7)^2 = 9$
	$\bar{x}_1 = 4.75$	$\bar{x}_2 = 7.75$	$\bar{x}_3 = 7.00$
	$SS_{w1} = 13.50$	$SS_{w2} = 21.50$	$SS_{w3} = 18.00$
$SS_w = 13.50 + 21.50 + 18.00 + 53.00$			

Sum of Squares Participants

Understanding SS_p

Recall that some of the observed variation within conditions is due to the fact that individual participants have different tendencies. This causes there to be variation in scores within each condition. Let's consider this concept with an example using two participants from Data Set 11.1. Suppose we want to know whether confidence is different under the three different conditions. Suppose that one participant (participant H) tends to have high confidence relative to other participants (such as participant D). Person H has a confidence score that is 4 units higher than person D in all conditions. Here we see the consistency that exists within each person in regards to the dependent variable: One generally has higher scores and the other generally has lower scores. These differences are what are measured as $SS_{\text{participants}}$ (SS_p); these are not caused by the experimental conditions and are not the focus of the hypothesis nor the corresponding ANOVA. Instead, we want to know whether there are differences in the conditions beyond this kind of individual variation in confidence.

Suppose that we find that, on average, both person H and person D were least confident in Condition 1 (before the class), but had higher confidence compared to themselves in both Condition 2 (immediately after the class) and Condition 3 (6 months after the class), even though they had different confidence from each other within each condition. The pattern we are observing here is that there are systematic differences in confidence across (i.e. between) conditions that are not simply due to individual differences; the differences across conditions are attributed to the conditions (via SS_b) and the difference between participants within conditions are a form of error known as SS_p .

When we partition variance in ANOVA, we are isolating amounts of variance attributed to each source so we can focus on the amount which can be attributed to conditions without or relative to other sources (i.e. relative to random error but not including individual differences). When F is computed in a repeated-measures ANOVA, SS_p is computed so that it can be removed from the denominator. This allows us to see the group by group differences (i.e. differences between groups) relative to random differences. Thus, after we compute SS_w , we compute SS_p so that it can be removed.

Computing SS_p

Computing SS_p is the third major step of using the repeated-measures ANOVA formula. The formula for computing SS_p is as follows:

$$SS_p = \sum \frac{P^2}{k} - \frac{G^2}{N}$$

We have a few new symbols so let's start by defining each of those and how to compute them.

- P refers to participant totals. This is the sum of raw scores across conditions computed separately for each participant.
- P^2 is found by squaring each participant total separately.

- G refers to the grand total. This is the sum of all raw scores across all conditions for all participants together.
- G^2 is found by squaring the total of all raw scores (i.e. squared G -value).

Let's walk through how to compute SS_p for Data Set 11.1, one sub-step at a time:

3. Find SS_p :

- Find each participant total known as P by summing scores across conditions for each participant separately.
- Square each P .
- Divide each P^2 by k (i.e. divide P^2 by the number of conditions).
- Sum the results of step 3c
- Find G by summing all the raw scores across all conditions.
- Square G .
- Divide G^2 by N (note: N in repeated-measures ANOVA represents the number of raw scores across all participants and conditions).
- Subtract the result of step 3g from the result of step 3d to get SS_p .

The computations for steps 3a through 3h are shown in the bottom of Table 11.4. The final result is $SS_p = 48.00$.

Table 11.4. Sum of Squares Participants Computations ($n = 8$)

Participant ID	Condition 1	Condition 2	Condition 3	P (Participant Totals)	P^2	$\frac{P^2}{k}$
A	6	10	8	24	576	$576 \div 3 = 192$
B	5	8	8	21	441	$441 \div 3 = 147$
C	5	9	7	21	441	$441 \div 3 = 147$
D	3	6	6	15	225	$225 \div 3 = 75$
E	3	6	6	15	225	$225 \div 3 = 75$
F	4	6	5	15	225	$225 \div 3 = 75$
G	5	7	6	18	324	$324 \div 3 = 108$
H	7	10	10	27	729	$729 \div 3 = 243$
Column Totals	38	62	56			$\Sigma \frac{P^2}{k} = 1,062$
$G = 38 + 62 + 56 = 156$ $G^2 = 24,336$ This is the total of all raw scores. This is the squared total of all raw scores.						
$k = 3$ $N = 24$ This is the number of conditions. This is the number of raw scores across conditions.						
$SS_p = \Sigma \frac{P^2}{k} - \frac{G^2}{N}$ $SS_p = 1,062 - \frac{24,336}{24}$ $SS_p = 1,062 - 1,014$ $SS_p = 48.00$						

Sum of Squares Error

Understanding and Computing SS_e

SS_e refers to the random error that occurs which is not attributed to differences between participants. It is the focus of the denominator of the F -formula but is not computed directly. Instead, SS_e is one of two sources of error that make up the SS_w and must be parsed from it; SS_w is the total of SS_e and SS_p . Thus, the sum of squared deviations within the conditions (SS_w) minus

the sum of squares between participants (SS_p) yields the sum of squares error (SS_e). This represents the otherwise unaccounted for error and is used for the denominator for the repeated measures ANOVA formula. The formula for SS_e is as follows:

$$SS_e = SS_w - SS_p$$

Thus, finding SS_e is simple once both SS_w and SS_p have been calculated. Let's walk through how to compute SS_e for Data Set 11.1:

4. Find SS_e by subtracting SS_p from SS_w . We found SS_p and SS_w in prior steps so we can now simply plug them in and find SS_e as follows:

$$\begin{aligned} SS_e &= 53.00 - 48.00 \\ SS_e &= 5.00 \end{aligned}$$

The final result is $SS_e = 5.00$.

Degrees of Freedom

Degrees of Freedom Between (df_b)

df_b is the degrees of freedom between the conditions; it is the adjusted k -value. Recall that k refers to the number of conditions. The formula is as follows:

$$df_b = k - 1$$

df_b is always the number of conditions minus 1. For example, if three conditions were being compared, the df_b would be 2 but if four conditions were being compared the df_b would be 3, and so on. Let's walk through how to compute df_b for Data Set 11.1:

5. Find the df_b by subtracting 1 from the number of groups (k).

$$\begin{aligned} df_b &= k - 1 \\ df_b &= 3 - 1 \\ df_b &= 2 \end{aligned}$$

Degrees of Freedom Error (df_e)

Degrees of freedom for the error is equal to the adjusted k -value multiplied by the adjusted sample size. Let's walk through how to compute df_e for Data Set 11.1:

6. Find the df_e
 - a. Subtract 1 from the number of groups (k)
 - b. Subtract 1 from the sample size (i.e. the number of participants)
 - c. Multiplying the two resulting values to get df_e

$$\begin{aligned} df_e &= (k - 1)(n - 1) \\ df_e &= (3 - 1)(8 - 1) \\ df_e &= (2)(7) \\ df_e &= 14 \end{aligned}$$

The final result is $df_e = 14$.

Putting the Formula Together

Once the four components are calculated, their results are put into the ANOVA formula and used to solve for F .

$$F = \frac{MSS_b}{MSS_e}$$

The numerator asks for the mean sum of squares between conditions (MSS_b). The denominator asks for the mean sum of squares error (MSS_e). Calculating these requires dividing the respective SS by its df , thus, the formula can be rewritten as follows:

$$F = \frac{SS_b \div df_b}{SS_e \div df_e}$$

We computed the four necessary components in prior steps and can now proceed to plugging them in and calculating F . Let's walk through these remaining steps to compute F for Data Set 11.1:

7. Write the ANOVA formula with the four values found in the above steps (i. e. SS_b , df_b , SS_e , and df_e) plugged into their respective locations like so:

$$F = \frac{SS_b \div df_b}{SS_e \div df_e}$$

$$F = \frac{39.00 \div 2}{5.00 \div 14}$$

$$F = \frac{19.50}{0.3571 \dots}$$

$$F = 54.60$$

Repeated Measures ANOVA Computations Summary

The goal of repeated-measures ANOVA is to assess the ratio of treatment effects to that of error variance. Treatment effects (i.e. variability attributed to conditions) are included in SS_b . Error variance is measured as SS_e . These two sources of variance are the focus of the repeated measures ANOVA formula:

$$F = \frac{MSS_b}{MSS_e} = \frac{SS_b \div df_b}{SS_e \div df_e}$$

Numerator Calculations

SS_b is found using the formula:

$$SS_b = \sum n_i \left[(\bar{x}_i - \bar{x}_{\text{grand}})^2 \right]$$

1. Use the formula to find the SS_b for each condition.
2. Sum the SS_b s for all conditions to get overall SS_b .

df_b is found using the formula: $k - 1$

Denominator Calculations

The sum of squares error (SS_e) is calculated indirectly by removing the sum of squares attributed to pre-existing differences among participants (SS_p) from the sum of squares within (SS_w) using this formula:

$$SS_e = SS_w - SS_p$$

Therefore, both SS_w and SS_p must be computed before finding SS_e .

SS_w is found using the formula:

$$SS_w = \sum \left[\sum (x_i - \bar{x}_{\text{condition}})^2 \right]$$

1. Use the formula to find the SS_w for each condition.
2. Sum the SS_w s for all conditions to get overall SS_w .

SS_p is found using the formula:

$$SS_p = \sum \frac{P^2}{k} - \frac{G^2}{N}$$

- P refers to participant totals. This is the sum of raw scores across conditions computed separately for each participant.
- G refers to the grand total. This is the sum of all raw scores across all conditions for all participants together.

Then SS_e is found using the formula:

$$SS_e = SS_w - SS_p$$

Next, df_e should be calculated using the formula: $(N - k)(n - 1)$

F Calculations

MSS_b is found using the formula:

$$MSS_b = SS_b \div df_b$$

MSS_e is found using the formula:

$$MSS_e = SS_e \div df_e$$

F is found using the formula:

$$F = \frac{MSS_b}{MSS_e}$$

Reading Review 11.3

1. What is a grand mean and how is it calculated?
2. What is P in a repeated-measures ANOVA and how is it calculated?
3. What two sources of variance together make up the SS between for repeated-measures ANOVA?
4. What two sources of variance together make up the SS within for repeated-measures ANOVA?
5. What is being calculated and represented by the numerator of the repeated-measures ANOVA formula?
6. What is being calculated and represented by the denominator of the repeated-measures ANOVA formula?

This page titled [11.4: Computations with the Repeated-Measures ANOVA Formula](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by .